UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA HIDRÁULICA E AMBIENTAL
DOUTORADO EM ENGENHARIA CIVIL
SANEAMENTO AMBIENTAL

ALLAN CLEMENTE DE SOUZA

GESTÃO PROATIVA DE CIANOTOXINAS: AVALIAÇÃO DA BIODIVERSIDADE DE CIANOBACTÉRIAS NOS RESERVATÓRIOS DESTINADOS AO ABASTECIMENTO PÚBLICO E PROPOSTA DE SELEÇÃO DE BIOINDICADORES

FORTALEZA
2024

ALLAN CLEMENTE DE SOUZA

GESTÃO PROATIVA DE CIANOTOXINAS: AVALIAÇÃO DA BIODIVERSIDADE DE CIANOBACTÉRIAS NOS RESERVATÓRIOS DESTINADOS AO ABASTECIMENTO PÚBLICO E PROPOSTA DE SELEÇÃO DE BIOINDICADORES

Tese submetida à Coordenação do curso de Pós-graduação em Engenharia Civil, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Doutor em Engenharia Civil. Área de Concentração: Saneamento Ambiental.

Orientador: Prof. Dr. José Capelo Neto.
Coorientador: Prof. Dr. Allan Amorim Santos.

FORTALEZA

2024

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

S713g Souza, Allan Clemete de.
Gestão proativa de cianotoxinas : avaliação da biodiversidade de cianobactérias nos reservatórios destinados ao abastecimento público e proposta de seleção de bioindicadores / Allan Clemete de Souza. – 2025.
860 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia Civil: Saneamento Ambiental, Fortaleza, 2025.
Orientação: Prof. Dr. José Capelo Neto.
Coorientação: Prof. Dr. Allan Amorim dos Santos.

1. Cianotoxinas. 2. Aprendizagem de Máquina. 3. Decisão baseada em dados. 4. Modelagem estatística . 5. Qualidade de água. I. Título.

CDD 628

ALLAN CLEMENTE DE SOUZA

GESTÃO PROATIVA DE CIANOTOXINAS: AVALIAÇÃO DA BIODIVERSIDADE DE CIANOBACTÉRIAS NOS RESERVATÓRIOS DESTINADOS AO ABASTECIMENTO PÚBLICO E PROPOSTA DE SELEÇÃO DE BIOINDICADORES

Tese submetida à Coordenação do curso de Pós-graduação em Engenharia Civil, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Doutor em Engenharia Civil.

Aprovada em: __/__/___

BANCA EXAMINADORA

____

Prof. Dr. José Capelo Neto (Orientador)
Universidade Federal do Ceará (UFC)

____

Prof. Dr. Allan Amorim dos Santos (Coorientador)
Universidade Federal do Rio de Janeiro (UFRJ)

____

Prof.a Dra. Ana Bárbara de Araújo Nunes
Universidade Federal do Ceará (UFC)

____

Prof. Dr. José Carlos Alves Barroso Júnior
Universidade Federal do Ceará (UFC)

____

Dr. Mario Ubirajara
Companhia de Gestão dos Recursos Hídricos do Ceará (COGERH)

____

Prof. Dr. Ismael Carloto
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

A Deus.
Aos meus pais, José de Souza Neto e Vandir Clemente de Souza e meus irmãos Thiago, Vanessa, Raquel, Davi e Tomás.

AGRADECIMENTOS

Ao CNPq (Conselho Nacional de Pesquisa), pelo apoio financeiro com a manutenção da bolsa de auxílio. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Ao Prof. Dr. José Capelo Neto, meu Orientador, pela sua atenciosa orientação na realização deste trabalho, bem como pela sua presteza e prontidão.

Ao Prof. Dr. Allan Amorim dos Santos, pela excelente coorientação e direcionamento na elaboração deste trabalho.

Aos professores participantes da banca examinadora, Ana Bárbara de Araújo Nunes, José Carlos Alves Barroso Júnior, Ismael Carloto e ao Dr. Mario Ubirajara, pelo tempo pelas valiosas colaborações e sugestões.

A todos os professores que tive o privilégio de ter ao longo do doutorado.

A Shirley e a Neusa por toda ajuda nos assuntos administrativos ao longo de todo o curso.

Aos amigos e colegas da CAGECE pelas ajudas na obtenção dos dados de pesquisa.

Em memória do meu pai, José de Souza Neto, pelos ensinamentos, amor e dedicação, que me inspiraram a sonhar em me tornar engenheiro.

Em especial à minha mãe, Vandir Clemente de Souza, exemplo de garra, força, coragem e fé em Deus, por me ensinar a lutar e nunca desistir dos meus objetivos, mesmo diante das inúmeras dificuldades.

Eia, agora, vós que dizeis: Hoje ou amanhã, iremos a tal cidade, e lá passaremos um ano, e contrataremos, e ganharemos. Digo-vos que não sabeis o que acontecerá amanhã. Porque, que é a vossa vida? É um vapor que aparece por um pouco e depois se desvanece.

Tiago *4:13–14*

RESUMO

No semiárido cearense, a presença ou a elevada quantidade de cianobactérias potencialmente tóxicas tornam as informações geradas pelo monitoramento de rotina das companhias de saneamento insuficientes para uma gestão antecipada de eventos críticos de toxinas. Diante dessa limitação, o presente estudo teve como objetivo testar a hipótese de que o uso de bioindicadores, baseados nas chances de risco de toxinas, é mais adequado que o monitoramento tradicional para subsidiar a tomada de decisão antecipada dos gestores diante do risco de elevação de toxinas na água bruta. Esta abordagem, até o momento, não foi amplamente explorada na literatura. Para testar essa hipótese, foi realizada uma revisão sistematizada de 1549 artigos que utilizaram dados de monitoramento regular, com uma estrutura similar à das companhias de saneamento, a fim de identificar práticas que possam complementar o monitoramento de rotina. Além disso, foi avaliada a performance de modelos lineares generalizados na modelagem de dados biológicos em experimentos de campo, que, assim como o monitoramento de rotina, apresentam limitações como poucas observações, alta dispersão e heterocedasticidade. Também foi realizada uma investigação sobre a biodiversidade dos 82 açudes de abastecimento humano, com o objetivo de compreender suas particularidades e selecionar bioindicadores de alerta de toxinas com maior precisão. Foi observado que o processo de homogeneização biótica, presente em todos os reservatórios estudados, favoreceu a dominância de cianobactérias potencialmente tóxicas. Entretanto, apesar das densidades celulares elevadas, foi observada uma baixa frequência de eventos críticos de toxinas. Isso possibilitou a identificação de bioindicadores relacionados tanto ao aumento quanto à redução das chances de eventos críticos em apenas metade dos reservatórios onde tais eventos ocorreram. Portanto, utilizando a infraestrutura já existente nas companhias de saneamento, foi possível desenvolver uma ferramenta complementar de gestão que permite uma transição de um monitoramento reativo para uma abordagem mais proativa.

Palavras-chave: Cianotoxinas. Cilindrospermopsina. Microcistina. Saxitoxina. Regressão logística. Odds Ratio. Aprendizagem de Máquina. Bioindicadores. Decisão baseada em dados.

ABSTRACT

In the semi-arid region of Ceará, the abundance of potentially toxic cyanobacteria makes the information generated by routine monitoring from sanitation companies insufficient for proactive management of critical toxin events. Given this limitation, the present study aimed to test the hypothesis that the use of bioindicators, based on toxin risk likelihood, is more appropriate than traditional monitoring to support early decision-making by managers in response to the risk of increasing toxin levels in raw water. This approach has not been widely explored in the literature to date. To test this hypothesis, a systematic review of 1,549 articles was conducted, analyzing regular monitoring data structured similarly to that of sanitation companies, in order to ident practices that could complement routine monitoring. Additionally, the performance of generalized linear models was evaluated for modeling biological data in field experiments, which, like routine monitoring, have limitations such as few observations, high dispersion, and heteroscedasticity. A biodiversity study of 82 water supply reservoirs was also conducted to understand their specific characteristics and select toxin alert bioindicators more accurately. It was observed that the process of biotic homogenization, present in all studied reservoirs, favored the dominance of potentially toxic cyanobacteria. However, despite high cellular densities, a low frequency of critical toxin events was observed. This allowed for the identification of bioindicators associated with both an increase and a decrease in the likelihood of critical events in only half of the reservoirs where such events occurred. Therefore, by utilizing the existing infrastructure of sanitation companies, it was possible to develop a complementary management tool that allows for a shift from reactive monitoring to a more proactive approach.

Keywords: Cyanotoxins. Cylindrospermopsin. Microcystin. Saxitoxin. Logistic Regression. Odds Ratio. Machine Learning. Bioindicators. Data-driven Decision-Making.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

Acc	Accuracy
AdaBoost	Adaptive Boosting
Adjust-R^2	Adjusted -R^2
AE	Absolute error
AICc	Akaike Information Criterion with a correction
ALGAE	Information of abundance or composition other than cell density or biovolume
ANN	Artificial neural networks
ANOVA	Analysis of variance
ARIMA	autoregressive integrated moving average
AS	Self-Organizing Map
ATX-a	Anatoxin-A
BEFORE	Cylindrospermopsin
BN	Beveridge-Nelson Decomposition
BiLSTM	bidirectional Long Short-Term Memory
BIO	Biovolume of phytoplankton
BCyano	Cyanobacteria biovolume
BM	Bayesian network models or others Bayesian Models
Bonferroni	Bonferroni test
CA	Correspondence Analysis
Ca	Calcium
CCA	Canonical Correlation Analysis
CD	Cell density
Chl-a	Chlorophyll-a
Chlorides	Chlorides
CLM	Critical Load Model OCDE
CNN	Convolutional Neural Networks
COAMING	Silicate
COD	Chemical oxygen demand
CyanoHab	Cyanobacterial harmful algal blooms
DCA	Detrended Correspondence Analysis
DIN/TP	Dissolved inorganic nitrogen and total phosphorus ratio
DIP	Dissolved inorganic phosphorus
DNN	Deep Neural Network
DOC	Dissolved Organic Carbon
DRIED	Secchi Depth
DS	Descriptive statistics
DTN	Total dissolved nitrogen
DTP	Total Dissolved Phosphorus
Dunn	Dunn’s Test
EBDI	Entropy-Based Dynamic Imputation
EC	Electrical conductivity

EE	Elliptical Envelope
Entropy	Entropy
F1	F1 Score
Fe	Iron
FRIEND	Spectral form Algorithms - Aphanizomenon-Microcystis Index
FROM	Alkalinity
FROM	Dissolved inorganic nitrogen
FUZZY	Fuzzy theory
GA	Hybrid Genetic Algorithm
GAM	Generalized additive models
GC	Gini coefficient
GEO	Geosmin
GERMAN	Decision Trees
HCO3	Bicarbonate
HI	Information Collected from Hyperspectral Imaging
HIS	Sensitivity analysis
HRT	Hydraulic detention time
IN	Affluent flow
KN	Kjeldahl Nitrogen
KW	Kruskal-Walli's test
LDA	Linear Discriminant Analysis
LEVEL	Reservoir water level
LR	Lenar regression
LSTM	Long Short-Term Memory
MAE	mean absolute error
MK	Mann-Kendall test
MAPE	Mean Absolute Percentage Error
MC	Microcystin
MCR-ALS	Multivariate curve resolution–alternating least squares
MEM	Mixed effects models
MIB	2-Methylisoborneol
MNB	Normalized Bias
MRM	Multivariate regression MODELS
MSE	Mean Squared Error
neo-STX	Saxitoxin
NH4	Ammoniacal nitrogen
NMDS	Non-metric multidimensional scaling
NO2	Nitrite
NO3	Nitrate
NOD	Nodularin
OF	Dynamic Imputation
OF	Dissolved oxygen
OUT	Effluent flow
PC	Phycocyanin

Pearson	Pearson’s correlation
Pettitt	Pettitt method
pH	pH
p-DNNR	Deep neural network regression based on pixels
PLSR	Partial least square regression
PO_4	Phosphate
PREC	Precipitation
pseudo-R^2	Pseudo-R^2
PSO	Particle Swarm Optimization
Q	Flow
Quantile	Quantile regression
r	Correlation coefficient
R^2	Determination coefficient
RDA	Redundancy Analysis
RE	Relative error
RF	Random Forest
RL-OLS	Linear regression - Ordinary Least Squares
RMSE	Root Mean Square Error
RP	Relative peak error
RPD	Residual Prediction Deviation
RU	Relative humidity
RV	Relative volume error
SEA	mean absolute relative error
SR	Remote sensing data from satellites or not
SMOTE	Synthetic minority oversampling technique
SO4	Sulfate
Spearman	Spearman correlation
SR	Solar radiation
SRP	soluble reactive phosphorus
SS	Suspended solids
ST	Seasonal Kendall Test
STX	Saxitoxin
SVC	Support vector machine
SVR	Support Vector Regression
TDN	Total dissolved nitrogen

SUMÁRIO

1 CAPÍTULO 1: INTRODUÇÃO

1.1 Contextualização e justificativa do trabalho

Mesmo com os avanços tecnológicos no tratamento de água, as cianobactérias continuam sendo um dos principais problemas enfrentados pelas empresas de saneamento em todo o mundo (DEVI *et al.*, 2021; KAKADE *et al.*, 2021; MAITY *et al.*, 2021; SERRÀ *et al.*, 2021; YANG *et al.*, 2021). Diante dessa situação, legislações e normas internacionais recomendam que as companhias de saneamento realizem um monitoramento regular da qualidade da água bruta em fontes destinadas ao consumo humano. Esse monitoramento inclui a coleta de amostras e análises laboratoriais periódicas (com frequência semanal, mensal, semestral, etc.) para verificar a presença desses organismos e de substâncias dentro dos limites seguros para a saúde pública (BRAZIL, 2021; WHO, 2022).

A maioria das companhias de saneamento no mundo utiliza técnicas tradicionais de microscopia e análises laboratoriais, o que limita tanto a frequência das análises quanto o número de pontos de monitoramento. Esses procedimentos geram séries de dados de muitos anos, mas com poucos registros. Além disso, essas técnicas demandam muito tempo de análise, recursos financeiros e profissionais qualificados, especialmente quando comparadas a outras abordagens, como fluorescência in situ e sensoriamento remoto (ROUSSO *et al.*, 2020).

Paralelamente, no que se refere à qualidade da água, explorar novas metodologias de análise não é uma tarefa fácil para pesquisadores e, principalmente, para gestores e tomadores de decisão do setor de saneamento. Isso ocorre porque, além de gerarem dados a partir de técnicas que exigem rotinas laboriosas, esses profissionais precisam monitorar e cumprir uma extensa lista de padrões de qualidade estabelecidos pela legislação de cada país. Assim, embora cumpram as exigências legais e gerem muitos dados, a quantidade de informação útil para a tomada de decisão é, na maioria das vezes, limitada.

Quando se trata de gerar informações, a qualidade dos dados é fundamental. Isso exige uma atenção especial, que vai desde a calibração e precisão dos sinais de sensores e sondas (OLSSON, 2012), até a implementação de um banco de dados robusto e um sistema de gerenciamento de informações eficiente e confiável (NOGUEIRA VILANOVA; FILHO; PERRELLA BALESTIERI, 2014). Além disso, para transformar dados em informações úteis, são necessárias habilidades em programação, estatística e conhecimento de algoritmos de aprendizado de máquina. No entanto, essas competências ainda são pouco difundidas no contexto das companhias de saneamento, tornando essa falta de aptidão um desafio adicional a ser superado em um curto espaço de tempo (HUANG *et al.*, 2021).

Diante desse cenário, o processo de transformação de dados em informações pode ser dividido em dois grupos principais: modelos baseados em processos (MBP) e modelos baseados em dados (MBD). Os MBP, também conhecidos como modelos mecanicistas ou de simulação, descrevem o comportamento do sistema por meio de hipóteses sobre as relações físicas, químicas ou biológicas entre variáveis. Para gerar informações úteis, esses modelos requerem uma série de parâmetros específicos para caracterizar o sistema, que geralmente são obtidos experimentalmente e muitas vezes não fazem parte da rotina de monitoramento existente. Embora os MBP sejam valiosos para prever novos dados e avaliar a causalidade entre variáveis, eles exigem um esforço significativo e tempo considerável para seu desenvolvimento (URRUTIAGUER, 2016).

Considerando as práticas de monitoramento das empresas de saneamento, a implementação dos MBP pode ser particularmente desafiadora. Isso ocorre devido à necessidade de conhecimento especializado em diversas áreas e à grande quantidade de parâmetros que precisam ser calibrados. Esses fatores aumentam a complexidade e os custos desse tipo de modelagem, principalmente pela necessidade de medir novos parâmetros, o que pode resultar em modelos complexos, com possíveis incertezas e erros nos resultados. Além disso, os MBP exigem um conhecimento prévio aprofundado do sistema, enquanto os MBD se destacam por serem mais acessíveis, amplamente utilizados e melhor adaptados à realidade operacional das companhias de saneamento (ROUSSO *et al.*, 2020).

Os MBD, por sua vez, estabelecem relações entre variáveis sem a necessidade de um conhecimento prévio do comportamento do sistema, partindo do princípio de que a informação necessária para compreender o fenômeno está contida nos dados, descrita pelas relações entre as variáveis. Esses modelos necessitam de um número suficiente de observações para retratar com precisão a condição do sistema, e incluem tanto modelos estatísticos quanto técnicas de aprendizado de máquina (LEE *et al.*, 2023). No entanto, os MBD são sensíveis à quantidade de dados disponíveis, especialmente aqueles que utilizam aprendizado de máquina, devido ao problema de sobreajuste (*overfitting*), que pode comprometer a precisão e a generalização dos resultados (AHMED *et al.*, 2024).

Assim, na realidade das companhias de saneamento, mesmo sendo menos complexos do que os MBP, o processo de tomada de decisão baseado em dados utilizando MBD não é uma tarefa trivial. Por exemplo, considerando que a coleta de dados ao longo do tempo é o método mais tradicional de monitoramento, é comum observar que o valor de uma observação em um dado instante apresente alta correlação com resultados anteriores, fenômeno conhecido como autocorrelação. Isso significa que dados com essas características violam a suposição de erros independentes, um pressuposto fundamental para métodos de regressão baseados em mínimos quadrados, amplamente utilizados para previsões (JAMES *et al.*, 2014; SHEATHER, 2009). Além disso, é importante considerar problemas que dados biológicos podem apresentar, como a heterocedasticidade, ou seja, a variabilidade não constante, e a assimetria, que podem afetar tanto a inferência quanto as previsões em modelos de regressão (ABBOTT; GUTGESELL, 1994). Como consequência, a violação da homogeneidade de variância pode resultar em aumentos de erros de falso positivo (erro Tipo I) (KNIEF; FORSTMEIER, 2021).

Outro desafio para a implementação dos MBD no contexto das companhias de saneamento, além da baixa frequência e das características dos dados gerados no monitoramento (i.e., autocorrelação, heterocedasticidade, assimetria, entre outros), são os problemas relacionados às cianobactérias e cianotoxinas no contexto do semiárido. A biodiversidade característica da região revela a presença quase constante de elevadas densidades de cianobactérias potencialmente tóxicas que dominam os açudes da região (BARROS *et al.*, 2017a; BRASIL *et al.*, 2016; CAMPOS *et al.*, 2024; FONSECA *et al.*, 2015; MARENGO ORSINI *et al.*, 2018; MENDES *et al.*, 2022; ROCHA *et al.*, 2024). Mesmo assim, os eventos críticos de toxinas relatados na água bruta em séries temporais são pontuais e estudos que avaliem séries longas são escassos. Diante desse contexto, o uso de técnicas de aprendizado de máquina e de estatísticas tradicionais torna-se inviável, especialmente para predições. Esse cenário torna o desafio de mudar o paradigma tradicional de tomada de decisão reativa particularmente complexo, principalmente para as empresas do semiárido que utilizam técnicas de monitoramento convencionais de baixa resolução.

Portanto, é essencial dispor de ferramentas que complementem o monitoramento das companhias de saneamento para alcançar uma gestão proativa. A simples presença e quantidade de cianobactérias nos açudes não oferecem informações suficientes para que os gestores tomem ações preventivas contra eventos críticos de toxinas, sem recorrer a técnicas analíticas mais complexas, que geralmente não são parte da rotina do monitoramento tradicional. Assim, torna-se necessária uma investigação detalhada de técnicas alternativas de MBD, acompanhada pela validação dessas metodologias em condições semelhantes às observadas no monitoramento regular. Além disso, uma análise aprofundada da biodiversidade dos 82 açudes destinados ao consumo humano ao longo de uma década é fundamental. Com essas informações, a seleção de organismos associados a eventos críticos, cuja presença pode aumentar ou diminuir as chances de elevação de toxinas na água bruta, poderá ser mais precisa, permitindo que os bioindicadores de cada açude ofereçam um monitoramento mais eficaz, direcionado e antecipado, respeitando as particularidades do ecossistema de cada manancial.

1.2 Estrutura do trabalho: perguntas, hipóteses e objetivos

O presente trabalho foi organizado em 5 capítulos. O primeiro capítulo apresenta os aspectos gerais, incluindo a justificativa, as hipóteses e os objetivos dos capítulos subsequentes. Do capítulo 2 ao capítulo 4, adotou-se uma abordagem sistemática, na qual as hipóteses, formuladas a partir das perguntas iniciais, foram testadas. Dessa forma, os objetivos de cada capítulo foram claramente definidos. Cada um desses capítulos foi estruturado no formato de artigo científico, contendo introdução, materiais e métodos, resultados, discussão, conclusões e sugestões para trabalhos futuros. Esses artigos serão submetidos para publicação, razão pela qual foram redigidos em inglês, conforme as exigências dos periódicos. Nos capítulos subsequentes, as informações dos capítulos 2 a 4 são detalhadas. No capítulo 5, apresenta-se uma conclusão final. Para evitar duplicidade, todas as referências foram reunidas ao final do trabalho, assim como os apêndices, que contêm os materiais suplementares dos artigos. Vale ressaltar que, como foram avaliados 82 açudes ao longo de 10 anos de monitoramento, bases de dados e mais de 12 milhões de registros foram analisadas. Consequentemente, centenas de gráficos e tabelas extensas foram elaborados para permitir a compreensão dos padrões gerais, resultando em materiais suplementares extensos, reportados no Apêndice. Ainda assim, optou-se por mantê-los, pois essas informações comprovam ou corroboram os resultados apresentados nos capítulos subsequentes. Os materiais suplementares dos capítulos 2 a 4 estão nos Apêndices de A a C, respectivamente.

1.2.1 Capítulo 2: soluções de baixo custo para CyanoHabs: revisão sistemática (2010-2024) de métodos alternativos de coleta e análise de dados adequados às condições usuais do monitoramento das companhias de saneamento

O Capítulo 2 apresenta uma revisão sistemática sobre as alternativas disponíveis para a coleta e análise de dados no monitoramento regular da qualidade da água pelas companhias de saneamento. Essa abordagem permite identificar lacunas de conhecimento no estado da arte relacionadas à gestão de florações de cianobactérias (CyanoHABs) entre os períodos de 2010 até 2024. Foram considerado os pontos:

a) **rotinas de monitoramento:** as legislações e normas internacionais geralmente recomendam o monitoramento regular da qualidade da água bruta nos mananciais destinados ao consumo humano. Esse monitoramento envolve a coleta de amostras em diferentes pontos e a realização de análises laboratoriais para identificar e quantificar organismos fitoplanctônicos, como as cianobactérias, e metabólitos, como toxinas e compostos que podem comprometer as características organolépticas da água tratada (BRAZIL, 2021; CHORUS, I., & WELKER, 2021; WHO, 2022);
b) **ferramentas complementares para tomada de decisão:** nos últimos 15 anos, houve um aumento nas pesquisas que indicam metodologias capazes de tornar a tomada de decisão em relação a eventos deletérios, como florações de cianobactérias, proativa. Essas metodologias utilizam dados gerados em campo, durante rotinas de monitoramento semelhantes às realizadas pelas companhias de saneamento, além de dados meteorológicos oficiais ou provenientes de instituições de monitoramento ambiental. (CAMPOS *et al.*, 2024; CHANG; IMEN; VANNAH, 2015; CHEUNG; LIANG; LEE, 2013; O'GRADY *et al.*, 2021; PANTELIĆ *et al.*, 2013; ROUSSO *et al.*, 2020; ZAMYADI *et al.*, 2016; ZANCHETT; OLIVEIRA-FILHO, 2013);

Diante desse cenário, surgem as seguintes questões:

a) o monitoramento regular realizado por empresas oficiais fornece dados suficientes para uma tomada de decisão antecipada em relação aos eventos incertos de floração de cianobactérias?

b) nos últimos 15 anos, quais metodologias de geração, coleta e análise de dados produziram resultados que podem ser aplicados pelas companhias de saneamento na gestão de florações tóxicas e metabólitos secundários?
c) quais dessas metodologias e técnicas são de baixo custo e poderiam complementar a rotina de monitoramento existente, especialmente para companhias de saneamento em países com recursos limitados para investimento?

Diante desses questionamentos, a hipótese avaliada foi:

Hipótese: Os dados gerados pelo monitoramento regular realizado por companhias de saneamento, quando complementados por metodologias de baixo custo e técnicas alternativas de análise de dados, podem aprimorar a capacidade de gestão antecipada de eventos críticos relacionados a florações tóxicas de cianobactérias e metabólitos secundários, especialmente em regiões com recursos financeiros limitados, tais como o semiárido cearense.

Portanto, os objetivos deste capítulo foram :

Objetivo geral: Identificar metodologias de coleta e análise de dados compatíveis com as condições usuais de monitoramento das companhias de saneamento.

Objetivos específicos: (i) selecionar técnicas que possibilitem aprimorar a gestão antecipada de eventos críticos relacionados a florações tóxicas de cianobactérias e metabólitos secundários; (ii) avaliar soluções de baixo custo adequadas a regiões com recursos financeiros limitados, como o semiárido cearense.

1.2.2 Capítulo 3: Impacto da pior seca na região do semiárido desde 1910: aumento da densidade de cianobactérias e homogeneização em reservatórios para abastecimento humano, com efeitos variáveis sobre toxinas

Após identificar as principais limitações e boas práticas para a modelagem de dados ambientais que geram os melhores resultados a partir de dados de monitoramento de rotina no Capítulo 2, a construção de uma ferramenta de apoio à decisão exige uma compreensão aprofundada da biodiversidade do semiárido sob a perspectiva das mudanças climáticas, conforme abordado neste capítulo. Adicionalmente, é importante avaliar o desempenho da modelagem de séries temporais de índices ecológicos, que permitem maximizar o aproveitamento das informações disponíveis a partir da massa de dados coletados.

Dessa forma, é necessário considerar:

a) **definir e medir biodiversidade:** compreender a biodiversidade local é um desafio devido à variedade de métricas e definições existentes (HURART, 1971; JOST, 2009). A biodiversidade, aqui considerada como sinônimo de diversidade biológica ou simplesmente diversidade, é um conceito multifacetado que se refere à variabilidade entre organismos vivos e seus complexos ecológicos, incluindo variações em atributos taxonômicos, genéticos, fenotípicos, filogenéticos e funcionais. Além disso, engloba mudanças na abundância (riqueza de espécies) e distribuição de espécies (equitabilidade) ao longo do tempo (diversidade alfa temporal) e espaço (DÍAZ *et al.*, 2015). Para uma compreensão mais completa da biodiversidade, é necessário vê-la como algo além do simples número de espécies (HAMILTON, 2005). Uma visão intuitiva é essencial para entender corretamente a biodiversidade. Nesse contexto, o uso de índices que medem a diversidade verdadeira ou o número efetivo de espécies, como os números de Hill (HILL, 1973), baseados na ideia original de MacArthur (MACARTHUR, 1965), fornece resultados mais compreensíveis do que os valores de entropia obtidos por índices tradicionais, como os de Shannon (SHANNON, 1948) e Simpson (SIMPSON, 1949). Além disso, a biodiversidade é melhor avaliada considerando a escala temporal. A diversidade beta temporal captura as mudanças na composição das comunidades ao longo do tempo, refletindo processos ecológicos de substituição de espécies (turnover) e perda sem reposição (aninhamento), onde as comunidades atuais são subconjuntos das anteriores. Ressalta-se que a diversidade beta temporal oferece mais informações sobre mudanças temporais na biodiversidade do que apenas avaliar a diversidade alfa temporal, que considera apenas a dinâmica de abundância e distribuição de espécies (LINDHOLM *et al.*, 2021).
b) **impacto das mudanças climáticas, ação antrópica e homogeneização biótica:** Nos últimos 50 anos, a atividade humana acelerou as mudanças na natureza de forma sem precedentes. Alterações no uso da terra, mudanças climáticas e poluição figuram como os exemplos de impulsionadores desse rápido declínio dos ecossistemas. Como resultado,

a abundância média de espécies nativas caiu pelo menos 20% e as comunidades biológicas estão se tornando cada vez mais semelhantes. Tudo isso repercute na perda da biodiversidade local (IPBES, 2019). Embora os impactos das pressões antropogênicas estejam bem documentados em algumas espécies de vertebrados(GAINSBURY; SANTOS; WIEDERHECKER, 2022; LINES *et al.*, 2023), florestas (BARRETO *et al.*, 2023; HAYES *et al.*, 2023; MAURE *et al.*, 2023) ou insetos (LIU; LIU; YANG, 2023), pouco se sabe sobre a biodiversidade do filo *Cyanobacteria* na água captada em mananciais superficiais destinados ao consumo humano, especialmente em regiões críticas como o semiárido.

Diante dessa conjuntura, indaga-se:

a) como a biodiversidade de cianobactérias nos 82 açudes destinados ao abastecimento humano no semiárido cearense variou ao longo de 10 anos de monitoramento?
b) quais foram os impactos das atividades antrópicas, das mudanças climáticas e do processo de homogeneização biótica nesses reservatórios ao longo da última década?
c) quais são os principais efeitos dessas mudanças no contexto do tratamento de água na região?

Nesse cenário, a hipótese foi:

Hipótese: as mudanças climáticas, aliadas às atividades antrópicas e ao processo de homogeneização biótica, têm contribuído para o aumento da densidade de cianobactérias e a perda de diversidade nos reservatórios do semiárido cearense destinados ao abastecimento humano. Essa combinação de fatores resulta em efeitos variáveis sobre a produção de toxinas, afetando diretamente a qualidade da água e a eficiência do tratamento.

Portanto, os objetivos do capítulo 3 foram:

Objetivo Geral: investigar as variações na biodiversidade de cianobactérias ao longo de 10 anos de monitoramento nos 82 açudes do semiárido cearense destinados ao abastecimento humano, avaliando os impactos das atividades antrópicas, mudanças climáticas e homogeneização biótica;

Objetivos específicos: (i) avaliar os impactos do maior evento de seca desde 1910 nas séries temporais da diversidade alfa e beta temporal e nas componentes de turnover e aninhamento nos 82 açudes avaliados; (ii) analisar como essas mudanças influenciam a produção de toxinas e a densidade celular no contexto do tratamento de água na região.

1.2.3 Capítulo 4: Como identificar bioindicadores de riscos elevados de toxinas em reservatórios dominados por cianobactérias potencialmente tóxicas com alto turnover usando dados de monitoramento de rotina

Diante da constatação de que elevadas concentrações de cianobactérias potencialmente tóxicas dominam todos os 82 reservatórios e de que, embora o turnover seja a principal componente da diversidade beta temporal, o aninhamento ainda intensifica o processo de homogeneização biótica em vários açudes, conforme constatado no Capítulo 3, faz-se necessário definir quais organismos são raros e quais são comuns. Isso é fundamental porque, para selecionar táxons bioindicadores associados a níveis elevados de toxinas na água bruta, deve-se escolher um conjunto representativo de organismos que respeite as particularidades do ecossistema de cada região. Por outro lado, embora técnicas de seleção de bioindicadores como o Valor Indicador (IndVal), como proposto por (LEGENDRE, 2013), sejam amplamente utilizadas e de fácil implementação e interpretação, elas não quantificam e direcionam essa associação com base na chance de ocorrência. Mesmo não sendo uma predição, essa abordagem permite uma gestão antecipada ao estimar um intervalo de confiança para a razão de chances. Assim, com base nisso, tem-se:

a) **regressão logística:** A regressão logística é uma técnica estatística usada para modelar a relação entre uma variável dependente binária (como a presença ou ausência de um evento crítico) e uma ou mais variáveis independentes (como fatores ambientais ou cianobactérias comuns em um dado reservatório). No contexto do monitoramento da qualidade da água, ela é particularmente útil para identificar associações entre fatores ambientais e a ocorrência de eventos críticos, como florações tóxicas, permitindo prever a probabilidade desses eventos e estimar as razões de chances (odds ratio - OR). A OR é uma medida de associação que indica a chance de um evento ocorrer em um grupo em comparação com outro. Em termos de cianobactérias, uma OR pode mostrar, por exemplo, como a presença de um táxon específico (denominado bioindicador) influencia a probabilidade de um evento crítico de toxina, em comparação à ausência desse táxon. De modo geral, uma OR > 1 sugere um aumento na chance de ocorrência, enquanto uma OR < 1 indica uma redução. As ORs ajudam a quantificar o efeito e a direção das associações, e, embora não sejam previsões determinísticas, podem fornecer uma ferramenta importante para decisões antecipadas, ajudando a definir estratégias de monitoramento e mitigação com base em dados observados (CHANDLER, 2020; JAMES *et al.*, 2013; NELDER; WEDDERBURN, 1972; SHEATHER, 2009).

b) **Algoritmo a priori**: sendo amplamente aplicado na análise de cestas de mercado (Market Basket Analysis), O algoritmo a priori é uma técnica de aprendizagem de máquina usada para identificar padrões frequentes de coocorrência em grandes conjuntos de dados. No contexto ecológico, pode complementar o IndVal ao detectar associações entre espécies e condições ambientais, ajudando a prever eventos críticos, como florações tóxicas, ao analisar padrões de presença conjunta. Sua principal limitação reside no fato de que o algoritmo a priori se baseia apenas na frequência das associações, sem quantificar a força ou a direção das relações, o que pode levar a interpretações menos precisas em situações em que é necessário compreender a probabilidade de ocorrência de um evento (como faz a regressão logística, por exemplo, ao fornecer odds ratio). Dessa forma, ele complementa o IndVal, que identifica bioindicadores com base em sua importância em um determinado habitat, mas não substitui técnicas mais robustas de modelagem preditiva (LEOTE *et al.*, 2020).

c) **bioindicadores:** são organismos que atuam como indicadores naturais de mudanças ambientais, refletindo as condições de um ecossistema. No contexto do monitoramento de qualidade da água, especialmente em reservatórios do semiárido cearense dominados por cianobactérias potencialmente tóxicas, os bioindicadores são fundamentais para a detecção precoce de eventos críticos de toxinas. Eles podem ajudar a identificar a presença de níveis perigosos de toxinas na água bruta, antecipando possíveis riscos para o abastecimento humano (SIDDIG *et al.*, 2016). Para serem eficazes, os bioindicadores precisam ser sensíveis a variações ambientais, facilitando a detecção de condições adversas antes que os problemas se agravem(CHANDEL *et al.*, 2023). No caso das cianobactérias, bioindicadores bem selecionados no conjunto de organismo comuns de cada reservatório podem ajudar a estimar a probabilidade de eventos críticos, permitindo às companhias de saneamento agirem de forma proativa. Além disso, é importante que esses bioindicadores sejam fáceis de monitorar e identificar, reduzindo tempo e custos, aproveitando as rotinas de monitoramento existentes.

Diante disso, pergunta-se:

a) como selecionar um conjunto representativo de organismos para cada um dos 82 reservatórios avaliados, considerando as particularidades ecológicas locais e a dinâmica da biodiversidade de cada reservatório?
b) a regressão logística oferece resultados mais consistentes e é menos suscetível a vieses em comparação com outras técnicas de seleção de bioindicadores, como IndVal e algoritmos a priori?
c) de que maneira a razão de chances (OR) e seus intervalos de confiança ajudam a identificar bioindicadores associados a níveis elevados de toxinas na água bruta, permitindo o desenvolvimento de uma ferramenta complementar para gestão antecipada desses eventos?

A hipótese desse capítulo ficou:

Hipótese: A seleção de organismos bioindicadores, quando realizada levando em conta as particularidades ecológicas de cada reservatório e utilizando regressão logística, proporciona resultados mais robustos e menos suscetíveis a vieses em comparação com técnicas tradicionais como IndVal e algoritmos a priori. A análise baseada na razão de chances (OR) e seus intervalos de confiança permite identificar de forma eficaz bioindicadores associados a níveis elevados de toxinas, oferecendo uma ferramenta complementar para a gestão antecipada de eventos críticos. Essa abordagem se mostra mais eficaz do que considerar apenas a presença ou as elevadas densidades celulares de cianobactérias potencialmente tóxicas.

Portanto, os objetivos foram:

Objetivo geral: identificar bioindicadores associados a níveis elevados de toxinas na água bruta, usando dados de monitoramento regular, oferecendo uma ferramenta complementar para a gestão antecipada de eventos críticos pelas companhias de saneamento, superando a simples dependência da presença ou densidade elevada de cianobactérias potencialmente tóxicas.

Objetivos específicos: (i) avaliar uma metodologia para a seleção de bioindicadores de cianobactérias potencialmente tóxicas, levando em conta as particularidades ecológicas de cada reservatório e utilizando regressão logística. (ii) demonstrar que essa abordagem proporciona resultados mais robustos e menos suscetíveis a vieses do que técnicas tradicionais, como IndVal e algoritmos a priori.

2 SOLUÇÕES DE BAIXO CUSTO PARA CYANOHABS: REVISÃO SISTEMÁTICA (2010-2024) DE MÉTODOS ALTERNATIVOS DE COLETA E ANÁLISE DE DADOS ADEQUADOS ÀS CONDIÇÕES USUAIS DO MONITORAMENTO DAS COMPANHIAS DE SANEAMENTO

LOW-COST SOLUTIONS FOR CYANOHABS: SYSTEMATIC REVIEW (2010-2024) OF ALTERNATIVE METHODS FOR DATA COLLECTION AND ANALYSIS SUITABLE FOR THE USUAL MONITORING CONDITIONS OF SANITATION COMPANIES

RESUMO

Esta revisão sistemática que contempla publicações dos últimos 15 anos explora como técnicas de baixo custo para coleta e análise de dados podem melhorar a gestão de florações nocivas de cianobactérias (CyanoHabs) nas companhias de saneamento (CS). Os resultados indicam que, para aprimorar a resolução espaço-temporal do monitoramento, ajustes nos calendários de coleta podem viabilizar o uso do sensoriamento remoto (SR), enquanto dispositivos de monitoramento de alta resolução (HRM) proporcionam uma caracterização mais precisa das CyanoHabs. A sinergia entre essas técnicas e o monitoramento regular pode ampliar a geração de dados, mas é necessária uma estrutura laboratorial mínima para validações, calibrações e identificação do fitoplâncton. Diante da baixa frequência de monitoramento, técnicas estatísticas predominaram sobre aprendizado de máquina (ML), com modelos que predizem probabilidade e que utilizam dados defasados mostrando bons resultados. A validação de algoritmos utilizado no RS em regiões tropicais e o desenvolvimento de novas metodologias de identificação do fitoplâncton com HRM foram identificadas como novas oportunidades de pesquisa. Além disso, deve-se popularizar o uso de índices biológicos e técnicas estatísticas com menos pressupostos, pois as condições usuais das CS ainda limitam alguns algoritmos de ML.

Palavras-chave: Sensoriamento remoto; Sondas multiparamétricas; Sondas de fluorescência; Aprendizado de Máquina, Inteligência Artificial, Modelagem estatística; Modelos baseados em dados; tomada de decisão baseada em dados.

ABSTRACT

This systematic review explores how low-cost data collection and analysis techniques can improve the management of Cyanobacteria harmful blooms (CyanoHabs) in sanitation companies (SC). The results indicate that to enhance the spatiotemporal resolution of monitoring, adjustments in collection schedules can enable the use of remote sensing (RS), while high-resolution monitoring (HRM) devices provide a more accurate characterization of CyanoHABs. The synergy between these techniques and regular monitoring can expand data generation, but a minimal laboratory structure is necessary for validations, calibrations, and phytoplankton identification. Given the low monitoring frequency, statistical techniques predominated over machine learning (ML), and models that predicted probabilities using lagged data yielded good results. Validating RS algorithms in tropical regions and developing new methodologies for phytoplankton identification with HRM were identified as new research opportunities. Additionally, it is important to promote the use of biological indices and statistical techniques with fewer assumptions, as the usual conditions of SC still limit some ML algorithms.

Keywords: Remote sensing; Multiparameter probes; Fluorescence probes; Machine learning; Artificial intelligence; Statistical modeling; Data-driven models; Data-driven decision-making.

Highlights:

a) SLR identified low-cost alternative techniques suitable for regular monitoring of CyanoHabs management.
b) remote sensing cheaply expands spatiotemporal monitoring and should be explored worldwide
c) adopting high-resolution monitoring (HRM) improves prediction accuracy and enables proactive management with automatic alerts
d) data-driven models can overcome the limitations of traditional monitoring.
e) biological indices and statistical techniques with fewer assumptions should be prioritized.
f) low-frequency data from water utilities limits the use of some machine learning algorithms.

2.1 Introduction

Despite technological advancement in the water sector, cyanobacteria and cyanotoxins remain the most concerning aspects faced by sanitation companies (SC) worldwide (DEVI *et al.*, 2021; KAKADE *et al.*, 2021; MAITY *et al.*, 2021; SERRÀ *et al.*, 2021; YANG *et al.*, 2021). Enhancing decision-making based on routine phytoplankton monitoring is challenging in this context due to the increasingly frequent and unpredictable occurrences of harmful cyanobacterial blooms (CyanoHABs). Traditionally, this sector has been applying a reactive outlook, concentrating on compliance with regulatory standards, while strategic or proactive decisions are more limited. Consequently, improving the decision-making process based on monitoring data is a challenge given the typical organizational structure of companies.

International legislation and standards recommend that water supply and sanitation companies regularly monitor the quality of raw water intended for human consumption by water sampling and conducting periodic analyses (BRAZIL, 2021; WHO, 2022). Although these companies have the infrastructure to produce large-scale data, decision-making based on this information can be more complex for CyanoHABs. At least two factors may explain this: (i) algal proliferation is generally heterogeneous and ephemeral (GOYENS *et al.*, 2022) and traditional structures are not entirely suitable for this; (ii) besides sampling, the process also involves data analysis and interpretation, which is often not performed within sanitation companies.

In developing countries and most SC worldwide, traditional microscopy techniques and laboratory analyses are primarily used to monitor cyanobacteria (AGUILERA *et al.*, 2023). This approach limits both the frequency of analyses and the number of monitoring points. Consequently, data collection capacity is reduced, impacting decision-making, and resulting in data series with few records over the years. Additionally, compared to other techniques such as in situ fluorescence and remote sensing, these procedures require significant analysis time, financial resources, and qualified technicians (ROUSSO *et al.*, 2020).

A way to overcome the limitations of traditional monitoring is by remote sensing (RS) information, which includes the processing of aerial photographs and/or analysis of satellite images. Generally, this information allows for the qualitative and quantitative identification of the spatiotemporal distribution of the electromagnetic spectrum. Consequently, these systems provide data across the entire spectrum, enabling a wide range of applications in environmental monitoring, including the assessment of

CyanoHABs, which can complement regular laboratory monitoring (COFFER *et al.*, 2021; HORNING, 2008).

Besides the chlorophyll-a (Chl-a), cyanobacteria synthesize additional pigments such as phycocyanin (PC), which is related to the blue-green aspect of a cyanobacterial bloom being emitted and absorbed around 620-630 nm (WHITTON; POTTS, 2012). Both pigments are widely used in CyanoHabs identification in water systems (BRESCIANI, 2011; COFFER *et al.*, 2021; GOYENS *et al.*, 2022; LUNETTA *et al.*, 2015; MEDINA-COBO *et al.*, 2014; REYNOLDS *et al.*, 2023; SCHAEFFER *et al.*, 2018b; ZHANG; WANG; CHANG, 2019). Furthermore, cyanobacteria pigments have been associated as a suggestive indication of cyanotoxins (HANDLER *et al.*, 2023) being estimated by data analysis of multispectral (PAMULA *et al.*, 2023) or hyperspectral images (GOYENS *et al.*, 2022; O'GRADY *et al.*, 2021). Regarding this, while multispectral images collect data in discrete bands of specific wavelengths within the electromagnetic spectrum, hyperspectral data provide reflectance information in narrower spectral bands, offering higher resolution. This higher resolution expands the capabilities of RS (KUDELA *et al.*, 2015).

Abiotic and biotic water quality parameters beyond pigments (Chl-a and PC) can also be estimated by buoys, probes, or sensors that perform continuous monitoring with high temporal resolution (HRM). These devices, integrated with systems like the Internet of Things (IoT) or Supervisory Control and Data Acquisition (SCADA), enable more precise and automated monitoring of water resources (KWON *et al.*, 2023). They complement traditional monitoring programs and enhance the capacity for water resource analysis (WARDROPPER; BROOKFIELD, 2022). Although they require frequent calibration and maintenance, these devices provide real-time data and are user-friendly, thereby improving the management of CyanoHABs. However, their accuracy can be influenced by factors such as cell size and morphology, high turbidity, cell aggregation, the presence of other organisms, or rapid changes in species composition (BERTONE; BURFORD; HAMILTON, 2018; ZAMYADI *et al.*, 2016).

Regarding data analysis and interpretation, the process of transforming data into information can be categorized into two groups according to the modeling: process-based models (PBM) and data-based models (DBM). PBMs, also known as mechanistic or simulation models, describe the environment through hypotheses of physical, chemical, or biological relationships between variables. The generation of information in these models requires a range of parameters for the environment characterization. These parameters are typically obtained experimentally and less explored in standard monitoring programs by water and sanitation companies. PBMs are useful for predicting new data and assessing causality between variables, but they require significant effort and time (URRUTIAGUER, 2016). Conversely, data-based models (DBMs) establish relationships between variables without prior knowledge of the environment, operating under the assumption that the information needed to understand the phenomenon is present in the data, as described by the relationships between variables. These models require enough observations to accurately represent the phenomena' conditions. Statistical models and machine learning (ML) techniques fall into this category (LEE *et al.*, 2023). DBMs present high sensitivity to the amount of available data, especially when using ML mainly due to the overfitting information (AHMED *et al.*, 2024).

Considering the monitoring practices of SC, the implementation of process-based models (PBMs) can be challenging. This is due to the need for specialized knowledge in several fields and the large number of parameters that require calibration. These factors increase the complexity and costs associated with this modeling, particularly due to the need for measuring new parameters. As a result, PBMs can become complex, potentially leading to uncertainties and errors in the outcomes. Additionally, PBMs require extensive prior knowledge of the environment, in contrast to data-based models (DBMs), which are more widely known and suitable for the realities of these companies (ROUSSO *et al.*, 2020).

Although less complex, converting data into useful information using DBMs requires skills in programming, statistics, and proficiency with ML algorithms (LEE *et al.*, 2023; ROUSSO *et al.*, 2020). Thus, considering the realities of SC and the complexities of CyanoHABs modeling, exploring new analytical methodologies may not be an easy task for managers in this sector. This can be attributed to the need for data generation through effort-intensive analytical techniques, coupled with the requirement for professionals to monitor and adhere to an extensive set of quality standards by official legislation. However, these skills may not be widely disseminated, making the potential lack of expertise an additional challenge to overcome in a short time frame (HUANG *et al.*, 2021). Therefore, while meeting legal requirements and generating extensive data, the amount of useful information for decision-making is limited.

In this context, although numerous studies and reviews address alternative methods to improve data collection and generation (AGUILERA *et al.*, 2023; BERTONE *et al.*, 2019; BERTONE; BURFORD; HAMILTON, 2018; MEZZANOTTE *et al.*, 2011; O'GRADY *et al.*, 2021; PAMULA *et al.*, 2023; PARK *et al.*, 2015; ZAMYADI *et al.*, 2016), besides presenting innovative techniques for effectively modeling CyanoHABs (AHMED *et al.*, 2024; LEE *et al.*, 2023; ROUSSO *et al.*, 2020), these reports are generic within contexts that frequently do not align with the realities of SC, especially in countries with limited financial and technological resources. From this, not all strategies presented in these studies are feasible. In these companies, both monitoring and data analysis, besides the interpretation capabilities, are not specifically designed for CyanoHAB monitoring but rather for other water quality parameters. Therefore, reviews that consider these aspects are still needed, representing a gap that must be addressed.

Therefore, this review aims to identify alternative methods for obtaining, analyzing, and interpreting CyanoHAB data in a way that enhances decision-making based on regular monitoring and its feasibility for sanitation companies. A systematic literature review was conducted for screening publications from the last 15 years to test the hypothesis that decision-making can be improved by adopting low-cost alternative methods or utilizing the existing infrastructure of water and sanitation companies. Consequently, promising and potentially applicable results are presented clearly and directly, including within the context of countries with limited resources.

2.2 Methodology

A systematic literature review (SLR) was carried out to assess the state of the art about knowledge applied to enhance the decision-making aspects for sanitation companies based on monitoring programs of CyanoHABs. The studies selected for this SLR were exclusively scientific papers available in the Web of Science and Compendex databases. Two general features were examined: (i) alternative methods for obtaining data, and (ii) methodologies for analysis and interpretation to obtain useful information. To achieve this, the essential steps outlined in Figure 1 followed predetermined recommendations for SLR (Higgins et al., 2020; Pickering and Byrne, 2014).

Figure 1 - Phases of the Systematic Literature Review (SLR) applied in this current study.

Source: prepared by the author

An SLR establishes clear, transparent, and reproducible criteria for the inclusion or exclusion of publications. It begins by identifying data sources and defining search criteria. Publications are selected or excluded based on relevant criteria to the research objective. Subsequently, relevant data are extracted from the full text and summarized. Although the SLR captures a significant sample, not all relevant references may be included. Therefore, in the discussion section, some specific references were added to complement or corroborate the respective findings, even if they are not part of the validated reference sample, following a similar procedure described by Rousso et al., 2020.

2.2.1 Planning

Following the exploratory analysis, a research protocol was developed for this phase. This document, Table S1 in Supplementary Materials (SM) - APPENDIX A, outlined the key elements of the SLR. To avoid bias, the authors were divided into two groups. One group followed the protocol shown in Table S1 before a set of published papers, while the other group validated this initial set also based on Table S1. Only selected papers in both groups were included in this SLR.

2.2.2 Execution

Based on Table S1, a search for studies was carried out in databases using specific strings (Table S2 in SM - APPENDIX A) to obtain the initial sample of publications. Several search strings were tested, and we chose those that showed the best trade-off between the number of publications found and the number of publications remaining after applying the inclusion and exclusion criteria. Two search strings were considered. Subsequently, studies were selected based on the inclusion and exclusion criteria, following a sequential order (Title, Abstract, and Full Text), as defined in the protocol. This process resulted in the final sample of publications.

2.2.3 Summary and Validation

After obtaining the final sample, three general aspects were evaluated to determine whether the selected publications were sufficiently representative to achieve the overarching aim of this review, which follows:

a) Overview of methodologies for data acquisition and collection: this aspect evaluated the following parameters: the country where the monitoring was conducted; the type of water system (artificial, lakes, rivers, varied, not specified); the trophic state (eutrophic, other, varied, and not specified); predominant climate and climatic zone (tropical, temperate, subtropical, semi-arid, Mediterranean, and unspecified); socioeconomic level of the country where the study was conducted (developed or developing); data acquisition methodology (remote sensing - RS, high-resolution monitoring – HRM; traditional laboratory monitoring - TM, combinations of these methods, or not specified); and data source (regular monitoring and/or official programs, specific data collection only during the study, a combination of both, or not specified);
b) Overview of alternative methodologies for data analysis and interpretation: This aspect assessed the primary techniques of data-based modeling (statistical: STAT or machine learning: ML); the models or methodologies that have been used and the performance criteria adopted (both categorized into the two classes ML or STAT). The list of all models and performance criteria with their respective acronyms can be found in Table S3 and S4 in SM, respectively.
c) Overview of the independent and dependent variables modeled - The variables were categorized into six groups (complete list in Table S5 and S6 in SM):
   - remote sensing: hyperspectral or multispectral images obtained or not by satellite.
   - toxins and other metabolites: including toxins and taste-and-odor compounds (T&O) such as 2- methylisoborneol (MIB) and geosmin.
   - hydrobiology data: indices, concentration and other information about the main microorganisms identified in the monitoring program.
   - pigments: chlorophyll-a (Chl-a) and phycocyanin (PC) data.
   - climatic variables: variables related to the climate conditions evaluated by the meteorological services.
   - general physical-chemical parameters: other abiotic variables used in the monitoring programs that were not set in the previous groups.

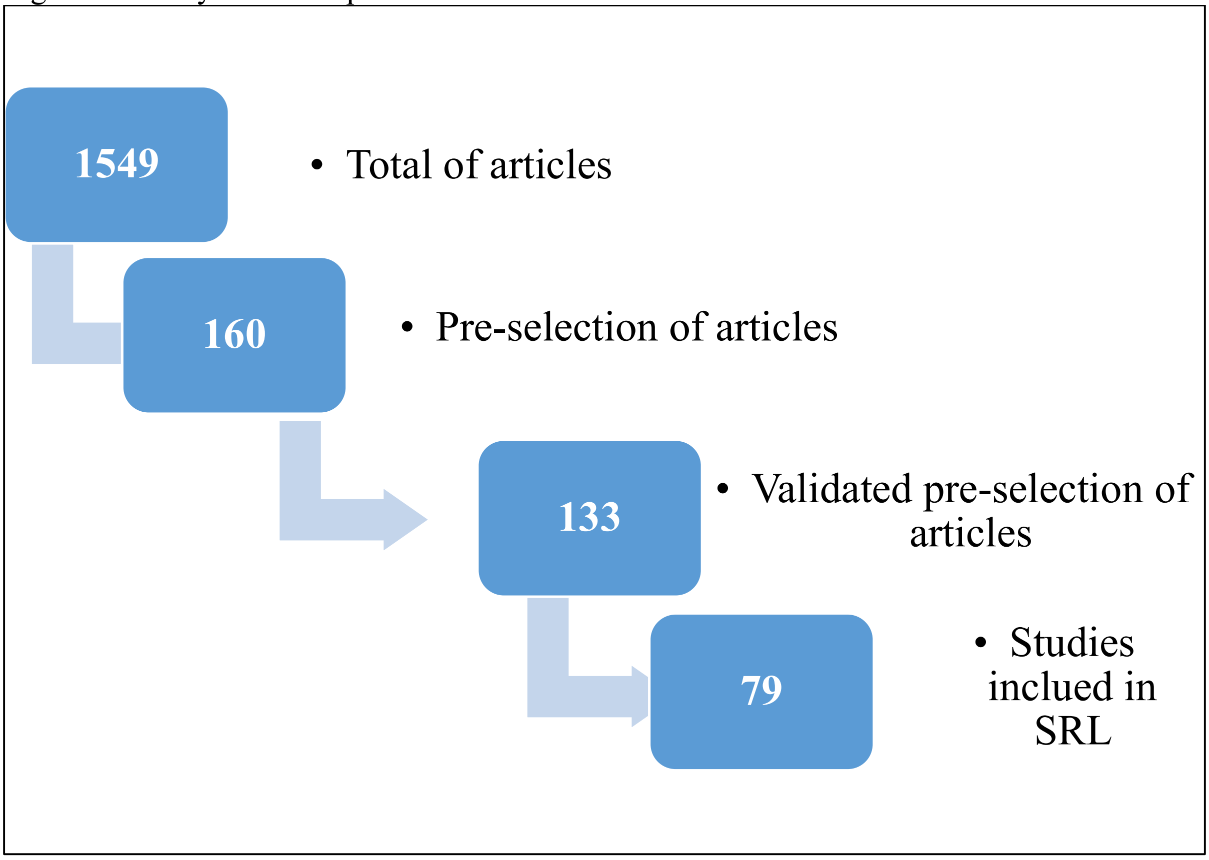
2.2.4 Discussion

After summarizing, the final set of validated publications was evaluated, and the key alternative methods for data collection, analysis, and interpretation related to CyanoHABs, which enhance decision-making based on regular monitoring, were described and presented. Gaps for future research were also identified. Additionally, primary and secondary studies deemed necessary by the authors but not retrieved by the search strings or excluded based on the criteria established in Table S1 - APPENDIX A. This is further detailed in the Discussion section.

2.3 Results

After applying the protocol presented in Table S1, 79 papers were selected for the validation stage as shown in Figure 2 and Table S3 in SM - APPENDIX A.

Figure 2 - Study inclusion processes in SLR

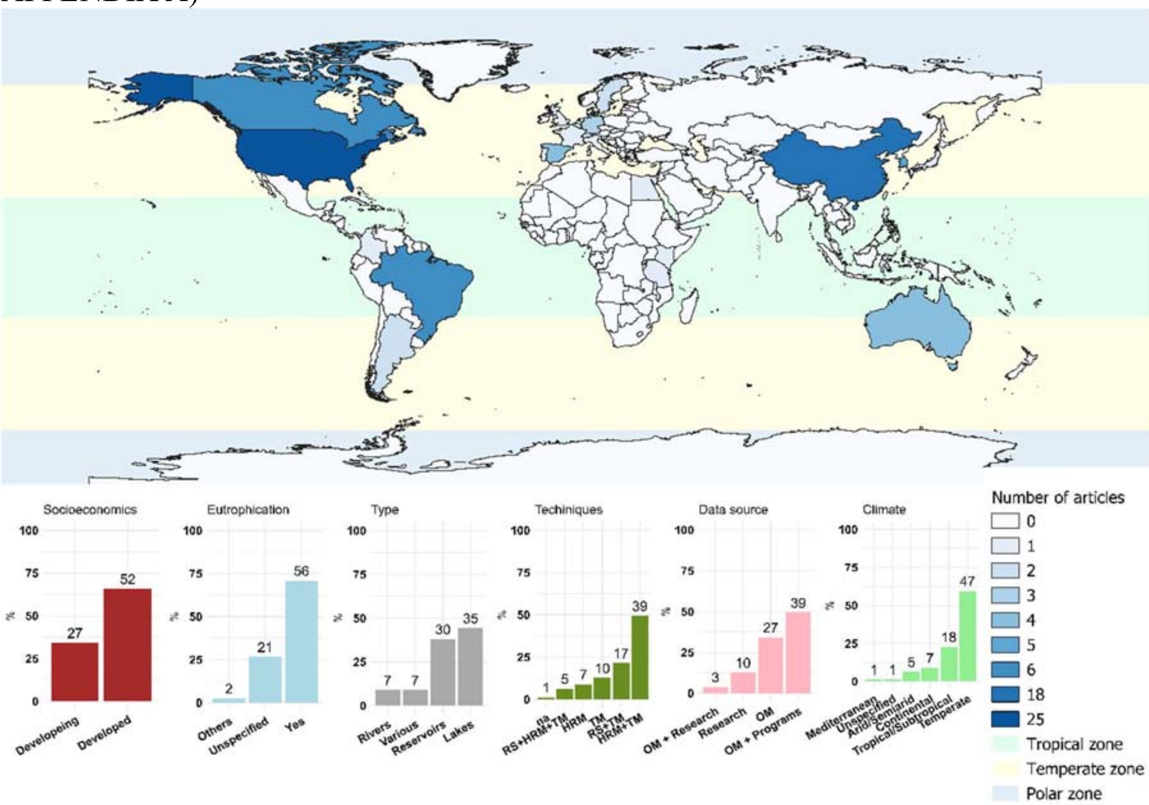


Source: prepared by the author

Based on this group of papers, Figure 3 shows that the SLR found a greater number of studies in developed countries within temperate zones (≈ 66%), which supports the high scientific output from these regions. Additionally, there was a predominance of studies carried out in lakes or artificial reservoirs (≈ 83%), mostly eutrophic (≈ 71%).

Although less susceptible to nutrient inputs, with greater water mixing and reduced retention time, 7% of the filtered papers found in this SLR reported CyanoHAB events in rivers. Furthermore, based on the findings of the reviewed papers, the frequent use of monitoring data from official programs or sanitation agencies (≈ 87%), traditionally generated in laboratories and supplemented by HRM or RS, underscores the potential of these data not only for monitoring but also for decision-making, prediction, and proactive management of these scenarios. Therefore, considering that the publications were produced across different climatic zones, water system circulation features (lentic and lotic water bodies), trophic conditions, and various socioeconomic contexts, the final set of filtered papers was deemed representative and appropriate for achieving the review's objectives.

Figure 3 - Summary of published papers used in this SLR, including the quantitative aspects and geographical distribution. (For better visualization, see Figure S1 in the SM-APPENDIX A)

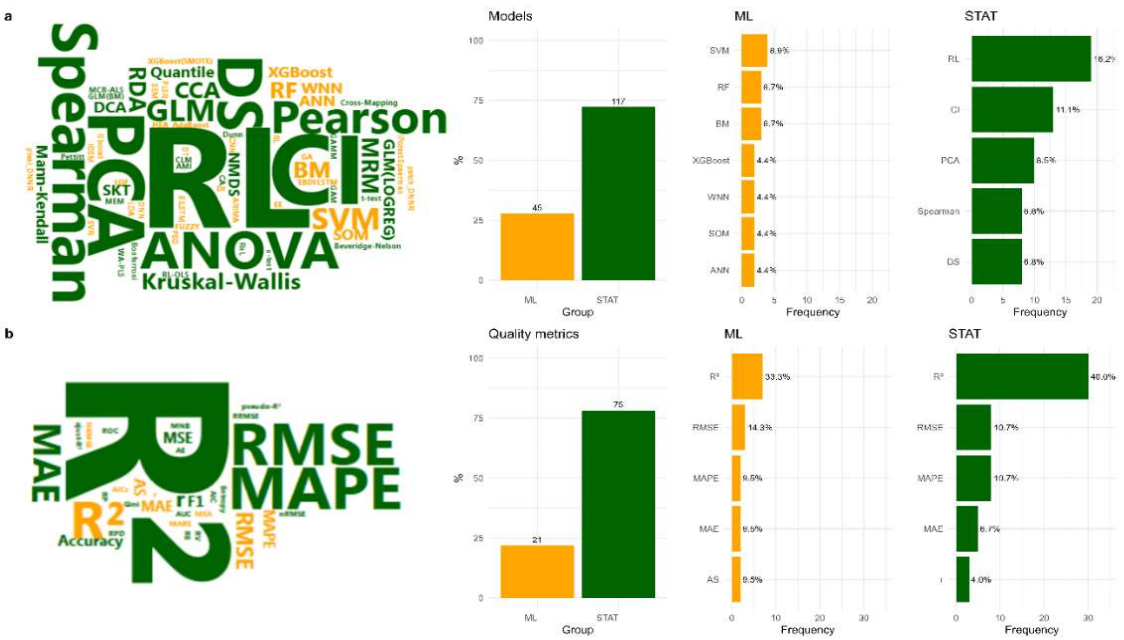


Prepared by the author

Regarding data analysis and interpretation, Figure 4 shows that 67% of selected papers applied supervised methodologies, such as regression or classification modeling, while 37% applied non-supervised methodologies and exploratory analysis. This may be related that both goodness-of-fit R^2 and RMSE were found in approximately

50% of the filtered papers, where 74% used primarily statistical techniques (36 models - Table S3) whereas 26% used predominantly ML (34 algorithms - Table S3). Regression techniques for predicting or estimating cyanobacteria using RS data accounted for more than 27% of the filtered papers within the total papers that used statistical methods. Although SVM had the highest percentage among studies that used ML (≈ 9%), there were no significant differences in the usage frequency of other algorithms. This balance can be attributed to the fact that most publications in this SLR evaluated multiple algorithms to select the one with the best fit, based on metrics such as those presented in Figure 4b.

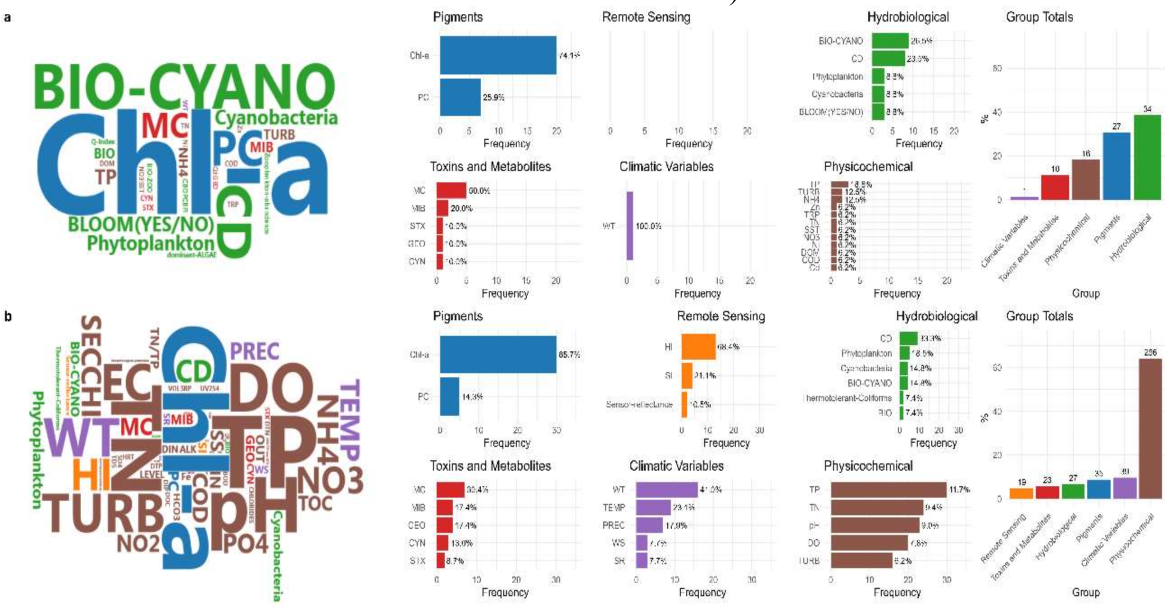
Figure 4 - Overview of models (a) and their respective metrics (b) obtained by a word cloud and exploratory analysis of the frequency of acronyms found in the papers retrieved by this SLR. The percentages on the horizontal bars were calculated based on the total for each group, indicated at the top of the vertically aligned bars of the same color. (For better visualization, see Figure S2. The list of acronyms is available in Tables S4 and S5. Both are included in the SM - APPENDIX A.)



Prepared by the author

According to Figure 5, about 77% of filtered papers in this SLR used regression or classification models to estimate hydrobiology variables, such as biovolume or cell density of cyanobacteria, besides pigments including Chl-a or PC. Although they represent the main threat to public health and water supply, toxins and T&O metabolites were the focus of only about 13% of the publications where 50% addressed microcystins (MC) and 20% addressed 2-methylisoborneol (MIB). Meanwhile, the other toxins and geosmin were nearly equally represented in the remaining publications within this group. Among the independent variables, total phosphorus (TP) and total nitrogen (TN) from the physical-chemical group, and Chl-a from the pigments group, were the most frequently used in the water monitoring program. Despite being the differential indicator of CyanoHABs, PC accounted for only 25% of the total filtered papers that used pigments as predictors. Additionally, temperature and precipitation were the most relevant climatic variables and more than 68% of the papers utilized RS monitoring employed data from hyperspectral images.

Figure 5 – Overview of dependent (a) and independent variables (b) obtained through a word cloud and exploratory analysis of the frequency of acronyms found in the papers retrieved by this SLR. The percentages on the horizontal bars were calculated based on the totals for each group, indicated at the top of the vertically aligned bars of the same color (For better visualization, see Figure S3. The list of acronyms is available in Tables S6 and S7. Both are included in the SM - APPENDIX A.)



Prepared by the author

2.4 Discussion

Based on the information gathered from the final set of selected publications by the SLR, as well as from other publications that corroborate or expand upon these concepts, this section presents the key aspects relevant to the central objective of this review.

2.4.1 Collecting and generating data

2.4.1.1 Remote sensing

According to the review, various studies have shown promising results in monitoring CyanoHABs through remote sensing throughout the past 15 years. Just as Chl-a facilitates detailed mapping of phytoplankton spatial distributions, the spectral characteristics of PC allow for similar outcomes, specifically targeting cyanobacteria. For instance, using images captured by MERIS (Medium Resolution Imaging Spectrometer), one of the primary sensors aboard the Envisat mission launched in March 2002 by the European Space Agency (ESA, 2024), it was possible to estimate cyanobacterial biovolume in 23 freshwater ecosystems on the Iberian Peninsula dominated by various cyanobacterial species, without scum identification, utilizing PC values (MEDINA-COBO *et al.*, 2014).

Using an extended inversion model of inherent optical properties (IOP) for inland waters (IIMIW), low levels of PC ($\leq$ 50 mg m^{-3}) were determined from remote sensing reflectance based on data collected between 2005 and 2010 in American reservoirs. The required bands for the algorithms are also available from satellite sensors such as MERIS, Ocean and Land Colour Instrument (OLCI) available at Sentinel-3 (NIEKE *et al.*, 2015), and Hyperspectral Imager for the Coastal Ocean (HICO), which may improve the potential predictive warnings of CyanoHabs (LI; LI; SONG, 2015).

The ability to use irradiance measurements to accurately quantify the concentrations of PC and Chl-a as indicators of cyanobacterial in eutrophic waters, without the need for simultaneous field observations, was demonstrated in Missisquoi Bay, Lake Champlain. This was achieved with a spatial resolution of 300 m and a revisit time of 2 to 4 days using MERIS on Envisat combined with atmospheric correction and semi-analytical retrieval algorithms (WHEELER *et al.*, 2012).

An algorithm was developed to detect cyanobacteria blooms and chrysophytes representatives in Lake Idro (Italy) using remote sensing reflectance data derived from in situ measurements and MERIS data. Two wavelength ratios (620 nm/560 nm) were selected based on the pigment characteristics of the phytoplankton population (phycocyanin, β-carotene, fucoxanthin, and other xanthophylls), enabling the monitoring of blooms that, as often happens with cyanobacteria, may not be detected using only Chl-a concentrations (BRESCIANI, 2011).

From this, remote sensing can provide a valuable source of information for monitoring the water quality of multiple water systems used for human supply, allowing for the tracking of algal bloom development (MEZZANOTTE *et al.*, 2011; REYNOLDS *et al.*, 2023; SCHAEFFER *et al.*, 2018a; WHEELER *et al.*, 2012). Patterns that are not discernible from in situ data alone can be elucidated through remote monitoring, thereby enhancing and refining the understanding of otherwise elusive processes (SCHAEFFER *et al.*, 2018b).

A finer spatiotemporal scale is crucial to establishing more effective CyanoHabs monitoring and the development of an alert system for managing blooms in water bodies. Periodic monitoring, including daily, weekly or monthly programs at regional or national levels can significantly enhance the understanding of bloom dynamics, as phytoplankton proliferation is often heterogeneous and transient (GOYENS *et al.*, 2022). Although annual frequencies are useful for understanding long-term global trends in water bodies, they provide limited insight into CyanoHabs emergency states (COFFER *et al.*, 2021). Thus, high spatiotemporal resolution of data can be a limiting factor for CyanoHabs monitoring in inland water systems.

In this context, the availability of frequent data at suitable spatial resolutions for surface water bodies is frequently cited as an obstacle to effective RS monitoring (CLARK *et al.*, 2017; COFFER *et al.*, 2020; HURTADO *et al.*, 2022; KAKOUEI; KRAEMER; ADRIAN, 2022; REYNOLDS *et al.*, 2023). Satellite data may reveal the percentage of large lakes affected by blooms in the United States, both at national and regional scales. However, the applicability of these results to smaller lakes remains uncertain (COFFER *et al.*, 2020). Similarly, the Sentinel-3 OLCI sensors from the Copernicus program, which have a spatial resolution of 300 m, were able to detect cyanobacteria in U.S. inland water bodies for approximately 5.6% of their lakes and reservoirs. Conversely, higher spatial resolution sensors of 30 m (Sentinel-2) showed a lower capability of monitoring the water quality due to less frequent revisit cycles (CLARK *et al.*, 2017). Additionally, specific challenges in satellite remote monitoring, such as cloud cover, overpass rates, or the presence of scum, can also pose complications (MEDINA-COBO *et al.*, 2014; REYNOLDS *et al.*, 2023).

Considering these challenges, data analysis of hyperspectral cameras and sensors has shown promising potential. Due to their high temporal resolution, these sensors can detect minor variations in Chl-a, suspended particulate matter, and PC over time with minimal technical effort (GOYENS *et al.*, 2022). Moreover, in the satellite context, Sentinel-3B and Sentinel-2B hold the promise of improving temporal coverage of water bodies in the U.S., enabling the evaluation of all its lakes and reservoirs (CLARK *et al.*, 2017).

As demonstrated in Missisquoi Bay, Lake Champlain, it is possible to quantify PC and Chl-a concentrations as indicators of cyanobacterial cell density in eutrophic waters solely using irradiance measurements, without the need for simultaneous field analysis (WHEELER *et al.*, 2012). Nonetheless, no single method is entirely without limitations. Thus, considering the potential risks associated with CyanoHabs, a more secure management approach involves combining satellite data, in-situ probes, field sampling, and modeling efforts. This is because these methods provide observations at different spatial and temporal scales (HANDLER *et al.*, 2023). Additionally, in cases of data anomalies or failures (e.g., images of areas with high cloud cover), one method can complement the others.

Compared to traditional laboratory monitoring, the use of RS is appealing to water utility managers due to its reduced costs and ease of implementation, allowing for spatiotemporal analyses that would be nearly impossible to carry out solely through fieldwork (FOURNIER *et al.*, 2024; MEDINA-COBO *et al.*, 2014; MUNOZ *et al.*, 2021; ROUSSO *et al.*, 2020). Despite following a comprehensive list of parameters and conducting laborious and costly field monitoring routines to comply with their respective national regulations, water utilities and official environmental agencies often overlook the available satellite images. Although algorithms used to detect CyanoHabs abundance (cells mL^{-1}) have shown good validation to the traditional cyanobacterial cell counting within ±7 days of satellite overpasses in the U.S (CLARK *et al.*, 2017; COOK *et al.*, 2023; LUNETTA *et al.*, 2015) following an accurate estimation of water temperature in U.S. lakes by Landsat 5 and 7 scenes (>10%) within ±3 days of coincident in-situ measurements (SCHAEFFER *et al.*, 2018b), small adjustments in sampling schedules in water sources could enhance monitoring capabilities without additional infrastructure investments. The adoption of these two methodologies may advance the data-driven decision-making capacity of CyanoHabs management (PYO *et al.*, 2021).

Thus, despite their misuse by many companies in the water sector worldwide, research demonstrates that remote sensing is a potent tool for managing cyanobacteria in water bodies designated for human supply. The ability to effectively monitor these blooms has profound implications for public health, water quality, and the sustainability of water resources. From this, integrating RS into existing monitoring strategies is not just beneficial but essential for addressing future challenges in CyanoHabs management and water quality.

2.4.1.2 High-resolution monitoring (HRM)

In approximately 90% of the reviewed studies, a combination of laboratory analyses and the use of fluorescent or multiparameter probes for regular monitoring was observed. These HRM devices enable the assessment of water quality parameters at hourly or shorter intervals, providing rapid and accurate quantification, which would be impractical in laboratories due to high costs. Although the confirmatory results from laboratory analyses are recommended, they are not always necessary, as observed by (ALMUHTARAM; ZAMYADI; HOFMANN, 2021; XIAO *et al.*, 2017). Therefore, to recognize more complex patterns, laboratory data can be supplemented with real-time HRM. This efficiency is a primary reason for the use of HRM in official monitoring programs and water utilities.

Using SCADA (Supervisory Control and Data Acquisition) systems or IoT (Internet of Things), HRM can improve water quality management by enabling automatic alerts for critical parameters. These systems can integrate into the information ecosystem for CyanoHabs, optimizing the laboratory infrastructure and incorporating remote sensing imagery. For example, it has been noted that neural networks-based systems such as SCADA could more efficiently manage emergencies by using a series of pre-alert thresholds for parameters correlated with Chl-a (MEZZANOTTE *et al.*, 2011). Furthermore, in a more recent approach, deep learning, LSTM (*Long Short-Term Memory*), and IoT monitoring have shown significant potential for detecting and quantifying CyanoHabs with high accuracy (KWON *et al.*, 2023).

Considering the high volume of generated data and the costs associated with expanding laboratory infrastructure, the use of probes can offer a more economical approach to enhancing CyanoHabs management by water utilities. This allows for a more precise description of the complex patterns of these events. For example, using a simple system that represents about 14% of the cost of a typical automatic monitoring system, it was possible to improve the prediction of cyanobacteria and phytoplankton blooms in the Siling Reservoir (China) and Lake Winnebago (USA) using daily data from a single parameter (cell density or Chl-a) obtained by a buoyant fluorescent probe (XIAO *et al.*, 2017).

According to researchers, high-resolution in-situ monitoring can replace the traditional laboratory analyses of cyanobacterial blooms, which are laborious and costly for water utilities. High-frequency data (hourly and every quarter hour) from continuous monitoring by PC and Chl-a analysis in situ, for example, have accurately detected elevated cyanobacteria activity, eliminating the need for laboratory-confirmed cell counts or corresponding biovolumes (ALMUHTARAM; ZAMYADI; HOFMANN, 2021)

In this context, molecular techniques such as Real-time Polymerase Chain Reaction (qPCR) and real-time sensors can enhance the modeling and prediction of cyanobacterial blooms. Real-time measurements of nutrients and cyanobacteria can reduce the nearly exclusive reliance on predictive models based on discrete monitoring. Thus, adopting continuous monitoring with advanced sensors is crucial for more accurate predictions and less susceptibility to biases, mainly when associated with low-resolution monitoring (HARRIS; GRAHAM, 2017).

While in-situ HRM is promising, laboratory infrastructure remains essential for addressing specific aspects of cyanobacteria monitoring as demonstrated by Prestigiacomo et al., 2023. Moreover, the ineffectiveness of in-situ sensor data, as observed in the Greenup pool in 2019 on the Ohio River during the 2019 CyanoHabs, due to spatial heterogeneity, led to the exclusion of these data from predictive models (NIETCH *et al.*, 2022). These examples highlight that maintaining a streamlined laboratory setup combined with an efficient network of in-situ probes and sensors for various water quality parameters is strategic. Although measurements from multiparametric fluorescence probes are more direct, simultaneous, and cost-effective compared to traditional laboratory monitoring, they can be affected by interferences, resulting in biased measurements of Chl-a from eukaryotic green algae. In addition to technological advancements, simultaneous measurements of turbidity and dissolved organic carbon can improve the accuracy of these devices (COURTOIS *et al.*, 2018). Nevertheless, both strategies have their limitations, but when used together, their synergistic effect provides a more robust and reliable approach to water quality monitoring.

Using a large temporal series of hydrological data (25 years) from the Ohio River (USA), researchers developed a web application for real-time prediction of CyanoHabs, conversely to the cyanobacteria indicators collected through regular monitoring and in-situ sensors (NIETCH *et al.*, 2022). The authors suggest that incorporating laboratory information on potentially toxin-producing taxa could enhance results, despite the sampling effort and analysis time required, which can range from hours to many days. Thus, developing HRM for these parameters could improve the monitoring programs which may be more accessible to stakeholders from the water industry. However, further research in this area is still needed.

Despite the advancements in digitalization towards a smart green planet, tools like IoT, big data management, and artificial intelligence are accessed unevenly between rich and poor countries. This disparity creates a digital divide, potentially leading to anomalous results and complicating decision-making (MONDEJAR *et al.*, 2021). Consequently, besides the historical and internal challenges faced by developing countries in achieving sustainable development, researchers and decision-makers in water utilities face at least two additional daunting tasks: (i) seeking new analysis and monitoring techniques and (ii) enhancing data-driven decision-making, often based on small and limited datasets.

Regarding analysis and monitoring techniques, beyond the inclusion of appropriate RS and HRM data, research shows promising advances in fluorescence scanning. This technique highlights the substantial amount of data generated, for which researchers have already developed analytical methods, resulting in a well-developed characterization of dissolved organic matter for example (Li et al., 2020).

To improve decision-making, as technological advances are not yet sufficient to reduce the disparities between developed and developing countries, researchers and managers must explore alternative modeling techniques. These techniques need to be resilient to the limiting conditions of monitoring and capable of describing complex environmental phenomena, such as those observed in water quality. This is particularly relevant for cyanobacteria and their metabolites, which, despite technological advances in water treatment, continue to pose significant challenges for water utilities worldwide. Some of these techniques are discussed in the following sections.

2.4.2 Data analysis and interpretation

2.4.2.1 Data-based statistics and modeling

Several works provide useful insights into the analysis of ecological data (JENNIONS; MØLLER, 2003; RAMETTE, 2007; VERNIEST; GREULICH, 2019) and CyanoHabs (PARK; PATEL; LEE, 2024; ROUSSO *et al.*, 2020). According to them, there is an almost unanimous consensus on the complexity of modeling cyanobacterial blooms or their metabolite production using conventional statistical methods.

This difficulty, beyond the inherent non-linearity of these processes, may be partially associated with the specific characteristics of biological data, which often present challenges such as heteroscedasticity (unequal variance of the errors) and skewness. These issues affect both inference techniques and the predictions made through some regression methodologies (ABBOTT; GUTGESELL, 1994).

Violating the homogeneity of variance can result in increased false positive errors (Type I errors) as exemplified by Knief and Forstmeier, (2021), who reported that applying the t-test - a widely known and used statistical method that assumes equal variances in both compared groups - on data with different variances resulted in approximately 23% Type I error rates.

Moreover, as long-series data collection is a key feature of monitoring by water companies, it is common to observe high correlations between values at a given time point and previous results, an effect known as autocorrelation. Among other aspects, data with these characteristics violate the assumption of independent errors, a fundamental prerequisite for least squares regression methods, which are widely used in predictions (SHEATHER, 2009).

The presence of correlated errors can significantly impact the prediction accuracy and p-values. If such a correlation exists, one error term can influence one or more subsequent terms. This can lead to underestimated standard errors, resulting in confidence intervals and p-values that are smaller than they should be. The combination of these effects can lead to unjustified overconfidence in models adjusted under these conditions (JAMES *et al.*, 2014).

This situation requires analysts to take additional precautions in modeling to ensure that the specific assumptions of techniques like those previously used (BHATT *et al.*, 2023; FERNANDEZ - FIGUEROA *et al.*, 2021; HARRIS; GRAHAM, 2017; MÂNICA; DE LIMA ISAAC, 2023; XU *et al.*, 2015). However, given the nature of regular monitoring data and the characteristics of CyanoHabs, this task may not be easily solved. For instance, in Ohio (USA), a multiple linear regression model was used to quickly estimate the probability of exceeding an action threshold for microcystin concentrations with an accuracy greater than 80%, providing valuable information for decision-makers at recreational sites and water treatment plants (FRANCY *et al.*, 2020).

The authors tested various transformations of the explanatory variables (log, inverse, square, square root, and fourth root) to ensure linearity and considered all other assumptions associated with ordinary least squares regression.

After confirming the absence of autocorrelation, ANOVA was applied to study changes in phytoplankton and the aquatic environment of Lake Hongze in China, using ten years of monitoring data (WEI *et al.*, 2023). Among other aspects, normality and homoscedasticity of data were confirmed before modeling data using multiple regression models to predict total phosphorus in a reservoir located in the semi-arid region of Brazil, employing Shapiro-Wilk and Levene tests, respectively (GOES; BARROS; NETO, 2023). Instead of Student's t-test, environmental variables that did not meet the normality and/or homogeneity requirements were compared between years using the Mann-Whitney U test in a study at the El Limón reservoir (Argentina), which found that temperature and precipitation variations affected cyanobacteria abundance, influencing bloom events (ALVAREZ DALINGER *et al.*, 2023).

The assumptions for using linear mixed models to evaluate the association between the categorized relative abundance of cyanobacteria and each water body assessed in New York State (USA) were tested visually through residual QQ plots (Quantile-Quantile) and residuals versus predicted plots (GORNEY *et al.*, 2023). Data were log-transformed, and tested for normality using Kolmogorov-Smirnov tests, while residual plots were visually inspected for homoscedasticity before fitting mixed-effect models to describe the seasonal dynamics of phosphorus, guiding effective management of Grand Lake St. Marys (USA). The study revealed that external P loading during winter and spring, along with internal P loading from sediments during summer, drives extremely high summer P concentrations that sustain dense cyanobacteria blooms (FILBRUN; CONROY; CULVER, 2013).

Although these procedures are well recognized in academia, they may pose an additional challenge for decision-makers in water utilities when implementing data-driven decision-making. The legal requirements and complex routines demand intense monitoring of various parameters, consuming a lot of time in data generation. Consequently, the process of obtaining actionable insights is compromised.

Therefore, given the specificities of CyanoHabs monitoring data produced by water utilities, analysis techniques less dependent on assumptions may represent a way to overcome these limitations. Some promising examples of these techniques were addressed in this review and will be detailed in the subsequent sections.

2.4.2.2 Alternative techniques with low requirements of assumptions

In the studies reviewed in this SLR, non-parametric techniques have shown promising results. This can be attributed, among other factors, to their low requirement for assumptions, such as the assumption of normality, which is useful for datasets with multiple sampling points (CLARKE, 1993; HELSEL; FRANS, 2006). For instance, the Seasonal Kendall test was employed to evaluate monotonic trends in CyanoHabs data series in water bodies in New York State, USA (GORNEY *et al.*, 2023). The Mann-Kendall Trend test revealed that only Lake Lillsjön, among other temperate lakes analyzed in Sweden, exhibited significant temporal trend variations in the average cyanobacteria biovolume before 2003, from 2003 to 2013, and after 2014, based on monitoring data from the last 23 years (LI *et al.*, 2021). In Lake Hallwil, Switzerland, the emergence and dominance of *Planktothrix rubescens* were studied using the Seasonal Kendall tests, the Sen slope — an estimate of the rate of trend change when significant — and the Pettitt method to determine the year in which the series changed the respective trends. These methods are less influenced by outliers, missing values, and heteroscedasticity and do not require a specific data distribution, characteristics often found in monitoring data (SUAREZ *et al.*, 2023).

Non-parametric multiple comparison tests, such as the Kruskal-Walli, were employed to combine samples from different locations and identify significant differences in CyanoHabs indicators (eg. Chl-a, microcystin, Secchi depth, and turbidity). Differences were analyzed concerning different depths (surface, epilimnion, top of the thermocline, hypolimnion) over time (weeks and months). This method was used to study the patterns and impacts of cyanobacteria in Lake Canandaigua, Canada, which is thermally stratified (PRESTIGIACOMO *et al.*, 2023). Another example of comparison using different parameters with the Kruskal-Wallis and the Mann-Whitney tests can be seen in Ludolf Gomes et al., (2012).

Non-parametric correlation coefficients (Spearman's correlation) were calculated to identify associations between microcystin concentrations and model factors in real-time at recreational and water treatment facilities in Ohio, USA, between 2016 and 2017 (FRANCY *et al.*, 2020). This same coefficient allowed for the verification of correlations between determining factors, which were regularly monitored over a decade, and different groups of phytoplankton in a highly urbanized subtropical lake in China (WEI *et al.*, 2023). Using monitoring data from urban reservoirs used for drinking water supply in southern China, it was found that cylindrospermopsin concentration was significantly correlated ($p < 0.01$) with the biomass of *Cylindrospermopsis* sp. (currently known as *Raphidiopsis* sp.), total cyanobacteria biomass, and environmental variables (nitrite and Secchi depth), but not with microcystin concentration, which only correlated with the biomass of *Microcystis* sp. (LEI *et al.*, 2014).

As an alternative, generalized linear models (GLMs) can offer a way to circumvent the difficulties of fitting traditional regression models, which have many assumptions. These models were successfully used to predict CyanoHabs events and MIB production using regular monitoring data. In the Ohio River, USA, hydrodynamic monitoring data allowed not only for prediction but also the development of a real-time alert, available via a web application, for elevated cyanobacteria concentrations, using logistic regression and Bayesian hierarchical modeling (NIETCH *et al.*, 2022). Quantile Regression and GLMs were used to identify MIB-producing species and assess odor risk in the QCS reservoir, located in the Yangtze river basin (China) (SU *et al.*, 2021). Quantile regression also yielded good results in studying nutrient influence on cyanobacteria abundance in Swedish lakes and in setting nutrient targets to sustain recreational and drinking water services, providing different levels of precaution for decision-making by the Swedish Agency for Marine and Water Management (LI *et al.*, 2021).

Although this review presented a significant sample of methods that do not require stringent assumptions, it is essential to encourage further research that provides suitable alternatives for modeling monitoring data. Considering the reality of water utility companies, where a significant portion of resources is dedicated merely to data generation rather than exploring new modeling methodologies, there is a need to focus efforts on methodological innovations in this context. Thus, collaboration between the academic sector and water utility companies can be an effective strategy to overcome these limitations and promote significant advances in water quality modeling and decision-making based on monitoring data. Furthermore, another challenge in modeling monitoring data is that, due to the nature of CyanoHabs, the frequency of regular monitoring data is rarely uniform. This fact restricts the use of certain techniques in a time series, a topic to be addressed in the next section.

2.4.2.3 The effect of frequency irregularity

CyanoHabs monitoring data typically exhibits irregular frequencies. This irregularity can be attributed to at least two factors: legal requirements or international recommendations, which mandate increased monitoring frequency when cyanobacteria concentrations reach potentially dangerous levels (i.e., 20,000 cells mL^{-1}) or when potentially toxin-producing taxa are detected (BRAZIL, 2021; WHO, 2022) and the high cost of maintaining consistently high frequencies of detailed phytoplankton analyses in the absence of health risks (BARRINGTON *et al.*, 2014).

As a result, to apply certain time series analysis methodologies, it is crucial to standardize the frequency of monitoring data. This can be achieved through resampling processes, known as downsampling or upsampling. In downsampling, the number of observations in the dataset is reduced. This process may involve summarizing the data by using the mean or sum of the values, effectively lowering the data frequency, as demonstrated by Kröger et al., 2023. These authors sampled irregular time series to obtain regularized series with monthly aggregated samples of biovolume data, quality parameters, and temporal beta diversity. Additionally, they obtained annual series of taxon incidences and average annual biomass per taxon. Among other analyses, after detrending the monthly time series using the Beveridge-Nelson decomposition method, the authors performed convergent cross-mapping to test causal relationships between biomass, total phosphorus, and soluble reactive phosphorus.

Alternatively, up-sampling addresses issues of irregular time series in monitoring data by increasing the data frequency. This process involves inserting records into the database using methodologies ranging from linear interpolation to machine learning techniques, similar to the process of obtaining daily data from monthly data as carried out by Zhang et al., (2015), which is detailed in the following sections.

Although resampling techniques can address problems with irregular time series, they must be used carefully. Down-sampling may result in a multi-year time series with sparse records, limiting the ability to represent complex phenomena such as CyanoHabs. Conversely, up-sampling can increase the risk of introducing bias or distortions if the interpolation method does not accurately raise the reality of the studied phenomenon. Given this dilemma, techniques that tolerate irregular frequencies and lagged data can yield good results.

For instance, using lagged data of 30 to 180 days and optimized multiple regression models, it was possible to predict total phosphorus concentrations in a reservoir by utilizing data from upstream reservoirs in a cascade system in the semi-arid region of Ceará, Brazil (GOES; BARROS; NETO, 2023). Logistic and Bayesian hierarchical regression using lagged flow data from the Ohio River (USA) enabled the development of real-time alerts and the prediction of CyanoHabs events (NIETCH *et al.*, 2022). Additionally, practical models using logistic regression were developed to predict the report cases of 2-MIB in Lake Ogawara (Japan) several months in advance with high accuracy, sometimes using only two water quality parameters (SHIZUKA *et al.*, 2021).

2.4.2.4 Multivariate techniques

Multivariate techniques were also employed in this SLR. This is largely attributed to the capability of these methods to simultaneously evaluate multiple variables and visualize complex patterns and trends. Principal component analysis (PCA) was applied to identify the primary parameters correlated with Chl-a and to develop pre-alerts in a SCADA system for real-time water quality monitoring in Lake Chao, China (MEZZANOTTE *et al.*, 2011). A PCA was also performed using different multivariate datasets of phytoplankton from Lake Stechlin, Germany, including monthly time series data on total biomass, covariates (temperature, Secchi transparency, TP, soluble reactive phosphorus, $NO2^-$, $NO3^-$, and $NH4^+$), and temporal beta diversity along with its nesting and replacement components (KRÖGER *et al.*, 2023). This analysis revealed that three principal components accounted for 82% of the original information from the 23-year monitoring database. Among other findings, it indicated that the increase in cyanobacteria was associated with the rise in TOC and Fe in the water, which are potential contributors to the deterioration of water quality in some temperate lakes in Sweden (LI *et al.*, 2021). Other studies also used PCA to identify factors associated with cyanobacteria (BERTRAND *et al.*, 2022; HWAN *et al.*, 2023; LUDOLF GOMES *et al.*, 2012; ZHANG; LIU, 2021).

In addition to PCA and Generalized Additive Mixed Models (GAMM), Redundancy Analysis (RDA) was used to reveal relationships between environmental variables, functional group biovolume, and operational periods of the Prospect Reservoir in Sydney, Australia. This analysis concluded that changes in hydrodynamics and extreme events, such as wildfires, impact water quality and phytoplankton dynamics (LUONG *et al.*, 2024). Canonical Correlation Analysis (CCA), a widely used multivariate ordination technique in ecology, correlated Actinobacteria and Cyanobacteria with the presence of T&O compounds in Lake Paldang, South Korea (HWAN *et al.*, 2023). Similar findings using ordination techniques can be found in Clercin and Druschel, 2019.

Another ordination technique, Non-metric multidimensional scaling (NMDS) based on the Bray-Curtis dissimilarity matrix was used to visualize changes in cyanobacterial communities in Lake Taihu, China, over a century, highlighting an increase in toxic cyanobacteria over the past 30 years (ZHANG *et al.*, 2023). Further multivariate ordination analysis in the Tianmuhu Reservoir, Southeast China, revealed that the concentration of T&O compounds varied significantly with phytoplankton succession and was influenced by hydrometeorological processes (WU *et al.*, 2022).

Alternative multivariate regression techniques can also be utilized to identify associations between monitored environmental variables and cyanobacteria taxa. For example, using Multivariate Curve Resolution–Alternating Least Squares (MCR-ALS), a form of multivariate linear regression, it was found that among 36 regularly monitored variables (22 physicochemical parameters and 14 cyanobacteria taxa), nutrient patterns (NH3, NO_2^-, NO_3^-, soluble reactive phosphorus) and the vertical temperature gradient within the water column were the most influential attributes determining the composition of cyanobacterial communities in the Grahamstown Reservoir, Australia (GOLSHAN *et al.*, 2020). Generalized additive models (GAMs) demonstrated the relationships of maximum cyanobacteria cell densities with transformed variables (z-scores) such as precipitation, surface temperatures in June, and total Kjeldahl nitrogen (TKN) in 20 seasonally stratified reservoirs in the United States (SMUCKER *et al.*, 2021). Among other results, the authors reported that lower cyanobacterial cell densities were associated with below-average summer precipitation, while higher densities were linked to average to above-average summer precipitation, warmer water temperatures in June, and higher TKN concentrations. Additionally, multiple linear regressions also yielded positive results in Goes et al., 2023 and Golshan et al., 2020.

2.4.2.5 Algorithm for estimating cyanobacteria

Occasionally, cyanobacteria may not be detected using only Chl-a concentrations. This limitation has driven the search for alternative methodologies to detect CyanoHabs using RS data by in situ measurements, being validated or not. According to this SLR, significant results in evaluating CyanoHabs were primarily obtained using characteristic pigments such as phycocyanin (PC), phycoerythrin, β-carotene, fucoxanthin, and other xanthophylls. These pigments are produced in situ and captured by satellite or specialized cameras. However, to extract useful information, these data must estimate some metric of cyanobacteria (e.g., cell density or biovolume) through specific algorithms.

Promising results have been observed with the use of a cyanobacteria estimation algorithm or Cyanobacteria Index (CI) as proposed by Lunetta et al., 2015a and Wynne et al., 2008. Generally, this algorithm utilizes spectral bands centered at 665 nm, 681 nm, and 709 nm to assess the raw biomass, while bands centered at 620 nm, 665 nm, and 681 nm are used as exclusion criteria to avoid quantifying other phytoplankton blooms. The efficiency of the CI algorithm in evaluating CyanoHabs also allowed for the observation of a stronger alignment at monthly frequencies than annual ones between RS and in situ data for Lake Okeechobee (USA) (REYNOLDS *et al.*, 2023).

The same algorithm was accurately used to estimate the “low” (10.000–109.999 cell mL^{-1}) and “very high” (> 1.000.000 cell mL^{-1}) cyanobacterial cell count in freshwater lakes (> 100 ha) in the eastern U.S by MERIS monitoring. Meanwhile, the “medium” (110.000–299.999 and “high” (300.000–1.000.000) ranges underperformed, possibly due to the lack of sensitivity for the algorithm, which is primarily attributed to the lack of taxonomic description from the traditional monitoring within these ranges (LUNETTA *et al.*, 2015). Moreover, a similar approach was developed to enable satellite data and large-scale field monitoring to assess CyanoHabs and microcystin risk in more than two thousand lakes across the U.S (HANDLER *et al.*, 2023).

Similarly, other studies have yielded good results based on Clark et al., 2017 and Coffer et al., 2020, including algorithms of distinct spectra for differentiating cyanobacterial taxa. For instance, the algorithm used to differentiate Aphanizomenon and Microcystis (Aphanizomenon-Microcystis Index - IAM) was based on the identification of cyanobacteria species using MASTER (Moderate Resolution Imaging Spectroradiometer - MODIS / Advanced Spaceborne Thermal Emission and Reflection Radiometer - ASTER) and HICO data (KUDELA *et al.*, 2015). In this case, IAM was used for the operational monitoring of small water bodies and facilitated the development of an early warning system in Pinto Lake, California (USA). Other algorithms for estimating Chl-a and PC concentrations adjusted for American lakes can be seen in Cook et al., 2023.

Comparatively, data from 23 aquatic ecosystems in the Iberian Peninsula that pointed out for the dominance by several cyanobacterial species were used to validate the MERIS images to estimate biovolume over a wide area with different trophic characteristics in freshwater under non-bloom conditions (Gómez et al., 2011; Medina-Cobo et al., 2014). Semianalytical algorithms for predicting Chl-a and C-PC concentrations in the eutrophic waters of Missisquoi Bay, Lake Champlain (Canada), were developed based on narrowband spectral data from MERIS aboard the ENVISAT satellite (WHEELER *et al.*, 2012). These authors used algorithms based on Gons (2004) for Chl-a and Simis et al., (2005) for C-PC. Despite highlighting the need for validation in other water sources due to the effects of cyanobacteria stratification along the water column, three central reservoirs in Indiana (USA)—Eagle Creek, Geist Reservoir, and Morse Reservoir—were used to validate and demonstrate an extension of the algorithm based on studies by Li et al., 2013. This algorithm aimed to improve remote prediction of PC (< 50 mg.m-3) and was developed to monitor cyanobacteria using RS techniques (LI; LI; SONG, 2015). A simple algorithm for detecting a bloom composed of cyanobacteria and chrysophytes in Lake Idro (Italy) from RS reflectance data derived from in situ measurements and MERIS data was developed based on the ratio 620 nm/560 nm, chosen according to the pigment characteristics of the local algal population (BRESCIANI, 2011).

Besides the previous examples, other publications presented accurate estimations of Chl-a and PC concentrations for temperate European lakes (PÉREZ-GONZÁLEZ *et al.*, 2021). Along with the satellite reports, other algorithms could be noticed to enable the use of hyperspectral data with very high temporal resolution (15 to 30 minutes) for CyanoHabs monitoring (Goyens et al., 2022). According to the review conducted in this study, algorithms enabling cyanobacteria estimates by hyperspectral (or multispectral) data are well advanced in traditional developed countries in the temperate zone compared to developing ones.

Nonetheless, along with the Western developed countries, China has been increasing its technology to monitor CyanoHabs. Chl-a values in a cyanobacteria-dominated region of Lake Dianchi have been obtained through remote monitoring using HJ-1A satellite data (ZHANG; WANG; CHANG, 2019). According to the authors, good estimates were made using monthly routine monitoring data conducted by the Kunming Environmental Monitoring Center, related to an algorithm developed by Yang et al., (2014), which combines different bands (b4-b3)/(b4+b3) as an indicator for Chl-a concentration in the same lake. Similarly, although using algorithms developed by Oyama et al., (2015), it was observed that more than 60% of the lakes and all reservoirs in the Yangtze River basin showed a trend of increasing cyanobacterial blooms influenced by climate change and human activities. This observation was carried out by using Landsat TM/ETM+/OLI satellite images, NDVI (Normalized Difference Vegetation Index) and EVI (Enhanced Vegetation Index), indices related to the greenish apparency of vegetation, Water-related Spectral Indices: LSWI (Land Surface Water Index), mNDWI (Modified Normalized Difference Water Index), Floating Algal Index: FAI (Float Algal Index), specifically developed to detect cyanobacterial blooms (ZONG *et al.*, 2019). In addition to this example, the use of multispectral images from the Landsat-8 OLI satellite and empirical multivariate regression models demonstrated reasonable accuracy for routine water quality monitoring in reservoirs, focusing on parameters such as turbidity, total suspended solids, and Chl-a in the Two Rivers Dam reservoir in Kenya (OMONDI *et al.*, 2023). The authors based their methods on procedures carried out by Ouma et al., (2018).

The lack of publications in arid, semi-arid, and tropical regions is not limited to cyanobacteria estimation algorithms. An evaluation of eutrophication using Chl-a and total phosphorus values was conducted in the Itaparica Reservoir, located in the Brazilian semi-arid region, using the Critical Load Model developed by the Organization for Economic Co-operation and Development in the Northern Hemisphere (GUNKEL; SOBRAL, 2013). Furthermore, the Carlson Trophic State Index, developed in temperate climate reservoirs, showed limited applicability for identifying the trophic state in reservoirs in Taiwan, with a tropical/subtropical climate (LIN *et al.*, 2022).

Therefore, this review suggests redirecting future research to developing regions, including tropical regions and arid and semi-arid areas. These regions often experience an intensification of CyanoHabs events due to abundant sunlight, high temperatures for most of the year, generally poor sanitation, and water scarcity. This shift could provide new tools for local sanitation companies and help mitigate the impacts of cyanobacteria on water supplied to vulnerable populations in these regions.

2.4.2.6 The Use of Machine Learning

Machine learning (ML) and artificial intelligence (AI) algorithms have recently been explored in managing CyanoHabs over the past 15 years. As highlighted in Table S2, at least two aspects justify this trend: i) the ability to handle complex non-linear relationships and ii) the unstructured data. These aspects enable the recognition of intricate patterns, a task that traditional statistical models struggle to accomplish, including the multiple linear regression and autoregressive integrated moving average (ARIMA) models, for example. As a result, more accurate descriptions of CyanoHabs dynamics are possible, typically non-linear and non-stationary phenomena due to the complex interaction of physical, chemical, and biological parameters (PARK; PATEL; LEE, 2024; RECKNAGEL *et al.*, 2013; ROUSSO *et al.*, 2020). However, the implementation of these techniques in Smart Cities (SC) depends on a robust monitoring system capable of providing data in adequate, regular, and enough monitoring programs. This may explain why this type of modeling was predominant in approximately 27% of the publications.

Almuhtaram et al. (2021) reliably detected cyanobacterial activity using high-frequency PC and Chl-a data as well as ML algorithms from 2014 to 2019 without the need for laboratory confirmation. Similarly, real-time online monitoring systems at Siling Reservoir (China) and Lake Winnebago (USA), which provided daily average values of cyanobacterial density and chlorophyll-a concentration, demonstrated high accuracy in predicting cyanobacterial and phytoplankton blooms (XIAO *et al.*, 2017). Compared to other forecasting methods, such as ARIMA models and other artificial neural networks (ANN), authors also noted that the wavelet neural network (WNN) model provided better forecasts using only one parameter (daily data of cell density or Chl-a, for example) obtained by a fluorescent probe. In a similar approach, WNN models proved promising for short-term forecasting of Chl-a concentration in Daecheong Reservoir (South Korea), outperforming other ANN models in terms of accuracy (KIM; SHON; SHIN, 2013).

As predicted by Almuhtaram et al., (2021) and Xiao et al., (2017), a considerable volume of data is required to model CyanoHabs. However, traditional monitoring programs by sanitation companies often lack appropriate high-resolution monitoring (HRM) devices. Besides reducing the ability to model complex patterns, this also compromises strategic decision-making. For example, using laboratory data alongside an HRM device (multiparameter probes), Bayesian network models assisted stakeholders and decision-makers at sanitation companies by providing data to cancel the development of nutrient pre-treatment facilities, saving approximately \$12 million in capital cost and about \$1 million in annual operating costs at the Grahamstown Dam, Australia (WILLIAMS; COLE, 2013). According to the authors, these probability-based models account for the inherent uncertainty in complex systems with limited data.

Similarly, a Bayesian network approach enabled the integration of future scenarios, process-based model outputs, monitoring data, and the national lake assessment system, providing a decision-making tool for Lake Vansjø (Norway), which has a history of eutrophication and cyanobacterial blooms (MOE; HAANDE; COUTURE, 2016). The authors reported that the use of these networks highlighted the importance of including more biological indicators in modeling environmental health, as including cyanobacterial biomass could lower the ecological diversity compared to assessments based solely on phytoplankton biomass. Thus, the Bayesian methodology offers a robust and reliable approach to addressing challenges in water quality management, often lacking sufficient data for precise point estimates. Also, this typically predicts the probability of different states, combined with the use of ecological indicators and indices to complement relevant biodiversity information.

Regarding the uncertainties in the monitoring and management of CyanoHabs Chen et al., (2014) integrated a self-organizing map (SOM) and diffuse information theory to classify the status of cyanobacteria blooms in Lake Taihu (China) into different categories across a wide spatial-temporal scale: none, light, moderate and severe blooms. Moreover, the authors concluded that the respective results could provide a deeper and more comprehensive view of the risk and uncertainty of algae blooms to the government and decision-makers, which can be useful in shaping policies and practices for monitoring and managing the risk of CyanoHabs. In addition to SOM, Hwang et al., (2023) evaluated eleven other algorithms to classify dominant algae in Juam and Tamjin Lakes in the Republic of Korea.

Apart from the limitation of data quantity, Bayesian models allowed for overcoming the impact of missing data in water quality analyses, a common issue in sanitation company databases. In the case study of the Limia River (Spain), the variability of missing data hindered the identification of underlying causes of aquatic cyanobacterial blooms. However, thanks to a methodological approach based on dynamic Bayesian imputation, it was possible to reliably approximate the missing values generated by HRM from sensors, thus improving water quality prediction (PAZO *et al.*, 2024).

Considering that critical states of high toxin concentration or blooms are exceptions, regular monitoring databases are generally unbalanced. This imbalance can compromise the training phase, limiting the generalization capacity of adjusted models. Thus, the Synthetic Minority Over-sampling Technique (SMOTE), which has shown superior performance in various applications, was used to mitigate data imbalance issues (JEONG *et al.*, 2022). According to the authors, the RF and XGBoost models used to predict CyanoHabs occurrences in Korean water supply reservoirs improved after applying SMOTE.

It is generally agreed that high-resolution data, such as hourly or daily frequency, are better for model development. From this, Zhang et al. (2015) linearly interpolated monthly monitoring data from 31 sampling sites in Lake Taihu between 2008 and 2012 to estimate daily values. These new data were used in a hybrid evolutionary algorithm, and the authors observed that spatially explicit models like the generic model were suitable for early warning of cyanobacterial blooms at most sampling sites, specifying the understanding of environmental conditions that promote cyanobacterial growth in Lake Taihu.

Another way to complement traditional monitoring is the use of unstructured data from hyperspectral images, enabled by deep learning techniques. This allowed for accurate estimation of optically inactive water quality parameters, a challenge in water quality monitoring using RS techniques, in the Guanhe River (China) (NIU *et al.*, 2021). The authors reported that combining laboratory data with hyperspectral images resulted in accurate spatial distribution maps, which can be useful for monitoring water pollution and assisting government departments in making decisions to prevent and control pollution in inland water bodies.

A significant advancement in water quality monitoring and management was observed in Daecheong reservoir (South Korea). The combination of real-time monitoring using IoT and deep learning through LSTM models, an ANN capable of "remembering" lagged information for long periods, demonstrated remarkable potential. This innovative approach to integrating water quality data, meteorological, and hydrological data, was effective in detecting and quantifying CyanoHabs due to the accuracy in simulating Chl-a and PC data (KWON *et al.*, 2023).

As the LSTM algorithm, which is capable of learning and retaining information from time series, the lag effects in modeling produced better results when using the Support Vector Machine (SVM) algorithm. Evaluating monthly monitoring data collected by the sanitation company, Xie et al., (2012) observed that 15 out of 23 water quality parameters showed a correlation greater than 0.3 with phytoplankton abundance in Macau's main storage reservoir (China). Using the SVM algorithm, the authors adjusted forecast models containing lagged data of up to three months and prediction models without lagged data. The authors reported that the short-term forecast model performed better than the long-term prediction model. This implies that CyanoHabs is a complex non-linear dynamic system affected not only by the water variables measured in a specific month but also by those from some previous months. Furthermore, compared to ANN, SVM showed superior forecasting power and similar predictive power, including when integrating support vector regression (SVR) with particle swarm optimization (PSO) (LOU *et al.*, 2015).

Regarding ecological metrics, such as abundance and taxonomic richness, only Kakouei et al., (2022) used them to increase the predictability of random forest models. Besides this, promising results were also obtained using indices and biological, climatic, physical, and chemical data collected from Lake Müggelsee in Germany during a long-term monitoring period from 2000 to 2017 (KRÖGER *et al.*, 2023). However, given the exceptional nature of the time interval and the resolution of the respective data, the authors emphasized that using ML tools with weekly biological sampling resolution data might be sufficient for making more robust and reliable ecological predictions.

Considering the potential of the presented results, the use of indices or ecological metrics in biodiversity studies offers a promising perspective for monitoring data in the evaluation of CyanoHabs, which still requires further exploration. Therefore, a holistic understanding of biodiversity is essential since it is a multifaceted concept that goes beyond the simple number of species (HAMILTON, 2005; HURART, 1971; JOST, 2009). In addition to changes in abundance (species richness), it is crucial to consider species distribution (evenness) over time and space (DÍAZ *et al.*, 2015). From this, data-driven decision-making in SC can consider the use of indices that measure true diversity, such as Hill numbers (HILL, 1973), based on the original idea of diversity measurement proposed by Macarthur (1965). Unlike the Shannon (H) and Simpson (D) indices (Shannon, 1948; Simpson, 1949), the Hill number qD can be used to quantify the loss or gain of biological diversity due to external factors, such as anthropogenic disturbances (Barbosa et al., 2020; Forlani et al., 2021). Additionally, by measuring α, β, and γ diversity at multiple scales, patterns of variation in true diversity can be compared to those obtained through commonly used indices (JOST, 2006).

Furthermore, in the evaluation of biodiversity, alpha, beta, and gamma diversity are fundamental for understanding the spatial context of biodiversity, as introduced by Whittaker (1972). Alpha diversity refers to species richness and evenness in a small area. Beta diversity represents the variation in diversity between different habitats. Gamma diversity is the total diversity in a large area or representative region (AKIRA S. MORI, FOREST ISBELL, 2018; JOST, 2007). However, this isolated approach can lead to the misconception that ecological communities are static, disregarding that all natural systems undergo constant compositional changes (MAGURRAN *et al.*, 2019). In this context, biodiversity is better assessed when the temporal scale is also considered, using temporal alpha and beta diversity, which can provide more information about changes in the monitored regions (LINDHOLM *et al.*, 2021). Therefore, based on our findings from this SLR, no studies applying Hill numbers were identified. Additionally, similar to the work performed by Kröger et al., (2023), the use of alpha and beta diversity at spatial or temporal scales to evaluate CyanoHabs in the Smart Cities context has been scarcely explored, representing an important gap that still needs to be addressed.

2.5 Concluding remarks and perspectives

In a scenario of water quality degradation due to CyanoHabs, which is expected to be intensified by climate change, water utility companies must minimize risks and ensure a safe supply. To achieve this, water and sanitation companies have access to regular monitoring data and infrastructure, including laboratories, sensors, and probes. Therefore, this infrastructure should be utilized innovatively, beyond traditional applications. In this context, a review of the primary alternative methods used for the collection, analysis, and interpretation of data on CyanoHabs identified several gaps for further research. Based on the reviewed studies, we can conclude and shed light on:

a) Complementary role of remote sensing (RS): RS can expand the spatiotemporal capacity of traditional monitoring with minimal additional resources. Alongside free image databases, adjustments in collection schedules will help the validation of algorithms using the existing routine of regular monitoring. However, the scarcity of studies from regions other than in the temperate zone may limit the accuracy of CyanoHabs estimation algorithms, similar to how Carlson trophic state indices are less effective in tropical regions when used indiscriminately. This lack of specific algorithms is a gap that is likely to be filled in the coming years, highlighting the need for further research in water management.

b) Importance of high-resolution monitoring (HRM): The adoption of HRM is crucial for achieving more accurate predictions and reducing biases associated with low-resolution monitoring. Furthermore, it allows for proactive management with automated alerts, integrating SCADA systems or IoT. While maintaining laboratories is advisable, an efficient ecosystem of sensors and probes can offer a more streamlined structure for water utilities. When used together, these resources provide a more robust and reliable approach to water quality monitoring.

c) Challenges of traditional techniques: The characteristics of non-linearity, heteroscedasticity, skewness, and autocorrelation in monitoring data can violate the assumptions of traditional techniques, such as linear regression, ANOVA, or t-tests, potentially leading to invalid or misleading information. To address this, data transformations or methods that do not require strict assumptions, such as non-parametric techniques, generalized linear models, and multivariate ordination methods, can be employed. However, due to the intensive routine of traditional monitoring, these practices present an additional challenge for companies and are less common in this sector than in academia.

d) Issues with irregular time series: Resampling techniques may correct issues related to irregular time series although reduce the accuracy of describing complex phenomena like CyanoHabs (downsampling) or increase the risk of bias and distortions if the interpolation does not reflect the reality of the phenomenon (up-sampling). An alternative is to use statistical and machine learning (ML) techniques that tolerate irregular frequencies or lagged data, allowing modeling to consider important information from previous periods.

e) Need for large data sets in ML: To minimize the risk of overfitting or underfitting, ML algorithms, such as neural networks or deep learning, typically require large amounts of data for training and validating the models, which often only HRM can provide. This may explain why, despite the ability to recognize complex patterns and use unstructured data, the number of publications primarily utilizing ML was lower than those employing statistical methods.

f) Emerging Equipment and Technologies: Data such as biovolume, cell density, phytoplankton composition, toxins and other metabolites are typically produced in laboratories. Therefore, new HRM equipment capable of monitoring these parameters in situ should be explored as an innovative perspective in this field.

g) Until these new technologies become widely available in the water sector, ecological indices can help describe biodiversity in scenarios with limited data. Additionally, statistical or ML models that predict probabilities can be an efficient solution to overcome the limitations of data quantity generated by traditional monitoring.

3 IMPACTO DA MAIOR SECA NO SEMIÁRIDO DESDE 1910: AUMENTO DA DENSIDADE DE CIANOBACTÉRIAS E HOMOGENEIZAÇÃO EM RESERVATÓRIOS PARA ABASTECIMENTO HUMANO, COM EFEITOS VARIÁVEIS NAS TOXINAS

IMPACT OF THE WORST DROUGHT IN THE SEMI-ARID REGION SINCE 1910: INCREASED CYANOBACTERIA DENSITY AND HOMOGENIZATION IN RESERVOIRS FOR HUMAN SUPPLY, WITH VARIABLE EFFECTS ON TOXINS

RESUMO

Regiões vulneráveis como o semiárido brasileiro são severamente impactadas por eventos extremos, como secas prolongadas. Neste estudo, analisou-se o efeito do pior evento de seca desde 1910 na biodiversidade de cianobactérias em 82 açudes destinados ao consumo humano no Ceará, Brasil. Para isso, utilizou-se números de Hill e séries temporais de 10 anos (2009-2019) de monitoramento rotineiro para descrever a biodiversidade, incluindo diversidade alfa e beta, além dos componentes de turnover e aninhamento. Observamos uma quebra estrutural (teste Zivot-Andrew) nas séries que alterou significativamente a biodiversidade de cianobactérias, resultando em homogeneização e aumento da densidade celular após os eventos extremos. Embora a maioria dos breakpoints não tenha afetado as concentrações de toxinas, algumas quebras nas séries da dinâmica composicional levaram a elevações em CYN e MC. O panorama revelado é preocupante, dado a grande presença de táxons potencialmente tóxicos, o aumento da densidade celular e a simplicidade das estações de tratamento.

Palavras-chave: CyanoHabs; NMDS; PERMANOVA; PERMDISP; Series temporais; Floração de Cianobactérias; Tratamento de água; Mudanças Climáticas; Crise Hídrica.

ABSTRACT

Vulnerable regions such as the Brazilian semi-arid area are severely impacted by extreme events like prolonged droughts. In this study, we analyzed the effect of the worst drought event since 1910 on cyanobacterial biodiversity in 82 reservoirs intended for human consumption in Ceará, Brazil. We used Hill numbers and a 10-year (2009-2019) time series of routine monitoring to describe biodiversity, including temporal alpha and beta diversity, as well as turnover and nestedness components. We observed a structural break (Zivot-Andrew's test) in the series that significantly altered cyanobacterial biodiversity,

resulting in homogenization and increased cell density following the extreme events. Although most breakpoints did not affect toxin concentrations, some breaks in the compositional dynamics series led to increases in CYN and MC. The revealed scenario is concerning, given the high presence of potentially toxic taxa, the increase in cell density, and the simplicity of the treatment stations.

Keywords: Water management; Climate change; Water crisis; CyanoHabs; Data analysis; Time series.

HIGHLIGHTS

a) A decade of monitoring of 82 dams in Ceará evaluated the biodiversity of cyanobacteria in human supply reservoirs.
b) There was an increase in the abundance and diversity of cyanobacteria, with few species dominating (most potentially toxin) indicating biotic homogenization.
c) The worst prolonged drought since 1910 has altered the biodiversity of cyanobacteria, with more relevant interannual effects than intra-annual ones;
d) The increase in cyanobacterial cell density after breakpoints requires more appropriate water treatment processes.
e) Most of the dams did not show significant variations in toxins after breakpoints, but there were elevations in some cases in CYN and MC.
f) Management tools based on the identification of bodies indicative of critical conditions are essential for the region.

3.1 Introduction

Over the past 50 years, anthropic activity has accelerated changes in nature, with land-use alterations, climate change intensification, and pollution driving a rapid decline in ecosystems. This resulted in a 20% drop in the abundance of native species and the homogenization of biological communities, contributing to the loss of local biodiversity (IPBES, 2019). Although the impacts of anthropogenic pressures are well documented in some vertebrate species (LINES *et al.*, 2023), forests (MAURE *et al.*, 2023), rotifers (LI *et al.*, 2022), insects (LIU; LIU; YANG, 2023) and fish (YANG *et al.*, 2024), little is known about the effects on microbial communities. A critical aspect is Cyanobacteria in water systems used for human supply, especially in vulnerable environments such as semiarid regions, which are particularly sensitive to climate change and face water scarcity scenarios. Moreover, this gap is most evident in long-term studies, such as decade-long continuous monitoring time series on a regional scale.

About 18.5% of the world's population resides in hyper-arid, arid and semiarid regions, which account for approximately one-third of the earth's surface. Ceará state, located in the northeast of Brazil, is one of these regions and its area is completely inserted in the Brazilian semiarid region. The region, which is home to approximately 8.7 million people, faces constant water scarcity and low levels of basic sanitation (IBGE, 2023). By 2021, only 37.1% of the sewage produced was treated, more than half of the urban solid waste (52%) was sent to 7 landfills and 29.5% of the municipalities did not have a rainwater drainage system (SNISB, 2023). These conditions, combined with the water scarcity in arid regions, restrict the ability to dilute waste (MILLENNIUM ECOSYSTEM ASSESSMENT, 2005) and urge great efforts for water quality and distribution.

Semiarid locals generally have fewer species than other tropical regions. Small changes in the quantity or distribution of organisms can have a significant impact on the biodiversity of their respective biomes (MAESTRE *et al.*, 2021). Therefore, considering the risks to the population and the potential deleterious effects on water treatment, understanding the dynamics of cyanobacteria in water bodies used for human supply is important from a One Health perspective (SVIRČEV *et al.*, 2014; WHO, 2022). This information can be valuable for sanitation companies in these regions, helping in the production and distribution of safe water and maintaining its quality.

Understanding local biodiversity is challenging due to the variety of existing metrics and definitions (JOST, 2009), that can be considered synonymous with biological diversity or simply diversity, being a multifaceted concept. It refers to the variability between living organisms and their ecological complexes, including variations in their taxonomic, genetic, phenotypic, phylogenetic, and functional attributes. Changes in abundance (species richness) and species distribution (evenness) in spatiotemporal perception are also considered (DÍAZ *et al.*, 2015). Therefore, to truly understand biodiversity, it is necessary to consider it as something more than just the number of species (HAMILTON, 2005).

An intuitive understanding is key to properly understanding biodiversity. In this context, the use of indices that measure true diversity or the effective number of species, such as Hill number (HILL, 1973), based on MacArthur's original idea (MACARTHUR, 1965), provides more comprehensible results than the entropy values obtained by some indices, such as the traditional Shannon/Shannon-Wiener (SHANNON,

1948) and Simpsons (SIMPSON, 1949) since Hill numbers have a set of mathematical properties that provides better expectations of diversity measurement and more adequately than entropies. The measurement of biodiversity is not determined by the specific functional form of the index but instead becomes dependent solely on the weighting given to rare or common species, which is characterized by the "order q" of the Hill number (JOST, 2006).

Still concerning the assessment of biodiversity, alpha, beta and gamma diversity are fundamental to understanding the spatial context of biodiversity, as presented by (WHITTAKER, 1972). Alpha diversity refers to the richness of species in a small area. Beta diversity is the variation in diversity between different habitats. Gamma diversity is the total diversity in a large area or region (AKIRA S. MORI, FOREST ISBELL, 2018; JOST, 2007). However, this isolated approach can lead to the misperception that ecological communities are static, disregarding the fact that nature undergoes constant compositional changes (MAGURRAN *et al.*, 2019). In this context, biodiversity is better assessed when the temporal scale is also considered, through temporal alpha and temporal beta diversity, including the components of turnover (rotation and replacement of species) and nestedness (loss of species without replacement) that contribute with more information about changes in the monitored regions (LINDHOLM *et al.*, 2021).

Over the years, several reports have associated the presence of cyanobacteria and their metabolites around the world with a series of adverse effects on human health (CHORUS, I., & WELKER, 2021) or the interruption of water supply services to the population, as occurred for approximately 500,000 inhabitants in Toledo in 2014 (Ohio, USA) (MASSEY *et al.*, 2018). These conditions are largely caused by concentrations of potentially dangerous cyanobacteria or cyanotoxins. In this context, Ceará state, as well as other arid and semi-arid regions, potentially toxin-producing strains can be found in most of the cyanobacterial taxa detected in the monitoring (BARROS *et al.*, 2017a).

Here, we present the first detailed investigation of a decade of cyanobacterial community monitoring in the region of Ceará, northeastern Brazil. Ceará, like other areas of the Brazilian semiarid region, faced severe droughts from 2010 to 2017 (MARENGO; CUNHA; ALVES, 2016; MARENGO ORSINI *et al.*, 2018), with the period from 2012 to 2016 being the worst drought since 1910 (FUCEME, 2016). To mitigate the effects caused by water scarcity, the state has an extensive network of artificial reservoirs that have as their main objective, but not exclusive by law (BRASIL, 2005), the supply of drinking water. Of these reservoirs, 157 are monitored by the Water Resources Management Company (COGERH), which classified, over the period analyzed, most of these reservoirs as (hyper)eutrophic (COGERH, 2023). In this context, the objectives of this study were: (1) to evaluate the structure and biodiversity of the cyanobacteria community in 82 dams used for human supply; (2) analyze the short- and long-term trends of this community; (3) investigate whether the worst drought event since 1910 significantly altered the trend of cyanobacterial biodiversity; and (4) examine the effects on cell density and cyanotoxin production before and after these extreme events and their implications for the region's water treatment. Therefore, the findings of this study expand our understanding of the dynamics of cyanobacterial communities in reservoirs intended for human consumption, considering the prospects of climate change in an environment marked by constant water deficit. In addition, the results provide an overview and valuable insights into how extreme weather events can influence the composition, dispersal, and toxic potential of these communities in semiarid regions.

3.2 Materials and Methods

3.2.1 Field of study

Data from 82 reservoirs intended for human consumption in Ceará, Brazil, were analyzed from 2009 to 2019. The state was divided into 5 macro-regions based on the administrative mesoregions defined by the Brazilian Institute of Geography and Statistics (IBGE, 2017). Each macro-region is home to the meteorological stations of the National Institute of Meteorology, responsible for official meteorological data (INMET, 2023). The region is part of the Brazilian northeastern semi-arid region, presents climatic diversity from tropical savannah (Aw) to arid-hot steppe (BSh) (PEEL; FINLAYSON; MCMAHON, 2007). Throughout the period analyzed, maximum temperatures remained high, and the total rainfall varied between 300 and 1200 mm/year. Historical hydrological dynamics reveal a rainy season from January to May and a dry season from June to December. The environmental variables show a moderate gradient throughout the state (see Section 1 of the Supplementary Material (SM) - APPENDIX B).

3.2.2 Data

The data were divided into three groups: climate data, monitoring data, and hydrobiological data. The manipulation was performed using RStudio (R CORE TEAM, 2023).

3.2.2.1 Climate data

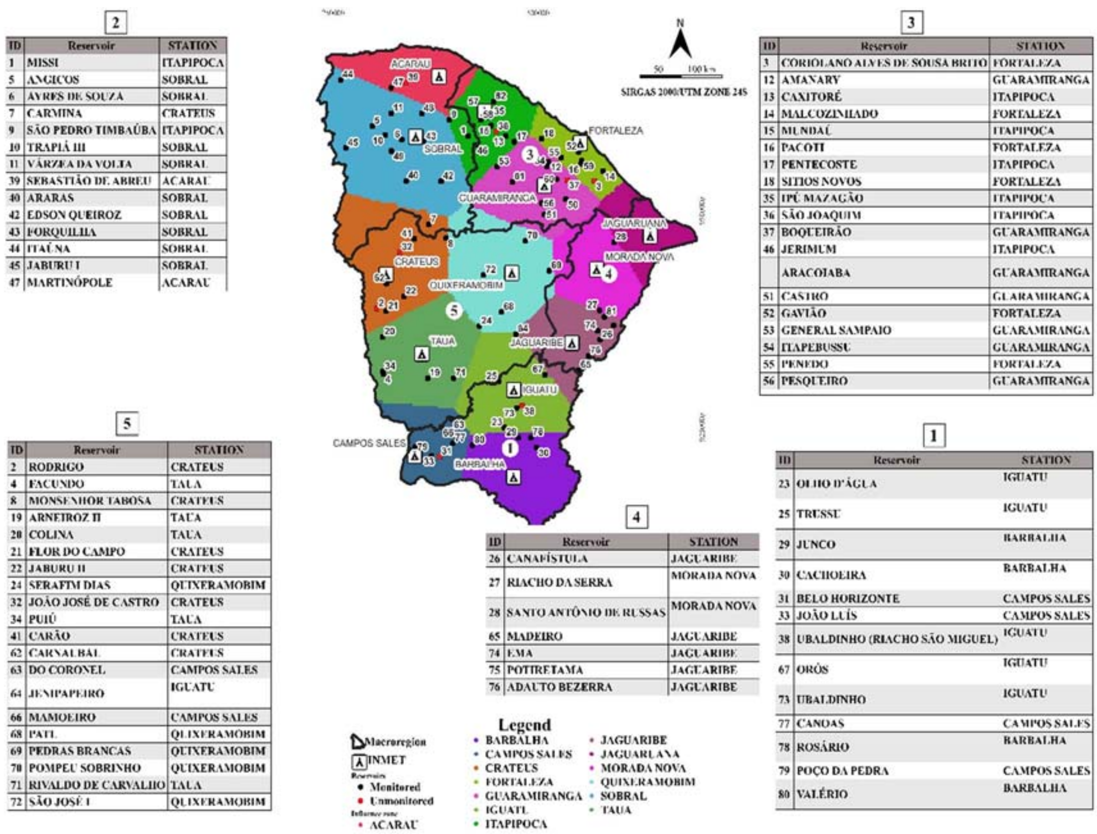
The climatic data were used to obtain the environmental gradient of maximum temperature (ºC), thermal amplitude (ºC), global radiation (kj m^2), wind speed (m s^{-1}) and precipitation (mm) and were obtained from the National Institute of Meteorology (INMET, 2023). For more details, refer to Section 1 of SM - APPENDIX B.

3.2.2.2 Monitoring data

Monitoring data was considered only for the reservoirs monitored by COGERH (Figure 6). Therefore, percentage data of total volume (% Volume), which includes 73 out of 82 dams studied (COGERH, 2023).

Figure 6 - Location and definition of 14 official meteorological stations of INMET and the list of reservoirs intended for water supply (monitored or not by COGERH) in these areas . (For better visualization, see Section 1 in the SM - APPENDIX B)

.



Source: prepared by the author

3.2.2.3 Hydrobiological data

For the hydrobiological data, the routine monitoring data for the 82 reservoirs studied was performed by the Water and Sewage Company of the State of Ceará were considered. General details of the dams and water treatment plants (WTP) can be found in Section 2 of SM. The water sampling was considered closest to the water uptake, according to the company's specific routine for each WTP. Cell counts were conducted with an inverted microscope (Zeiss Axio, A1), counting at least 100 individuals or colonies per sample. Phytoplankton were identified to the species level when possible. The protocol followed (APHA-AWWA-WEF, 2005) and Brazilian regulations, which also defined the analysis frequency based on cyanobacteria density and required toxin monitoring when necessary (BRAZIL, 2021). More details and the list of the main organisms identified throughout the decade is available in Section 2. Enzyme-linked immunosorbent assays (ELISA) were used to quantify total microcystin (MC), saxitoxin (STX) and cylindrospermopsin (CYN) in raw water.

3.2.3 Biodiversity assessment

3.2.3.1 Rarefaction curves

Rarefaction curves were used to evaluate the adequacy of the sampling effort in the evaluated reservoirs for cyanobacterial biodiversity analysis. For each reservoir, two graphs were built, considering the accumulated richness (y-axis) as a function of the number of available observations (x-axis) and the accumulated richness (y-axis) as a function of the number of individuals (x-axis). In both curves, the sampling effort was considered adequate when:

a) The first part of the curve showed a rapid rise, corresponding to the discovery of new individuals. This part was considered to be up to the first quartile (Q1).
b) The second part of the curve shows a stabilizing trend, indicating a reduction in the rate of discovery of new individuals. Data from Q1 were considered for this part of the curve.

A stabilization trend was considered when 100 observations on average were needed for the addition of a maximum of 5 units in the accumulated richness. This was evidenced by the formation of a plateau or discrete growth on the charts. The accumulated richness values were obtained by specaccum script in the Vegan package (OKSANEN *et al.*, 2022) following its 95% confidence intervals from 100 permutations. The numbers of new individuals were determined by linear regression models and all the curves of the 82 dams are available in Section 3 of SM - APPENDIX B.

3.2.3.2 Hill numbers

Hill Numbers (HILL, 1973) were used to calculate the real diversity of the reservoirs, according to Equation 1 (JOST, 2006):

$$^{q}D = \left(\sum_{i=1}^{S} p_i^q\right)^{1/(1-q)} \qquad \text{(Eq. 1)}$$

Where:

a) q: parameter known as the order of diversity, which means variation according to the frequency of the individuals identified in each sample.

b) S: sample richness, defined as the number of distinct taxa found in each monitoring effort.

c) *Pi*: relative abundance of individuals from each taxon, calculated by the proportion of individuals identified from a given taxa by the total number of organisms identified in the monitoring.

Based on this, we have:

a) q = 0 : in order 0, the frequency of the collected individuals is not considered. Then, the number of distinct individuals in the sample is quantified, which represents the richness of the sample.

b) q = 1: in order 1, each individual is given weight based on the proportion of its frequencies, this order is equivalent to the exponential Shannon entropy (JOST, 2006; SHANNON, 1948), also known as the Shannon–Wiener index transformation.

c) q = 2: in order 2, higher weights are assigned to the most common individuals, which represents the inverse of Simpson's entropy (JOST, 2006; SIMPSON, 1949), also called the Gini-Simpson index transformation.

d) q > 2: the importance of rare species decreases as the order of q increases > 2. With this, it is possible to evaluate the behavior of biodiversity loss by reducing the importance of rare organisms.

Thus, when the order of diversity is q ≤ 1, rare species are favored, due to the relevance given their low relative frequencies. On the contrary, when the order of diversity is q > 1, more common species gain relevance. The results were presented in curves with the biodiversity (Hill Number) as a function of the order q. To evaluate how the reduction occurred, that is, whether the decrease in biodiversity was abrupt or not, the curves were drawn up to order 5.

Given that 82 reservoirs were studied during the period 2009 to 2019, the curves were summarized in 3 different periods. In each case, these periods allowed us to assess the existence of patterns in:

a) general behavior: evaluates in a general way the average of all available observations.

b) annual behavior: evaluates long-term behavior by averaging all observations of a given year.

c) monthly behavior: evaluates the intra-annual or seasonal effects through the average of a given month over all available years.

Mostly, annual and monthly behavior curves, a pattern of uniformity or homogeneity of the distribution of abundance of individuals in the sample was observed. To do so, the percentage of diversity loss between order 0 and orders > 0 were compared.

In the graphs of annual and monthly behaviors, three patterns were observed:

a) Reduction: which occurred when all Hill numbers showed a significant reduction trend during the period evaluated. This behavior was observed when both richness (order 0) and evenness decreased (order > 0).
b) Elevation – occurred when all Hill numbers showed a significant growth trend during the evaluated period. This has been verified when both richness and evenness increase.
c) Inconclusive - occurred when Hill numbers of order 0 or order greater than 0 do not show a significant trend.

Increasing tendencies were identified when the slopes of the lines, adjusted by linear regression, were positive and significant up to 10% (α). On the other hand, when the coefficients were negative and significant, it was interpreted as a downward trend. Hill numbers were calculated using the *Renyi*() Package *Vegan* R (KINDT; OKSANEN, 2022) and the linear regressions were adjusted using the *Lm*() Package *Stats* (more details in Section 4 of SM - APPENDIX B).

3.2.3.3 Temporal alpha and beta diversity

Changes in biodiversity, i.e., how the abundance and taxonomic identities of cyanobacteria varied between 2009 and 2019 in the 82 dams, were evaluated through temporal α diversity, temporal β diversity, and its components of Turnover and Nesting (MAGURRAN *et al.*, 2019), also called "the biodiversity series". To quantify the changes in the cyanobacteria community related to abundance and diversity, represented by the number and distribution of species, temporal α diversity was used (DORNELAS *et al.*, 2014). To understand the changes related to the dynamics of the compositional variation of the assemblies and their hierarchical structures, the temporal β diversity and its components of turnover and nesting were used. This was done to improve the understanding of biodiversity because, although widely adopted, Hill numbers may not be the best solution for all biodiversity analyses, especially where ecological processes are non-linear, such as the potential effects of environmental changes caused by long periods of drought (RICOTTA; FEOLI, 2024). To avoid constraining the understanding of ecological complexity by enforcing linearity in all diversity metrics, the time series representing biodiversity were analyzed.

All analyses were considered on the monthly scale, summarized by the mean for the months with more than one observation. The order values 1 of the Hill series were used to quantify the temporal α diversity component (JOST, 2007). The components of temporal beta diversity, turnover, and nesting were obtained through the *beta.multi* of the R betapart package (BASELGA *et al.*, 2023; BASELGA; ORME, 2012), similar to the works reviewed by MAGURRAN et al. (2019). For this, matrices of the presence and absence of cyanobacteria composition and the Sørensen dissimilarity index family were used. The value of the turnover component was measured by Simpson's dissimilarity. The nesting component was calculated by the difference between the Sørensen and Simpson indices. The time series graphs describing biodiversity can be seen in Sections 5 to 7 of SM - APPENDIX B.

3.2.4 Statistical analysis

During the analyzed period, it was found that some reservoirs dried up (% Volume = 0). The monthly mean % Volume for the 73 monitored reservoirs was used to identify these periods. After determining the moments when % Volume = 0, a data imputation process was applied using the moving average technique with 6 lags, focusing only on periods where % Volume > 0. After imputation, "time series that describes biodiversity" (temporal alpha diversity, temporal beta diversity, and the turnover and nestedness components) without missing values were evaluated.

3.2.4.1 Zivot-Andrew test: evaluation of the time series that describes biodiversity

To investigate the impact of the prolonged dry season on the biodiversity of cyanobacteria in reservoirs intended for human supply, the Zivot-Andrew (ZA) test was applied (ZIVOT; ANDREWS, 1992). This test made it possible to identify whether the time series that describes biodiversity contained a unit root with structural change, as indicated in other studies (LIN; ULLAH, 2024; QUAN *et al.*, 2024). The test evaluated whether there was a significant change in the structure of the series during the drought period. According to the ZA test, time series can be described as: (i) Growth trend: when the trend coefficients (γ or β) are significant ($p < 0.05$) and greater than zero; (ii) Decreasing trend: when the coefficients are significant and less than zero; (iii) Random behavior: when no significant effects are observed in either the base level or the trend, before or after the breakpoint. Details of the coefficients and equations are provided in Section 8 of SM - APPENDIX B.

Thus, according to the change in the behavior observed after the breakpoint, the time series that describe the biodiversity of cyanobacteria were subdivided according to the observed effects, as shown in Table 1.

Table 1 - List of observed effects on biodiversity series behavior after the breakpoint. The coefficients α, β, δ, and γ refer to the ZA test, and NS means “No significant”

Trend	ZA test	Breakpoint effect on trend or baseline level	Effect
Increasing α or γ > 0	β > 0 or δ > 0	Elevation in the magnitude of an increasing trend or baseline	Elevation
	β < 0 or δ < 0	Reduction in the magnitude of an increasing trend or baseline	Reduction
	β or δ → NS	Retention of a similar increasing trend or baseline level as before the breakpoint	Retention
	α or γ → NS β > 0 or δ > 0	Initialization of an increasing trend or baseline level after the breakpoint	Initialization
Decreasing α or γ < 0	β < 0 or δ < 0	Elevation in the magnitude of a decreasing trend or baseline	Elevation
	β > 0 or δ > 0	Reduction in the magnitude of a decreasing trend or baseline	Reduction
	β or δ → NS	Retention of a similar decreasing trend or baseline level as before the breakpoint	Retention
	α or γ → NS β > 0 or δ > 0	Initialization of a decreasing trend or baseline level after the breakpoint	Initialization
Random	α, β, γ or δ →NS	No significant effect	Non-significant

Source: prepared by author

‘

The series were also decomposed additively into a trend (t_t), annual seasonality (s_t) and error (e_t), according to Equation 2. The estimate was made by LOESS *(Locally Estimated Scatterplot Smoothing*) using local polynomials. The estimate of s_t also uses LOESS, but for an detrended series, i.e., original data from the series subtracted from t_t. Only monthly variations during the year were considered. The e_t estimate was made by the difference between the value of the series and the estimates of t_t and s_t. The ZA test was performed using the *ur.za* Package *Urca* (PFAFF, 2008). The decomposition was carried out by the function *STL* Package *feasts* (O’HARA-WILD; HYNDMAN; WANG, 2023). The quality of the models was evaluated by the graphs of the residuals, Residual I for LOESS (Equation 2) and Residual II for ZA Test (Equation 2 in Section 5 of SM). The time series graphs describing biodiversity can be seen in Sections 6 to 8 of SM - APPENDIX B.

$$y_t = t_t + s_t + e_t \quad \text{(Eq. 2)}$$

3.2.4.2 Consequence of breakpoints in the species composition matrix, cell density and cyanotoxins

To evaluate whether the breakpoints caused significant changes in the composition and structure of the cyanobacterial community, considering cell density and toxin concentrations in the raw water, the values of the species composition matrix, the cell density data and the CYN, MC and STX concentrations were divided into two groups: i) before and ii) after the breakpoints identified in the time series that describe the cyanobacterial biodiversity in each of the Dams. Based on these groups, the procedures described in the following sections were carried out.

3.2.4.2.1 Effect on the species composition matrix

To identify the differences between these groups in the species composition matrices, PERMANOVA and PERMDISP tests were performed using the *Adonijah2 and Permutest,* respectively*,* by *vegan* package (OKSANEN *et al.*, 2022). PERMANOVA and PERMDISP were used to identify the effects of localization vs. dispersion, respectively, in the space of the chosen similarity measurement (Bray–Curtis dissimilarity). Before all the analyses, to reduce the negative effects of the imbalance, the amount of data in the larger group was approximated to that of the smaller group (with more than 5 observations) from the breakpoint (ANDERSON; WALSH, 2013). In these analyses, location can be understood as the mean (or centroid) of the groups, i.e., the mean of the cyanobacteria composition. Dispersion refers to variation within groups, that is, whether there has been a reorganization in the structure and dispersion of communities. Thus, when both tests were significant ($p < 0.05$) the structural break in the series significantly altered both the composition (Location) and the structure of the cyanobacteria communities (Dispersions) and the weir was classified as "both" for the evaluated component. When not, the weir was classified as "Composition" or "Dispersions" in each component, if only PERMANOVA or PERMDISP was significant. A non-metric multidimensional scale (NMDS) was used to visualize the patterns of each group in two dimensions. Although with some distortions, graphs with stress values around 0.2 were considered to evaluate the general patterns and differences between the groups (see Section 9 of SM - APPENDIX B).

3.2.4.2.2 Effect on cell density and toxins

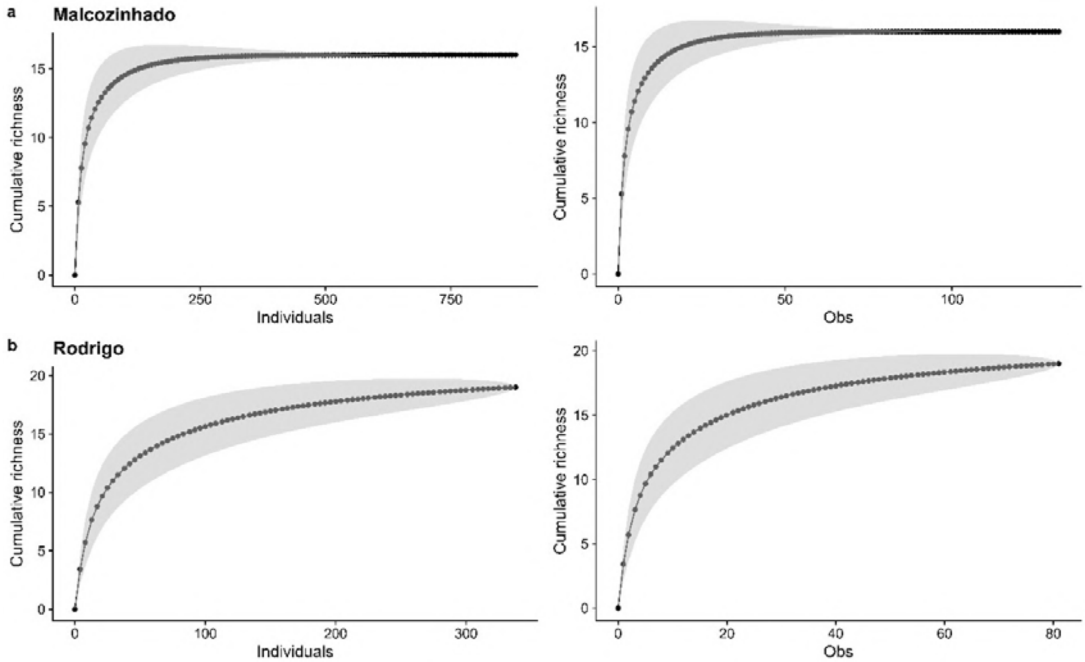
At a significance level of 5%, the two-tailed Wilcoxon rank sum test was performed to verify significant non-directional differences, since both increases and decreases occurred between the groups. For this, the *wilcox.test* function of the *stats* package was used. The increases or decreases following the breakpoint in significantly different groups were determined using boxplot analysis and comparisons of the mean and median values for each group. All results for all weirs are found in SM - Section 10 - APPENDIX B.

3.3 Results

4.3.1 Assessing biodiversity by Hill numbers

The sampling effort carried out during the monitoring period of Cyanobacteria was sufficient to adequately represent the biodiversity of the raw water in all 82 reservoirs studied, considering the taxonomic classification by genera level. This is because a stabilizing trend was observed in the two patterns identified from the rarefaction curves (Figure 7). Thus, similar to the one seen for the “Malcozinhado” dam (Figure 7a), there was plateau formation for a larger group of 61 dams (74.4%). For the other 21 reservoirs (25.6%), there was a slight growth (Figure 7b), similar to the “Rodrigo” dam. Additionally, as the sampling effort was different, the analyses were carried out separately in each reservoir to avoid disproportionate comparisons between the water sources studied (see Section 3 of SM - APPENDIX B).

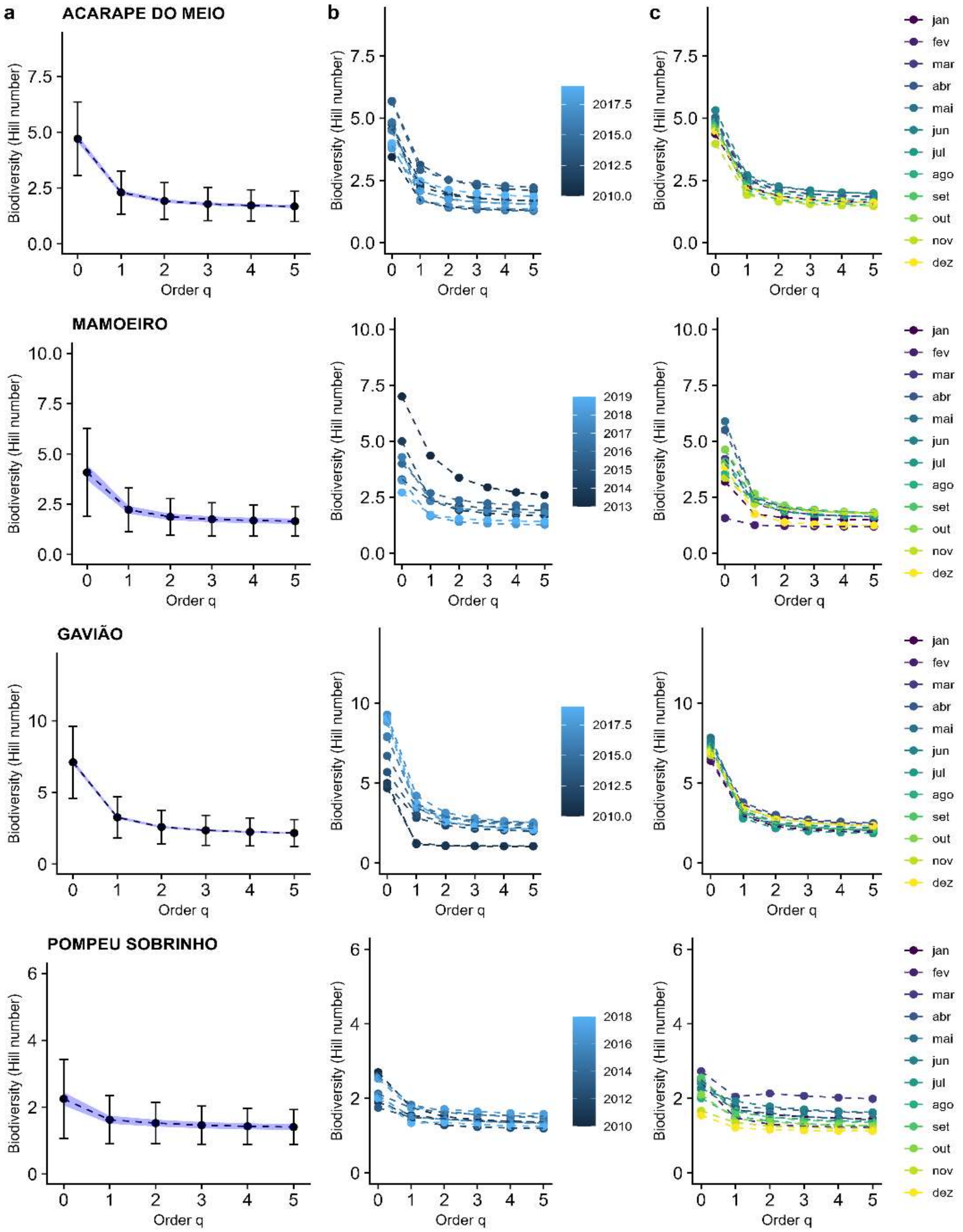
Figure 7 - : Patterns observed in the rarefaction curves of the dams in Ceará used for human supply. The pattern in (a) shows the formation of a plateau, while (b) shows a slight increase, both for curves of accumulated cyanobacterial richness as a function of the number of individuals, and for accumulated richness of cyanobacteria as a function of the number of observations used in the analyses. The accumulated richness was obtained by averaging 100 random permutations of the data and in gray is shown its 95% confidence interval.



Source: prepared by the author

In the evaluation of both species richness (order 0) and evenness (order > 0), the Hill numbers for the 82 dams in Ceará were similar to those observed in the Figure 8 (check the other reservoirs in Section 4 of SM - APPENDIX B). In global mean (Figure 8a), monthly (Figure 8b) and annual (Figure 8c) there was a reduction of around 50% of biodiversity from order 0 to orders greater than or equal to 1 in the Hill Series. This may be associated with the uneven distribution observed in cyanobacterial taxa of raw water used for human supply. This reduction occurred mainly due to two characteristics of the species composition matrices obtained in the routine monitoring of all water sources: (i) the presence of a substantial number of rare taxa, or those with only a single occurrence, increased the Hill number values of order 0, representing cyanobacterial taxa richness, but had minimal impact on orders greater than or equal to 1, which reflect evenness. (ii) the existence of a few common taxa with high relative frequencies. These organisms have a greater influence on the evenness represented by orders greater than 0, but as the number of taxa with these characteristics is limited, the observed reduction in biodiversity occurred, which stabilizes from order 1. Therefore, based on these results, there is a strong indication of the dominance of frequent cyanobacterial taxa in the raw water collected from the 82 studied dams.

Figure 8 - Cyanobacterial Biodiversity in 4 of the 82 Dams Used for Human Consumption Based on Hill Numbers: (a) General Means, (b) Annual Means, and (c) Monthly Means.



Source: prepared by the author

Additionally, as presented in Section 4 of SM, trends in the means were observed on different time scales. In the long term, the interannual mean of the Hill numbers indicated a trend of increase in cyanobacterial biodiversity in 16 reservoirs (19.5%), a decrease in 3 (3.7%) and, in the remaining 63 (76.8%), no significant trend was found ($p > 0.1$). In the intra-annual period, the seasonal trend, represented by the monthly averages, revealed a reduction in biodiversity in only 4 springs (4.9 %) and, in the other 78 reservoirs (95.1 %) no significant trends were observed ($p > 0.1$).

On average each year, species richness increased approximately 0.3 to 0.7 in the 16 reservoirs with a significant increase in cyanobacteria biodiversity, while in the 3 reservoirs with a significant reduction trend, ranged from 0.3 to 1. For the monthly variation, the richness of cyanobacteria decreased by around 0.08 to 0.18 each month. For evenness (orders > 0), both the interannual and intra-annual behavior were similar, with the values of orders > 0 remaining around 50% of the values of order 0. In addition, no evident pattern of distribution of dams was observed with a significant trend along the environmental gradients studied. Therefore, considering that more than 95% of the dams did not show a significant intra-annual trend, there are indications that the long-term effects were more relevant for the cyanobacterial biodiversity in the 82 reservoirs than the seasonal ones.

3.3.1 Assessment of biodiversity by series describing biodiversity

The analysis of the quantitative variation (dynamics of abundance and relative frequency) of cyanobacteria was performed in the 82 dams by the time series of the temporal alpha diversity. Despite attempts to correct it, missing data were still found in 23 reservoirs (28.1%). It is believed that the most likely cause for this is the fact that these weirs have dried up, since % Volume = 0 in the monitored weirs (Section 5 of SM). Thus, significant changes in temporal alpha diversity were evaluated in 59 of the 82 dams (71.9%).

In the study of the compositional dynamics of biodiversity, 41.5% of the dams (34 out of 82) were evaluated. In these reservoirs, 34 series of temporal beta diversity, 34 of the turnover components and 34 of the nestedness component were evaluated, totaling 102 series that described the behavior of the cyanobacterial compositional dynamics of these dams. In addition to the periods in which the dams dried up (% Volume = 0), the same data imputation process resulted in a lower number of reservoirs evaluated than the temporal alpha diversity. This was due to periods with an insufficient number of observations in the taxon composition matrices for the estimation of temporal beta diversity and its components (Sections 6 and 7 of SM). Therefore, a total of 161 series were analyzed. In Table 2, the results of the ZA test for the "Acarapé do Meio" dam, one of the 82 dams studied can be found, and the other dams are found in Section 8 of SM.

Table 2: Effects on the base level and trend of the alpha and beta temporal diversity series and the turnover and nestedness components in the Acarape do Meio dam before and after their respective breakpoints (BP). The coefficients α, β, γ and δ were obtained by the ZA test. In parentheses, the 95% confidence interval. Bold values are non-meaningful.

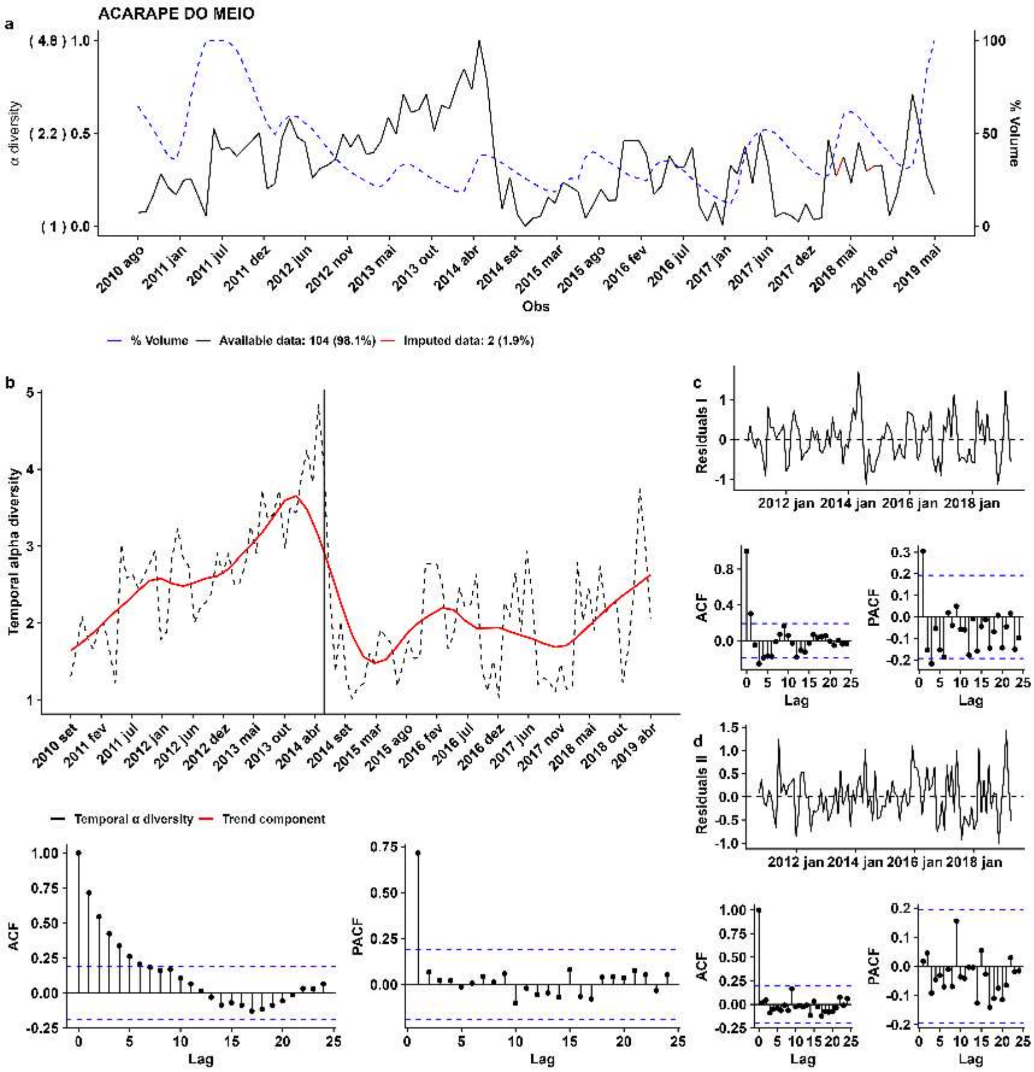
Time Series	BP	Baseline level: Before (BP) α	Baseline level: After (BP) β	Trend: Before (BP) γ	Trend: After (BP) δ	Series Behavior Breakpoint effect: Baseline	Series Behavior Breakpoint effect: Trend
Temporal α diversity	2014 Jun.	1.01 (0.61,1.42)	-1.47 (-1.99, -0.95)	0.03 (0.02,0.04)	-0.02 (-0.04, -0.01)	Increasing Reduction	Increasing Reduction
Temporal β diversity	2016 Feb.	0.18 (0.12,0.24)	0.12 (0.04,0.20)	0.003 (0.002,0.004)	-0.004 (-0.006, -0.001)	Increasing Elevation	Increasing Reduction
Turnover component	2016 ago.	***0.022 (-0.03,0.08)***	-0.14 (-0.24, -0.04)	0.005 (0.003,0.006)	***0.001 (-0.003,0.006)***	*Decreasing Initialization*	*Increasing Retention*
Nestedness component	2016 Jul.	0.123 (0.07,0.20)	0.158 (0.07,0.24)	***0 (-0.002,0.001)***	-0.004 (-0.008, -0.001)	Increasing *Initialization*	*Decreasing Initialization*

Source: prepared by the author

Figure 9 presents the time series of the "Acarape do Meio" dam, where both alpha temporal diversity and reservoir level (% Volume) were evaluated (Figure 9).

Following the breakpoint, marked by the black vertical line (Figure 9b), a decrease in the baseline magnitude of the series is observed (see $\alpha > 0$ and $\beta < 0$ in Table 2). The baseline here refers to the average behavior of the values before and after the breakpoint. Additionally, a clear reduction in the growth pattern of the series is noted (see $\gamma > 0$ and $\delta < 0$ in Table 2), meaning the magnitude of the growth trend diminished. Thus, for temporal alpha diversity, the evidence suggests that external factors, occurring around June 2014, significantly altered the elevation behavior of the series.

Figure 9 - In (a) the series of temporal alpha diversity and %Volume is presented. The values have been rescaled to improve the visualization of the series. In (b) the series of the temporal alpha diversity (dotted black line), the trend component obtained by the decomposition using LOESS (red line), the breakpoint (vertical black line) and the ACF and PACF graphs of the raw data are shown. In (c) the residues obtained by the LOESS decomposition are presented and in (d) the residues after the adjustment of the ZA test, both with their respective ACF and PACF graphs.



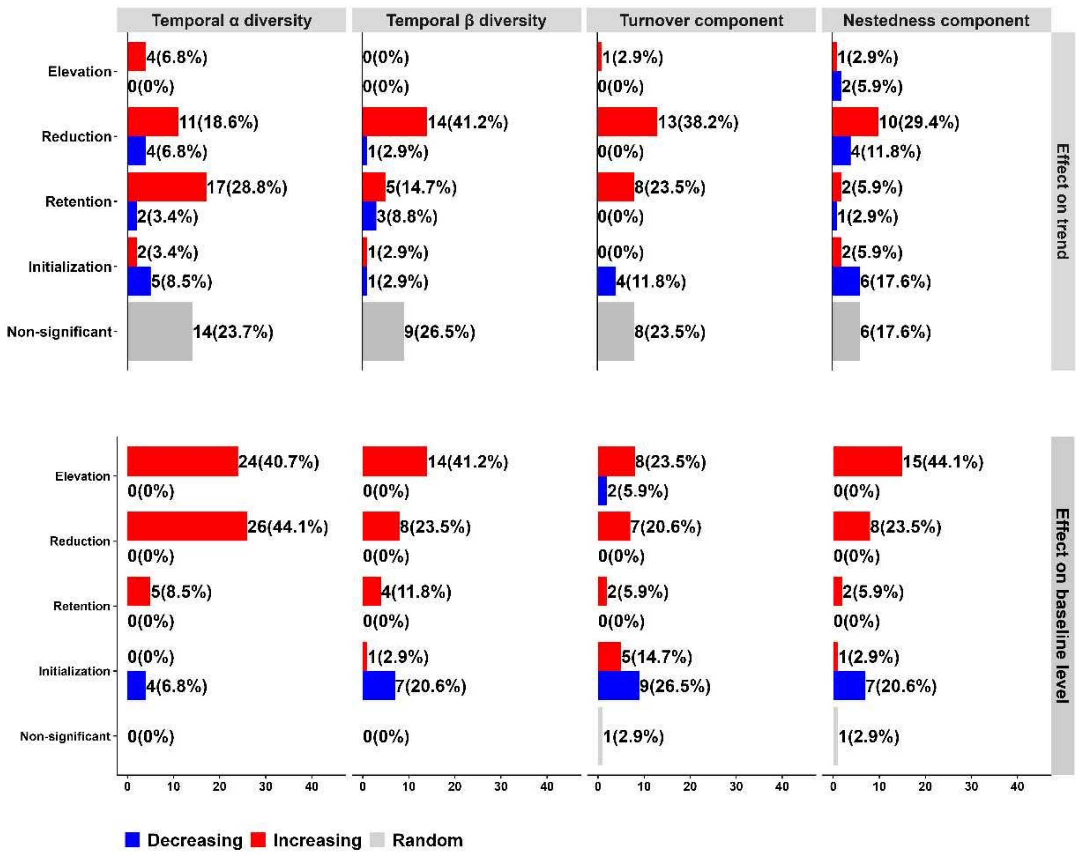
Source: prepared by the author

Similarly, breakpoints caused by exogenous factors also significantly altered the behavior of the series that evaluates the compositional dynamics. For the series of the temporal beta diversity of the “Acarape do Meio” dam, the breakpoint produced two different effects on the growth behavior. For the baseline level, there was an increase, whereas for the trend the effect was a reduction in the magnitude of growth. For the turnover component, the series only showed a reduction in the magnitude of the base level after the breakpoint and the upward trend was not significantly changed. For the nestedness series, the effects were an increase in the base level and an upward trend was only started after the breakpoint. All graphics similar to Figure 9 for all 161 series that were presented in Sections 5, 6, 7 of SM may be interpreted similarly. Therefore, even occurring in different periods, as more than 95% of the breakpoints occurred between the second half of 2011 and the second half of 2017 (see Section 8 of SM - APPENDIX B), there are strong indications that the worst period of prolonged drought since 1910 that occurred in Ceará significantly changed the behavior of the cyanobacteria community, affecting both the compositional dynamics and diversity of cyanobacteria in the studied dams.

In addition, considering the correlogram of the residuals, the autocorrelation (see the ACF and PACF graphs of the Figure 9b) was better removed by the adjustments of the ZA test (Residual II in Figure 9d) than when both the trend and seasonality components were considered using LOESS (Residual I in Figure 9c). This means that a large part of the behavior of the series is due to the existence of deterministic tendency, not seasonality. In general, this also occurred for the other reservoirs, which corroborates the fact that only 4 of the 82 reservoirs showed an intra-annual trend in cyanobacterial biodiversity in the previous section (Section 4 of SM - APPENDIX B).

Throughout the analyzed period, the dynamics of cyanobacterial richness and abundance (alpha-temporal diversity) and the changes in both the composition of taxa and the hierarchical structure of cyanobacterial communities (beta-temporal diversity and the components of Turnover and Nestedness) were characterized by systematic and non-random changes that were significantly altered by external factors. Regarding the stationarity of the 161 series, for the alpha temporal diversity, only 4 of the 59 reservoirs (Do Coronel, Forquilha, Martinópole and Riachão) were considered non-stationary. In addition, only the series of temporal beta diversity of the “Forquilha” dam and the series of the turnover components of the “Madeiro” and “Forquilha” dams 3 out of the 102 series that describe the compositional dynamics of cyanobacteria were non-stationary, totaling 7 in 161 series. The lack of observations in these series probably contributed to the non-stationarity observed (Sections 5-8 of SM - APPENDIX B). Therefore, for the other 154 series, the mean and variance tend to remain constant over time, with their values oscillating around a baseline level. However, due to the existence of a deterministic tendency, changes resulting from external factors over time have defined the pattern of reduction or increase in temporal alpha and beta diversity and its components. This pattern was significantly altered by the breakpoints that occurred during the long dry period. Figure 10 summarizes the effects of breakpoints previously presented in Table 1 for all 161 series describing the dynamics of cyanobacterial biodiversity. In general, the growth behavior was the most observed in all the evaluated series.

Figure 10 - Summary of the effects of the 161 breakpoints on the behavior of the series. The percentages were calculated based on the total of the series of alpha-temporal (59 reservoirs), beta-temporal (34 reservoirs) and the turnover (34 reservoirs) and nestedness (34 reservoirs) components evaluated



Source: prepared by the author

3.3.2 Consequence of biodiversity changes caused by breakpoints

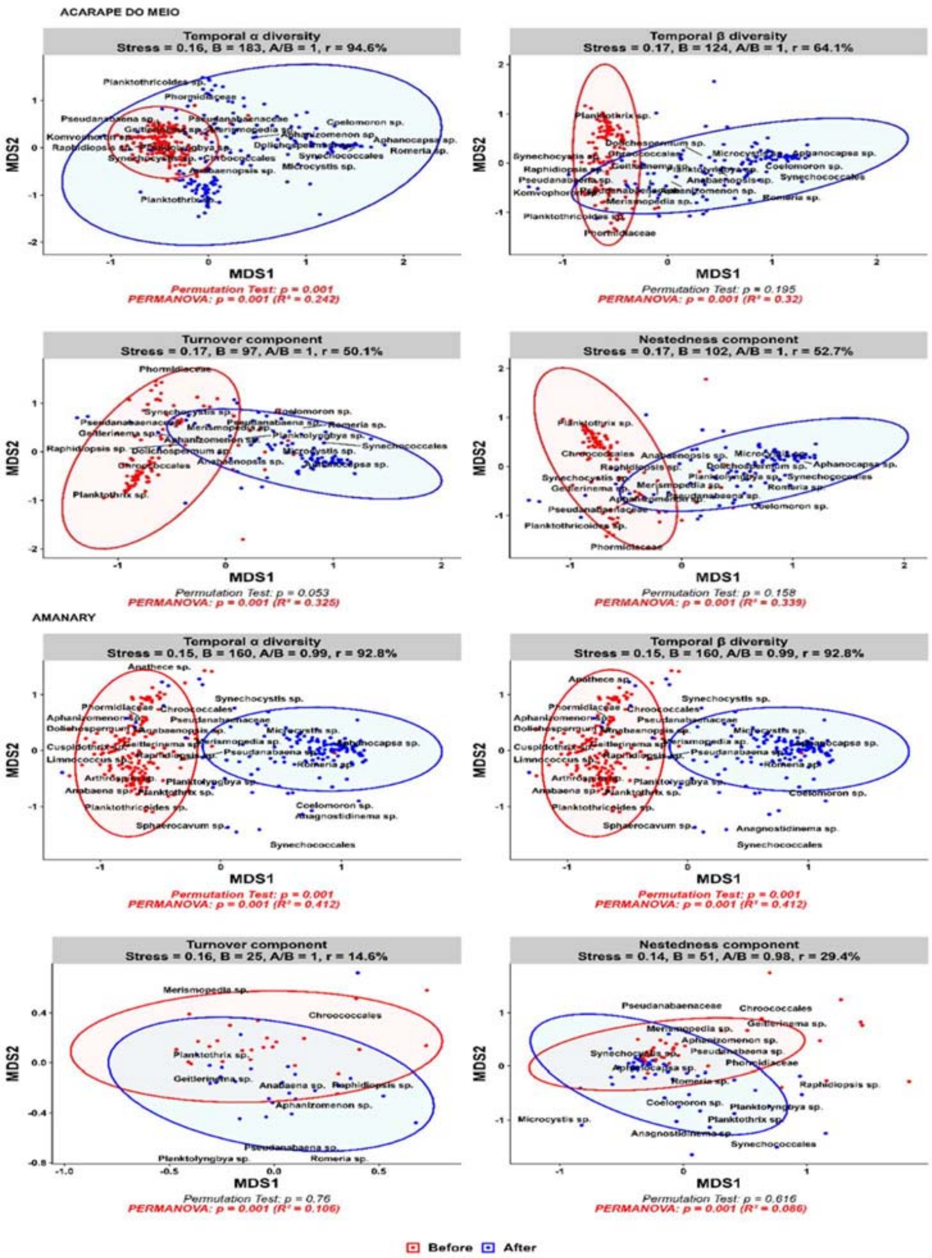
3.3.2.1 Effects on species composition and dispersal

Figure 11 shows the similarity of the cyanobacterial communities in two dams before and after the breakpoint in the biodiversity series simultaneously to the other dams presented in Section 9 of SM - APPENDIX B. Although the NMDS indicates evidence of differences between the communities before and after the breakpoint (see α diversity and temporal β of the "Amanary" Dam), it does not specify the type of difference, whether in the general composition of the communities of each group (localization), in the dispersion or variability of the taxa (dispersion) or in both. In addition, even with acceptable stress values (≈ 0.20), the use of only two dimensions was not enough to represent the complexity of some dams (see temporal α diversity of the "Acarape do Meio" dam).

Therefore, as in most dams, the groups were nearly balanced (B/A $\cong$ 1) and most of the data were used in the analyses ($r > 50\%$), the PERMANOVA and PERMDISP tests improved the understanding of the effect of breakpoints on the species composition matrices. This is because, unlike omnibus tests such as the Mantel test or ANOSIM (Similarity Analysis), PERMANOVA is robust to heterogeneity and is not affected by differences in the correlation structure. In addition, PERMDISP exclusively tests for changes in dispersion, which made it possible to separately evaluate the effects on dispersion and localization by PERMANOVA.

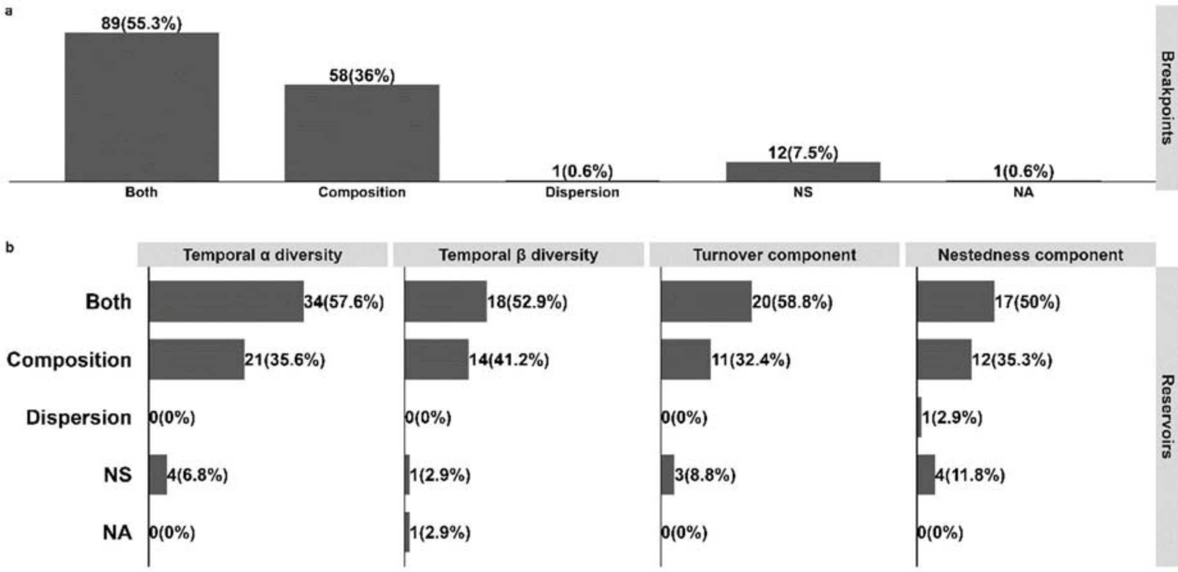
Regardless of the series (Figure 12a), in more than half of the evaluated reservoirs, the breakpoint altered both the composition and the dispersion of the species. This indicates that environmental factors during the prolonged dry period not only modified the species that were presented but also how they were distributed among the reservoirs. In nearly a third of the series, changes were observed only in composition, suggesting that, despite the replacement of the species, the dispersion among the dams remained relatively constant. Notably, only the breakpoint in the turnover component of the "Martinópole" weir caused significant changes exclusively in dispersion. Thus, considering the low number of reservoirs in which breakpoints did not have significant effects (<10%), the data indicate that the drought had a substantial impact on cyanobacteria communities in dams intended for human consumption. Figure 12b shows the results by grade.

Figure 11 - NMDS representation of the effect on the composition and structure of the cyanobacteria community after the occurrence of breakpoints in the series that describe the cyanobacterial biodiversity (two-dimensional NMDS). Each point on the graph represents a sample of routine monitoring in the new NMDS plane, and the distance between the points indicates the dissimilarity between them (Bray-Curtis dissimilarity). Similar samples are close together, while different samples are farther away. Cyanobacterial taxa, positioned based on the average abundance in the samples, are close to the points where these species were most abundant. "B" represents the number of observations of the Before group, "A/B" is the ratio of the number of observations of the After and Before groups, indicative of the balance level. "r" represents the percentage of the data analyzed concerning the total. The p-value and R^2 PERMANOVA test and the PERMSIM p-value (permutation test) below each plane in red were significant ($p < 0.05$) and not in black. A total of 999 permutations were performed in both cases and better visualization is available in Section 9 of SM- APPENDIX B.



Source: prepared by the author

Figure 12 - Summary of the effects on the groups obtained with the cyanobacteria composition matrices. The groups were formed with the Before and After observations of the breakpoints of the series that described biodiversity. In (a) the general effects and in (b) the effects are presented separately by series. The "Both" category represents significant differences ($p < 0.05$) in both composition (PERMANOVA) and dispersion (PERMDISP). The exclusive significant differences (i.e. only in "Composition" or only in "Dispersion") were represented in the categories of the same name. Non-significant differences were indicated as "NS", and "NA" refers to the reservoirs where testing was not possible due to the limited amount of data (< 5 observations).

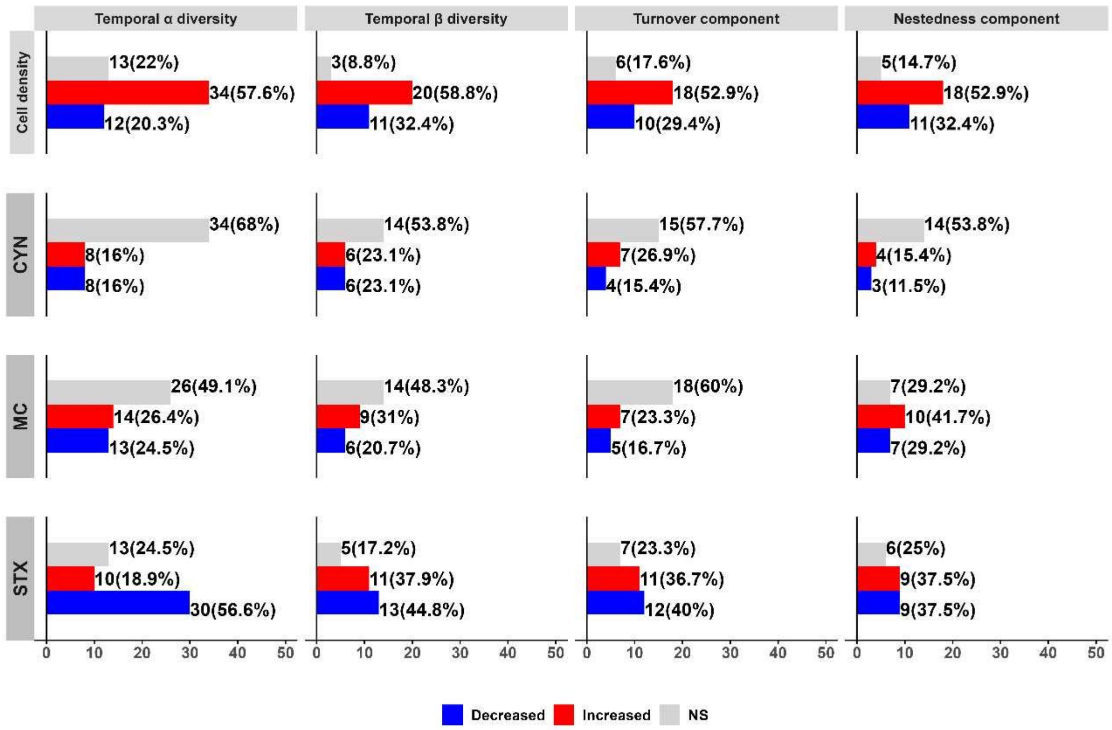


Source: prepared by the author

3.3.2.2 Effect on cell density and cyanotoxin concentration

Figure 13 shows how structural breaks in the series describing biodiversity affected cell density and toxin concentrations. Comparing the data before and after the breakpoints, there were no significant variations in cell density in 27 out of 161 cases analyzed (16.7%). In the other cases, there was an increase in 90 (55.9%) and a reduction in 44 (27.4%). In general, most breakpoints resulted in increased cell density. As for toxins, the number of evaluated breakpoints was lower due to changes in legislation (see Section 2 in SM). For CYN concentration, 77 out of 123 cases (62.6%) did not show significant changes, while there were increases in 25 (20.3%) and decreases in 21 (17.1%). For MC, 65 out of 136 cases (47.8%) did not show significant variations, where 40 showed an increase of 29.4% and 31 a reduction of 22.8%. In the case of STX, 31 out of 136 cases (22.8%) did not show significant changes, while there was an increase in 41 (30.1%) and a decrease in 64 (47.1%). In the case of toxins, contrary to what was observed in cell density, the increase was not the predominant effect. For CYN and MC, most breakpoints did not cause significant changes in concentration. However, for STX, the reduction was the most common effect, followed by the increase.

Figure 13 - Effects on cell density and toxin concentrations (CYN, MC and STX) observed after the breakpoints of the time series describing cyanobacterial biodiversity. The percentages were calculated based on the total number of reservoirs available in each series for each of the evaluated parameters.



Source: prepared by the author

Analyzing the biodiversity series separately, the percentages of increase and decrease in cell density followed a similar pattern being the increase as the predominant effect. For CYN, the percentages of increase and decrease were similar, except in the turnover and nestedness components, where the increase was slightly higher. For MC, the number of reservoirs with an increase was always greater than the number of reductions, especially in the series that describe the dynamics of species composition and replacement. For STX, the reduction was considerably greater in temporal alpha diversity, while in the turnover and nesting components, the percentages were more balanced although the reduction still prevails.

Therefore, the results indicate that the cell density of cyanobacteria tends to increase after changes in both abundance and dispersion patterns (temporal α diversity) and composition patterns (temporal β diversity and its turnover and nesting components).

In reservoirs where breakpoints caused significant changes, the predominant increase in cyanotoxins, such as CYN and MC, was more related to changes in composition than in abundance and diversity. On the other hand, for STX, the reduction, which was the most observed effect, was mainly associated with changes in abundance and diversity than in composition.

3.4 Discussion

3.2.3 Overview of cyanobacterial biodiversity in the 82 weirs addressed for human consumption.

Climate change and extreme events such as droughts and floods have significant impacts on water quality, altering the concentration of pollutants and the dynamics of nutrients. In addition, these phenomena can cause an increase in temperature and promote the growth of phytoplankton and cyanobacteria, which can negatively impact other aquatic species and water quality (XIA *et al.*, 2015).

During periods of drought, decreased water flow can weaken the dilution effects of some pollutants, resulting in higher concentrations of contaminants. In addition, a worsening of water quality can be caused by a decrease in dissolved oxygen and an increase in mineralization and salinization of water bodies. The sum of these effects contributes to changes in water quality that can negatively affect aquatic ecology, including ecosystem health and local biodiversity (CHATURVEDI *et al.*, 2021).

At the same time, the process of urbanization in arid regions of emerging countries also deserves to be highlighted. In these regions, rapid urban growth is increasing water demand. In addition, more than 80% of wastewater is discharged into the environment without suitable treatment, contaminating the already scarce water resources and affecting public health. In light of this, many cities are on the verge of a water crisis due to inadequate infrastructure to support population growth (KOOKANA *et al.*, 2020). For example, urbanization, considered one of the main causes of widespread biodiversity loss in freshwater ecosystems, has reduced fish diversity, in a process of biotic homogenization and reduction in environmental heterogeneity, with taxonomic and functional diversity being significantly lower in urbanized subtropical Chinese rivers compared to forested rivers (YANG et al., 2024).

Although Ceará has a significant regeneration of natural vegetation, especially in the central region of the state, the Caatinga, one of the main biomes of the semiarid Northeast, land use in recent years has shown a high conversion of natural vegetation for pasture and other agricultural activities (CABALLERO *et al.*, 2023). Anthropogenic actions, such as agriculture and soil exposure, common in the surroundings of the studied reservoirs, were associated with a significant increase in the amount of mobile phosphorus in the sediment, representing a risk for eutrophication in the semiarid region of northeastern Brazil (ROCHA JUNIOR; ARAÚJO; BECKER, 2024). Similar effects were also observed in the US, where the presence of cropland was associated with an increase in the proportion of water bodies covered by cyanobacterial blooms, while forested regions showed a reduction (MARION *et al.*, 2017). In addition, it was reported that in 98 springs in the semi-arid region of northeastern Brazil, where water quality is mainly affected by anthropogenic land use near the margins of water bodies and by rainfall variability, the effects were more pronounced in small and shallow springs (NOBRE *et al.*, 2020) similar to the 82 reservoirs studied here (Supplementary Material - Section 2).

The combination of these factors contributed to the trend of cyanobacterial biodiversity to be mostly increasing. Even though it was significantly altered by exogenous factors, this behavior persisted in most reservoirs, reflecting both growth in species quantity and diversity and significant variations in the composition of these species including species replacement (turnover) and loss without replacement (nestedness). However, the observed reductions of the Hill number in order 0 to higher orders, followed by the stabilization of biodiversity after order 1, suggest a process of biotic homogenization. This indicates that despite the emergence of a growing number of cyanobacterial species, considering the more restrictive conditions of the dams throughout the decade, in the general context, few species end up prevailing and dominating the environment. Some studies have also identified a process of homogenization of species associated with the effects of climate change, eutrophication, or land use and occupation (MONCHAMP *et al.*, 2018; MORI *et al.*, 2015; ZUO *et al.*, 2024).

Regarding series describing the dynamics of community composition, a previous study showed that spatial variation in β diversity was mainly controlled by the substitution component, while variation in temporal beta diversity was mainly controlled by the nestedness component under the impact of climate change (Lima et al., 2019). A study focused on stream invertebrate communities demonstrated that nestedness dominated under unpredictable seasonality, but turnover contributed mainly to temporal β diversity under predictable seasonal conditions (TONKIN *et al.*, 2017). In addition to the lifespan, the degree of temporal turnover may be related to the characteristics of the organism, since larger organisms with active mobility had slower temporal turnover than smaller organisms, suggesting that the rotation of species in time is jointly driven by several ecological, physical and geographic factors in aquatic ecosystems, and is not yet uniform among taxonomic groups (KORHONEN; SOINOUS; HILLEBRAND, 2010). We believe that this justifies the variability in the importance of the turnover and nesting components presented during the decade analyzed. In general, cycles of change in environmental conditions were interspersed with new levels of stability. Since sensitive species, with narrower environmental tolerances, can be replaced by tolerant species under eutrophication stress (ZHANG *et al.*, 2019) , the nesting process tends to dominate during periods of change - often near the breakpoints in the compositional dynamics - leading to homogenization and biodiversity loss. After the periods that the breakpoints occurred, a new environmental condition is established, and turnover has become more relevant. This is potentially concerning, given that many of the remaining dominant species identified during routine monitoring have already been reported as potential producers of cyanotoxins and compounds that alter the taste and odor of water, such as 2-methylisoborneol and geosmin (WHO, 2022).

Trends observed in both temporal alpha and temporal beta diversity and their components of cyanobacterial turnover and nestedness in reservoirs suggest that long-term climate change has a more significant impact on biodiversity than seasonal. This underscores the vulnerability of semi-arid regions to climate change, where even small changes in rainfall patterns can dramatically affect water quality and ecosystem health (NOBRE *et al.*, 2020), similar to other developing regions as highlighted by the Intergovernmental Panel on Climate Change (IPCC, 2023). It was identified that the long-term (interannual) effects were more significant for cyanobacterial biodiversity in the 82 reservoirs than the seasonal (intra-annual) effects, as detailed in Section 4 of the Supplementary Material. Similarly, authors have observed a more pronounced interannual variability in planktonic ciliate communities than seasonality (ABDULLAH AL *et al.*, 2023) However, other studies have identified the most relevant intra-annual variability represented by seasonal effects, for example (LI *et al.*, 2022; TONKIN *et al.*, 2017). We believe that the time and frequency of sampling, the type of organisms, and the environmental conditions justify these differences (KORHONEN; SOINOUS; HILLEBRAND, 2010). Although previous findings have indicated that the assembly of the ciliated plankton community was predominantly controlled by stochastic processes (ABDULLAH AL *et al.*, 2023), our findings suggest otherwise. We observed the presence of a deterministic tendency, and in a similar way to the deterministic processes that contributed to the formation of prokaryotic communities in shallow layers (≥85 m) dominated by cyanobacteria (HUANG *et al.*, 2023), we believe that the biodiversity of cyanobacteria in the evaluated reservoirs, which have lower depths and are dominated by these organisms, it follows a similar behavior. Thus, analogous to the breakpoints caused by drought periods, both the intensification and reduction of the exogenous factors that defined the current state of cyanobacteria in the evaluated dams have the potential to alter the biodiversity of these organisms. Thus, management actions should be encouraged so that external factors change this biodiversity to a less worrying condition.

Also, according to Section 4 of the Supplementary Material, although we did not observe a clear pattern of distribution of dams with a significant trend along the environmental gradients, previous studies indicate that the composition of cyanobacteria in semiarid regions is strongly associated with several environmental factors. For example, research has revealed that the reduction in wind speed increases the stability of the water column, resulting in hypoxia and increased release of phosphorus and nitrogen from sediments, conditions that favor the growth of cyanobacteria, such as *Microcystis spp.*, especially during spring and summer in a shallow lake in China (DENG *et al.*, 2018). In another shallow Chinese lake, it was observed that cyanobacteria dominated under conditions of high concentrations of total suspended solids and low light intensity, and non-nitrogen-fixing species (*Microcystis* and *Oscillatoria*) outperformed nitrogen-fixing species (*Anabaena* and *Aphanizomenon*) when the N:P ratio exceeded 20 (LIU *et al.*, 2021). In addition to the anthropogenic use of land near the banks of small and shallow springs located in the Brazilian semi-arid region, water quality was mainly affected by rainfall variability (NOBRE *et al.*, 2020). Besides that, other factors such as nutrient availability, long hydraulic detention times, and complementary environmental conditions were also associated with cyanobacteria community dynamics in semiarid reservoirs (BARROS *et al.*, 2017b; DE OLIVEIRA; DANTAS, 2019; MENDES *et al.*, 2022).

Therefore, given that the detailed identification of these patterns was beyond the scope of this study, which aimed to provide an overview of the effect of droughts on cyanobacterial dynamics, we suggest that future research focus on identifying the specific factors influencing cyanobacterial biodiversity in each pond individually. We believe that, similar to what was observed in generalist models (those that analyze multiple species and multiple sites) which have lower predictive capabilities when compared to more specialized models focused on a single species in freshwater environments (ROUSSO *et al.*, 2020), this more specific approach may provide more accurate results than the generalist approach used in this study. Thus, mitigation actions can be carried out in a more personalized way for each source, increasing the chance of success.

3.4.1 Effect of breakpoints on cell density, cyanotoxin concentration and the implications in water treatment

The risk scenario, characterized by the increase in the abundance and diversity of cyanobacteria, alongside a reduction in evenness and the consequent dominance of a few potentially toxic species, was significantly affected by the most severe period of drought. As a result, according to, most reservoirs not only altered the composition of the species present but also modified the way these species were distributed before and after breakpoints. In just over a third of the cases, changes were restricted to species composition and a small number of weirs did not detect significant variations.

In practice, this suggests that the environmental conditions or ecological factors mentioned above favored the presence of certain species to the detriment of others, possibly due to changes caused by the long drought that promoted the emergence of new species or the extinction of others. A change in dispersal, in turn, may indicate that cyanobacterial communities are becoming more uniform or more varied within groups.

For example, a reduction in dispersal may warn for a scenario of communities becoming more similar to each other, possibly due to the dominance of a few species adapted to the new conditions, assuming a characteristic of breakdown in stable states of phytoplankton dynamics. In contrast, an increase in dispersal may reflect a greater diversity in communities' responses to environmental conditions, suggesting a more unstable environment or one with multiple emerging ecological niches. Therefore, the final balance of the emergence or extinction of species and the increase or reduction of diversity before and after periods of drought resulted in biotic homogenization revealed by Hill's numbers and in the series that describe biodiversity. Consequently, the cell density of cyanobacteria increased after changes in the patterns of all series describing biodiversity in most dams (Supplementary Material - Section 10 - APPENDIX B).

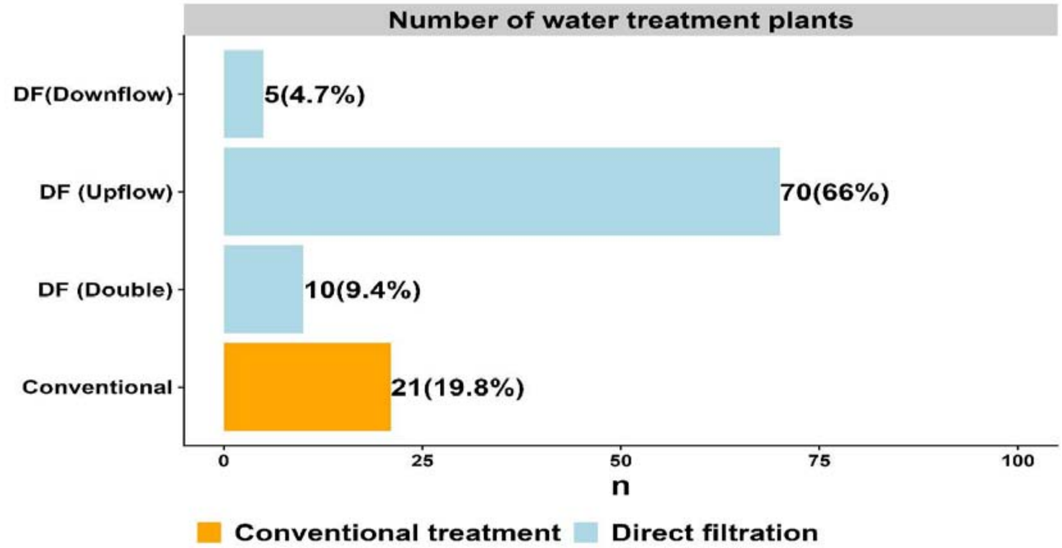
The presence of cyanobacteria in high concentrations can lead to several problems that impact both the success and cost of treatment, such as the need for pre-oxidation (LIN *et al.*, 2018), reduced efficiency and increased consumption of chemicals in coagulation (XIE *et al.*, 2016), premature clogging of filter, release of compounds capable of altering the organoleptic characteristics of water (ZAMYADI *et al.*, 2015), production of potentially carcinogenic precursor compounds of disinfection by-products (RIZZO, 2014), making it impossible to reuse supernatants from flocculation and decanter sludge (PESTANA *et al.*, 2016) besides the potential releasing of cyanotoxins (WHO, 2022)

The water systems of the 82 dams studied here are predominantly treated by direct filtration, especially upward filtration (Figure 14) . Generally, these stations have as the only stage of turbidity removal, sand filters with a single layer, preceded by a stage of coagulation, addition of polymer and pre-oxidation with chlorine. After these processes, the filtered water is disinfected with chlorine and goes to distribution. In the region, the preference for simpler and potentially less expensive treatment methods compared to conventional treatment, which may require more steps (coagulation, flocculation, clarification, filtration and disinfection) and larger quantities of chemical products, is associated with cost issues, difficulty of qualified operation, maintenance and the low requirement of the quality standard required in the first ordinances at the time of implementation of the stations (BRAZIL, 2004). However, considering the critical scenario described, where there is an increase in the density and diversity of cyanobacteria, it is possible that these direct filtration technologies, especially when associated with significant environmental changes, may not offer the necessary resilience to effectively struggle against the high presence of cyanobacteria and their byproducts, such as cyanotoxins.

Among other problems, these technologies can compromise the structure of organisms, especially filamentous species *Raphidiopsis sp*. or *Planktothrix sp*. that constantly dominate the studied dams (PESTANA *et al.*, 2019). In addition, the required hydraulic operating conditions for these technologies might lead to cell lysis

(CLEMENTE *et al.*, 2020; OLIVEIRA *et al.*, 2021), compromising both the safety and the quality of the drinking water (PESTANA *et al.*, 2018).

Figure 14 - Technologies for the treatment of the plants that make the water collected from the 82 weirs evaluated



Source: prepared by the author

Unlike the observed effect on cell density in the reservoirs where the breakpoints caused significant changes, most of the dams did not present significant variations in the concentrations of the cyanotoxin. In the case of STX, the reduction was the most frequent effect after breakpoints, especially after breaks in the alpha diversity series, which attenuates the risk outlook. This becomes particularly relevant considering the simplicity of treatment at the stations, where chlorine is the main agent used to ensure safe water for the region without additional steps.

However, the elevations observed after breakpoints in the series related to composition dynamics for CYN and MC indicate the need to maintain a constant state of warning. As advanced DNA sequencing analyses, capable of accurately detecting genes associated with the production of toxins, are not yet part of the routine of sanitation companies in developing regions, it is essential to develop management tools that help in making early decisions for potentially dangerous events. These tools should use the available monitoring infrastructure, associating the presence of frequent taxa with critical toxin events, and would be particularly valuable for managers in critical regions of the semi-arid region, similar to those observed in this study.

3.5 Conclusion

A decade of monitoring cyanobacteria in the 82 dams intended for human consumption in Ceará, located in the semi-arid region of northeastern Brazil, revealed that over time:

a) Cyanobacterial biodiversity was marked by a predominant trend of increase in both abundance and diversity (alpha-temporal diversity) and compositional dynamics (beta-temporal diversity and its turnover and nestedness components). However, when considering the reductions followed by stabilization between order 0 and higher orders, which reflect the evenness in Hill Numbers, a process of biotic homogenization was identified, in which a few species, many of them potentially toxic, dominate the environments.

b) Although the increase still prevailed, the most severe period of prolonged drought in the Brazilian semiarid region since 1910 significantly altered the trend of cyanobacterial biodiversity. In addition, intra-annual effects (seasonality) were less relevant than interannual effects and practically all series describing biodiversity showed a deterministic trend. From the perspective of climate change, these results point to a greater vulnerability of the region to exogenous factors, such as the effects reported here of extreme drought events.

c) The main consequence observed after the breakpoints in the biodiversity series was the increase in the cell density of cyanobacteria. Given the predominant use of simple technologies for water treatment in the region (direct filtration and conventional treatment), it is imperative to adopt more appropriate processes. The reduced efficiency of these traditional methods, e.g. shortening of filtration rows and increased frequency of washing, can aggravate the water distribution and result in lower quality water, posing a significant risk to the health of the population.

d) After breakpoints, most of the dams did not show significant variations in CYN and MC concentrations, with the reduction of STX being the most observed effect. However, elevations were identified in MC and CYN, especially after the breaks in the series of compositional dynamics. Therefore, it is crucial to implement continuous monitoring systems to quickly detect potentially dangerous conditions related to cyanotoxins.

In the absence of sophisticated methods that allow the precise identification of cyanotoxin-producing species or genes during toxic blooms, management tools based on the identification of organisms indicative of critical cyanotoxin conditions (or any other critical conditions, such as blooms or turbidity elevations of treated water, for example) can be integrated into the companies' protocols. This will enable quick or early adjustments in treatment processes and help prevent the distribution of water inappropriate for human consumption.

4 COMO IDENTIFICAR BIOINDICADORES PARA RISCOS ELEVADOS DE TOXINAS EM RESERVATÓRIOS DOMINADOS POR CIANOBACTÉRIAS POTENCIALMENTE TÓXICAS COM ALTA ROTATIVIDADE USANDO DADOS DE MONITORAMENTO DE ROTINA

HOW TO IDENTIFY BIOINDICATORS FOR ELEVATED TOXIN RISKS IN RESERVOIRS DOMINATED BY POTENTIALLY TOXIC CYANOBACTERIA WITH HIGH TURNOVER USING ROUTINE MONITORING DATA

RESUMO

Uma ferramenta foi desenvolvida para permitir que empresas de saneamento ajam de forma proativa em resposta a altas concentrações de toxinas na água bruta. Utilizando regressão logística (RL), IndVal e o algoritmo a priori, foram analisados 10 anos de dados de monitoramento de cianobactérias de 82 reservatórios na região semiárida, destinados ao abastecimento público. Bioindicadores foram identificados e associados à chance de ocorrência de eventos críticos. Um processo robusto de pré-seleção, considerando a biodiversidade local, permitiu a definição de bioindicadores para mais de 50% dos reservatórios com eventos críticos. A RL provou ser eficaz no desenvolvimento de uma ferramenta de tomada de decisão, sem a necessidade de investimento adicional. Além disso, essa metodologia superou outras abordagens na seleção de bioindicadores, sendo menos afetada pelo desbalanceamento da variável resposta e pelo número limitado de observações, com uma complexidade de ajuste comparável ao algoritmo a priori, mas superior ao IndVal.

Palavras-chave: Razão de chances; MBA; CyanoHabs; eutrofização; decisão baseada em dados; cianotoxinas; CYN; MC; STX

ABSTRACT

A tool was developed to enable sanitation companies to act proactively in response to high concentrations of toxins in raw water. Using logistic regression (LR), IndVal, and the a priori algorithm, 10 years of cyanobacteria monitoring data from 82 reservoirs in the semi-arid region, designated for public supply, were analyzed. Bioindicators were identified and associated with the odds of critical events. A robust pre-selection process, considering the local biodiversity, allowed bioindicators to be defined for over 50% of the reservoirs with critical events. LR proved to be effective in developing a decision-making tool, without the need for additional investment. Furthermore, this methodology outperformed other approaches for selecting bioindicators, as it was less affected by the imbalance of the response variable and the limited number of observations, offering a fitting complexity comparable to the a priori algorithm, but superior to the IndVal.

Keywords: Odds Ratio; MBA; CyanoHabs; eutrophication; data-driven decision; Cyanotoxins, CYN; STX; MC

HIGHLIGHTS

a) Biodiversity of cyanobacteria in 82 reservoirs for public supply in Brazil's semi-arid region was assessed using a decade of monitoring data.
b) Two groups of cyanobacteria—common and rare—were identified in each reservoir.
c) Logistic regression (LR), IndVal, and the a priori algorithm were used to pre-select cyanobacteria indicators (CyanoBio) from the common taxa.
d) Based on CyanoBio, odds ratios (OR) were defined, identifying bioindicators associated with toxin risk in raw water for over 50% of critical reservoirs.
e) Using standard monitoring infrastructure from sanitation companies, a tool was developed to enable proactive decision-making for toxin-related critical events without additional investment.

4.1 Introduction

Even with international recommendations suggesting alert states for microcystins (MC), saxitoxins (STX), and cylindrospermopsins (CYN) in the presence of certain cyanobacteria (CHORUS, I., & WELKER, 2021), in regions such as the semiarid, where water bodies often have high densities of potentially toxic cyanobacteria, the mere detection or elevated concentration of species traditionally associated with toxic blooms may not suffice for managers and operators of treatment plants to make appropriate decisions in risk situations. Considering this scenario, selecting indicator species, henceforth referred to as bioindicators, for critical toxin states in raw water could offer a more effective management tool.

Bioindicators should be characterized by specificity (occurring predominantly in specific locations or conditions) and fidelity (present in most conditions or locations within this group). This duality ensures the quality of the indicator, which is of great interest in ecological studies (LEGENDRE, 2013). These species are crucial for monitoring ecosystem health and identifying changes caused by environmental or human factors, aiding in conservation and ecological management (SIDDIG *et al.*, 2016). Statistical analyses, such as the Indicator Value Index (IndVal), are used to select them, evaluating the species' abundance, occurrence, and significance in different locations (LEGENDRE, 2013).

While the IndVal is easy to interpret, statistically robust and applicable across ecosystems, it may overestimate the number of indicator species, especially when site classifications are not independent of species composition data (LEGENDRE, 2013). Additionally, in large datasets, IndVal may require pre-selection of species, introducing biases and limiting the representativeness of indicators. This limitation can be circumvented by using the a priori algorithm, commonly applied in techniques like Market Basket Analysis (MBA), which handles large databases better without the need for pre-selection. However, MBA can generate extensive lists of indicators, necessitating a refinement step (LEOTE *et al.*, 2020). This pruning step can also be prone to partiality. While both methods effectively identify associations between species and environmental conditions, they do not provide detailed information on the direction or magnitude of these associations. Moreover, they cannot calculate the probability of the occurrence of bioindicator species in response to specific environmental conditions, such as the presence of high concentrations of toxins in raw water.

In this context, logistic regression (LR) emerges as a potential alternative. This Generalized Linear Models (GLM) technique allows estimating the probability of binary events, such as the presence of toxins in raw water, by converting independent variables into probabilities from 0 to 1. The odds ratio (OR), calculated from LR, quantifies the magnitude and direction of the relationship between environmental variables and the event of interest (DOBSON; BARNETT, 2019), which methods like IndVal and MBA cannot do. For instance, using LR and OR, it was possible to quantify how gradient velocity and activated carbon presence influenced the odds of cell lysis in *Microcystis agardhii*, *Dolichospermum circinalis*, and *Raphidiopsis raciborskii* (OLIVEIRA *et al.*, 2021). In another application, LR models were used to create an early warning system for cyanobacterial blooms based on hydraulic parameters in the Ohio River, USA (NIETCH *et al.*, 2022). Therefore, LR shows great potential for selecting and quantifying the association between bioindicator species and high toxin concentration events in raw water, contributing to proactive management instead of the traditional reactive approach in sanitation companies.

Understanding local biodiversity is essential for effective bioindicator selection, particularly when using techniques like LR. Although it addresses issues of association magnitude and direction found in IndVal and MBA, it still faces challenges in selecting independent variables. Biodiversity, multifaceted and complex, includes taxonomic, genetic, phenotypic, and functional aspects, in addition to species abundance (richness) and distribution (evenness) over time and space (HAMILTON, 2005). Tools like Hill numbers (HILL, 1973), which measure true diversity, offer a more accurate view of biodiversity variations compared to traditional entropy-based indices such as Shannon and Simpson indices (CAZZOLLA GATTI; AMOROSO; MONACO, 2020). Furthermore, the concepts of alpha, beta, and gamma diversity are crucial for a more comprehensive analysis (JOST, 2007).

In bioindicator selection, considering these metrics from a temporal perspective is essential, as choosing appropriate predictors is directly tied to understanding the ever-changing ecological dynamics. Temporal beta diversity, with its turnover (species rotation) and nestedness (loss without replacement) components, can provide valuable information during biodiversity assessment (LINDHOLM *et al.*, 2021; MAGURRAN *et al.*, 2019; PEREIRA *et al.*, 2023). Thus, a detailed study of biodiversity using these concepts can provide less biased information in the bioindicator pre-selection process.

Considering cost and complexity, monitoring a single population or a group of populations as a bioindicator is often straightforward. Parameters such as abundance, density, age/size structure, reproduction rate, and growth rate are relatively easy to measure and sensitive to environmental changes, such as drought or high toxin concentrations, making population monitoring an effective process (CHANDEL *et al.*, 2023; SIDDIG *et al.*, 2016). This process is similar to what sanitation companies perform when monitoring phytoplankton in surface reservoirs for public water supply.

Brazil has regulations to ensure water quality for the population, aligned with international recommendations (BRAZIL, 2021; CHORUS, I., & WELKER, 2021). However, like elsewhere, cyanobacteria and toxin monitoring are contingent on organism quantity, with the established limit at 20,000 cells/mL for weekly frequency. Additionally, only recently have mandatory weekly analyses included CYN (CAMPOS *et al.*, 2024).

Despite laborious and costly laboratory routines, they are often underutilized and performed merely for legal compliance, overlooking their potential for providing valuable information for bioindicator selection.

In the semiarid region of northeastern Brazil, the proliferation of cyanobacteria has caused serious incidents. In Paulo Afonso (Bahia), 88 deaths were linked to toxic blooms of *Dolichospermum* and *Microcystis*. In Caruaru (Pernambuco), 50 deaths were associated with the presence of MC. Between 2015 and 2016, neurological complications caused by the Zika virus were aggravated by the presence STX, as reported by (CAMPOS *et al.*, 2024). Therefore, these events raise a warning signal for water utility managers.

Although studies indicate cyanobacteria as an indicator of eutrophication and toxin presence, the given focus is not fully utilized in managing critical events by sanitation companies in the semiarid. This is due to the persistent reality of high densities of potentially toxic cyanobacteria in the water bodies used for regional water treatment (CAMPOS *et al.*, 2024; KUCZYŃSKA-KIPPEN; KOZAK; CELEWICZ, 2024; MENDES *et al.*, 2022). Consequently, studies that integrate data from these companies' routine monitoring to select cyanobacteria bioindicators (CyanoBio), measuring and directing their associations with high toxin concentration events in raw water, are rare. This gap is particularly evident in regional-scale and long-term studies, with a decade of information.

Therefore, to fill this gap, this study aimed to: (i) identify the most suitable set of individuals to be used as CyanoBio, through biodiversity assessment of reservoirs for public supply; (ii) compare LR, IndVal, and a priori algorithm methods in pre-selecting CyanoBio for critical toxin events; and (iii) determine the magnitude and direction of the association between CyanoBio, the total cell density of cyanobacteria, and the periods of high toxin concentrations in raw water, considering both high and low concentrations, and define a final list of species and conditions that most significantly influence the risk of toxin alerts for each studied water source. The region, home to approximately 8.7 million people, is entirely within the semiarid and faces prolonged droughts, with the period from 2012 to 2016 being the most severe since 1910 (FUCEME, 2016; MARENGO ORSINI *et al.*, 2018). This context highlights the need for effective management tools to support sanitation company managers, leveraging existing monitoring systems without the need for significant additional investments.

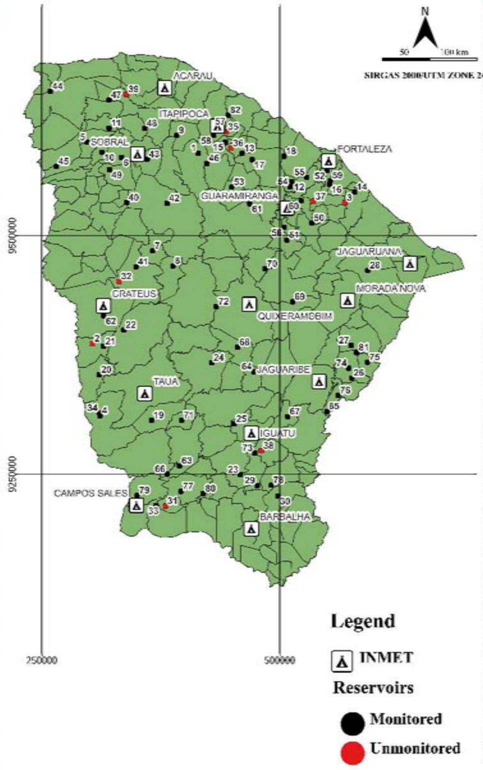
4.2 Materials and Methods

4.2.1 Study area

Data from 82 reservoirs designated for human water supply in Ceará, Brazil, were analyzed for the period from 2009 to 2019 (Figure 15). Located in the Brazilian semi-arid Northeast, the region exhibits a climatic diversity ranging from tropical savanna (Aw) to hot arid steppe (BSh) (PEEL; FINLAYSON; MCMAHON, 2007). Throughout the period analyzed, environmental variables displayed a moderate gradient across the state. According to data from the 16 meteorological stations of the National Institute of Meteorology (INMET), maximum temperatures averaged between 21 and 32 ºC; thermal amplitude ranged from 5 to 16 ºC; global radiation from 1000 to 2100 $kJ.m^2$; wind speed from 1 to 5 m/s; and total annual precipitation varied from 300 to 1300 mm (INMET, 2023). Historically, the region experiences a rainy season from January to May and a dry season from June to December. Additional details are available in Section 1 of the supplementary material (SM) - APPENDIX C. To mitigate the effects of persistent water scarcity, the state has an extensive network of artificial reservoirs, primarily, but not exclusively, aimed at supplying drinking water, as mandated by law. Of these reservoirs, 157 are monitored by the Water Resources Management Company (COGERH), which classified most of these reservoirs as eutrophic throughout the analyzed period (COGERH, 2023).

Figure 15 – List of the 82 monitored and unmonitored reservoirs analyzed, along with their closest INMET meteorological monitoring stations. For further details, see Section 1 of the SM - APPENDIX C.

ID	Reservoir	STATION
1	MISSI	ITAPIPOCA
2	RODRIGO	CRATEUS
3	CORIOLANO ALVES DE SOUSA BRITO	FORTALEZA
4	FACUNDO	TAUA
5	ANGICOS	SOBRAL
6	AYRES DE SOUZA	SOBRAL
7	CARMINA	CRATEUS
8	MONSENHOR TABOSA	CRATEUS
9	SÃO PEDRO TIMBAÚBA	ITAPIPOCA
10	TRAPIÁ III	SOBRAL
11	VÁRZEA DA VOLTA	SOBRAL
12	AMANARY	GUARAMIRANGA
13	CAXITORÉ	ITAPIPOCA
14	MALCOZINHADO	FORTALEZA
15	MUNDAÚ	ITAPIPOCA
16	PACOTI	FORTALEZA
17	PENTECOSTE	ITAPIPOCA
18	SITIOS NOVOS	FORTALEZA
19	ARNEIROZ II	TAUA
20	COLINA	TAUA
21	FLOR DO CAMPO	CRATEUS
22	JABURU II	CRATEUS
23	OLHO D'ÁGUA	IGUATU
24	SERAFIM DIAS	QUIXERAMOBIM
25	TRUSSU	IGUATU
26	CANAFÍSTULA	JAGUARIBE
27	RIACHO DA SERRA	MORADA NOVA
28	SANTO ANTÔNIO DE RUSSAS	MORADA NOVA
29	JUNCO	BARBALHA
30	CACHOEIRA	BARBALHA
31	BELO HORIZONTE	CAMPOS SALES
32	JOÃO JOSÉ DE CASTRO	CRATEUS
33	JOÃO LUÍS	CAMPOS SALES
34	PUIÚ	TAUA
35	IPÚ MAZAGÃO	ITAPIPOCA
36	SÃO JOAQUIM	ITAPIPOCA
37	BOQUEIRÃO	GUARAMIRANGA
38	UBALDINHO (RIACHO SÃO MIGUEL)	IGUATU
39	SEBASTIÃO DE ABREU	ACARAU
40	ARARAS	SOBRAL
41	CARÃO	CRATEUS



ID	Reservoir	STATION
42	EDSON QUEIROZ	SOBRAL
43	FORQUILHA	SOBRAL
44	ITAÚNA	SOBRAL
45	JABURU I	SOBRAL
46	JERIMUM	ITAPIPOCA
47	MARTINÓPOLE	ACARAU
48	ACARAÚ MIRIM	SOBRAL
49	TAQUARA	SOBRAL
	ARACOIABA	GUARAMIRANGA
51	CASTRO	GUARAMIRANGA
52	GAVIÃO	FORTALEZA
53	GENERAL SAMPAIO	GUARAMIRANGA
54	ITAPEBUSSU	GUARAMIRANGA
55	PENEDO	FORTALEZA
56	PESQUEIRO	GUARAMIRANGA
57	POÇO VERDE	ITAPIPOCA
58	QUANDÚ	ITAPIPOCA
59	RIACHÃO	FORTALEZA
60	ACARAPE DO MEIO	GUARAMIRANGA
61	SÃO DOMINGOS	GUARAMIRANGA
62	CARNAUBAL	CRATEUS
63	DO CORONEL	CAMPOS SALES
64	JENIPAPEIRO	IGUATU
65	MADEIRO	JAGUARIBE
66	MAMOEIRO	CAMPOS SALES
67	ORÓS	IGUATU
68	PATU	QUIXERAMOBIM
69	PEDRAS BRANCAS	QUIXERAMOBIM
70	POMPEU SOBRINHO	QUIXERAMOBIM
71	RIVALDO DE CARVALHO	TAUA
72	SÃO JOSÉ I	QUIXERAMOBIM
73	UBALDINHO	IGUATU
74	EMA	JAGUARIBE
75	POTIRETAMA	JAGUARIBE
76	ADAUTO BEZERRA	JAGUARIBE
77	CANOAS	CAMPOS SALES
78	ROSÁRIO	BARBALHA
79	POÇO DA PEDRA	CAMPOS SALES
80	VALÉRIO	BARBALHA
81	FIGUEIREDO	JAGUARIBE
82	GAMELEIRA	ITAPIPOCA

Source: prepared by the author

4.2.2 Data Sources

Routine monitoring data generated by the Water and Sewerage Company of the State of Ceará, which oversees the treatment plants drawing water from the 82 studied reservoirs, were considered (see Section 2 of SM - APPENDIX C). According to reports, 73 out of the 82 reservoirs are managed by the COGERH (COGERH, 2023). General details about the reservoirs and treatment plants are provided in Section 2.1 of the SM. The collection points nearest to the intake was chosen, and at least 100 individuals or colonies were counted using an inverted microscope (Zeiss Axio, A1) with a Sedgewick-Rafter chamber. Phytoplankton was identified at the species level when possible. The protocol followed regulatory guidelines, with analysis frequency and toxin testing based on cell density. Further details can be found in Section 2.2 of the SM. The list of primary organisms identified over the decade is available in Section 2.3 of the material. Enzyme-Linked Immunosorbent Assays (ELISA; Abraxis) were employed to quantify total MC, STX, and CYN in raw water.

4.2.3 Biodiversity analysis for Cyanobio pre-selection

Some details from the following sections are clarified in Section 3 of the SM - APPENDIX C.

4.2.3.1 Hill Numbers

The Hill Numbers were used to estimate the true biodiversity in the reservoirs (HILL, 1973), as shown in Equation 1 (JOST, 2006):

$$^{q}D = \left(\sum_{i=1}^{S} p_i^q\right)^{1/(1-q)} \qquad \text{(Eq. 1)}$$

Where:

a) **q**: indicates the order of biodiversity, which varies with the frequency of individuals.
b) **S**: sample **richness**, defined as the number of distinct taxa found in each monitoring campaign.
c) $\mathbf{p_i}$: relative **abundance** of individuals of each taxon, calculated as the proportion of individuals of a specific taxon to the total number of organisms identified in the monitoring process.

Based on this:

a) **q = 0**: At order 0, the frequency of collected individuals is not considered, which corresponds to the sample richness.
b) **q = 1**: At order 1, individuals are weighted according to their frequencies, which is equivalent to the exponential of Shannon entropy (JOST, 2006).
c) **q = 2**: At order 2, higher weights are given to common individuals, analogous to the inverse of Simpson entropy (JOST, 2006).

Lower orders emphasize rare species, while higher orders favor more common species. Results were presented in the series, with biodiversity (Hill Number) summarized as the monthly mean. This approach allows not only for assessing biodiversity reduction among different orders but also for examining how these series behaved over the decade evaluated. The Hill Numbers were calculated using the *renyi()* function from the R vegan package (KINDT; OKSANEN, 2022).

4.2.3.2 Temporal beta diversity

To evaluate changes in cyanobacterial assemblages between 2009 and 2019 across 82 reservoirs, temporal β diversity and its turnover and nestedness components were used (MAGURRAN *et al.*, 2019). This approach offers a more detailed understanding of biodiversity, particularly in ecosystems where Hill numbers may be limited, as they do not always capture the complexities of nonlinear processes, such as environmental impacts caused by prolonged droughts (RICOTTA; FEOLI, 2024). The analyses were conducted at monthly and annual scales, with results summarized as mean. Temporal beta diversity and its components were calculated using the *beta.multi* function from the R package *betapart* (BASELGA *et al.*, 2023; BASELGA; ORME, 2012), following methodologies reviewed by (MAGURRAN *et al.*, 2019). Matrices of cyanobacterial presence and absence and the Sørensen dissimilarity index family were used. Turnover was measured using Simpson dissimilarity, while nestedness was determined by the difference between Sørensen and Simpson indices.

4.2.3.3 Unconstrained ordination methods: CA, PCA, and DCA

Unconstrained ordination methods were used to simultaneously represent sample and species data in a reduced space. Rarefaction curves were evaluated before the analyses to verify if the sampling effort adequately represented the cyanobacteria. Differences in the effort prevented comparisons between different reservoirs and motivated individual analyses (see Section 3.1 of the SM - APPENDIX C). For visualizing the data in two-dimensional graphs, biplots were employed. In these graphs, the proximity of symbols representing samples (points) and organisms ("*" or arrows) indicates similarity in relative frequency and high abundance, respectively. Based on preliminary biplot analyses and the relative frequency of the presence/absence matrix, rare and common organisms were identified. The criterion for rare organisms was established as those with at most 10% of the frequency of the most frequent organism(s). These individuals were identified, separated into a set of rare organisms, and ordered in ascending frequency, starting with those with only a single occurrence. Subsequently, a three-step iterative process was conducted:

a) removal of low-frequency organisms.
b) performing a new ordination analysis.
c) evaluating the biplot and total explained inertia.

After these iterative processes, the species and sample distribution patterns were analyzed through the obtained graphs. For each of the 82 studied reservoirs, a graph with five biplots and a scree plot was generated. The biplots represent ordination analyses performed with all individuals, excluding single-occurrence organisms, and the remaining three exclude organisms from the three quartiles of the rare species set. During testing, the best results consistently appeared in one of these five scenarios for all studied water bodies. The scree plot displays how inertia varied as a function of the relative frequency of rare species. The number of times each rare species was identified in samples was also presented, enabling an estimate of the number of samples from each analysis. Correspondence analysis (CA) and principal component analysis (PCA) were performed using the CA and PCA functions from the FactoMineR package, and their biplots were generated using the fviz_ca_biplot and fviz_pca_biplot functions from the factoextra package (KASSAMBARA; MUNDT, 2020; LÊ; JOSSE; HUSSON, 2008). Detrended correspondence analysis (DCA) was conducted using the decorana function from the vegan package (OKSANEN *et al.*, 2022). For more details, refer to Section 3.2 of the SM - APPENDIX C.

4.2.4 Pre-selection and identification of bioindicators

Since the data analyzed were obtained in an operational monitoring context focused on meeting legal regulations, temporal series gaps occurred due to the reservoirs drying up and discontinuation of reservoir use by strategic decisions of the companies. Given these conditions, which would not allow for a robust predictive approach, and the absence of detailed analytical techniques capable of identifying the taxa responsible for critical toxin events, the decision was made to explore the associations between cyanobacteria and critical toxin events.

Two groups were formed: one for days when the concentrations of MC, STX, and CYN exceeded the alert level, and another for days with concentrations below the limit. More details on the analyses are found in Section 4 of the SM. Based on guidelines from the World Health Organization (CHORUS, I., & WELKER, 2021) and local legislation (BRAZIL, 2021), alert levels were defined for raw water concentrations exceeding 90% of the safe standards for long-term exposure (see Section 4.1 of the SM). The selection process began with common organisms defined in the previous section. From this set, the following steps were carried out for each reservoir:

a) **Pre-selection of CyanoBio associated with critical toxin levels using LR:** The data were categorized into two levels based on the median. For each organism, observations with cell.mL^{-1} concentrations above the median were categorized as high concentration (H), while the rest were categorized as low concentration (L). In some cases, even though the organism was not classified as rare, the threshold value was set to 0 when the organism's frequency was low. Concentrations below the thresholds presented in Section 4.2 of the SM were considered as the reference category. More detailed information about the fitting process is provided in Section 4.3 of the SM - APPENDIX C.
b) **Pre-selection based on IndVal:** A strict criterion was established to select taxa indicative of critical toxin conditions to ensure robust selection. Only species with an IndVal greater than 0.7 and both A and B values above 0.5 were considered, provided they were statistically significant ($p < 0.05$). More details in Section 4.4 of the SM - APPENDIX C.
c) **Pre-selection using an a priori algorithm:** Based on organisms associated with those selected in steps (1) and (2). Equation 1 was considered:

$$\{A\} \Rightarrow \{B\} \qquad \text{Eq. 1}$$

Where A, referred to as antecedents (LHS – left-hand side), represents the organisms associated with those pre-selected in steps (1) and (2); and B, known as consequents (RHS – right-hand side), represents the organisms pre-selected by LR and IndVal in steps (1) and (2), respectively. In this representation, given any reservoir, the probability of species B occurring is influenced (increased) by A. Only rules that met the conditions of Section 4.5 of the SM were considered. The ratio between the total number of records and the number of times both genus A and genus B occurred together was calculated. Based on this value, for reservoirs with more than one rule for the same pre-selected organism, the rule with the highest percentage was chosen. If more than one rule still existed, preference was given to the one with the highest leverage, support, and confidence, respectively. Further details can be found in Section 4.5 of the SM - APPENDIX C.

d) **Definition of the final set of pre-selected CyanoBio:** The final set was obtained by combining all taxa pre-selected by LR, IndVal, and the a priori algorithm (see Section 4.6 of the SM - APPENDIX C).

e) **Estimation of the OR and its 95% confidence interval (CI):** Based on the full model adjusted with all organisms from step (4), plus the predictor "Total" (representing the total cyanobacteria cell density), a more parsimonious nested model (*stepwise model*) was obtained for OR and CI estimates. The predictor "Total" was included to test the hypothesis that total cyanobacteria cell density is also associated with the risk of reaching toxin alert levels in raw water. The intercept analysis of the *stepwise model* revealed the influence of low CyanoBio density and total cyanobacteria density on the same risk events. More details are available in Section 4.7 of the SM - APPENDIX C.

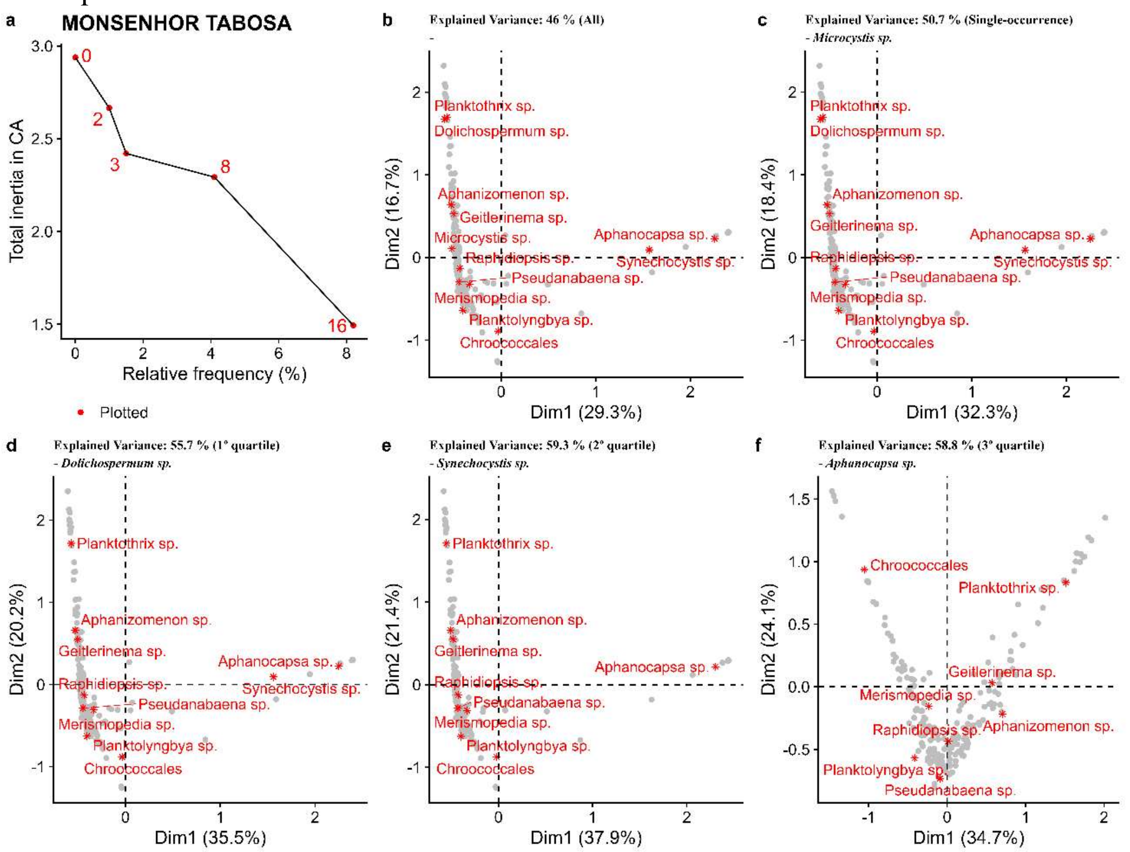
f) **Bioindicators significantly associated with toxin alert states:** CyanoBio, Total, or Intercept were considered bioindicators that significantly increase risk when OR > 1, with its CI excluding the value 1, preferably with the narrowest interval possible. Conversely, bioindicators with OR < 1 and a CI also excluding the value 1, equally narrow, were associated with a significant reduction in the odds of critical toxin events. When the CI included value 1, or the model values did not allow for the calculation of both OR and CI extremes (with infinite values or very wide confidence intervals), the bioindicator in question was considered inconclusive, and the predictor was not deemed an adequate bioindicator. In Sections 4.8 of the SM - APPENDIX C, a list of all bioindicators associated with toxin alert states in raw water can be found, separated by the toxin, and indicating the direction of the association (increasing or reducing the odds of critical events).In Section 4.9 of the SM - APPENDIX C, the results from Section 4.8 were graphically presented, along with the Hill numbers and temporal beta diversity series, which complement the biodiversity analysis of the 82 evaluated reservoirs.

4.3 Results

4.3.1 Set of common organisms used for the CyanoBio pre-selection

For a more precise selection of CyanoBio bioindicators, biplots from the CA of 82 reservoirs were analyzed, showing a pattern similar to that of the Monsenhor Tabosa reservoir in Figure 16 (see Section 3 of the SM - APPENDIX C). The χ^2 distance metric effectively addressed the issue of double zeros in abundance matrices; however, the presence of rare species distorted the graph (BORCARD; GILLET; LEGENDRE, 2011). Generally, an arching pattern was observed across all reservoirs, becoming clearer as taxa classified as rare were removed according to the established criteria.

Figure 16 – Ca biplot. CA biplots. In (a), the reduction in total variance after the removal of rare organisms (relative frequency < 10%) is shown. The numbers indicate the total number of samples in which this frequency was observed. The red numbers correspond to the conditions shown in graphs (b) to (f). In these graphs, the total explained variance and the removed species are indicated at the top. The removals were carried out without replacement, meaning that a species removed in one analysis does not reappear in subsequent ones.



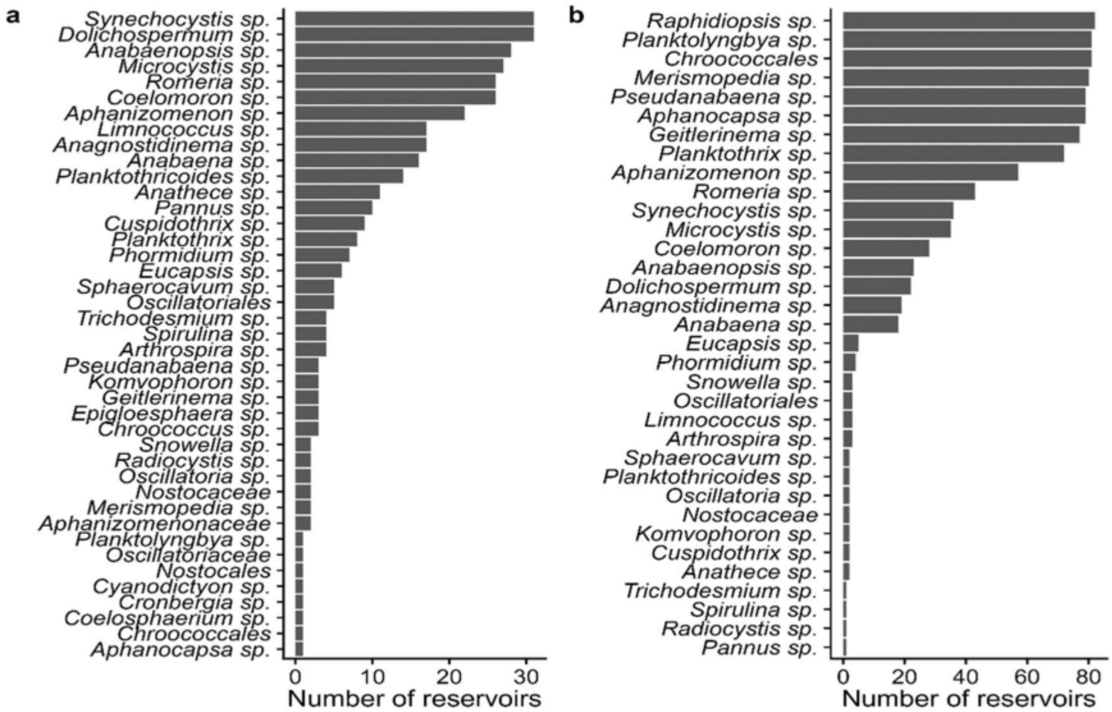
Source: prepared by the author

Additionally, a reduction in the total variance of the CA was noted, accompanied by an increase in the proportion of variance explained by two components (Figure 16a). In 63 reservoirs (76.8%), these two components explained between 23%

and 43% of the total variance, while in 16 reservoirs (19.5%), the explained variance ranged between 43% and 63%. The removal of rare organisms did not always improve the visualization of the biplots or the total variance explanation for all 82 reservoirs, similar to what was observed in items "b" to "f" of the Monsenhor Tabosa Figure 16. In 4 reservoirs (4.8%), the best results were obtained by removing taxa that appeared only once during the monitoring period. In another 4 reservoirs (4.8%), the best results were achieved by removing taxa with relative frequencies in the first quartile of the rare organisms set. The removal of taxa in the second and third quartiles produced the best results in 15 (18.3%) and 58 (70.7%) reservoirs, respectively. These results are available in Table from Section 3.3 of SM - APPENDIX C.

In the Malcozinhado reservoir, no rare organisms were identified, and the removal of taxa did not improve the CA. Table from Section 3.3 of SM also lists the 42 rare taxa and the 34 common taxa observed in all the studied reservoirs (Figure 17). Only *Raphidiopsis sp.* was classified as common in all reservoirs, while *Chroococcus sp.*, *Epigloesphaera sp.*, *Coelosphaerium sp.*, *Cronbergia sp.*, and *Cyanodictyon sp.* were classified as rare. Organisms from the *Aphanizomenonaceae*, *Nostocales*, and *Oscillatoriaceae* families were also identified only as rare.

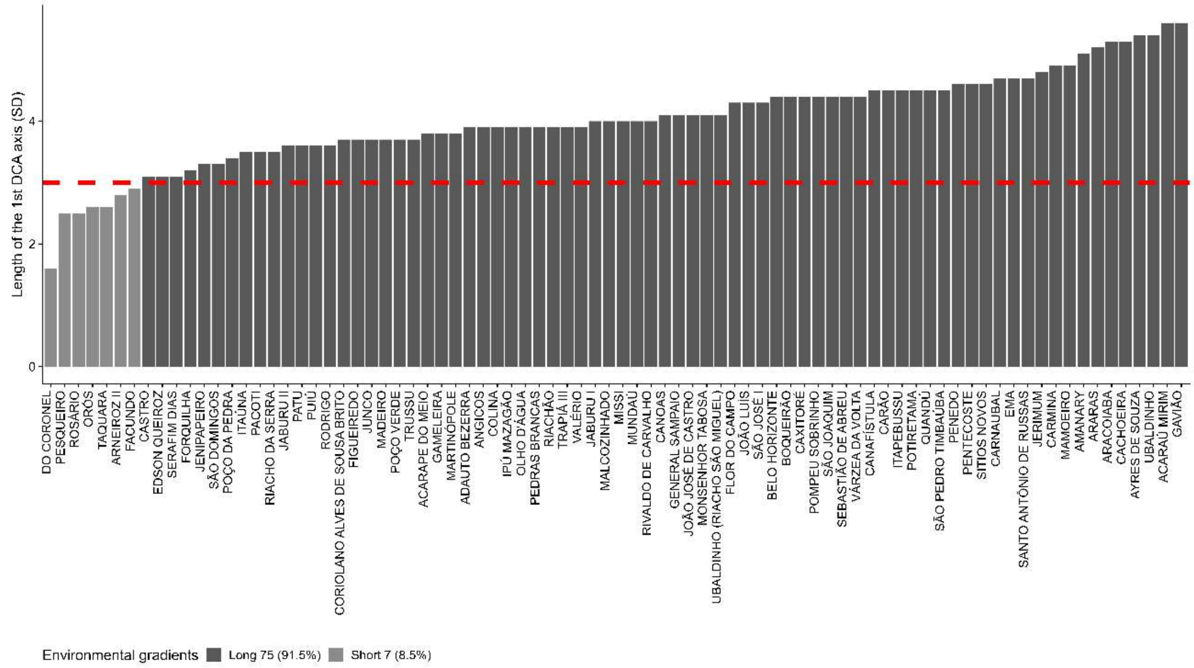
Figure 17 - List of rare organisms (a) and common organisms (b) found in the 82 reservoirs evaluated



Source: prepared by the author

Thus, the analysis of the 82 reservoirs in Ceará reveals notable heterogeneity, characterized by low percentages of total variance explained in two dimensions and the importance of rare organisms in the observed arch pattern. These organisms introduce additional variability that CA attempts to capture, possibly linked to high species turnover, typical of long or intense ecological gradients. The DCA supported this hypothesis, with the first axis of the DCA exceeding 3 standard deviations in 75 (91.5%) of the reservoirs, as shown in Figure 18. The DCA also reduced the arch effect in the CA biplots, improving visualization (see Section 3 of the SM - APPENDIX C).

Figure 18 - Length of the first axis of the DCA, reflecting the extent of the ecological gradient



Source: prepared by the author

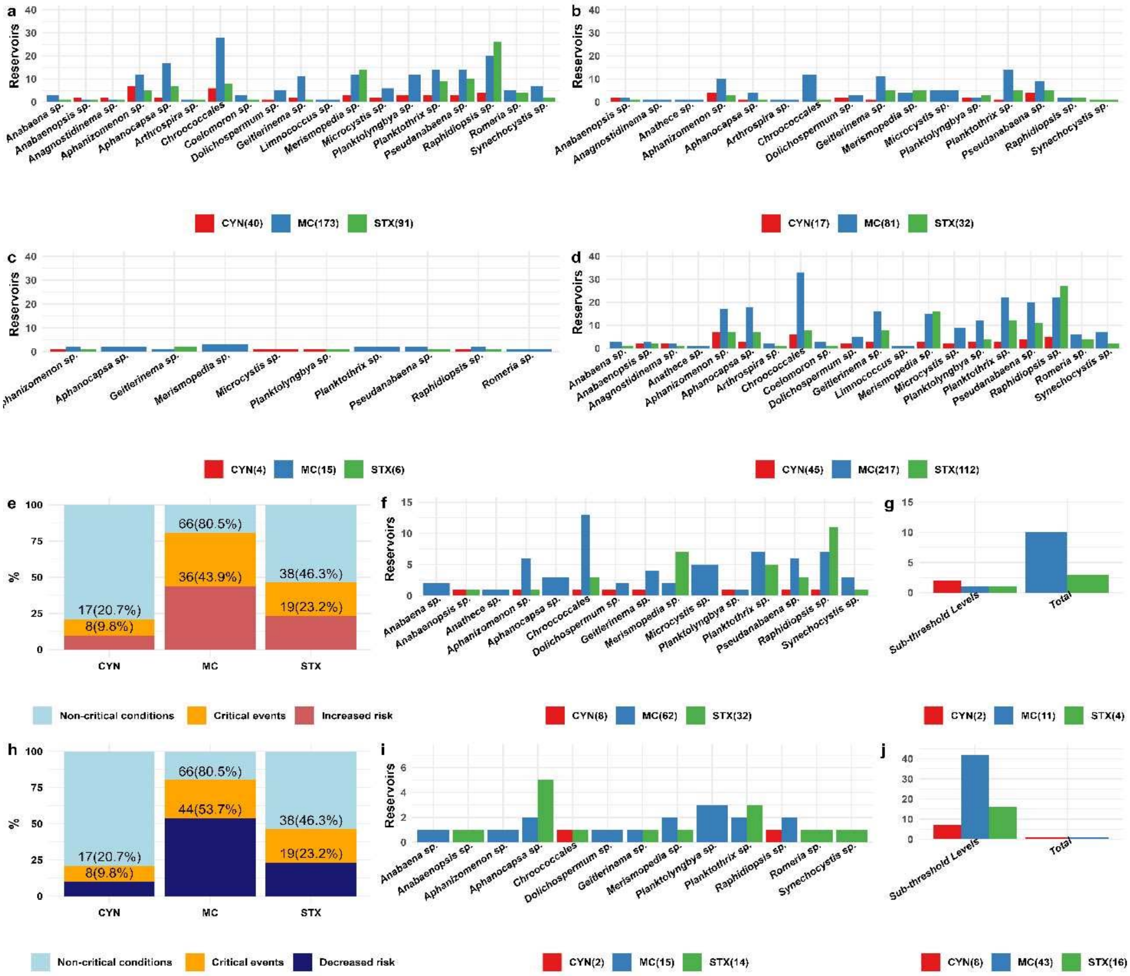
Only 7 reservoirs (8.5%) presented a short gradient, making PCA the most suitable technique for these cases (LEGENDRE; LEGENDRE, 2012). The comparative results between the CA, DCA, and PCA biplots are available in Section 3 of the SM - APPENDIX C. The overall pattern was consistent, with high-frequency taxa grouped in the central areas of the biplots, while rare organisms were dispersed. Therefore, the set of common organisms was used for the pre-selection of CyanoBio, ensuring more representative bioindicators for toxin alert states.

4.3.2 Overview of CyanoBio pre-selection process

Monitoring 82 reservoirs over a decade revealed that MC was detected above the reference limits in 66 reservoirs, STX in 38, and CYN in 17. To some extent, this difference can be explained by the regulations of the time, which recommended monitoring CYN only in the presence of species potentially producing this toxin, while STX and MC analyses were mandatory for cyanobacteria densities above 20,000 cells/mL, a situation more commonly observed during the period analyzed, as almost all reservoirs consistently exhibited high cyanobacterial cell densities throughout the decade (CAMPOS *et al.*, 2024).

According to Figure 19, LR pre-selected more species associated with critical toxin events than IndVal (Figure 19a and Figure 19b), and the a priori algorithm also resulted in a limited number of associated taxa (Figure 19c). Although several cyanobacteria were pre-selected and associated with critical events in various reservoirs (Figure 19d), bioindicators were identified in only half of these reservoirs, both those that increased risks (Figure 19e) and those that reduced them (Figure 19h). This can partly be attributed to the low number of critical events in most reservoirs, in addition to the biotic homogenization process, which will be discussed later.

Figure 19 – Overview of the bioindicator pre-selection process. In (a), (b), and (c), the totals of reservoirs where CyanoBio were pre-selected by LR, IndVal, and the a priori algorithm, respectively, are presented. In (e) and (h), the percentages of reservoirs where CYN, MC, and STX exceeded the critical levels and where bioindicators increased or reduced the chances of critical events, respectively, are shown. In (f) and (g), the numbers of reservoirs where bioindicators increased the risk of critical toxin events are shown, and in (i) and (j), those that reduced this risk. The legends present the totals per toxin. "Sub-threshold levels" indicate densities below the established thresholds, while "Total" refers to the total cyanobacterial cell density. The scales were adjusted for better visualization. (For better visualization, see Section 4 in SM)



Source: prepared by the author

In comparison to the exclusive use of IndVal and the a priori algorithm, the OR estimates increased the ability to select bioindicators associated with critical toxin events. This is because it allowed determination of both the magnitude and the direction of the association, meaning it was possible to distinguish whether the CyanoBio increased (Figure 19f) or decreased (Figure 19i) the odds of critical events, i.e., raised or lowered the risk of high toxin concentrations in raw water.

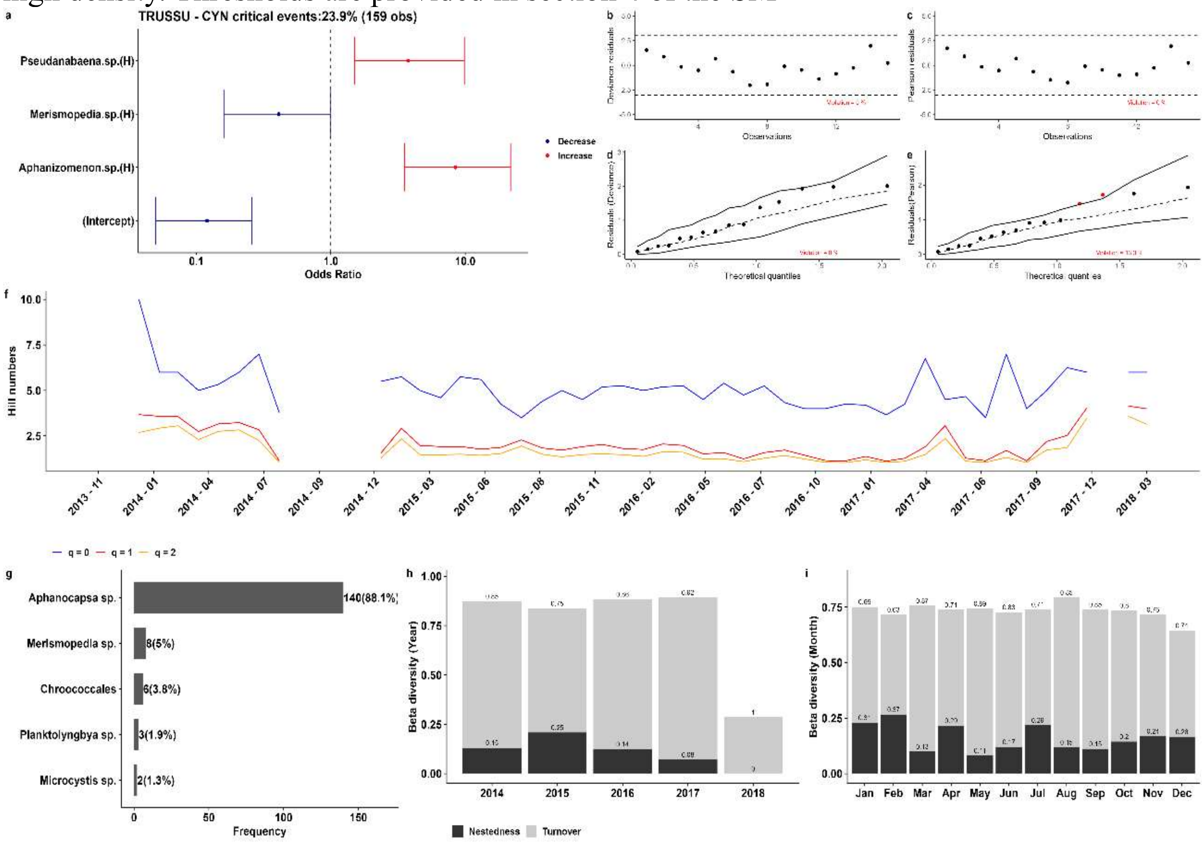
Considering the number of reservoirs, *Chroococcales, Planktothrix sp*., and *Raphidiopsis sp*. stood out in increasing the risk for MC, while *Planktolyngbya sp*. reduced the risk. *Raphidiopsis sp*. and *Merismopedia sp.* increased the risk of STX in more reservoirs, whereas *Aphanocapsa sp.* stood out for reducing it. In the case of CYN, all CyanoBio exhibited similar behavior.

LR also indicated that the total cell density of cyanobacteria and cell densities below the reference limits can serve as bioindicators, providing an additional advantage over IndVal. The total density of cyanobacteria was a relevant factor in increasing risk (Figure 19g), especially for MC, while cell densities below the established limits were more effective in reducing critical events for all the toxins evaluated (Figure 19j).

4.3.3 Bioindicators

Bioindicators were identified in most reservoirs with more than 4% of critical events, such as Trussu (23.9%) shown in Figure 20. The stepwise models presented significant OR and CI (Figure 20a), with good fits (Figures 20b to 20e). In the Hill numbers, the presence of rare taxa and few dominant species, as noted in the previous section, elevated order 0 compared to higher orders, both in Trussu and in other reservoirs (Figure 20f). The monitoring intermittence did not compromise the model fit for CYN in Trussu , as seen in other reservoirs like Jenipapeiro and EMA (see Section 4.8 of SM - APPENDIX C). When the Hill numbers were low (order 0 < 6 and other orders < 2) or the number of observations was limited (< 50), the models showed unsatisfactory fits, especially when the percentage of critical events was below 4%. In these cases, the OR and CI were indeterminate, and severe violations in the residual plots occurred, as observed in Angicos for CYN (see Sections 4.8 of the SM - APPENDIX C).

Figure 20 – Bioindicators of the Trussu reservoir. In (A), the OR (points) and CI (error bars) of bioindicators that increase (red) and decrease (blue) the odds of CYN alert states in raw water are shown. When the dotted vertical line is intercepted, the associations are not significant. In (b) and (c), the deviance and Pearson residuals plots are displayed, and in (d) and (e), the same residuals are shown in simulated envelopes. In (f), cyanobacterial biodiversity is represented by Hill numbers. In (g), the frequency of dominant taxa from 2009 to 2019 is shown. In (h) and (i), temporal beta diversity and its turnover and nestedness components are presented as annual and monthly averages, respectively. The "intercept" represents the low cell density state of the bioindicators, while "H" refers to high density. Thresholds are provided in section 4 of the SM



Source: prepared by the author

In the Trussu reservoir, the cell density of *Aphanizomenon sp.* above the established threshold (median cell density) increased the chances of a critical CYN event by more than 8 times, while *Pseudanabaena sp.* raised the risk by about 4 times. On the other hand, *Merismopedia sp.* significantly reduced the risk, decreasing the chances by 59%. Comparing Figures 20a and 20g, it is observed that most dominant taxa in Trussu were not associated with critical events, and some bioindicators did not even dominate the environment. This pattern is also repeated in other reservoirs, indicating that the simple presence or dominance of potentially toxin-producing taxa in routine monitoring is insufficient to predict the risk of toxin elevation without more sophisticated analyses.

In the reservoirs where CYN levels exceeded established limits, such as in Trussu, the total cyanobacterial density was not associated with critical events. However, for MC, 10 out of 66 reservoirs, and STX, 3 out of 38 reservoirs showed an association between cell density above the median and elevated toxin alert levels in raw water. These results suggest that total cyanobacterial density may not be a universally applicable risk indicator.

Conversely, cell densities below the threshold were consistently associated with a significant reduction in the risk of critical toxin events, as shown by the "intercept" in Figure 20a for the Trussu reservoir. This pattern was also observed in 6 out of 17 reservoirs for CYN (Gavião, Pacoti, among others), 42 out of 66 for MC (Aracoiaba, Orós, among others), and 16 out of 38 for STX (Boqueirão, Canafístula, among others), all detailed in Section 4.8 of the SM. For example, in Trussu, when the densities of *Pseudanabaena sp., Merismopedia sp.,* and *Aphanizomenon sp.* were below the median, the chance of a CYN alert in raw water was reduced by approximately 88%. This suggests that low cell densities were more relevant in defining toxin alert states than simply high concentrations of cyanobacteria in the water.

Across all reservoirs, turnover was identified as the dominant driver of temporal beta diversity, consistently observed in both monthly and annual means, as seen in the Trussu reservoir (Figure 20h and Figure 20i). The significant biodiversity loss, indicated by a sharp decline in Hill numbers from order 0 to higher orders, alongside the critical role of turnover in temporal beta diversity and the heterogeneity detected in the previous unconstrained ordination analyses, were key factors in identifying bioindicators in only 50% of the reservoirs where toxin levels exceeded the established thresholds.

This situation also led to a reduction in the number of pre-selected species due to the stricter thresholds applied in both the IndVal calculations and the a priori algorithm. Nonetheless, despite these limitations, it was still possible to accurately identify bioindicators associated with both increased and decreased odds of elevated toxin concentrations in raw water. The complete list of bioindicators is provided in Section 4.9 of SM - APPENDIX C.

4.4 Discussion

In the last decade, monitoring of the reservoirs revealed that rare taxa increased the Hill numbers of order 0, boosting cyanobacteria richness, but had little impact on orders 1 and 2, which measure evenness. The predominance of a few common taxa reduced evenness values, resulting in similar values in orders 1 and 2. Studies have also reported low ecological indices in reservoirs similar to those assessed. In Zimbabwe, the Malilangwe reservoir recorded an average richness of 4.7 to 7.5 cyanobacterial taxa during the warm-wet season (DALU; WASSERMAN, 2018). In Ceará, reservoirs affected by eutrophication, such as Acarape do Meio and Sítios Novos, showed Shannon indices between 1.03 and 0.6, reflecting the dominance of few species(BARROS *et al.*, 2017a). These results likely contributed to the bow effect observed in the 82 reservoirs assessed (BORCARD; GILLET; LEGENDRE, 2011). These phenomena may be linked to the combined impacts of global warming and eutrophication, which favor the predominance of cyanobacteria in lakes and reservoirs due to their ecophysiological adaptations, such as CO_2 concentration mechanisms, nitrogen fixation, buoyancy regulation, and adaptation to low light levels (ZUO *et al.*, 2024).

Studies indicate that specific environmental conditions directly influence cyanobacterial taxa in water sources (IBELINGS *et al.*, 2021). In the semi-arid region, factors such as eutrophication, high light and nutrient availability, low Secchi depth, prolonged hydraulic retention, and droughts contributed to the dominance and prevalence of common taxa found in the 82 reservoirs studied, as reported in other studies (BRASIL *et al.*, 2016; CAMPOS *et al.*, 2024; DE OLIVEIRA; DANTAS, 2019; FONSECA *et al.*, 2015; MARENGO ORSINI *et al.*, 2018; MENDES *et al.*, 2022)

Between 2012 and 2016, the tropical semi-arid region faced the worst drought event since 1910 (MARENGO ORSINI *et al.*, 2018). During this period, the reduced water volume in shallow reservoirs caused a significant increase in cyanobacterial biomass, especially filamentous heterocystous species, an effect related to increased heterocyst production (MENDES *et al.*, 2022). This scenario contributed to the dominance of these species in the analyzed reservoirs. In the region's reservoirs, the combination of high temperatures and phosphorus and sodium inputs favored the continuous proliferation of *Raphidiopsis raciborskii* and *Planktothrix agardhii* in stationary states (DE OLIVEIRA; DANTAS, 2019). This partly explains *Raphidiopsis sp.*, recognized for its phenotypic plasticity and global dispersal capacity(CAMPOS *et al.*, 2024), as a common organism in all the reservoirs studied, given the presence of these conditions in the 82 reservoirs analyzed.

The cyanobacteria dynamics in the reservoirs were primarily marked by turnover, reinforcing the heterogeneity observed in unconstrained ordination analyses, both on monthly and annual scales. The presence of nestedness indicates a loss of species without replacement, particularly in reservoirs such as Potiretama, Adauto Bezerra, and Do Coronel, when compared to Gavião, Amanary, and Araras. This phenomenon fuels the dominance process, and the effect of biotic homogenization has favored a group of taxa known to be potential toxin producers (CHORUS, I., & WELKER, 2021). In some reservoirs, this homogenization seems to be intensifying, given the more significant role of nestedness in temporal beta diversity. Consequently, these organisms have become more frequent and dominant in drinking water reservoirs, reducing the diversity of common taxa. Other studies have also shown that semi-arid Brazilian reservoirs host similar cyanobacterial communities (BARROS *et al.*, 2017a; BITTENCOURT-OLIVEIRA *et al.*, 2014), with up to three species dominating 80% of the total biomass for weeks without significant variation (DE OLIVEIRA; DANTAS, 2019). This scenario is similar to that observed in temperate regions of Europe, where eutrophication and climate change over the last century have expanded the distribution of toxic cyanobacteria(MONCHAMP *et al.*, 2018).

The presence of some species on both days with and without critical toxin levels made it difficult to associate these organisms with toxic events, allowing the identification of bioindicators in only 50% of the affected reservoirs. This suggests that the mere presence or dominance of toxic taxa, especially in the semiarid region, is insufficient to warn about toxin risks. Confirming toxic blooms require advanced analytical and genomic methods, which are not always available to water utilities. Therefore, it is essential to develop tools that enhance routine monitoring with existing infrastructure, enabling a more adequate response than conventional identification and quantification practices.

About 18.5% of the world's population resides in hyper-arid, arid, and semiarid regions, which together occupy one-third of the Earth's surface, including the state of Ceará. With the biotic homogenization process observed in recent years, which has favored potentially toxic cyanobacteria in the region's reservoirs, challenges to producing safe water have intensified. This highlights the challenges of producing safe water in these regions and underscores the need to develop tools that make routine monitoring by local sanitation companies more effective in managing and mitigating contamination risks.

In this context, water quality bioindicators are species whose presence or absence indicates the state of a water body (SINGH *et al.*, 2013). Plankton, due to their rapid response to environmental changes, are particularly valuable as early warning signs, providing an assessment of aquatic health at a relatively low cost. Their survival depends on abiotic factors such as temperature, dissolved oxygen, and pH, as well as biological interactions. With sensitivity to detect environmental changes and resilience to survive variations, plankton reflect the overall biological response (HOLT; MILLER, 2010).

When individuals are exposed to environmental conditions outside their tolerance range, their physiology and behavior are affected, leading to a decline in overall fitness (the ability to survive, reproduce, and compete effectively within a specific environment). This decline can destabilize population dynamics and cause cascading effects throughout the ecosystem (KAMBOJ *et al.*, 2022). Like plankton in general, cyanobacteria have the ability to respond to the specific variations of the semi-arid region and, therefore, can be used as bioindicators, providing valuable information on ecosystem health, especially concerning high toxin levels in water sources.

When using bioindicators, it must be considered that indicator populations can be influenced by factors such as diseases and predation, complicating the interpretation of the causes of changes, such as increased toxins in raw water (CHANDEL *et al.*, 2023). However, the cyanobacteria selected as bioindicators in this study are adapted to the semiarid reservoir conditions, minimizing these effects. Only organisms common to each water source were considered, respecting the particularities of each ecosystem.

No species can identify all disturbances in any environment, so it is essential to consider the particularities of each ecosystem when selecting bioindicators(CHANDEL *et al.*, 2023). These factors justified a rigorous pre-selection in each reservoir with critical toxin events. Some observed associations suggest the need for further investigation into future studies. An example is *Merismopedia sp.*, which was linked to reducing the risk of CYN in Trussu, while other studies indicate its association with the production of microcystins and nodularins (CHORUS, I., & WELKER, 2021). This scenario reinforces the importance of adjusting monitoring and the selection of bioindicators according to the specificities of each water source and toxin, as done in this study. In addition to influencing the selection of cyanobacteria, environmental conditions may also affect their ability to produce toxins, since the global distribution of a species does not guarantee genetic homogeneity. Regional genotype variations can alter the production of toxic metabolites. A clear example is *R. raciborskii*, which shows higher CYN production in Australian strains compared to Brazilian semi-arid strains (CHORUS, I., & WELKER, 2021; HOFF-RISSETI *et al.*, 2013).

Given the results, cyanobacteria, sensitive to disturbances and resistant to mortality, are suitable bioindicators in semiarid reservoirs, tolerating adverse conditions and reacting proportionally to the presence of toxins in raw water, as indicated by LR and OR. The selection excluded rare species, prioritizing frequent organisms and ensuring their applicability in different environmental conditions. Well-documented taxonomically and behaviorally, cyanobacteria can be considered good bioindicators of toxin alert states in raw water, as recommended (CHANDEL *et al.*, 2023; SIDDIG *et al.*, 2016). Therefore, the use of cyanobacteria as bioindicators in the semi-arid region proves to be an effective and low-cost strategy for monitoring and mitigating the risk of toxin contamination in reservoirs intended for human consumption, complementing existing routine monitoring practices and providing an approach adapted to the specific environmental conditions of the region.

4.5 Conclusion

The biodiversity of the semi-arid region, with low Hill numbers and beta diversity dominated by turnover, but with an important nestedness component in some reservoirs, highlights an advanced process of biotic homogenization. This, combined with the reduced number of critical toxin events and intermittent monitoring, limited the pre-selection of bioindicators. Nevertheless, LR, IndVal, and the a priori algorithm adequately selected CyanoBio for more than 50% of reservoirs with critical levels of CYN, MC, and STX.

LR selected a larger number of CyanoBio and enabled the analysis of total cyanobacterial density. This method also allowed for the evaluation of the effects of cell densities above or below established thresholds. Additionally, ORs with their respective CIs were generated, which, although not predictions, provide a probabilistic estimate of the chances of a critical event occurring in the future. This is because if a bioindicator increased the risk during the analyzed period in a given reservoir, it is reasonable to assume it might indicate a similar risk in the future, as it was selected from the common organisms of that reservoir.

Thus, with the monitoring infrastructure commonly observed in water utility companies, especially in the semi-arid region, it was possible to develop a tool that allows managers to act proactively and specifically in the impacted areas of a particular reservoir, intensifying monitoring or adjusting water treatment plant operations before the critical event occurs. In this regard, LR outperformed the other bioindicator selection methods, being less susceptible to response variable imbalance and the number of observations, with adjustment complexity similar to the a priori algorithm but superior to IndVal.

The identification of taxa responsible for critical toxin events, the unexpected association of some toxic species with reduced risks, the reasons why few reservoirs with high cell densities were associated with increased chances of critical events, and the observation that few critical events occurred despite high cell densities and the homogenization process promoting the dominance of potentially toxic taxa are issues that require further investigation in future studies of semi-arid reservoirs.

5 CONSIDERAÇÕES FINAIS E PROPOSTA DE TRABALHOS FUTUROS

A degradação da qualidade da água causada por CyanoHABs, intensificada pelas mudanças climáticas, exige que empresas de saneamento utilizem sua infraestrutura de monitoramento de forma mais eficiente. Alternativas como o sensoriamento remoto (SR) e o monitoramento de alta resolução (HRM) podem melhorar a precisão sem grandes investimentos. No entanto, os algoritmos para estimar CyanoHABs em regiões tropicais ainda são limitados, assim como o número de parâmetros monitorados pelo HRM, especialmente no que diz respeito à identificação de táxons de cianobactérias. Diante das limitações de resolução dos dados do monitoramento tradicional, que muitas vezes são os únicos recursos disponíveis, os métodos tradicionais de modelagem, com pressupostos difíceis de serem atendidos devido à irregularidade dos dados, tornam-se inadequados. Isso exige a adoção de técnicas mais robustas, como os modelos lineares generalizados (GLMs) e o aprendizado de máquina (ML), que permitem análises mais precisas e adaptadas a essas condições.

Embora o ML seja promissor para reconhecer padrões complexos, inclusive em dados não estruturados, a escassez de informações ainda é um obstáculo. Uma alternativa viável é o uso de modelos que tolerem frequências irregulares ou dados defasados, sem a necessidade de interpolação que possa distorcer os fenômenos observados. GLMs, como regressão logística (RL), foram capazes de lidar com conjuntos de dados menores ou menos regulares. Esses modelos ofereceram uma maior flexibilidade em comparação a métodos tradicionais, como ANOVA ou t-test, permitindo a análise de dados com heterocedasticidade ou autocorrelação, comuns em monitoramentos ambientais. Além disso, a RL pode prever probabilidades, oferecendo uma solução eficiente para superar as limitações dos dados gerados pelo monitoramento convencional, o que auxilia gestores na tomada de decisões antecipadas e mais assertivas, mesmo em cenários com dados limitados.

Em contextos de dados escassos, os índices ecológicos ampliaram a compreensão da biodiversidade utilizando dados de monitoramento regular. Isso possibilitou a análise dos 82 ecossistemas dos reservatórios estudados sem a necessidade de grandes volumes de dados, oferecendo uma ferramenta complementar para avaliar a dinâmica desses ecossistemas. Dessa forma, os gestores obtêm uma visão mais detalhada das tendências ecológicas e das consequências dos filtros ambientais no contexto das mudanças climáticas. Isso expandiu a perspectiva sobre a quantidade e a dinâmica das cianobactérias no tempo e no espaço, permitindo a elaboração de estratégias de gerenciamento mais eficazes.

Ao longo de uma década de monitoramento em 82 reservatórios, foi observado um aumento na abundância de cianobactérias. Esse crescimento, no entanto, veio acompanhado de um processo de homogeneização biótica, onde poucas espécies, muitas delas potencialmente tóxicas, passaram a dominar os ambientes. A seca mais severa desde 1910 segundo a FUNCEME alterou significativamente essa biodiversidade, com os efeitos interanuais se mostrando mais impactantes que os sazonais. Isso sugere que fatores exógenos, como eventos climáticos extremos, aumentam a vulnerabilidade da região semiárida às mudanças climáticas. Esses eventos podem gerar novos pontos de ruptura nas séries temporais que descrevem a diversidade, com consequente aumento da densidade celular de cianobactérias, exigindo tecnologias de tratamento de água mais avançadas.

A adoção de sistemas de monitoramento contínuo torna-se essencial para detectar condições potencialmente perigosas, como elevações de toxinas, possibilitando ajustes rápidos nos processos de tratamento para garantir a qualidade da água. Até que esses sistemas estejam disponíveis, é fundamental desenvolver metodologias que permitam uma gestão proativa dos eventos críticos, utilizando a infraestrutura atual. Diante desse cenário, a RL, o IndVal e o algoritmo a priori se mostraram eficazes na seleção de bioindicadores associados ao risco de eventos críticos de toxinas em mais de 50% dos reservatórios com níveis críticos de CYN, MC e STX, permitindo uma tomada de decisão proativa com base na infraestrutura usual de monitoramento das companhias de saneamento, sem nenhuma necessidade de investimento extra.

Paralelamente, a RL possibilitou a avaliação de como as densidades celulares acima ou abaixo dos limites estabelecidos influenciam o risco de eventos críticos, oferecendo estimativas probabilísticas das chances de ocorrência desses eventos no futuro. Assim, diante do processo de homogeneização biótica, a simples presença de táxons tóxicos ou elevadas densidades celulares não é suficiente para prever o risco, demandando, portanto, estratégias mais adequadas, como a proposta neste estudo, para aprimorar o monitoramento dos reservatórios no semiárido.

Portanto, mesmo não sendo previsões, as OR e os bioindicadores encontrados representam um primeiro passo na construção de uma ferramenta capaz de aprimorar a tomada de decisão dos gestores do saneamento, sem a necessidade de investimento adicional e utilizando a infraestrutura tradicionalmente encontrada na maioria das companhias de saneamento em países em desenvolvimento. Adicionalmente, na ausência de métodos sofisticados que permitam a identificação precisa de espécies ou genes produtores de cianotoxinas durante florações tóxicas, essas ferramentas de gestão, baseadas na identificação de organismos indicativos de condições críticas de cianotoxinas (ou outras condições críticas, como florações ou elevações de turbidez), podem ser integradas aos protocolos das companhias. Isso permitirá ajustes rápidos ou antecipados nos processos de tratamento, ajudando a prevenir a distribuição de água inadequada para o consumo humano. Dessa forma, essa abordagem oferece uma solução eficiente para gestão proativa, sendo uma base para futuros aprimoramentos que possam incluir novas tecnologias e métodos mais avançados de monitoramento.

Propostas para futuros trabalhos incluem:

a) Investigar detalhada dos táxons responsáveis pelos eventos críticos de toxinas e a inesperada associação de algumas espécies tóxicas com a redução dos riscos. Além disso, é essencial
b) explorar porque poucos reservatórios com altas densidades celulares foram associados a um aumento nas chances de eventos críticos. Além disso, em muitos casos eventos críticos não ocorreram ou foram raros (percentual menor do que 5%), apesar das elevadas densidades celulares e do processo de homogeneização que favorece a dominância de táxons potencialmente tóxicos.
c) Desenvolver um protocolo estruturado com base nos resultados obtido. Nesse protocolo deve-se incluir estratégias proativas capazes de prever eventos críticos em reservatórios, estabelecer rotinas claras de ações corretivas e garantir a continuidade e a segurança do abastecimento de água para a população.

Estudar essas questões em maior profundidade nos reservatórios do semiárido pode contribuir significativamente para o aprimoramento das práticas de monitoramento e gestão de riscos associados à qualidade da água em regiões críticas como a do semiárido cearense.

REFERÊNCIAS

ABBOTT, R. D.; GUTGESELL, H. P. Effects of heteroscedasticity and skewness on prediction in regression: Modeling growth of the human heart. **Methods in Enzymology**, v. 240, n. C, p. 37–51, 1994. Disponível em: https://doi.org/10.1016/S0076-6879(94)40042-3

ABDULLAH AL, M.; WANG, W.; JIN, L.; CHEN, H.; XUE, Y.; JEPPESEN, E.; MAJANEVA, M.; XU, H.; YANG, J. Planktonic ciliate community driven by environmental variables and cyanobacterial blooms: A 9-year study in two subtropical reservoirs. **Science of the Total Environment**, v. 858, n. November 2022, p. 159866, 2023. Disponível em: https://doi.org/10.1016/j.scitotenv.2022.159866

AGUILERA, A.; ALMANZA, V.; HAAKONSSON, S.; PALACIO, H.; BENITEZ, G. A.; BARROS, U. G.; URRUTIA, R.; AUBRIOT, L.; BONILLA, S. Cyanobacterial bloom monitoring and assessment in Latin America. v. 125, n. March, 2023. Disponível em: https://doi.org/10.1016/j.hal.2023.102429

AHMED, A. A.; SAYED, S.; ABDOULHALIK, A.; MOUTARI, S.; OYEDELE, L. Applications of machine learning to water resources management: A review of present status and future opportunities. **Journal of Cleaner Production**, v. 441, n. August 2023, p. 140715, 2024. Disponível em: https://doi.org/10.1016/j.jclepro.2024.140715

AKIRA S. MORI, FOREST ISBELL, R. S. β-Diversity, Community Assembly, and Ecosystem Functioning. **Trends in Ecology & Evolution,** v. 33, n. 7, p. 549–564, 2018. Disponível em: https://doi.org/https://doi.org/10.1016/j.tree.2018.04.012

ALMUHTARAM, H.; ZAMYADI, A.; HOFMANN, R. Machine learning for anomaly detection in cyanobacterial fluorescence signals. **Water Research**, v. 197, p. 117073, 2021. Disponível em: https://doi.org/10.1016/j.watres.2021.117073

ALVAREZ DALINGER, F. S.; LOZANO, V. L.; BORJA, C. N.; MORAÑA, L. B.; MARÍA MÓNICA, S. Short-Term Meteorological Conditions Explain Cyanobacterial Blooms in a Tropical Reservoir. **Water (Switzerland)**, v. 15, n. 2, p. 1–16, 2023. Disponível em: https://doi.org/10.3390/w15020302

ANDERSON, M. J.; WALSH, D. C. I. PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing? **Ecological Monographs**, v. 83, n. 4, p. 557–574, 2013. Disponível em: https://doi.org/10.1890/12-2010.1

APHA-AWWA-WEF. **Standard Methods for the Examination of Water and Wastewater**. 21th. ed. New York: *[s. n.]*, 2005.

ASSESSMENT, M. E. **Ecosystems and human well-being: wetlands and water**. *[S. l.]*: World resources institute, 2005. v. 5 Disponível em: https://doi.org/10.1080/17518253.2011.584217

BARRETO, J. R.; PARDINI, R.; METZGER, J. P.; SILVA, F. A. B.; NICHOLS, E. S. When forest loss leads to biodiversity gain: Insights from the Brazilian Atlantic Forest. **Biological Conservation**, v. 279, n. June 2022, p. 109957, 2023. Disponível em: https://doi.org/10.1016/j.biocon.2023.109957

BARRINGTON, D. J.; GHADOUANI, A.; SINANG, S. C.; IVEY, G. N. Development of a new risk-based framework to guide investment in water quality monitoring. **Environmental Monitoring and Assessment**, v. 186, n. 4, p. 2455–2464, 2014. Disponível em: https://doi.org/10.1007/s10661-013-3552-1

BARROS, M. U. G.; LOPES, I. K. C.; CARVALHO, S. M. de C.; CAPELO NETO, J. **Impact of filamentous cyanobacteria on the water quality of two tropical reservoirs**. *[S. l.: s. n.]* Disponível em: https://doi.org/10.1590/2318-0331.011716072

BARROS, M. U. G.; LOPES, I. K. C.; CARVALHO, S. M. de C.; CAPELO NETO, J. Impact of filamentous cyanobacteria on the water quality of two tropical reservoirs Impacto. **Rbrh**, v. 22, n. 0, 2017 b. Disponível em: https://doi.org/10.1590/2318-0331.011716072

BASELGA, A.; ORME, C. D. L. betapart: an R package for the study of beta diversity. **Methods in Ecology and Evolution**, v. 3, n. 5, p. 808–812, 2012. Disponível em: https://doi.org/10.1111/j.2041-210X.2012.00224.x

BASELGA, A.; ORME, D.; VILLEGER, S.; DE, B. J.; FABIEN, L.; MAXIME, L.; MARTINEZ-SANTALLA, S.; MARTIN-DEVASA , RAMIRO GOMEZ-RODRIGUEZ, C. **CRAN - Package betapart**. *[s. l.]*, 2023. Disponível em: https://cran.r-project.org/web/packages/betapart/index.html. Acesso em: 4 mar. 2024.

BERTONE, E.; BURFORD, M. A.; HAMILTON, D. P. Fluorescence probes for real-time remote cyanobacteria monitoring: A review of challenges and opportunities. **Water Research**, v. 141, p. 152–162, 2018. Disponível em: https://doi.org/10.1016/j.watres.2018.05.001

BERTONE, E.; CHUANG, A.; BURFORD, M. A.; HAMILTON, D. P. In-situ fluorescence monitoring of cyanobacteria: Laboratory-based quantification of species-specific measurement accuracy. **Harmful Algae**, v. 87, n. June, p. 101625, 2019. Disponível em: https://doi.org/10.1016/j.hal.2019.101625

BERTRAND, G. F.; DE PAIVA, A. L. R.; DE ARAÚJO FREITAS, J. B.; DA SILVA PEREIRA CABRAL, J. J.; VERAS ALBUQUERQUE, T. B.; DE CARVALHO FILHO, J. A. A. River bank filtration in tropical metropoles: integrated evaluation of physical, geochemical and biochemical interactions in Recife, NE Brazil. **International Journal of Environmental Science and Technology**, v. 19, n. 7, p. 5803–5818, 2022. Disponível em: https://doi.org/10.1007/s13762-021-03558-x

BHATT, P.; ENGEL, B. A.; REUHS, M.; SIMSEK, H. Cyanophage technology in removal of cyanobacteria mediated harmful algal blooms: A novel and eco-friendly method. **Chemosphere**, v. 315, n. January, p. 137769, 2023. Disponível em: https://doi.org/10.1016/j.chemosphere.2023.137769

BITTENCOURT-OLIVEIRA, M. D. C.; PICCIN-SANTOS, V.; MOURA, A. N.; ARAGÃO-TAVARES, N. K. C.; CORDEIRO-ARAÚJO, M. K. Cyanobacteria, microcystins and cylindrospermopsin in public drinking supply reservoirs of Brazil. **Anais da Academia Brasileira de Ciências**, v. 86, n. 1, p. 297–310, 2014. Disponível em: https://doi.org/10.1590/0001-3765201302512

BORCARD, D.; GILLET, F.; LEGENDRE, P. **Numerical Ecology with R**. *[S. l.: s. n.]*. Disponível em: https://doi.org/10.1007/978-1-4419-7976-6

BRASIL. **Resolução CONAMA nº 357, de 17 de março de 2005.** *[s. l.]*, 2005. Disponível em: http://www.siam.mg.gov.br/sla/download.pdf?idNorma=2747. Acesso em: 2 dez. 2022.

BRASIL, J.; ATTAYDE, J. L.; VASCONCELOS, F. R.; DANTAS, D. D. F.; HUSZAR, V. L. M. Drought-induced water-level reduction favors cyanobacteria blooms in tropical shallow lakes. **Hydrobiologia**, v. 770, n. 1, p. 145–164, 2016. Disponível em: https://doi.org/10.1007/s10750-015-2578-5

BRAZIL. **Portaria MS n.º 518**. *[S. l.: s. n.]*

BRAZIL. **PORTARIA GM/MS Nº 888, DE 4 DE MAIO DE 2021**. *[s. l.]*, 2021. Disponível em: https://bvsms.saude.gov.br/bvs/saudelegis/gm/2021/prt0888_07_05_2021.html. Acesso em: 13 out. 2023.

BRESCIANI, M. Recognizing harmful algal bloom based on remote sensing reflectance band ratio. **Journal of Applied Remote Sensing**, v. 5, n. 1, p. 053556, 2011. Disponível em: https://doi.org/10.1117/1.3630218

CABALLERO, C. B.; BIGGS, T. W.; VERGOPOLAN, N.; WEST, T. A. P.; RUHOFF, A. Transformation of Brazil's biomes: The dynamics and fate of agriculture and pasture expansion into native vegetation. **Science of the Total Environment**, v. 896, n. April, p. 166323, 2023. Disponível em: https://doi.org/10.1016/j.scitotenv.2023.166323

CAMPOS, T. G. V.; GAMA, W. A.; GERALDES, V.; YOON, J.; CRNKOVIC, C. M.; PINTO, E.; JACINAVICIUS, F. R. New records on toxic cyanobacteria from Brazil: Exploring their occurrence and geography. **Science of the Total Environment**, v. 931, n. September 2023, p. 172689, 2024. Disponível em: https://doi.org/10.1016/j.scitotenv.2024.172689

CAZZOLLA GATTI, R.; AMOROSO, N.; MONACO, A. Estimating and comparing biodiversity with a single universal metric. **Ecological Modelling**, v. 424, p. 109020, 2020. Disponível em: https://doi.org/10.1016/j.ecolmodel.2020.109020

CHANDEL, P.; MAHAJAN, D.; THAKUR, K.; KUMAR, R.; KUMAR, S.; BRAR, B.; SHARMA, D.; SHARMA, A. K. A review on plankton as a bioindicator: A promising tool for monitoring water quality. **World Water Policy**, n. April, p. 1–20, 2023. Disponível em: https://doi.org/10.1002/wwp2.12137

CHANDLER, R. E. Multisite, multivariate weather generation based on generalised linear models. **Environmental Modelling and Software**, v. 134, p. 104867, 2020. Disponível em: https://doi.org/10.1016/j.envsoft.2020.104867

CHANG, N. Bin; IMEN, S.; VANNAH, B. Remote sensing for monitoring surface water quality status and ecosystem state in relation to the nutrient cycle: A 40-year perspective. **Critical Reviews in Environmental Science and Technology**, v. 45, n. 2, p. 101–166, 2015. Disponível em: https://doi.org/10.1080/10643389.2013.829981

CHATURVEDI, A.; PANDEY, B.; YADAV, A. K.; SAROJ, S. **An overview of the potential impacts of global climate change on water resources**. *[S. l.]*: INC, 2021. *E-book*. Disponível em: https://doi.org/10.1016/B978-0-12-820200-5.00012-9

CHEN, Q.; RUI, H.; LI, W.; ZHANG, Y. Analysis of algal bloom risk with uncertainties in lakes by integrating self-organizing map and fuzzy information theory. **Science of the Total Environment**, v. 482–483, n. 1, p. 318–324, 2014. Disponível em: https://doi.org/10.1016/j.scitotenv.2014.02.096

CHEUNG, M. Y.; LIANG, S.; LEE, J. Toxin-producing cyanobacteria in freshwater: A review of the problems, impact on drinking water safety, and efforts for protecting public health. **Journal of Microbiology**, v. 51, n. 1, p. 1–10, 2013. Disponível em: https://doi.org/10.1007/s12275-013-2549-3

CHORUS, I., & WELKER, M. **Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring and Management**. *[S. l.]*: CRC Press., 2021. Disponível em: https://doi.org/https://doi.org/10.1201/9781003081449

CLARK, J. M.; SCHAEFFER, B. A.; DARLING, J. A.; URQUHART, E. A.; JOHNSTON, J. M.; IGNATIUS, A. R.; MYER, M. H.; LOFTIN, K. A.; WERDELL, P. J.; STUMPF, R. P. Satellite monitoring of cyanobacterial harmful algal bloom frequency in recreational waters and drinking water sources. **Ecological Indicators**, v. 80, n. November 2016, p. 84–95, 2017. Disponível em: https://doi.org/10.1016/j.ecolind.2017.04.046

CLARKE, K. R. Non-parametric multivariate analyses of changes in community structure. **Australian Journal of Ecology**, v. 18, n. 1, p. 117–143, 1993. Disponível em: https://doi.org/10.1111/j.1442-9993.1993.tb00438.x

CLEMENTE, A.; WILSON, A.; OLIVEIRA, S.; MENEZES, I.; GOIS, A.; CAPELO-NETO, J. The role of hydraulic conditions of coagulation and flocculation on the damage of cyanobacteria. **Science of the Total Environment**, v. 740, p. 139737, 2020. Disponível em: https://doi.org/10.1016/j.scitotenv.2020.139737

CLERCIN, N. A.; DRUSCHEL, G. K. Influence of Environmental Factors on the Production of MIB and Geosmin Metabolites by Bacteria in a Eutrophic Reservoir. **Water Resources Research**, v. 55, n. 7, p. 5413–5430, 2019. Disponível em: https://doi.org/10.1029/2018WR023651

COFFER, M. M.; SCHAEFFER, B. A.; DARLING, J. A.; URQUHART, E. A.; SALLS, W. B. Quantifying national and regional cyanobacterial occurrence in US lakes using satellite remote sensing. **Ecological Indicators**, v. 111, n. November 2019, 2020. Disponível em: https://doi.org/10.1016/j.ecolind.2019.105976

COFFER, M. M.; SCHAEFFER, B. A.; SALLS, W. B.; URQUHART, E.; LOFTIN, K. A.; STUMPF, R. P.; WERDELL, P. J.; DARLING, J. A. Satellite remote sensing to assess cyanobacterial bloom frequency across the United States at multiple spatial scales. **Ecological Indicators**, v. 128, p. 107822, 2021. Disponível em: https://doi.org/10.1016/j.ecolind.2021.107822

COGERH. **Ceará Hydrological Portal**. *[s. l.]*, 2023. Disponível em: http://www.hidro.ce.gov.br/. Acesso em: 14 dez. 2023.

COOK, K. V.; BEYER, J. E.; XIAO, X.; HAMBRIGHT, K. D. Ground-based remote sensing provides alternative to satellites for monitoring cyanobacteria in small lakes. **Water Research**, v. 242, n. May, p. 120076, 2023. Disponível em: https://doi.org/10.1016/j.watres.2023.120076

COURTOIS, S.; STEINMANN, D.; CAJON, A.; VAN DER LINDEN, L. Continuous monitoring of cyanobacterial blooms: benefits and conditions for using fluorescence probes. **Revue des sciences de l'eau**, v. 30, n. 2, p. 149–155, 2018. Disponível em: https://doi.org/10.7202/1042923ar

DALU, T.; WASSERMAN, R. J. Cyanobacteria dynamics in a small tropical reservoir: Understanding spatio-temporal variability and influence of environmental variables. **Science of the Total Environment**, v. 643, p. 835–841, 2018. Disponível em: https://doi.org/10.1016/j.scitotenv.2018.06.256

DE OLIVEIRA, F. H. P. C.; DANTAS, Ê. W. Sodium and nitrate favor the steady state of cyanobacteria in a semiarid ecosystem. **Acta Scientiarum - Biological Sciences**, v. 41, n. 1, p. 1–13, 2019. Disponível em: https://doi.org/10.4025/actascibiolsci.v41i1.47924

DENG, J.; PAERL, H. W.; QIN, B.; ZHANG, Y.; ZHU, G.; JEPPESEN, E.; CAI, Y.; XU, H. Climatically-modulated decline in wind speed may strongly affect eutrophication in shallow lakes. **Science of the Total Environment**, v. 645, p. 1361–1370, 2018. Disponível em: https://doi.org/10.1016/j.scitotenv.2018.07.208

DEVI, A.; CHIU, Y. T.; HSUEH, H. T.; LIN, T. F. Quantitative PCR based detection system for cyanobacterial geosmin/2-methylisoborneol (2-MIB) events in drinking water sources: Current status and challenges. **Water Research**, v. 188, p. 116478, 2021. Disponível em: https://doi.org/10.1016/j.watres.2020.116478

DÍAZ, S. *et al.* The IPBES Conceptual Framework - connecting nature and people. **Current Opinion in Environmental Sustainability**, v. 14, p. 1–16, 2015. Disponível em: https://doi.org/10.1016/j.cosust.2014.11.002

DOBSON, A. J.; BARNETT, A. G. **An Introduction to Generalized Linear Models**. *[S. l.]*: Chapman and Hall/CRC, 2019.

DORNELAS, M.; GOTELLI, N. J.; MCGILL, B.; SHIMADZU, H.; MOYES, F.; SIEVERS, C.; MAGURRAN, A. E. Assemblage Time Series Reveal Biodiversity Change but Not Systematic Loss. **Science**, v. 344, n. 6181, p. 296–299, 2014. Disponível em: https://doi.org/10.1126/science.1248484

ESA. **Envisat - Earth Online**. *[s. l.]*, 2024. Disponível em: https://earth.esa.int/eogateway/instruments/meris. Acesso em: 26 maio. 2024.

FERNANDEZ-FIGUEROA, E. G.; BULEY, R. P.; BARROS, M. U. G.; GLADFELTER, M. F.; MCCLIMANS, W. D.; WILSON, A. E. Carlson’s Trophic State Index is a poor predictor of cyanobacterial dominance in drinking water reservoirs. **AWWA Water Science**, v. 3, n. 2, 2021. Disponível em: https://doi.org/10.1002/aws2.1219

FILBRUN, J. E.; CONROY, J. D.; CULVER, D. A. Understanding seasonal phosphorus dynamics to guide effective management of shallow, hypereutrophic Grand Lake St. Marys, Ohio. **Lake and Reservoir Management**, v. 29, n. 3, p. 165–178, 2013. Disponível em: https://doi.org/10.1080/10402381.2013.823469

FONSECA, J. R.; VIEIRA, P. C. S.; KUJBIDA, P.; COSTA, I. A. S. da. Cyanobacterial occurrence and detection of microcystins and saxitoxins in reservoirs of the Brazilian semi-arid. **Acta Limnologica Brasiliensia**, v. 27, n. 1, p. 78–92, 2015. Disponível em: https://doi.org/10.1590/s2179-975x2814

FOURNIER, C.; QUESADA, A.; CIRÉS, S.; SABERIOON, M. Discriminating bloom-forming cyanobacteria using lab-based hyperspectral imagery and machine learning: Validation with toxic species under environmental ranges. **Science of the Total Environment**, v. 932, n. September 2023, 2024. Disponível em: https://doi.org/10.1016/j.scitotenv.2024.172741

FRANCY, D. S.; BRADY, A. M. G.; STELZER, E. A.; CICALE, J. R.; HACKNEY, C.; DALBY, H. D.; STRUFFOLINO, P.; DWYER, D. F. Predicting microcystin concentration action-level exceedances resulting from cyanobacterial blooms in selected lake sites in Ohio. **Environmental Monitoring and Assessment**, v. 192, n. 8, 2020. Disponível em: https://doi.org/10.1007/s10661-020-08407-x

FUCEME. **Ceará is experiencing the worst prolonged drought since 1910 - FUNCEME: Ceará Foundation for Meteorology and Water Resources**. *[s. l.]*, 2016. Disponível em: http://www.funceme.br/?p=1403. Acesso em: 11 mar. 2024.

GAINSBURY, A. M.; SANTOS, E. G.; WIEDERHECKER, H. Does urbanization impact terrestrial vertebrate ectotherms across a biodiversity hotspot? **Science of the Total Environment**, v. 835, n. October 2021, 2022. Disponível em: https://doi.org/10.1016/j.scitotenv.2022.155446

GOES, M. C. A.; BARROS, M. U. G.; NETO, I. E. L. Prediction of total phosphorus in reservoir cascade systems. **Environmental Monitoring and Assessment**, v. 195, n. 12, p. 1–12, 2023. Disponível em: https://doi.org/10.1007/s10661-023-12155-z

GOLSHAN, A.; EVANS, C.; GEARY, P.; MORROW, A.; MAEDER, M.; TAULER, R. Patterns of cyanobacterial abundance in a major drinking water reservoir: what 3 years of comprehensive monitoring data reveals? **Environmental Monitoring and Assessment**, v. 192, n. 2, 2020. Disponível em: https://doi.org/10.1007/s10661-020-8090-z

GÓMEZ, J. A. D.; ALONSO, C. A.; GARCÍA, A. A. Remote sensing as a tool for monitoring water quality parameters for Mediterranean Lakes of European Union water framework directive (WFD) and as a system of surveillance of cyanobacterial harmful algae blooms (SCyanoHABs). **Environmental Monitoring and Assessment**, v. 181, n. 1–4, p. 317–334, 2011. Disponível em: https://doi.org/10.1007/s10661-010-1831-7

GORNEY, R. M.; JUNE, S. G.; STAINBROOK, K. M.; SMITH, A. J. Detections of cyanobacteria harmful algal blooms (cyanoHABs) in New York State, United States (2012–2020). **Lake and Reservoir Management**, v. 39, n. 1, p. 21–36, 2023. Disponível em: https://doi.org/10.1080/10402381.2022.2161436

GOYENS, C.; LAVIGNE, H.; DILLE, A.; VERVAEREN, H. Using Hyperspectral Remote Sensing to Monitor Water Quality in Drinking Water Reservoirs. **Remote Sensing**, v. 14, n. 21, p. 5607, 2022. Disponível em: https://doi.org/10.3390/rs14215607

GUNKEL, G.; SOBRAL, M. Re-oligotrophication as a challenge for tropical reservoir management with reference to Itaparica Reservoir, São Francisco, Brazil. **Water Science and Technology**, v. 67, n. 4, p. 708–714, 2013. Disponível em: https://doi.org/10.2166/wst.2012.583

HAMILTON, A. J. Species diversity or biodiversity? **Journal of Environmental Management**, v. 75, n. 1, p. 89–92, 2005. Disponível em: https://doi.org/10.1016/j.jenvman.2004.11.012

HANDLER, A. M.; COMPTON, J. E.; HILL, R. A.; LEIBOWITZ, S. G.; SCHAEFFER, B. A. Identifying lakes at risk of toxic cyanobacterial blooms using satellite imagery and field surveys across the United States. **Science of the Total Environment**, v. 869, n. October 2022, p. 161784, 2023. Disponível em: https://doi.org/10.1016/j.scitotenv.2023.161784

HARRIS, T. D.; GRAHAM, J. L. Predicting cyanobacterial abundance, microcystin, and geosmin in a eutrophic drinking-water reservoir using a 14-year dataset. **Lake and Reservoir Management**, v. 33, n. 1, p. 32–48, 2017. Disponível em: https://doi.org/10.1080/10402381.2016.1263694

HAYES, W. M.; VOIGT, M.; ROSA, I.; CORT, K. A.; KOTLINSKI, N.; KALAMANDEEN, M.; DAVIES, Z. G.; BICKNELL, J. E. Predicting the loss of forests, carbon stocks and biodiversity driven by a neotropical 'gold rush'. **Biological Conservation**, v. 286, n. September, p. 110312, 2023. Disponível em: https://doi.org/10.1016/j.biocon.2023.110312

HELSEL, D. R.; FRANS, L. M. Regional Kendall Test for Trend. **Environmental Science & Technology**, v. 40, n. 13, p. 4066–4073, 2006. Disponível em: https://doi.org/10.1021/es051650b

HIGGINS, J.; THOMAS, J.; CHANDLER, J.; CUMPSTON, M.; LI, T.; PAGE, M.; WELCH, V. (editors). **Cochrane Handbook for Systematic Reviews of Interventions | Cochrane Training**. *[s. l.]*, 2020. Disponível em: https://training.cochrane.org/handbook/current. Acesso em: 27 jan. 2021.

HILL, M. O. Diversity and Evenness: A Unifying Notation and Its Consequences. **Ecology**, v. 54, n. 2, p. 427–432, 1973. Disponível em: https://doi.org/10.2307/1934352
HOFF-RISSETI, C.; DÖRR, F. A.; SCHAKER, P. D. C.; PINTO, E.; WERNER, V. R.; FIORE, M. F. Cylindrospermopsin and Saxitoxin Synthetase Genes in Cylindrospermopsis raciborskii Strains from Brazilian Freshwater. **PLoS ONE**, v. 8, n. 8, p. 35–39, 2013. Disponível em: https://doi.org/10.1371/journal.pone.0074238

HOLT, E. a.; MILLER, S. W. Bioindicators: using organisms to measure environmental impacts. **Nature Education Knowledge**, v. 3, n. 10, p. 1–8, 2010. Disponível em: https://www.nature.com/scitable/knowledge/library/bioindicators-using-organisms-to-measure-environmental-impacts-16821310/. Acesso em: 10 out. 2024.

HORNING, N. Remote Sensing. *In*: **Encyclopedia of Ecology**. *[S. l.]*: Elsevier, 2008. p. 2986–2994. *E-book*. Disponível em: https://doi.org/10.1016/B978-008045405-4.00237-8

HUANG, R.; MA, C.; MA, J.; HUANGFU, X.; HE, Q. Machine learning in natural and engineered water systems. **Water Research**, v. 205, n. September, p. 117666, 2021. Disponível em: https://doi.org/10.1016/j.watres.2021.117666

HUANG, Y.; FENG, J. C.; KONG, J.; SUN, L.; ZHANG, M.; HUANG, Y.; TANG, L.; ZHANG, S.; YANG, Z. Community assemblages and species coexistence of prokaryotes controlled by local environmental heterogeneity in a cold seep water column. **Science of the Total Environment**, v. 868, n. October 2022, p. 161725, 2023. Disponível em: https://doi.org/10.1016/j.scitotenv.2023.161725

HURART, S. H. The nonconcept of species diversity: a critique and alternative parameters. **Ecology**, v. 52, n. 4, p. 577–586, 1971.

HURTADO, I.; POUGET, L.; FERNÁNDEZ, S.; CASCALES, P. Monitoring and forecasting cyanobacteria risk for a drinking water plant in Spain. **Water Supply**, v. 22, n. 7, p. 6296–6307, 2022. Disponível em: https://doi.org/10.2166/ws.2022.246

HWAN, B. J.; KIM, H. N.; KANG, T. G.; KIM, B. H.; BYEON, M. S. Study of the cause of the generation of odor compounds (geosmin and 2-methylisoborneol) in the Han River system, the drinking water source, Republic of Korea. **Water Supply**, v. 23, n. 3, p. 1081–1093, 2023. Disponível em: https://doi.org/10.2166/ws.2023.037

HWANG, S. Y.; CHOI, B. W.; PARK, J. H.; SHIN, D. S.; CHUNG, H. S.; SON, M. S.; LIM, C. H.; CHAE, H. M.; HA, D. W.; JUNG, K. Y. Evaluating Statistical Machine Learning Algorithms for Classifying Dominant Algae in Juam Lake and Tamjin Lake, Republic of Korea. **Water (Switzerland)**, v. 15, n. 9, 2023. Disponível em: https://doi.org/10.3390/w15091738

IBELINGS, B. W.; KURMAYER, R.; AZEVEDO, S. M. F. O.; WOOD, S. A.; CHORUS, I.; WELKER, M. Understanding the occurrence of cyanobacteria and cyanotoxins. *In*: **Toxic Cyanobacteria in Water**. *[S. l.: s. n.]*. p. 213–294. Disponível em: https://doi.org/10.1201/9781003081449-4

IBGE. **IBGE map portal**. *[s. l.]*, 2017. Disponível em: https://portaldemapas.ibge.gov.br/portal.php#mapa221109. Acesso em: 19 set. 2023.

IBGE. **Ceará | Cities and States | IBGE**. *[s. l.]*, 2023. Disponível em: https://www.ibge.gov.br/en/cities-and-states/ce.html. Acesso em: 18 nov. 2023.

INMET. **National Meteorology Institute**. *[s. l.]*, 2023. Disponível em: https://portal.inmet.gov.br/dadoshistoricos. Acesso em: 19 set. 2023.

IPBES. **Summary for policymakers of the global assessment report on biodiversity and ecosystem services**. *[S. l.: s. n.]*. v. 45*E-book*. Disponível em: https://zenodo.org/record/3553579#.YfmYTerMI2w

IPCC. **Intergovernmental Panel On Climate Change - Climate Change 2022 – Impacts, Adaptation and Vulnerability**. *[S. l.]*: Cambridge University Press, 2023. *E-book*. Disponível em: https://doi.org/10.1017/9781009325844

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. **An introduction to statistical learning**. New York: Springer, 2013. Disponível em: https://doi.org/10.1080/24754269.2021.1980261

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. **An Introduction to Statistical Learning: with Applications in R**. New York: Springer, 2014. *E-book*. Disponível em: https://doi.org/10.1007/978-1-4614-7138-7

JENNIONS, M. D.; MØLLER, A. P. A survey of the statistical power of research in behavioral ecology and animal behavior. **Behavioral Ecology**, v. 14, n. 3, p. 438–445, 2003. Disponível em: https://doi.org/10.1093/beheco/14.3.438

JEONG, B.; CHAPETA, M. R.; KIM, M.; KIM, J.; SHIN, J.; CHA, Y. K. Machine learning-based prediction of harmful algal blooms in water supply reservoirs. **Water Quality Research Journal**, v. 57, n. 4, p. 304–318, 2022. Disponível em: https://doi.org/10.2166/wqrj.2022.019

JOST, L. Entropy and diversity. **Oikos**, v. 113, n. 2, p. 363–375, 2006. Disponível em: https://doi.org/10.1111/j.2006.0030-1299.14714.x

JOST, L. Partitioning diversity into independent Alpha and Beta components. **Ecology**, v. 88, n. 10, p. 2427–2439, 2007. Disponível em: https://doi.org/DOI: 10.1890/06-1736.1

JOST, L. **The New Synthesis of Diversity Indices and Similarity Measures**. *[s. l.]*, 2009. Disponível em: http://www.loujost.com/Statistics and Physics/Diversity and Similarity/DiversitySimilarityHome.htm. Acesso em: 17 nov. 2023.

KAKADE, A.; SALAMA, E. S.; HAN, H.; ZHENG, Y.; KULSHRESTHA, S.; JALALAH, M.; HARRAZ, F. A.; ALSAREII, S. A.; LI, X. World eutrophic pollution of lake and river: Biotreatment potential and future perspectives. **Environmental Technology and Innovation**, v. 23, p. 101604, 2021. Disponível em: https://doi.org/10.1016/j.eti.2021.101604

KAKOUEI, K.; KRAEMER, B. M.; ADRIAN, R. Variation in the predictability of lake plankton metric types. **Limnology and Oceanography**, v. 67, n. 3, p. 608–620, 2022. Disponível em: https://doi.org/10.1002/lno.12021

KAMBOJ, V.; KAMBOJ, N.; SHARMA, A. K.; BISHT, A. Phytoplankton communities as bio-indicators of water quality in a mining-affected area of the river Ganga, Haridwar, India. **Energy, Ecology and Environment**, v. 7, n. 4, p. 425–438, 2022. Disponível em: https://doi.org/10.1007/s40974-022-00238-5

KASSAMBARA, A.; MUNDT, F. **factoextra: Extract and Visualize the Results of Multivariate Data Analyses**. *[s. l.]*, 2020. Disponível em: https://cran.r-project.org/package=factoextra%7D. Acesso em: 24 dez. 2022.

KIM, M. E.; SHON, T. S.; SHIN, H. S. Forecasting algal bloom (chl-a) on the basis of coupled wavelet transform and artificial neural networks at a large lake. **Desalination and Water Treatment**, v. 51, n. 19–21, p. 4118–4128, 2013. Disponível em: https://doi.org/10.1080/19443994.2013.781108

KINDT, R.; OKSANEN, J. **R: Renyi and Hill Diversities and Corresponding Accumulation...** *[s. l.]*, 2022. Disponível em: https://search.r-project.org/CRAN/refmans/vegan/html/renyi.html. Acesso em: 13 nov. 2023.

KNIEF, U.; FORSTMEIER, W. Violating the normality assumption may be the lesser of two evils. **Behavior Research Methods**, 2021. Disponível em: https://doi.org/10.3758/s13428-021-01587-5

KOOKANA, R. S.; DRECHSEL, P.; JAMWAL, P.; VANDERZALM, J. Urbanisation and emerging economies: Issues and potential solutions for water and food security. **Science of the Total Environment**, v. 732, p. 139057, 2020. Disponível em: https://doi.org/10.1016/j.scitotenv.2020.139057

KORHONEN, J. J.; SOININEN, J.; HILLEBRAND, H. A quantitative analysis of temporal turnover in aquatic species assemblages across ecosystems. **Ecology**, v. 91, n. 2, p. 508–517, 2010. Disponível em: https://doi.org/10.1890/09-0392.1

KRÖGER, B.; SELMECZY, G. B.; CASPER, P.; SOININEN, J.; PADISÁK, J. Long-term phytoplankton community dynamics in Lake Stechlin (north-east Germany) under sudden and heavily accelerating eutrophication. **Freshwater Biology**, v. 68, n. 5, p. 737–751, 2023. Disponível em: https://doi.org/10.1111/fwb.14060

KUCZYŃSKA-KIPPEN, N.; KOZAK, A.; CELEWICZ, S. Cyanobacteria respond to trophic status in shallow aquatic ecosystems. **Science of the Total Environment**, v. 949, n. July, 2024. Disponível em: https://doi.org/10.1016/j.scitotenv.2024.174932

KUDELA, R. M.; PALACIOS, S. L.; AUSTERBERRY, D. C.; ACCORSI, E. K.; GUILD, L. S.; TORRES-PEREZ, J. Application of hyperspectral remote sensing to cyanobacterial blooms in inland waters. **Remote Sensing of Environment**, v. 167, p. 196–205, 2015. Disponível em: https://doi.org/10.1016/j.rse.2015.01.025

KWON, D. H.; HONG, S. M.; ABBAS, A.; PYO, J.; LEE, H. K.; BAEK, S. S.; CHO, K. H. Inland harmful algal blooms (HABs) modeling using internet of things (IoT) system and deep learning. **Environmental Engineering Research**, v. 28, n. 1, p. 0–1, 2023. Disponível em: https://doi.org/10.4491/eer.2021.280

LÊ, S.; JOSSE, J.; HUSSON, F. FactoMineR : An R Package for Multivariate Analysis. **Journal of Statistical Software**, v. 25, n. 1, p. 1–18, 2008. Disponível em: https://doi.org/10.18637/jss.v025.i01

LEE, D. Y.; LEE, D. S.; CHA, Y. K.; MIN, J. H.; PARK, Y. S. Data-driven models for predicting community changes in freshwater ecosystems: A review. **Ecological Informatics**, v. 77, n. June, p. 102163, 2023. Disponível em: https://doi.org/10.1016/j.ecoinf.2023.102163

LEGENDRE, P. Indicator Species: Computation. **Encyclopedia of Biodiversity: Second Edition**, v. 4, p. 264–268, 2013. Disponível em: https://doi.org/10.1016/B978-0-12-384719-5.00430-5

LEGENDRE, P.; LEGENDRE, L. **Numerical ecology**. *[S. l.]*: Elsevier, 2012.

LEI, L.; PENG, L.; HUANG, X.; HAN, B. P. Occurrence and dominance of Cylindrospermopsis raciborskii and dissolved cylindrospermopsin in urban reservoirs used for drinking water supply, South China. **Environmental Monitoring and Assessment**, v. 186, n. 5, p. 3079–3090, 2014. Disponível em: https://doi.org/10.1007/s10661-013-3602-8

LEOTE, P.; CAJAIBA, R. L.; CABRAL, J. A.; BRESCOVIT, A. D.; SANTOS, M. Are data-mining techniques useful for selecting ecological indicators in biodiverse regions? Bridges between market basket analysis and indicator value analysis from a case study in the neotropics. **Ecological Indicators**, v. 109, n. March 2019, p. 105833, 2020. Disponível em: https://doi.org/10.1016/j.ecolind.2019.105833

LI, J.; PERSSON, K. M.; PEKAR, H.; JANSSON, D. Evaluation of indicators for cyanobacterial risk in 108 temperate lakes using 23 years of environmental monitoring data. **Environmental Sciences Europe**, v. 33, n. 1, p. 1–13, 2021. Disponível em: https://doi.org/10.1186/s12302-021-00483-1

LI, L.; LI, L.; SONG, K. Remote sensing of freshwater cyanobacteria: An extended IOP Inversion Model of Inland Waters (IIMIW) for partitioning absorption coefficient and estimating phycocyanin. **Remote Sensing of Environment**, v. 157, p. 9–23, 2015. Disponível em: https://doi.org/10.1016/j.rse.2014.06.009

LI, L.; LI, L.; SONG, K.; LI, Y.; TEDESCO, L. P.; SHI, K.; LI, Z. An inversion model for deriving inherent optical properties of inland waters: Establishment, validation and application. **Remote Sensing of Environment**, v. 135, p. 150–166, 2013. Disponível em: https://doi.org/10.1016/j.rse.2013.03.031

LI, Y.; SHEN, R.; LIU, X.; SU, H.; WU, F.; CHEN, F. Impacts of nutrient reduction on temporal β-diversity of rotifers: A 19-year limnology case study on Lake Wuli, China. **Water Research**, v. 216, p. 118364, 2022. Disponível em: https://doi.org/10.1016/j.watres.2022.118364

LIN, B.; ULLAH, S. Effectiveness of energy depletion, green growth, and technological cooperation grants on CO2 emissions in Pakistan's perspective. **Science of the Total Environment**, v. 906, n. July 2023, p. 167536, 2024. Disponível em: https://doi.org/10.1016/j.scitotenv.2023.167536

LIN, J. L.; HUA, L. C.; HUNG, S. K.; HUANG, C. Algal removal from cyanobacteria-rich waters by preoxidation-assisted coagulation–flotation: Effect of algogenic organic matter release on algal removal and trihalomethane formation. **Journal of Environmental Sciences (China)**, v. 63, p. 147–155, 2018. Disponível em: https://doi.org/10.1016/j.jes.2017.02.007

LIN, J. L.; KARANGAN, A.; HUANG, Y. M.; KANG, S. F. Eutrophication factor analysis using Carlson trophic state index (CTSI) towards non-algal impact reservoirs in Taiwan. **Sustainable Environment Research**, v. 32, n. 1, 2022. Disponível em: https://doi.org/10.1186/s42834-022-00134-x

LINDHOLM, M.; ALAHUHTA, J.; HEINO, J.; TOIVONEN, H. Temporal beta diversity of lake plants is determined by concomitant changes in environmental factors across decades. **Journal of Ecology**, v. 109, n. 2, p. 819–832, 2021. Disponível em: https://doi.org/10.1111/1365-2745.13508

LINES, R.; JUGGERNAUTH, M.; PEVERLEY, G.; SIMPSON, T.; MOUSAVI-DERAZMAHALLEH, M.; BUNCE, M.; BERRY, T.; KEATING, J.; TAYSOM, A.; WHITTLE, P. A large scale temporal and spatial eDNA biodiversity survey of offshore marine vertebrates in Brazil following the upriver Fundão tailings dam failure. Manuscript in preparation. **Marine Environmental Research**, v. 192, n. July, p. 106239, 2023. Disponível em: https://doi.org/10.1016/j.marenvres.2023.106239

LIU, T.; LIU, H.; YANG, Y. Uncovering the determinants of biodiversity hotspots in China: Evidence from the drivers of multiple diversity metrics on insect assemblages and implications for conservation. **Science of the Total Environment**, v. 880, n. April, p. 163287, 2023. Disponível em: https://doi.org/10.1016/j.scitotenv.2023.163287

LIU, X.; CHEN, L.; ZHANG, G.; ZHANG, J.; WU, Y.; JU, H. Spatiotemporal dynamics of succession and growth limitation of phytoplankton for nutrients and light in a large shallow lake. **Water Research**, v. 194, 2021. Disponível em: https://doi.org/10.1016/j.watres.2021.116910

LOU, I.; XIE, Z.; UNG, W. K.; MOK, K. M. Integrating Support Vector Regression with Particle Swarm Optimization for numerical modeling for algal blooms of freshwater. **Applied Mathematical Modelling**, v. 39, n. 19, p. 5907–5916, 2015. Disponível em: https://doi.org/10.1016/j.apm.2015.04.001

LUDOLF GOMES, L. N.; ALVES CORRÊA OLIVEIRA, S. M.; GIANI, A.; VON SPERLING, E. Association between biotic and abiotic parameters and the occurrence of cyanobacteria in a Brazilian reservoir. **Environmental Monitoring and Assessment**, v. 184, n. 8, p. 4635–4645, 2012. Disponível em: https://doi.org/10.1007/s10661-011-2291-4

LUNETTA, R. S.; SCHAEFFER, B. A.; STUMPF, R. P.; KEITH, D.; JACOBS, S. A.; MURPHY, M. S. Evaluation of cyanobacteria cell count detection derived from MERIS imagery across the eastern USA. **Remote Sensing of Environment**, v. 157, p. 24–34, 2015. Disponível em: https://doi.org/10.1016/j.rse.2014.06.008

LUONG, H. A.; ROHLFS, A. M.; FACEY, J. A.; COLVILLE, A.; MITROVIC, S. M. Long-term study of phytoplankton dynamics in a supply reservoir reveals signs of trophic state shift linked to changes in hydrodynamics associated with flow management and extreme events. **Water Research**, v. 256, n. November 2023, p. 121547, 2024. Disponível em: https://doi.org/10.1016/j.watres.2024.121547

MACARTHUR, R. H. Patterns of species diversity. **Biol. Rev**, n. 40, p. 510–533, 1965.
MAESTRE, F. T. *et al.* Biogeography of global drylands. **New Phytologist**, v. 231, n. 2, p. 540–558, 2021. Disponível em: https://doi.org/10.1111/nph.17395

MAGURRAN, A. E.; DORNELAS, M.; MOYES, F.; HENDERSON, P. A. Temporal β diversity—A macroecological perspective. **Global Ecology and Biogeography**, v. 28, n. 12, p. 1949–1960, 2019. Disponível em: https://doi.org/10.1111/geb.13026

MAITY, S.; GUCHHAIT, R.; CHATTERJEE, A.; PRAMANICK, K. Co-occurrence of co-contaminants: Cyanotoxins and microplastics, in soil system and their health impacts on plant – A comprehensive review. **Science of the Total Environment**, v. 794, p. 148752, 2021. Disponível em: https://doi.org/10.1016/j.scitotenv.2021.148752

MÂNICA, A. N.; DE LIMA ISAAC, R. Seasonal dynamics and diversity of cyanobacteria in a eutrophied Urban River in Brazil. **Water Supply**, v. 23, n. 9, p. 3868–3880, 2023. Disponível em: https://doi.org/10.2166/ws.2023.216

MARENGO, J. A.; CUNHA, A. P.; ALVES, L. M. A seca de 2012-15 no semiárido do Nordeste do Brasil no contexto histórico. **Revista Climanálise**, v. 4, n. 1, p. 49–54, 2016. Disponível em: http://climanalise.cptec.inpe.br/~rclimanl/revista/pdf/30anos/marengoetal.pdf

MARENGO ORSINI, J. A.; ALVES, L. M.; ALVALA, R. C. S.; CUNHA, A. P.; BRITO, S.; MORAES, O. L. L. Climatic characteristics of the 2010-2016 drought in the semiarid northeast Brazil region. **Anais da Academia Brasileira de Ciencias**, v. 90, n. 2, p. 1973–1985, 2018. Disponível em: https://doi.org/10.1590/0001-3765201720170206

MARION, J. W.; ZHANG, F.; CUTTING, D.; LEE, J. Associations between county-level land cover classes and cyanobacteria blooms in the United States. **Ecological Engineering**, v. 108, n. April, p. 556–563, 2017. Disponível em: https://doi.org/10.1016/j.ecoleng.2017.07.032

MASSEY, I. Y.; YANG, F.; DING, Z.; YANG, S.; GUO, J.; TEZI, C.; AL-OSMAN, M.; KAMEGNI, R. B.; ZENG, W. Exposure routes and health effects of microcystins on animals and humans: A mini-review. **Toxicon**, v. 151, n. July, p. 156–162, 2018. Disponível em: https://doi.org/10.1016/j.toxicon.2018.07.010

MAURE, L. A.; DINIZ, M. F.; PACHECO COELHO, M. T.; MOLIN, P. G.; RODRIGUES DA SILVA, F.; HASUI, E. Biodiversity and carbon conservation under the ecosystem stability of tropical forests. **Journal of Environmental Management**, v. 345, n. September, p. 118929, 2023. Disponível em: https://doi.org/10.1016/j.jenvman.2023.118929

MEDINA-COBO, M.; DOMÍNGUEZ, J. A.; QUESADA, A.; DE HOYOS, C. Estimation of cyanobacteria biovolume in water reservoirs by MERIS sensor. **Water Research**, v. 63, p. 10–20, 2014. Disponível em: https://doi.org/10.1016/j.watres.2014.06.001

MENDES, C. F.; DOS SANTOS SEVERIANO, J.; MOURA, G. C. de; DOS SANTOS SILVA, R. D.; MONTEIRO, F. M.; BARBOSA, J. E. de L. The reduction in water volume favors filamentous cyanobacteria and heterocyst production in semiarid tropical reservoirs without the influence of the N:P ratio. **Science of the Total Environment**, v. 816, 2022. Disponível em: https://doi.org/10.1016/j.scitotenv.2021.151584

MEZZANOTTE, V. *et al.* Monitoring, environmental emergencies management and water treatment improvement of freshwater lakes in China: The Chao Lake case study. **Water Science and Technology: Water Supply**, v. 11, n. 4, p. 490–496, 2011. Disponível em: https://doi.org/10.2166/ws.2011.071

MOE, S. J.; HAANDE, S.; COUTURE, R.-M. Climate change, cyanobacteria blooms and ecological status of lakes: A Bayesian network approach. **Ecological Modelling**, v. 337, p. 330–347, 2016. Disponível em: https://doi.org/10.1016/j.ecolmodel.2016.07.004

MONCHAMP, M. E.; SPAAK, P.; DOMAIZON, I.; DUBOIS, N.; BOUFFARD, D.; POMATI, F. Homogenization of lake cyanobacterial communities over a century of climate change and eutrophication. **Nature Ecology and Evolution**, v. 2, n. 2, p. 317–324, 2018. Disponível em: https://doi.org/10.1038/s41559-017-0407-0

MONDEJAR, M. E. *et al.* Digitalization to achieve sustainable development goals: Steps towards a Smart Green Planet. **Science of the Total Environment**, v. 794, n. June, 2021. Disponível em: https://doi.org/10.1016/j.scitotenv.2021.148539

MORI, A. S.; OTA, A. T.; FUJII, S.; SEINO, T.; KABEYA, D.; OKAMOTO, T.; ITO, M. T.; KANEKO, N.; HASEGAWA, M. Biotic homogenization and differentiation of soil faunal communities in the production forest landscape: taxonomic and functional perspectives. **Oecologia**, v. 177, n. 2, p. 533–544, 2015. Disponível em: https://doi.org/10.1007/s00442-014-3111-7

MUNOZ, M. *et al.* Overview of toxic cyanobacteria and cyanotoxins in Ibero-American freshwaters: Challenges for risk management and opportunities for removal by advanced technologies. **Science of the Total Environment**, v. 761, 2021. Disponível em: https://doi.org/10.1016/j.scitotenv.2020.143197

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized Linear Models. **Journal of the Royal Statistical Society A**, v. 135, n. 3, p. 1–4, 1972. Disponível em: https://doi.org/https://doi.org/10.2307/2344614

NIEKE, J.; MAVROCORDATOS, C.; DONLON, C.; BERRUTI, B.; GARNIER, T.; RITI, J.; DELCLAUD, Y. **Optical Payloads for Space Missions**. *[S. l.]*: Wiley, 2015. *E-book*. Disponível em: https://doi.org/10.1002/9781118945179

NIETCH, C. T.; GAINS-GERMAIN, L.; LAZORCHAK, J.; KEELY, S. P.; YOUNGSTROM, G.; URICHICH, E. M.; ASTIFAN, B.; DASILVA, A.; MAYFIELD, H. Development of a Risk Characterization Tool for Harmful Cyanobacteria Blooms on the Ohio River. **Water (Switzerland)**, v. 14, n. 4, 2022. Disponível em: https://doi.org/10.3390/w14040644

NIU, C.; TAN, K.; JIA, X.; WANG, X. Deep learning based regression for optically inactive inland water quality parameter estimation using airborne hyperspectral imagery. **Environmental Pollution**, v. 286, n. May, p. 117534, 2021. Disponível em: https://doi.org/10.1016/j.envpol.2021.117534

NOBRE, R. L. G. *et al.* Precipitation, landscape properties and land use interactively affect water quality of tropical freshwaters. **Science of the Total Environment**, v. 716, p. 137044, 2020. Disponível em: https://doi.org/10.1016/j.scitotenv.2020.137044

NOGUEIRA VILANOVA, M. R.; FILHO, P. M.; PERRELLA BALESTIERI, J. A. Performance measurement and indicators for water supply management: Review and international cases. **Renewable and Sustainable Energy Reviews**, v. 43, p. 1–12, 2014. Disponível em: https://doi.org/10.1016/j.rser.2014.11.043

O’GRADY, J.; ZHANG, D.; O’CONNOR, N.; REGAN, F. A comprehensive review of catchment water quality monitoring using a tiered framework of integrated sensing technologies. **Science of the Total Environment**, v. 765, p. 142766, 2021. Disponível em: https://doi.org/10.1016/j.scitotenv.2020.142766

O'HARA-WILD, M.; HYNDMAN, R.; WANG, E. **feasts: Feature Extraction and Statistics for Time Series**. *[S. l.: s. n.]* Disponível em: https://cran.r-project.org/package=feasts

OKSANEN, j *et al.* **vegan: Community Ecology Package [R package version 2.6-4]**. *[s. l.]*, 2022. Disponível em: https://cran.r-project.org/package=vegan. Acesso em: 26 dez. 2023.

OLIVEIRA, S.; CLEMENTE, A.; MENEZES, I.; GOIS, A.; CARLOTO, I.; LAWTON, L.; CAPELO-NETO, J. Hazardous cyanobacteria integrity response to velocity gradient and powdered activated carbon in water treatment plants. **Science of the Total Environment**, v. 773, p. 145110, 2021. Disponível em: https://doi.org/10.1016/j.scitotenv.2021.145110

OLSSON, G. ICA and me - A subjective review. **Water Research**, v. 46, n. 6, p. 1585–1624, 2012. Disponível em: https://doi.org/10.1016/j.watres.2011.12.054

OMONDI, A. N.; OUMA, Y.; KOSGEI, J. R.; KONGO, V.; KEMBOI, E. J.; NJOROGE, S. M.; MECHA, A. C.; KIPKORIR, E. C. Estimation and mapping of water quality parameters using satellite images: a case study of Two Rivers Dam, Kenya. **Water Practice and Technology**, v. 18, n. 2, p. 428–443, 2023. Disponível em: https://doi.org/10.2166/wpt.2023.010

OUMA, Y. O.; WAGA, J.; OKECH, M.; LAVISA, O.; MBUTHIA, D. Estimation of Reservoir Bio-Optical Water Quality Parameters Using Smartphone Sensor Apps and Landsat ETM+: Review and Comparative Experimental Results. **Journal of Sensors**, v. 2018, p. 1–32, 2018. Disponível em: https://doi.org/10.1155/2018/3490757

OYAMA, Y.; MATSUSHITA, B.; FUKUSHIMA, T. Distinguishing surface cyanobacterial blooms and aquatic macrophytes using Landsat/TM and ETM + shortwave infrared bands. **Remote Sensing of Environment**, v. 157, p. 35–47, 2015. Disponível em: https://doi.org/10.1016/j.rse.2014.04.031

PAMULA, A. S. P.; GHOLIZADEH, H.; KRZMARZICK, M. J.; MAUSBACH, W. E.; LAMPERT, D. J. A remote sensing tool for near real-time monitoring of harmful algal blooms and turbidity in reservoirs. **JAWRA Journal of the American Water Resources Association**, v. 59, n. 5, p. 929–949, 2023. Disponível em: https://doi.org/10.1111/1752-1688.13121

PANTELIĆ, D.; SVIRČEV, Z.; SIMEUNOVIĆ, J.; VIDOVIĆ, M.; TRAJKOVIĆ, I. Cyanotoxins: Characteristics, production and degradation routes in drinking water treatment with reference to the situation in Serbia. **Chemosphere**, v. 91, n. 4, p. 421–441, 2013. Disponível em: https://doi.org/10.1016/j.chemosphere.2013.01.003

PARK, J.; PATEL, K.; LEE, W. H. Recent advances in algal bloom detection and prediction technology using machine learning. **Science of The Total Environment**, v. 938, n. May, p. 173546, 2024. Disponível em: https://doi.org/10.1016/j.scitotenv.2024.173546

PARK, Y.; CHO, K. H.; PARK, J.; CHA, S. M.; KIM, J. H. Development of early-warning protocol for predicting chlorophyll-a concentration using machine learning models in freshwater and estuarine reservoirs, Korea. **Science of the Total Environment**, v. 502, p. 31–41, 2015. Disponível em: https://doi.org/10.1016/j.scitotenv.2014.09.005

PAZO, M.; GERASSIS, S.; ARAÚJO, M.; MARGARIDA ANTUNES, I.; RIGUEIRA, X. Enhancing water quality prediction for fluctuating missing data scenarios: A dynamic Bayesian network-based processing system to monitor cyanobacteria proliferation. **Science of the Total Environment**, v. 927, n. January, p. 172340, 2024. Disponível em: https://doi.org/10.1016/j.scitotenv.2024.172340

PEEL, M. C.; FINLAYSON, B. L.; MCMAHON, T. A. Updated world map of the Köppen-Geiger climate classification. **Hydrology and Earth System Sciences**, v. 11, n. 5, p. 1633–1644, 2007. Disponível em: https://doi.org/10.5194/hess-11-1633-2007. Acesso em: 12 mar. 2024.

PEREIRA, A. L. A.; DE CARVALHO, P.; GRANZOTTI, R. V.; VIEIRA, L. C. G.; BINI, L. M. Temporal beta diversity increases with environmental variability in zooplankton floodplain communities. **Limnology**, v. 25, n. 1, p. 1–10, 2023. Disponível em: https://doi.org/10.1007/s10201-023-00724-7

PÉREZ-GONZÁLEZ, R.; SÒRIA-PERPINYÀ, X.; SORIA, J. M.; DELEGIDO, J.; URREGO, P.; SENDRA, M. D.; RUÍZ-VERDÚ, A.; VICENTE, E.; MORENO, J. Phycocyanin monitoring in some spanish water bodies with sentinel-2 imagery. **Water (Switzerland)**, v. 13, n. 20, p. 1–22, 2021. Disponível em: https://doi.org/10.3390/w13202866

PESTANA, C. J.; CAPELO-NETO, J.; LAWTON, L.; OLIVEIRA, S.; CARLOTO, I.; LINHARES, H. P. The effect of water treatment unit processes on cyanobacterial trichome integrity. **Science of the Total Environment**, v. 659, p. 1403–1414, 2019. Disponível em: https://doi.org/10.1016/j.scitotenv.2018.12.337

PESTANA, C. J.; NETO, J. C.; BARROS, M. U.; MENEZES, I.; GÓIS, A.; SANTOS, G. Consumer perception of water quality during an off-flavor event in Fortaleza-Brazil. **Journal of Water Supply: Research and Technology-Aqua**, p. 1–11, 2018. Disponível em: https://doi.org/10.2166/aqua.2018.077

PESTANA, C. J.; REEVE, P. J.; SAWADE, E.; VOLDOIRE, C. F.; NEWTON, K.; PRAPTIWI, R.; COLLINGNON, L.; DREYFUS, J.; HOBSON, P.; GAGET, V.; NEWCOMBE, G. Fate of cyanobacteria in drinking water treatment plant lagoon supernatant and sludge. **Science of the Total Environment**, v. 565, p. 1192–1200, 2016. Disponível em: https://doi.org/10.1016/j.scitotenv.2016.05.173

PFAFF, B. **Analysis of Integrated and Cointegrated Time Series with R**. Second ed. *[S. l.]*: Springer, 2008. *E-book*. Disponível em: https://www.pfaffikus.de

PICKERING, C.; BYRNE, J. The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. **Higher Education Research and Development**, v. 33, n. 3, p. 534–548, 2014. Disponível em: https://doi.org/10.1080/07294360.2013.841651

PRESTIGIACOMO, A. R.; GORNEY, R. M.; HYDE, J. B.; DAVIS, C.; CLINKHAMMER, A. Patterns and impacts of cyanobacteria in a deep, thermally stratified, oligotrophic lake. **AWWA Water Science**, v. 5, n. 2, p. 1–18, 2023. Disponível em: https://doi.org/10.1002/aws2.1326

PYO, J. C.; CHO, K. H.; KIM, K.; BAEK, S. S.; NAM, G.; PARK, S. Cyanobacteria cell prediction using interpretable deep learning model with observed, numerical, and sensing data assemblage. **Water Research**, v. 203, n. May, p. 117483, 2021. Disponível em: https://doi.org/10.1016/j.watres.2021.117483

QUAN, Z.; XU, X.; WANG, W.; JIANG, J.; GAO, S. Do industrial solid waste recycling and technological innovation promote low-carbon development in China? New insights from NARDL approach. **Science of the Total Environment**, v. 916, n. November 2023, 2024. Disponível em: https://doi.org/10.1016/j.scitotenv.2024.170446

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. *[s. l.]*, 2023. Disponível em: https://www.r-project.org/. Acesso em: 14 nov. 2023.

RAMETTE, A. **Multivariate analyses in microbial ecology**. *[S. l.]*: Wiley-Blackwell, 2007. Disponível em: https://doi.org/10.1111/j.1574-6941.2007.00375.x. Acesso em: 30 maio. 2021.

RECKNAGEL, F.; OSTROVSKY, I.; CAO, H.; ZOHARY, T.; ZHANG, X. Ecological relationships, thresholds and time-lags determining phytoplankton community dynamics of Lake Kinneret, Israel elucidated by evolutionary computation and wavelets. **Ecological Modelling**, v. 255, p. 70–86, 2013. Disponível em: https://doi.org/10.1016/j.ecolmodel.2013.02.006

REYNOLDS, N.; SCHAEFFER, B. A.; GUERTAULT, L.; NELSON, N. G. Satellite and in situ cyanobacteria monitoring: Understanding the impact of monitoring frequency on management decisions. **Journal of Hydrology**, v. 619, n. February, p. 129278, 2023. Disponível em: https://doi.org/10.1016/j.jhydrol.2023.129278

RICOTTA, C.; FEOLI, E. Hill numbers everywhere. Does it make ecological sense? **Ecological Indicators**, v. 161, n. March, p. 111971, 2024. Disponível em: https://doi.org/10.1016/j.ecolind.2024.111971

RIZZO, L. The Contribution of the Coagulation Process in Controlling Microbial Risk and Disinfection By-products Formation in Drinking Water. *In*: **The Role of Colloidal Systems in Environmental Protection**. Salerno: Elsevier B.V., 2014. p. 219–238. Disponível em: https://doi.org/10.1016/B978-0-444-63283-8.00009-0

ROCHA JUNIOR, C.; ARAÚJO, F.; BECKER, V. Influence of land use on spatial distribution of mobile phosphorus forms in the sediment of a tropical semi-arid reservoir. **Science of the Total Environment**, v. 914, n. August 2023, p. 169836, 2024. Disponível em: https://doi.org/10.1016/j.scitotenv.2023.169836

ROCHA, M. A. M.; BARROS, M. U. G.; COSTA, A. C.; DE ASSIS DE SOUZA FILHO, F.; LIMA NETO, I. E. Understanding the Water Quality Dynamics in a Large Tropical Reservoir Under Hydrological Drought Conditions. **Water, Air, and Soil Pollution**, v. 235, n. 1, p. 1–25, 2024. Disponível em: https://doi.org/10.1007/s11270-024-06890-3

ROUSSO, B. Z.; BERTONE, E.; STEWART, R.; HAMILTON, D. P. A systematic literature review of forecasting and predictive models for cyanobacteria blooms in freshwater lakes. **Water Research**, v. 182, p. 115959, 2020. Disponível em: https://doi.org/10.1016/j.watres.2020.115959

SCHAEFFER, B. A. *et al.* Mobile device application for monitoring cyanobacteria harmful algal blooms using Sentinel-3 satellite Ocean and Land Colour Instruments. **Environmental Modelling and Software**, v. 109, n. August, p. 93–103, 2018 a. Disponível em: https://doi.org/10.1016/j.envsoft.2018.08.015

SCHAEFFER, B. A.; IIAMES, J.; DWYER, J.; URQUHART, E.; SALLS, W.; ROVER, J.; SEEGERS, B. An initial validation of Landsat 5 and 7 derived surface water temperature for U.S. lakes, reservoirs, and estuaries. **International Journal of Remote**

Sensing, v. 39, n. 22, p. 7789–7805, 2018 b. Disponível em: https://doi.org/10.1080/01431161.2018.1471545

SERRÀ, A.; PHILIPPE, L.; PERREAULT, F.; GARCIA-SEGURA, S. Photocatalytic treatment of natural waters. Reality or hype? The case of cyanotoxins remediation. **Water Research**, v. 188, 2021. Disponível em: https://doi.org/10.1016/j.watres.2020.116543

SHANNON, C. E. A Mathematical Theory of Communication. **Bell System Technical Journal**, v. 27, n. 4, p. 623–656, 1948. Disponível em: https://doi.org/10.1002/j.1538-7305.1948.tb00917.x

SHEATHER, S. J. **A modern approach to regression with R**. Texas: Springer, 2009. v. 102 Disponível em: https://doi.org/10.1007/978-0-387-09608-7

SHIZUKA, K.; MAIE, N.; KAKINO, W.; TARUYA, H.; TANJI, H. Forecasting a 2-methylisoborneol outbreak in a brackish lake. **Environmental Monitoring and Assessment**, v. 193, n. 7, 2021. Disponível em: https://doi.org/10.1007/s10661-021-09140-9

SIDDIG, A. A. H.; ELLISON, A. M.; OCHS, A.; VILLAR-LEEMAN, C.; LAU, M. K. How do ecologists select and use indicator species to monitor ecological change? Insights from 14 years of publication in Ecological Indicators. **Ecological Indicators**, v. 60, p. 223–230, 2016. Disponível em: https://doi.org/10.1016/j.ecolind.2015.06.036

SIMIS, S. G. H.; PETERS, S. W. M.; GONS, H. J. Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. **Limnology and Oceanography**, v. 50, n. 1, p. 237–245, 2005. Disponível em: https://doi.org/10.4319/lo.2005.50.1.0237

SIMPSON, E. H. Measurment of Diversity. **Nature**, v. 688, n. 1943, p. 688, 1949. Disponível em: https://doi.org/10.1038/163688a0

SINGH, U. B.; AHLUWALIA, A. S.; SHARMA, C.; JINDAL, R.; THAKUR, R. K. Planktonic indicators: A promising tool for monitoring water quality (early-warning signals). **Ecology, Environment and Conservation**, v. 19, n. 3, p. 793–800, 2013.

SMUCKER, N. J.; BEAULIEU, J. J.; NIETCH, C. T.; YOUNG, J. L. Increasingly severe cyanobacterial blooms and deep water hypoxia coincide with warming water temperatures in reservoirs. **Global Change Biology**, v. 27, n. 11, p. 2507–2519, 2021. Disponível em: https://doi.org/10.1111/gcb.15618

SNISB. **National Information System on Dam Safety**. *[s. l.]*, 2023. Disponível em: https://www.gov.br/cidades/pt-br/acesso-a-informacao/acoes-e-programas/saneamento/snis/painel. Acesso em: 18 nov. 2023.

SU, M.; ZHU, Y.; JIA, Z.; LIU, T.; YU, J.; BURCH, M.; YANG, M. Identification of MIB producers and odor risk assessment using routine data: A case study of an estuary drinking water reservoir. **Water Research**, v. 192, p. 116848, 2021. Disponível em: https://doi.org/10.1016/j.watres.2021.116848

SUAREZ, E. L.; DE VENTURA, L.; STÖCKLI, A.; ORDÓÑEZ, C.; THOMAS, M. K.; IBELINGS, B. W.; MCGINNIS, D. F. The emergence and dominance of Planktothrix rubescens as an hypolimnetic cyanobacterium in response to re-oligotrophication of a deep peri-alpine lake. **Limnology and Oceanography**, v. 68, n. 6, p. 1346–1359, 2023. Disponível em: https://doi.org/10.1002/lno.12351

SVIRČEV, Z. B.; TOKODI, N.; DROBAC, D.; CODD, G. A. Cyanobacteria in aquatic ecosystems in Serbia: Effects on water quality, human health and biodiversity. **Systematics and Biodiversity**, v. 12, n. 3, p. 261–270, 2014. Disponível em: https://doi.org/10.1080/14772000.2014.921254

TONKIN, J. D.; BOGAN, M. T.; BONADA, N.; RIOS-TOUMA, B.; LYTLE, D. A. Seasonality and predictability shape temporal species diversity. **Ecology**, v. 98, n. 5, p. 1201–1216, 2017. Disponível em: https://doi.org/10.1002/ecy.1761

URRUTIAGUER, M. Management of the Ecological Impacts of Urban Land and Activities on Waterways. Issues Paper: Understanding the Science. **Melbourne Water Price Review 2016**, 2016.

VERNIEST, F.; GREULICH, S. Methods for assessing the effects of environmental parameters on biological communities in long-term ecological studies - A literature review. **Ecological Modelling**, v. 414, n. February, p. 108732, 2019. Disponível em: https://doi.org/10.1016/j.ecolmodel.2019.108732

WARDROPPER, C.; BROOKFIELD, A. Decision-support systems for water management. **Journal of Hydrology**, v. 610, 2022. Disponível em: https://doi.org/10.1016/j.jhydrol.2022.127928

WEI, J.; LI, Q.; LIU, W.; ZHANG, S.; XU, H.; PEI, H. Changes of phytoplankton and water environment in a highly urbanized subtropical lake during the past ten years. **Science of the Total Environment**, v. 879, n. March, p. 162985, 2023. Disponível em: https://doi.org/10.1016/j.scitotenv.2023.162985

WHEELER, S. M.; MORRISSEY, L. A.; LEVINE, S. N.; LIVINGSTON, G. P.; VINCENT, W. F. Mapping cyanobacterial blooms in Lake Champlain's Missisquoi Bay using QuickBird and MERIS satellite data. **Journal of Great Lakes Research**, v. 38, n. SUPPL. 1, p. 68–75, 2012. Disponível em: https://doi.org/10.1016/j.jglr.2011.06.009

WHITTAKER, R. H. EVOLUTION AND MEASUREMENT OF SPECIES DIVERSITY. **TAXON**, v. 21, n. 2–3, p. 213–251, 1972. Disponível em: https://doi.org/10.2307/1218190

WHITTON, B. A.; POTTS, M. Introduction to the Cyanobacteria. *In*: **Ecology of Cyanobacteria II**. Dordrecht: Springer Netherlands, 2012. p. 1–13. *E-book*. Disponível em: https://doi.org/10.1007/978-94-007-3855-3_1

WHO. **Guidelines for drinking-water quality**. fourth edi ed. *[S. l.]*: WHO TEAM, 2022. *E-book*. Disponível em: https://www.who.int/publications/i/item/9789240045064

WILLIAMS, B. J.; COLE, B. Mining monitored data for decision-making with a Bayesian network model. **Ecological Modelling**, v. 249, p. 26–36, 2013. Disponível em: https://doi.org/10.1016/j.ecolmodel.2012.07.008

WU, T.; ZHU, G.; WANG, Z.; ZHU, M.; XU, H. Seasonal dynamics of odor compounds concentration driven by phytoplankton succession in a subtropical drinking water reservoir, southeast China. **Journal of Hazardous Materials**, v. 425, n. August 2021, p. 128056, 2022. Disponível em: https://doi.org/10.1016/j.jhazmat.2021.128056

WYNNE, T. T.; STUMPF, R. P.; TOMLINSON, M. C.; WARNER, R. A.; TESTER, P. A.; DYBLE, J.; FAHNENSTIEL, G. L. Relating spectral shape to cyanobacterial blooms in the Laurentian Great Lakes. **International Journal of Remote Sensing**, v. 29, n. 12, p. 3665–3672, 2008. Disponível em: https://doi.org/10.1080/01431160802007640

XIA, X. H.; WU, Q.; MOU, X. L.; LAI, Y. J. Potential impacts of climate change on the water quality of different water bodies. **Journal of Environmental Informatics**, v. 25, n. 2, p. 85–98, 2015. Disponível em: https://doi.org/10.3808/jei.201400263

XIAO, X.; HE, J.; HUANG, H.; MILLER, T. R.; CHRISTAKOS, G.; REICHWALDT, E. S.; GHADOUANI, A.; LIN, S.; XU, X.; SHI, J. A novel single-parameter approach for forecasting algal blooms. **Water Research**, v. 108, p. 222–231, 2017. Disponível em: https://doi.org/10.1016/j.watres.2016.10.076

XIE, P.; CHEN, Y.; MA, J.; ZHANG, X.; ZOU, J.; WANG, Z. A mini review of preoxidation to improve coagulation. **Chemosphere**, v. 155, p. 550–563, 2016. Disponível em: https://doi.org/10.1016/j.chemosphere.2016.04.003

XIE, Z.; LOU, I.; UNG, W. K.; MOK, K. M. Freshwater algal bloom prediction by support vector machine in Macau storage reservoirs. **Mathematical Problems in Engineering**, v. 2012, 2012. Disponível em: https://doi.org/10.1155/2012/397473

XU, H.; PAERL, H. W.; QIN, B.; ZHU, G.; HALL, N. S.; WU, Y. Determining critical nutrient thresholds needed to control harmful cyanobacterial blooms in eutrophic Lake

Taihu, China. **Environmental Science and Technology**, v. 49, n. 2, p. 1051–1059, 2015. Disponível em: https://doi.org/10.1021/es503744q

YANG, B.; QU, X.; LIU, H.; YANG, M.; XIN, W.; WANG, W.; CHEN, Y. Urbanization reduces fish taxonomic and functional diversity while increases phylogenetic diversity in subtropical rivers. **Science of the Total Environment**, v. 908, n. July 2023, p. 168178, 2024. Disponível em: https://doi.org/10.1016/j.scitotenv.2023.168178

YANG, R.; YANG, K.; HONG, L.; YAN YANG. A study of inversion modeling of water quality parameters in the dianchi lake using CCD1 data of HJ-1A satellite. *In*: 2014, **2014 22nd International Conference on Geoinformatics**. : IEEE, 2014. p. 1–6. Disponível em: https://doi.org/10.1109/GEOINFORMATICS.2014.6950798

YANG, Y.; YU, G.; CHEN, Y.; JIA, N.; LI, R. Four decades of progress in cylindrospermopsin research: The ins and outs of a potent cyanotoxin. **Journal of Hazardous Materials**, v. 406, n. August 2020, p. 124653, 2021. Disponível em: https://doi.org/10.1016/j.jhazmat.2020.124653

ZAMYADI, A.; CHOO, F.; NEWCOMBE, G.; STUETZ, R.; HENDERSON, R. K. A review of monitoring technologies for real-time management of cyanobacteria: Recent advances and future direction. **TrAC - Trends in Analytical Chemistry**, v. 85, p. 83–96, 2016. Disponível em: https://doi.org/10.1016/j.trac.2016.06.023

ZAMYADI, A.; HENDERSON, R.; STUETZ, R.; HOFMANN, R.; HO, L.; NEWCOMBE, G. Fate of geosmin and 2-methylisoborneol in full-scale water treatment plants. **Water Research**, v. 83, p. 171–183, 2015. Disponível em: https://doi.org/10.1016/j.watres.2015.06.038

ZANCHETT, G.; OLIVEIRA-FILHO, E. C. **Cyanobacteria and cyanotoxins: From impacts on aquatic ecosystems and human health to anticarcinogenic effects**. *[S. l.: s. n.]* Disponível em: https://doi.org/10.3390/toxins5101896

ZHANG, J. *et al.* Ancient DNA reveals potentially toxic cyanobacteria increasing with climate change. **Water Research**, v. 229, n. September 2022, p. 119435, 2023. Disponível em: https://doi.org/10.1016/j.watres.2022.119435

ZHANG, S.; WANG, W.; CHANG, J. The contribution of cyanobacteria bloom decline to phosphorus in water column of Dianchi Lake, China. **Polish Journal of Environmental Studies**, v. 28, n. 5, p. 3513–3520, 2019. Disponível em: https://doi.org/10.15244/pjoes/94217

ZHANG, W.; LIU, Z. Daily water quality evaluation of reservoir and cyanobacteria pollution index calculation. **Water Science and Technology: Water Supply**, v. 21, n. 2, p. 836–847, 2021. Disponível em: https://doi.org/10.2166/ws.2020.350

ZHANG, X.; RECKNAGEL, F.; CHEN, Q.; CAO, H.; LI, R. Spatially-explicit modelling and forecasting of cyanobacteria growth in Lake Taihu by evolutionary computation. **Ecological Modelling**, v. 306, p. 216–225, 2015. Disponível em: https://doi.org/10.1016/j.ecolmodel.2014.05.013

ZHANG, Y.; CHENG, L.; LI, K.; ZHANG, L.; CAI, Y.; WANG, X.; HEINO, J. Nutrient enrichment homogenizes taxonomic and functional diversity of benthic macroinvertebrate assemblages in shallow lakes. **Limnology and Oceanography**, v. 64, n. 3, p. 1047–1058, 2019. Disponível em: https://doi.org/10.1002/lno.11096

ZIVOT, E.; ANDREWS, D. W. K. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. **Journal of Business and Economic Statistics**, v. 10, n. 3, p. 251–270, 1992. Disponível em: https://doi.org/10.1080/07350015.1992.10509904

ZONG, J. M.; WANG, X. X.; ZHONG, Q. Y.; XIAO, X. M.; MA, J.; ZHAO, B. Increasing outbreak of cyanobacterial blooms in large lakes and reservoirs under pressures from climate change and anthropogenic interferences in the middle-lower Yangtze River basin. **Remote Sensing**, v. 11, n. 15, p. 1–21, 2019. Disponível em: https://doi.org/10.3390/rs11151754

ZUO, J.; XIAO, P.; HEINO, J.; TAN, F.; SOININEN, J.; CHEN, H.; YANG, J. Eutrophication increases the similarity of cyanobacterial community features in lakes and reservoirs. **Water Research**, v. 250, n. November 2023, p. 120977, 2024. Disponível em: https://doi.org/10.1016/j.watres.2023.120977

APPENDIX A – SUPPLEMENTARY MATERIALS TO CHAPTER 2

Table S1 - Main steps in the elaboration of the research protocol

Item	Description	What was adopted
Objective	Specifies the main objective of the review	Identify alternative methods for collecting, analyzing, and interpreting CyanoHabs data that enhance decision-making based on regular monitoring and are feasible for sanitation companies
PICOC criteria	They define 5 key points that delimit the research universe: • Population – P: The group of articles obtained from the proposed systematization; • Intervention – I: What will be observed in the population; • Control – C: background used for comparison with the articles obtained from the systematization; • Output – O: a deep and comprehensive view of the purpose of the review; • Context – C: who will benefit;	• P: all articles obtained from the approved protocol; • I: alternative methods for collecting, analyzing, and interpreting data related to CyanoHabs; • C: review articles, books, and articles suggested by the review team; • O: a deep and comprehensive overview of alternative methods that can complement or enhance decision-making based on regular monitoring of CyanoHabs and be suitable for sanitation companies. • C: researchers, sanitation companies, etc.
Keywords	Definition of the search string	Terms and keywords obtained in Control. From this *string*, the general selection criteria were added. The search was performed identically in the databases, combining multiple keywords with Boolean operators AND, OR and NOT
Sources	Definition of the database	A comprehensive database, but focused on engineering (COMPEDEX) and a multidisciplinary platform that covers other areas of knowledge (Web of Science) were used in April 2024
Language	Setting the language of the found articles	English
Exclusion Criteria	Definition of how articles are excluded from review. Two criteria were used: • **General criteria**: those that were applied in the databases when available; • **Specific criteria**: applied after reading the title, abstract and full text in this order. The article will be deleted if at least one of the criteria is met. Repeated articles will be counted only once. The general criteria that could not be applied directly at the bases have been considered here.	• **General criteria:** 1) Articles outside the 2010-2024 range; 2) Non-peer-reviewed articles (when available); 3) Reviews, books, or secondary studies (where available); 4) Engineering area correlated to sanitation (when available); • **Specific criteria:** 1) Article that does not relate to water quality and CyanoHabs events; 2) Articles that do not use monitoring data from official agencies, that have not performed regular monitoring for at least 1 year, or that are mostly experimental; 5) Articles that did not use appropriate technology/methodology for sanitation companies, especially in the context of developing countries; 6) Articles that preferably used process-based models. 7) Articles with at least one source used for human supply;
Inclusion Criteria	Definition of how articles are included in the review. Where available, this criterion can be applied directly on the platform. The absence of these items does not necessarily exclude publication	• Articles that address technologies/methodologies relevant to sanitation companies; • Articles related to water quality and CyanoHabs events in water sources intended for human supply. • Studies that use modeling based on data and information from regular monitoring.

Source: prepared by the author

Table S2 – Strings and their respective results. The Symbols * where general exclusion criteria have been applied. (WOS: Web of Science and Comp: Compendex)

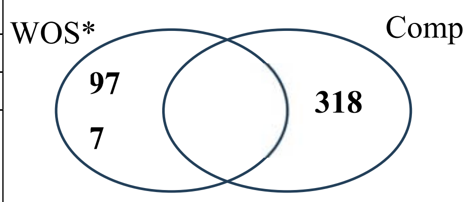
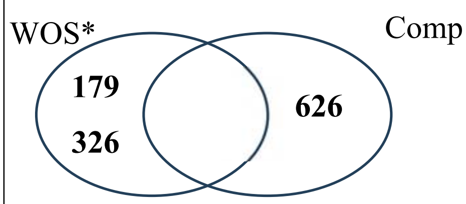
<table>
<tr><td colspan="2">String 1</td></tr>
<tr><td>“Monitoring data”</td><td rowspan="5">
</td></tr>
<tr><td>And</td></tr>
<tr><td>Cyanobacteria* or cyanotoxin* or toxin*</td></tr>
<tr><td>NOT</td></tr>
<tr><td>oceanography OR plants OR fish OR proteins OR shellfish OR mollusks OR genes OR forestry OR radiometers OR effluents OR “air quality” OR “vegetation”</td></tr>
<tr><td colspan="2">String 2</td></tr>
<tr><td>"Monitoring data" OR "water quality monitoring" OR "Mining monitored data" OR "environmental monitoring data" OR "regular monitoring"</td><td rowspan="7">
</td></tr>
<tr><td>AND</td></tr>
<tr><td>"Water suppl*" OR "human supply" OR "Public Water Supply for Human Consumption" OR "Environmental pollution" OR "drinking water reservoir*"</td></tr>
<tr><td>AND</td></tr>
<tr><td>"blue-green algae" OR "Anatoxin-a" OR "Harmful cyanobacterial bloom*" OR "Harmful algal bloom*" OR "eutrophication" OR "Cyanobacteria Control" OR "hab*" OR "cyanohab*" or "benthic cyanobacteria" OR "Harmful algal bloom" OR "Turbidity" OR "Chlorophyll-a" OR "algal pigment" OR "algal pigments" OR "MIB" OR "2-methylisoborneol" OR "Taste and odor" OR "Microcystis" OR "phytoplankton community" OR "total dissolved phosphorus" OR "phycocyanin" or "Cyanobacteria harmful bloom" OR "Cyanobacteria harmful blooms" OR "Chlorophyll a" OR "mesotrophi*" OR "oligotrophi*" OR "cyanobacteri*" OR "cyanotoxi*" OR "Microcysti*" OR "Cylindrospermopsi*" OR "Nodulari*" OR "bloo*" OR "phytoplankto*" or Geosmin</td></tr>
<tr><td>NOT</td></tr>
<tr><td>"Life Cycle" or "Blood" or "Irrigation" or "Wastewater Treatment" OR "Biogeochemistry" OR "Wetland*" OR "Effluent*” OR "Sediment*" OR "Agricultural Pollution" OR "Land Use" or Plant* or Forestry</td></tr>
<tr><td>String 1 ∩ String 2</td><td>24</td></tr>
<tr><td>Total</td><td>1549</td></tr>
</table>

Table S3 - List of articles in the systematic review of literature (Eut: Trophic state. Clim: Predominant climate; DS: data source; Tech: Monitoring technique predominantly used in the article; Meth: methodology predominantly used in modeling; Eva: Metric for evaluating the quality of the adjustments; Ind: independent variables; Dep: dependent variables when methodology was used or list of dependent variables when supervised methodology was used) The meanings of the acronyms are in the tables of supplementary materials.

Title	**country**	**Type**	**Had**	**Air conditioning .**	**Eco.**	**DS**	**Techs**	**Meth.**	**Models**	**Eva**	**Indep**	**Dep**
(XIAO *et al.*, 2017)	China, USA	Various	Yes	Temperate	Developed	OM + Programs	HRM	ML	WNN ANN ARIMA	r MARE	Chl-a	Chl-a
(SU *et al.*, 2015)	**China**	Reservoirs	Others	Temperate	Developeing	OM + Programs	HRM + TM	ML	GA SVM	AE RE ,RMSE R^2	TP TN COD VOL	Chl-a
(NI *et al.*, 2022)	**China**	Lakes	Yes	Temperate	Developeing	OM + Programs	HRM + TM	ML	CNN BiLSTM	MAE RMSE	WT pH EC TURB DO CD	Chl-a
(SCHAEFFER *et al.*, 2018)	**USA**	Lakes	Unspecified	Temperate	Developed	OM + Programs	RS+TM	STAT	DS+RL	MAP R^2	HI	WT
(CHEN *et al.*, 2014)	**China**	Lakes	Unspecified	Temperate	Developeing	OM + Programs	HRM + TM	ML	LIKE FUZZY IOSM	N/A	Chl-a WT COD TN TP	ULM/FDFA

(ZHANG *et al.*, 2023)	**China, USA, Canada**	Lakes	Unspecified	Temperate	Developeing	OM + Programs	HRM + TM	STAT	NMDS	N/A	Phytoplankton	ULM/FDFA
(KUDELA *et al.*, 2015)	**USA**	Lakes	Yes	Temperate	Developed	OM + Programs	RS+TM	STAT	CI AMI	R^2	HI PC	PC CD
(BRINOVCAR *et al.*, 2022)	**Canada**	Reservoirs	Yes	Continental	Developed	TO	HRM + TM	STAT	ANOVA Pearson	N/A	CD TP DIN	ULM/FDFA
(LUDOLF GOMES *et al.*, 2012)	**BRAZIL**	Reservoirs	Yes	Arid/Semiars	Developeing	OM + Research	TM	STAT	PCA	N/A	TEMP DO NH4 NO3 PO4 TP TN Cyanobacteria	ULM/FDFA
(FERNANDEZ-FIGUEROA *et al.*, 2021)	**USA**	Various	Unspecified	Tropical/Subtropical	Developed	Research	TM	STAT	Pearson anova	N/A	Chl-a PC MC STX CYN TN TP	ULM/FDFA
(WEI *et al.*, 2023)	**China**	Lakes	Yes	Tropical/Subtropical	Developeing	OM + Programs	HRM + TM	STAT	Mann-Kendall ANOVA Spearman	N/A	WT PREC IN OUT TP TN NO3 TN/TP pH WS SECCHI DO TURB COD	ULM/FDFA

(PATERSON *et al.*, 2017)	**USA, Canada**	Lakes	Unspecified	Temperate	Developed	OM + Programs	TM	STAT	WA-PLS	R^2 RMSE AS	Chl-a	HCMC
(MOE; HAANDE; COUTURE, 2016)	**Canada**	Lakes	Yes	Temperate	Developed	OM + Programs	HRM + TM	ML	BM	N/A	WT SECCHI TP Chl-a	BIO-CYANO
(SCHEIFHACKEN; HORN; PAUL, 2010)	**Germany**	Reservoirs	Unspecified	Temperate	Developed	OM + Programs	HRM + TM	STAT	PCA de Pearson RL	R^2	Phytoplankton TURB Chl-a MorpHological-predictors	ULM/FDFA
(COURTOIS *et al.*, 2018)	**France**	Rivers	Yes	Temperate	Developed	Research	HRM + TM	STAT	RL	R^2	CD	TURB PC Chl-a DOM
(LORENZI *et al.*, 2018)	**BRAZIL**	Various	Yes	Arid/Semiars	Developeing	TO	HRM + TM	STAT	DS Pearson	R^2	MC CYN STX neo-STX ATX-a	ULM/FDFA
(LOAIZA-GONZÁLEZ; RUBIO-CLEMENTE;	**Colombia**	Lakes	Yes	Temperate	Developed	OM + Programs	HRM + TM	STAT	CCA DCA RDA	R^2	Chl-a TN	MC

PEÑUELA, 2024)												
(ZHANG; LIU, 2021)	**China**	Reservoirs	Yes	Temperate	Developeing	Research	TM	STAT	Pearson ,PCA	N/A	Chl-a TURB CD TP DO	ULM/FDFA
(NIU *et al.*, 2021)	**China**	Rivers	Unspecified	Temperate	Developeing	Research	RS+TM	ML	pixel_DNNR SVR patch_DNNR PLSR	R^2 MSE MAE MNB RPD	HI	Zn Cd Ni NH4 COD TN TP
(MELARAM; LOPEZ-DUEÑAS, 2022)	**USA**	Lakes	Unspecified	Temperate	Developed	OM + Programs	TM	STAT	Kruskal-Wallis	N/A	MC NOD	ULM/FDFA
(GORNEY *et al.*, 2023)	**USA**	Lakes	Unspecified	Temperate	Developed	OM + Programs	HRM + TM	STAT	ST	R^2	Chl-a	Chl-a
(XU *et al.*, 2015)	**China**	Lakes	Yes	Temperate	Developeing	Research	HRM + TM	STAT	ANOVA	N/A	Chl-a TN TDN NH4 NO3 NO2 TP TDP SRP	ULM/FDFA
(BARRINGTON *et al.*, 2014)	**Australia**	Lakes	Unspecified	Mediterranean	Developed	TO	N/A	STAT	DS	N/A	"Defining the hazard" "Timescale Hazard frequency"	Cyanobacteria

											"Tolerance" "monitoring frequency"	
(NIETCH *et al.*, 2022)	**USA**	Rivers	Unspecified	Temperate	Developed	OM + Program s	HRM + TM	STAT	GLM (LOGREG) GLM(BM)	N/A	IN OUT	BLOOM(YES/ NO)
(PAZO *et al.*, 2024)	**Spain**	Rivers	Unspecified	Temperate	Developed	OM + Program s	HRM	ML	BM DI EBDI SEM	RMSE ,NRMSE ,R^2 ,MAE (TABLE 4)	Q TEMP DO pH TURB EC WT LEVEL NH4 PO4	Chl-a
(OMONDI *et al.*, 2023)	**Kenya, Tanzania**	Reservo irs	Unspecified	Tropical/Subt ropical	Developeing	TO	RS+HRM + TM	STAT	MRM,CA	R^2,r MAE nRMSE	Sensor-reflectance	TURB SST Chl-a
(MEDINA-COBO *et al.*, 2014)	**Spain**	Various	Unspecified	Unspecified	Developed	TO	RS+TM	STAT	CI RL	R^2	YES	BIO-CYANO
(LIN *et al.*, 2022)	**Taiwan**	Reservo irs	Yes	Tropical/Subt ropical	Developed	TO	HRM + TM	STAT	PCA	N/A	Chl-a TP SECCHI SS TURB NH4 COD	ULM/FDFA

(HWANG *et al.*, 2023)	**Republic of Korea**	Lakes	Yes	Continental	Developed	TO	HRM + TM	ML	SOM Spearman AdaBoost Gboost XGBoost LDA DNN	Gini MSE Entropy	COD BOD TN TP TOC SS EC pH DO TURB TEMP Chl-a Phytoplankton	dominant-ALGAE
(LUNETTA *et al.*, 2015)	**USA**	Various	Unspecified	Temperate	Developed	OM + Programs	RS+TM	STAT	DS ,CI	N/A	HI	BLOOM(YES/NO)
(LI *et al.*, 2021)	**Sweden**	Lakes	Yes	Temperate	Developed	OM + Programs	HRM + TM	STAT	Mann Kendall ,PCA Quantile	R^2	BIO-CYANO Chl-a SECCHI TP TN TOC pH Fe Si ALK EC	BIO-CYANO
(FERNÁNDEZ; ESTRADA ; PARODI, 2015)	**Argentina**	Reservoirs	Yes	Tropical/Subtropical	Developeing	TO	HRM + TM	STAT	ANOVA Bonferroni Kruskal-Wallis Dunn.	N/A	Chl-a SR PO4	CD Chl-a
(SHIZUKA *et al.*, 2021)	**Japan**	Lakes	Yes	Temperate	Developed	OM + Programs	HRM + TM	STAT	GLM (LOGREG)	pseudo-R^2	SR WT SAL DO pH SECCHI TN DIN TP DIP TN/TP DIN/TP Chl-a	MIB

(KIM; SHON; SHIN, 2013)	**Republic of Korea**	Reservoirs	Yes	Continental	Developed	TO	HRM	ML	ANN WNN	RP RV RMSE r	WT pH DO TOC TP TN IN OUT	Chl-a
(XIE *et al.*, 2012)	**China**	Reservoirs	Yes	Tropical/Subtropical	Developeing	TO	HRM + TM	ML	SVM	R^2 MAE (TABLE 2)	TURB WT pH EC CHLORIDES SO4 SIL ALK HCO3 DO NH4 NO3 NO2 TN TP SS TOC UV254 Fe PO4 TOC HRT LEVEL PREC Phytoplankton	Phytoplankton
(COOK *et al.*, 2023)	**USA**	Lakes	Unspecified	Temperate	Developed	OM + Programs	RS+TM	ML	RL RF.	RMSE R^2	HI	PC Chl-a
(SU *et al.*, 2021)	**China**	Reservoirs	Yes	Tropical/Subtropical	Developeing	TO	RS+TM	STAT	MLG RLE Qunatile	R^2	CD	MIB
(HANDLER *et al.*, 2023)	**USA**	Various	Unspecified	Temperate	Developed	OM + Programs	RS+TM	STAT	THERE	MAP R^2	HI	Cyanobacteria MC
(ZONG *et al.*, 2019)	**China**	Rivers	Yes	Temperate	Developeing	TO	RS+HRM + TM	STAT	RL GLM	AICc	YES	cyanobacterial bloom area

												percentage cyanobacterial bloom frequency index
(SMUCKER *et al.*, 2021)	**USA**	Reservoirs	Yes	Temperate	Developed	OM + Programs	HRM + TM	STAT	GAM	R^2	PREC WT KN	CD
(CLERCIN ; DRUSCHEL, 2019)	**USA**	Reservoirs	Yes	Temperate	Developed	TO	HRM + TM	STAT	CCA and Spearman	N/A	TEMP EC DO pH Chl-a PC TN TP TN/TP NO2 NO3 MIB GEO MC TDS	ULM/FDFA
(KWON *et al.*, 2023)	**Republic of Korea**	Reservoirs	Yes	Continental	Developed	Research	HRM	ML	LSTM	r RMSE	WT pH EC DO TURB PREC WS IN OUT	Chl-a PC
(LOU *et al.*, 2015)	**China**	Reservoirs	Yes	Tropical/Subtropical	Developeing	TO	HRM + TM	ML	SVM PSO	R^2 MAE	TURB TEMP pH EC CHLORIDES SO4 Si ALK HCO3 DO NH4 NO2 NO3 TN TP SS TOC UV254 Fe PO4	Phytoplankton

											HRT LEVEL PREC Phytoplankton	
(KRÖGER *et al.*, 2023)	**Germany**	Lakes	Yes	Temperate	Developed	OM + Programs	HRM + TM	STAT	PCA Beveridge-Nelson Cross-Mapping	R^2	TP SRP	CD
(LUONG *et al.*, 2024)	**Australia**	Reservoirs	Yes	Tropical/Subtropical	Developed	TO	TM	STAT	GAMM,PCA DCA RDA	N/A	WT DO pH EC SECCHI	BIO Q-Index
(ALMUHTARAM; ZAMYADI; HOFMANN, 2021)	**Canada, USA**	Lakes	Yes	Temperate	Developed	OM + Programs	HRM	ML	LOF SVM,EE iForest	F1	PC Chl-a	PC Chl-a
(JEONG *et al.*, 2022)	**Republic of Korea**	Reservoirs	Yes	Continental	Developed	TO	HRM + TM	ML	RF XGBoost XGBoost (SMOTE)	Accuracy AUC F1	Chl-a TN TP	CD
(WHEELER *et al.*, 2012)	**Canada, USA**	Lakes	Yes	Temperate	Developed	OM + Programs	RS+TM	STAT	THERE	MAP R^2	HI	Cyanobacteria

(WILLIAMS; COLE, 2013)	**Australia**	Reservoirs	Yes	Temperate	Developed	OM + Programs	HRM	ML	BM	N/A	Chl-a TN TP	BLOOM(YES/NO)
(HURTADO *et al.*, 2022)	**Spain**	Reservoirs	Yes	Temperate	Developed	OM + Programs	HRM + TM	ML	GERMAN	Acc ROC	TEMP Chl-a	Chl-a BIO-CYANO MC EX STX
(MEZZANOTTE *et al.*, 2011)	**China**	Lakes	Yes	Temperate	Developeing	OM + Programs	RS+HRM + TM	STAT	DS PCA	N/A	Chl-a NH4 HCMC	Chl-a
(LI; HANSSON; PERSSON, 2018)	**Sweden**	Lakes	Yes	Continental	Developed	TO	TM	STAT	Spearman	N/A	TN DIN TP TN/TP BIO-CYANO	ULM/FDFA
(LEI *et al.*, 2014)	**China**	Reservoirs	Yes	Tropical/Subtropical	Developeing	TO	HRM + TM	STAT	Spearman	N/A	CYN MC BIO-CYANO BIO	ULM/FDFA
(MOHAMED *et al.*, 2016)	**Egypt**	Reservoirs	Yes	Arid/Semiars	Developeing	Research	HRM + TM	STAT	ANOVA Spearman	N/A	DO EC DOC TN TP pH MC CD	ULM/FDFA
(PRESTIGIACOMO *et al.*, 2023)	**USA**	Lakes	Others	Tropical/Subtropical	Developed	TO	HRM + TM	STAT	DS Kruskal-Wallis	N/A	HCMC Chl-a MC SECCHI	ULM/FDFA

(GOLSHAN *et al.*, 2020)	**Australia**	Reservoirs	Yes	Tropical/Subtropical	Developed	TO	HRM	STAT	MRM MCR-ALS	N/A	Cyanobacteria PHYSICOCHEMICAL	ULM/FDFA
(PÉREZ-GONZÁLEZ *et al.*, 2021)	**Spain**	Reservoirs	Yes	Temperate	Developed	OM + Programs	RS+HRM + TM	STAT	TABle 2	R^2 RRMSE RMSE	Sensor-reflectance	CD BIO-CYANO Chl-a PC
(HARRIS; GRAHAM, 2017)	**USA**	Reservoirs	Yes	Temperate	Developed	OM + Programs	HRM + TM	STAT	RL RnL Table 2	RMSE R^2	TEMP pH TN TP COD DO Chl-a TURB	Cyanobacteria MC GEO
(FRANCY *et al.*, 2020)	**USA**	Lakes	Unspecified	Temperate	Developed	OM + Programs	HRM + TM	STAT	RL-OLS	R^2 AIC	WT pH DO EC Chl-a PC Cyanobacteria	MC
(GOES; BARROS; NETO, 2023)	**BRAZIL**	Reservoirs	Yes	Arid/Semiars	Developeing	TO	HRM + TM	STAT	MLR	R^2	VOL PREC TP TN Chl-a CD SECCHI	HCMC
(COFFER *et al.*, 2020)	**USA**	Lakes	Unspecified	Temperate	Developed	OM + Programs	RS+TM	STAT	THERE	MAP R^2	HI	BIO-CYANO
(BRESCIANI, 2011)	**Italy**	Lakes	Yes	Temperate	Developed	Research	RS+TM	STAT	THERE	MAP R^2	HI	BIO-CYANO

(LI; LI; SONG, 2015)	**USA**	Reservoirs	Yes	Temperate	Developed	OM + Programs	RS+TM	STAT	CI RL	R^2 RMSE	HI	PC Chl-a
(GUNKEL; SOBRAL, 2013)	**BRAZIL**	Reservoirs	Yes	Tropical/Subtropical	Developeing	TO	HRM + TM	STAT	CLM RL	R^2	HCMC	Chl-a
(BERTRAND *et al.*, 2022)	**BRAZIL**	Rivers	Yes	Arid/Semiars	Developeing	OM + Research	HRM + TM	STAT	PCA	N/A	pH TEMP EC TURB Thermotolerant-Coliforms Ca NO3 HCO3	ULM/FDFA
(REYNOLDS *et al.*, 2023)	**USA**	Lakes	Yes	Tropical/Subtropical	Developed	TO	RS+HRM + TM	STAT	DS CI	WHAT I	YES	Chl-a
(CLARK *et al.*, 2017)	**USA**	Various	Unspecified	Temperate	Developed	OM + Programs	RS+TM	STAT	THERE	MAP R^2	HI	BIO-CYANO
(COFFER *et al.*, 2021)	**USA**	Lakes	Unspecified	Temperate	Developed	OM + Programs	RS+TM	STAT	THERE	MAP R^2	HI	BIO-CYANO
(MÂNICA; DE LIMA ISAAC, 2023)	**BRAZIL**	Rivers	Yes	Tropical/Subtropical	Developeing	TO	TM	STAT	ANOVa Spearman	N/A	CD Chl-a BOD COD DO NH4 NO3 NO2 pH TDS TP TURB	ULM/FDFA

											Thermotolerant-Coliforms WT	
(WU *et al.*, 2022)	**China**	Reservoirs	Yes	Temperate	Developeing	OM + Research	HRM + TM	STAT	NMDS,PCA RDA Pearson	N/A	WT DO pH EC TURB pH Chl-a SS DTN DTP TP TN NH4 NO3 NO2 PO4 MIB GEO	ULM/FDFA
(ALVAREZ DALINGER *et al.*, 2023)	**Argentina**	Reservoirs	Yes	Tropical/Subtropical	Developeing	TO	TM	STAT	t-test u-test,Spearman	N/A	WT DO pH EC TURB pH Chl-a SS DTN DTP TP TN NH4 NO3 NO2 PO4 MIB GEO	ULM/FDFA
(ZHANG *et al.*, 2015)	**China**	Lakes	Yes	Tropical/Subtropical	Developeing	Research	HRM + TM	ML	GOOD	R^2 RMSE AS	Chl-a WT SECCHI pH NO3 NH4 PO4 CD	CD
(HWAN *et al.*, 2023)	**Republic of Korea**	Lakes	Yes	Continental	Developed	TO	TM	STAT	Spearman CCA	N/A	GEO MIB ALGAE Cyanobacteria	ULM/FDFA
(ZHANG; WANG;	**China**	Lakes	Yes	Tropical/Subtropical	Developeing	TO	RS+TM	STAT	THERE	MAP R^2	YES	Chl-a

CHANG, 2019)												
(SUAREZ *et al.*, 2023)	**Switzerland**	Lakes	Yes	Temperate	Developed	OM + Programs	RS+TM	STAT	SKT Pettitt	MAP, R^2	BIO BIO-CYANO DIP TURB	ULM/FDFA
(ELLIOTT, 2010)	**UK**	Lakes	Yes	Temperate	Developed	OM + Programs	HRM + TM	STAT	RL	R^2	Chl-a	Chl-a
(FILBRUN; CONROY; CULVER, 2013)	**USA**	Lakes	Yes	Temperate	Developed	OM + Programs	HRM + TM	STAT	MEM	N/A	Factors(Reservoir ConnectivityTreatment Time (days or months), Interaction (Treatment and Time))	TRP NH4 NO3
(GOYENS *et al.*, 2022)	**Belgium**	Reservoirs	Yes	Temperate	Developed	Research	RS+TM	STAT	THERE	MAP R^2	HI	CD
(KAKOUEI; KRAEMER; ADRIAN, 2022)	**Germany**	Lakes	Yes	Temperate	Developed	OM + Programs	HRM + TM	ML	RF	adjustment-R^2	TEMP SR RU PREC WS IN Chl-a DO EC pH TN TP DOC	Phytoplankton Zooplankton-abundance BIO BIO-ZOO

Table S4 - List of model acronyms

Acronym	Description	Acronym	Description
AdaBoost	Adaptive Boosting	Kruskal-Wallis	Kruskal-Walli's test
FRIEND	Spectral form Algorithms - Aphanizomenon-Microcystis Index	LDA	Linear Discriminant Analysis
ANN	Artificial neural networks	PRAISE	Local Outlier Factor
ANOVA	Analysis of variance	LR	Lenar regression
ARIMA	autoregressive integrated moving average	LSTM	Long Short-Term Memory
Beveridge-Nelson	Beveridge-Nelson Decomposition	Mann-Kendall	Mann-Kendall test
BiLSTM	bidirectional Long Short-Term Memory	MCR-ALS	Multivariate curve resolution–alternating least squares
BM	Bayesian network models or others Bayesian Models	MEM	Mixed effects models
Bonferroni	Bonferroni test	MRM	Multivariate regression MODELS
CA	Correspondence Analysis	NMDS	Non-metric multidimensional scaling
CCA	Canonical Correlation Analysis	Pearson	Pearson's correlation
THERE	Cyanobacteria index	Pettitt	Pettitt method
CLM	Critical Load Model OCDE	pixel_DNNR	Deep neural network regression based on pixels
CNN	Convolutional Neural Networks	PLSR	Partial least square regression
Cross-Mapping	Convergent Cross-Mapping	PSO	Particle Swarm Optimization
DCA	Detrended Correspondence Analysis	Quantile	Quantile regression
OF	Dynamic Imputation	RDA	Redundancy Analysis
DNN	Deep Neural Network	RF	Random Forest
DS	Descriptive statistics	RL-OLS	Linear regression - Ordinary Least Squares
GERMAN	Decision Trees	WITHOUT	Structural Expectation Maximization
Dunn	Dunn's Test	ST	Seasonal Kendall Test
EBDI	Entropy-Based Dynamic Imputation	SMOTE	Synthetic minority oversampling technique
EE	Elliptical Envelope	AS	Self-Organizing Map
FUZZY	Fuzzy theory	Spearman	Spearman correlation
GA	Hybrid Genetic Algorithm	SVC	Support vector machine

GAM	Generalized additive models	SVR	Support Vector Regression
GAMM	generalized additive mixed models	t-test	Student’s t-test
Gboost	Gradient Boosting	u-test	Mann–Whitney U test
GLM	Generalized Linear Model	WA-PLS	Weighted averaging partial least squares regression
GOOD	Hybrid Evolutionary Algorithms	WNN	Wavelet neural networks
iForest	Isolation Forest	XGBoost	Extreme Gradient Boosting
IOSM	interior-outer-set models		

Table S5 - List of Fit Metrics acronyms

Acronym	Description
r	Correlation coefficient
THERE ARE	mean absolute error
SEA	mean absolute relative error
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Square Error
HIS	Sensitivity analysis
R^2	Determination coefficient
MSE	Mean Squared Error
MNB	Normalized Bias
RPD	Residual Prediction Deviation
pseudo-R^2	Pseudo-R^2
F1	F1 Score
Acc	Accuracy
RP	Relative peak error
RV	Relative volume error
GC	Gini coefficient
Entropy	Entropy
AE	Absolute error
RE	Relative error
adjust-R^2	Adjusted -R^2
AICc	Akaike Information Criterion with a correction

Table S6 - List of acronyms for independent variables

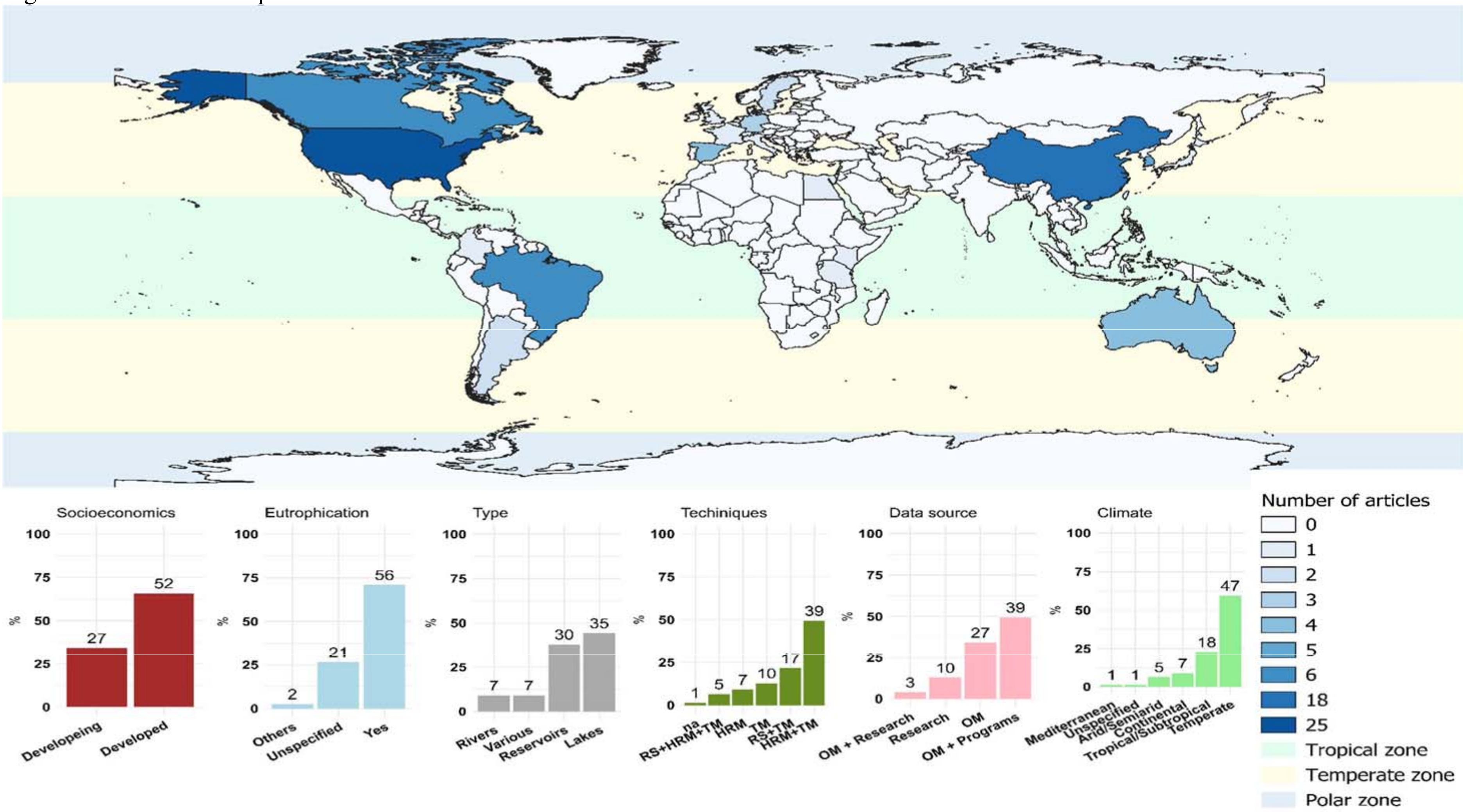
Variable	Description	Variable	Description
ALGAE	Information of abundance or composition other than cell density or biovolume	MC	Microcystin
FROM	Alkalinity	MIB	2-Methylisoborneol
ATX-a	Anatoxin-A	Morphological-predictors	Morphological predictors
BIO	Biovolume of phytoplankton	neo-STX	Saxitoxin
BIO-CYANO	Cyanobacteria biovolume	NH4	Ammoniacal nitrogen
POINT	Biochemical oxygen demand	NO2	Nitrite
Ca	Calcium	NO3	Nitrate
CD	Cell density	NOD	Nodularin
Chl-a	Chlorophyll-a	OUT	Effluent flow
CHLORIDES	Chlorides	PC	Phycocyanin
COD	Chemical oxygen demand	pH	pH
Cyanobacteria	Information of abundance or composition other than cell density or biovolume	PHYSICOCHEMICAL	Varied physicochemical variables: see the publication for more details
BEFORE	Cylindrospermopsin	Phytoplankton	Information of abundance or composition other than cell density or biovolume
FROM	Dissolved inorganic nitrogen	PO4	Phosphate
DIN/TP	Dissolved inorganic nitrogen and total phosphorus ratio	PREC	Precipitation
DIP	Dissolved inorganic phosphorus	Q	Flow
OF	Dissolved oxygen	RU	Relative humidity
DOC	Dissolved Organic Carbon	WILL	Salinity
DTN	Total dissolved nitrogen	DRIED	Secchi Depth
DTP	Total Dissolved Phosphorus	Sensor-reflectance	Remote sensing data from satellites or not
EC	Electrical conductivity	YES	Remote sensing data from multispectral or hyperspectral satellites
Fe	Iron	Yes	Silicon
GEO	Geosmin	COAMING	Silicate
HCO3	Bicarbonate	SO4	Sulfate

HI	Information Collected from Hyperspectral Imaging	SR	Solar radiation
HRT	Hydraulic detention time	SRP	soluble reactive phosphorus
IN	Affluent flow	SS	Suspended solids
KN	Kjeldahl Nitrogen	STX	Saxitoxin
LEVEL	Reservoir water level	TDN	Total dissolved nitrogen

Table S7 - List of acronyms of dependent variables

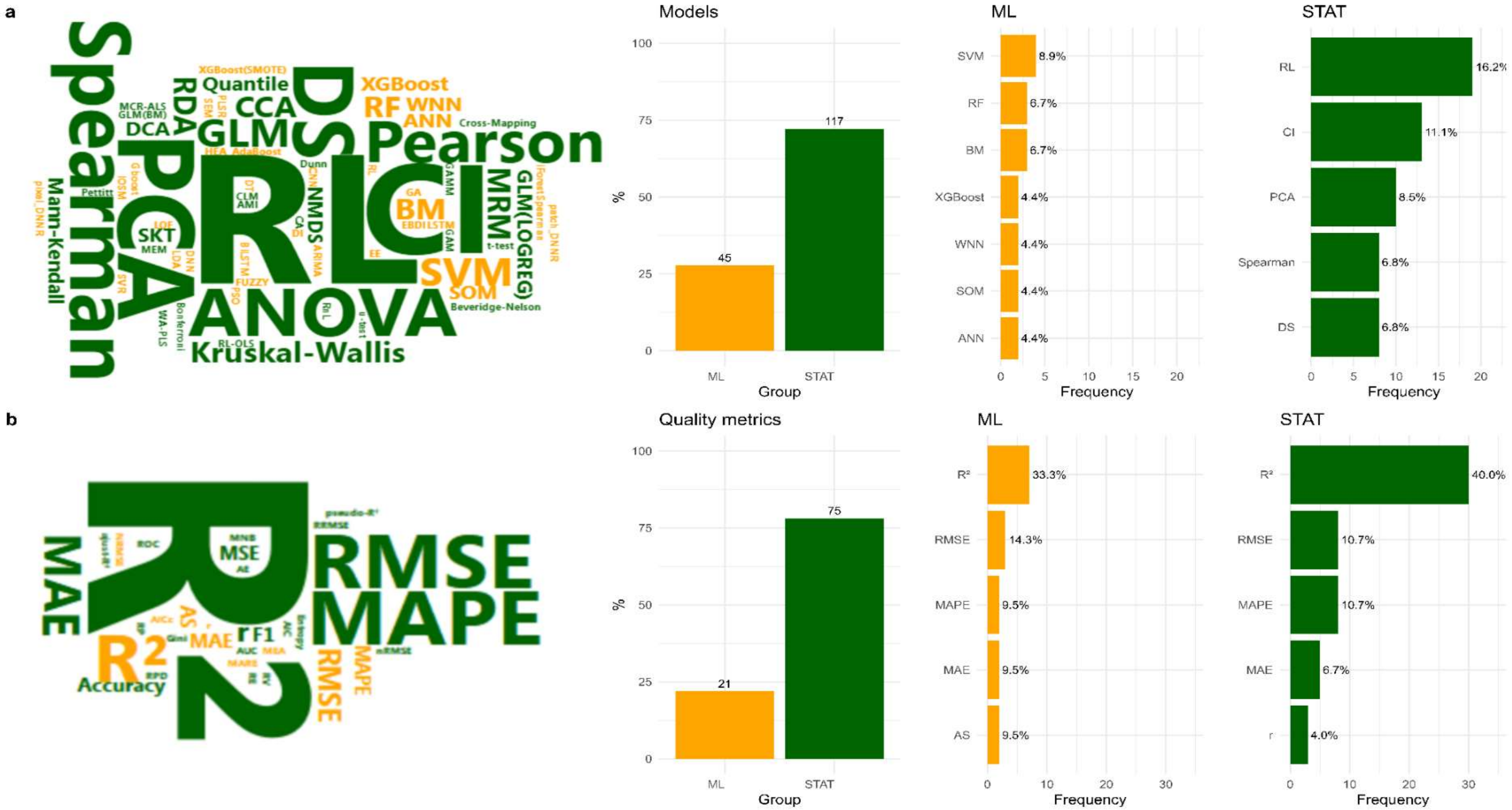
Variable	Description
Chl-a	Chlorophyll-a
BIO-CYANO	Cyanobacteria Biovolume
CD	Cell density
PC	phycocyanin
MC	Microcystin
BLOOM(YES/NO)	Bloom's state (factor)
Cyanobacteria	Mediated by different abundance of biovolume or cell density
Phytoplankton	Mediated by different abundance of biovolume or cell density
HCMC	Total phosphorus
BIO	biovolume of phytoplankton
MIB	2-Methylisoborneol
NH4	ammoniacal nitrogen
TURB	Turbidity
BIO-ZOO	Zooplankton biovolume
CBD	Cyanobacterial bloom area
COD	Chemical oxygen demand
BEFORE	Cylindrospermopsin
Cd	Cadmus
HOUSE	Dissolved organic matter
GEO	Geosmin
NO3	Nitrate
Nor	Nickel
CVITP	Percentage cyanobacterial bloom frequency index
Q-Index	Q-Index
SST	Total suspended solids
STX	Saxitoxin
TN	Total Nitrogen
TRP	Total reactive phosphorus
WT	Water temperature
Zn	Zinc
Zooplankton-abundance	Mediated by different abundance of biovolume or cell density
Dominant-ALGAE	Dominant algae
ULM/FDFA	Unsupervised learning methodology/Exploratory data analysis

Figure S1 – Overview of publications selected for RSL



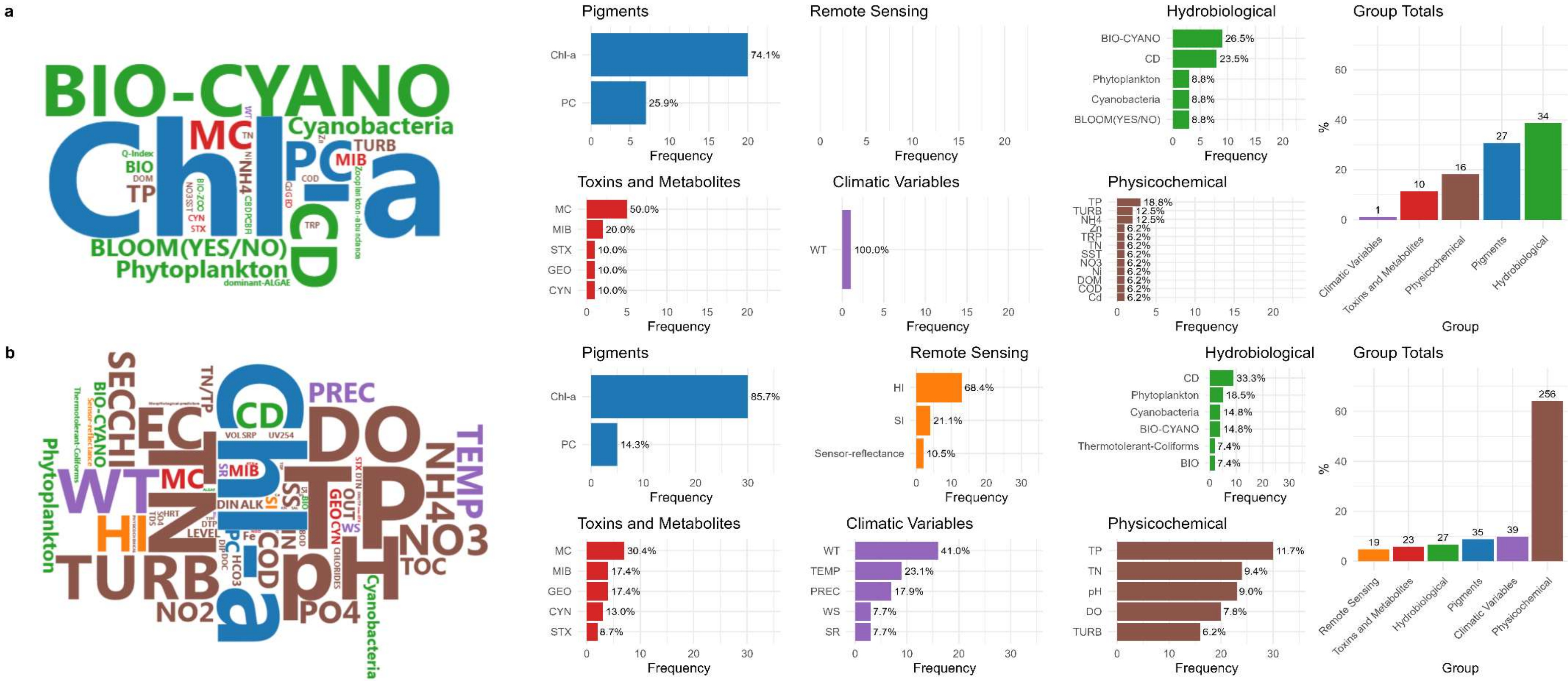
Source: prepared by the author.

Figure S2 - Overview of the models (a) and their respective metrics (b) obtained through the word cloud and exploratory analysis of the frequency of the acronyms found in the articles obtained by RSL. The percentages in the horizontal bars were calculated based on the totals of each group indicated at the top of the vertical bars of the same color.



Source: prepared by the author.

Figure S3 – List of dependent (a) and independent (b) variables.



Source: prepared by the author.

REFERENCES

ALMUHTARAM, H.; ZAMYADI, A.; HOFMANN, R. Machine learning for anomaly detection in cyanobacterial fluorescence signals. **Water Research**, v. 197, p. 117073, 2021. Disponível em: https://doi.org/10.1016/j.watres.2021.117073

ALVAREZ DALINGER, F. S.; LOZANO, V. L.; BORJA, C. N.; MORAÑA, L. B.; MARÍA MÓNICA, S. Short-Term Meteorological Conditions Explain Cyanobacterial Blooms in a Tropical Reservoir. **Water (Switzerland)**, v. 15, n. 2, p. 1–16, 2023. Disponível em: https://doi.org/10.3390/w15020302

BARRINGTON, D. J.; GHADOUANI, A.; SINANG, S. C.; IVEY, G. N. Development of a new risk-based framework to guide investment in water quality monitoring. **Environmental Monitoring and Assessment**, v. 186, n. 4, p. 2455–2464, 2014. Disponível em: https://doi.org/10.1007/s10661-013-3552-1

BERTRAND, G. F.; DE PAIVA, A. L. R.; DE ARAÚJO FREITAS, J. B.; DA SILVA PEREIRA CABRAL, J. J.; VERAS ALBUQUERQUE, T. B.; DE CARVALHO FILHO, J. A. A. River bank filtration in tropical metropoles: integrated evaluation of physical, geochemical and biochemical interactions in Recife, NE Brazil. **International Journal of Environmental Science and Technology**, v. 19, n. 7, p. 5803–5818, 2022. Disponível em: https://doi.org/10.1007/s13762-021-03558-x

BRESCIANI, M. Recognizing harmful algal bloom based on remote sensing reflectance band ratio. **Journal of Applied Remote Sensing**, v. 5, n. 1, p. 053556, 2011. Disponível em: https://doi.org/10.1117/1.3630218

BRINOVCAR, C.; CRUMB, J.; TAYLOR, K.; CLARKE, P.; HABASH, M. Assessing Trends of Physical and Chemical Parameters During Elevated Cyanobacteria Growth Periods in a Drinking Water Reservoir in Ontario, Canada (2011–2016). **Water, Air, & Soil Pollution**, v. 233, n. 7, p. 274, 2022. Disponível em: https://doi.org/10.1007/s11270-022-05728-0

CHEN, Q.; RUI, H.; LI, W.; ZHANG, Y. Analysis of algal bloom risk with uncertainties in lakes by integrating self-organizing map and fuzzy information theory. **Science of the**

Total Environment, v. 482–483, n. 1, p. 318–324, 2014. Disponível em: https://doi.org/10.1016/j.scitotenv.2014.02.096

CLARK, J. M.; SCHAEFFER, B. A.; DARLING, J. A.; URQUHART, E. A.; JOHNSTON, J. M.; IGNATIUS, A. R.; MYER, M. H.; LOFTIN, K. A.; WERDELL, P. J.; STUMPF, R. P. Satellite monitoring of cyanobacterial harmful algal bloom frequency in recreational waters and drinking water sources. **Ecological Indicators**, v. 80, n. November 2016, p. 84–95, 2017. Disponível em: https://doi.org/10.1016/j.ecolind.2017.04.046

CLERCIN, N. A.; DRUSCHEL, G. K. Influence of Environmental Factors on the Production of MIB and Geosmin Metabolites by Bacteria in a Eutrophic Reservoir. **Water Resources Research**, v. 55, n. 7, p. 5413–5430, 2019. Disponível em: https://doi.org/10.1029/2018WR023651

COFFER, M. M.; SCHAEFFER, B. A.; DARLING, J. A.; URQUHART, E. A.; SALLS, W. B. Quantifying national and regional cyanobacterial occurrence in US lakes using satellite remote sensing. **Ecological Indicators**, v. 111, n. November 2019, 2020. Disponível em: https://doi.org/10.1016/j.ecolind.2019.105976

COFFER, M. M.; SCHAEFFER, B. A.; SALLS, W. B.; URQUHART, E.; LOFTIN, K. A.; STUMPF, R. P.; WERDELL, P. J.; DARLING, J. A. Satellite remote sensing to assess cyanobacterial bloom frequency across the United States at multiple spatial scales. **Ecological Indicators**, v. 128, p. 107822, 2021. Disponível em: https://doi.org/10.1016/j.ecolind.2021.107822

COOK, K. V.; BEYER, J. E.; XIAO, X.; HAMBRIGHT, K. D. Ground-based remote sensing provides alternative to satellites for monitoring cyanobacteria in small lakes. **Water Research**, v. 242, n. May, p. 120076, 2023. Disponível em: https://doi.org/10.1016/j.watres.2023.120076

COURTOIS, S.; STEINMANN, D.; CAJON, A.; VAN DER LINDEN, L. Continuous monitoring of cyanobacterial blooms: benefits and conditions for using fluorescence probes. **Revue des sciences de l’eau**, v. 30, n. 2, p. 149–155, 2018. Disponível em: https://doi.org/10.7202/1042923ar

ELLIOTT, J. A. The seasonal sensitivity of Cyanobacteria and other phytoplankton to changes in flushing rate and water temperature. **Global Change Biology**, v. 16, n. 2, p. 864–876, 2010. Disponível em: https://doi.org/10.1111/j.1365-2486.2009.01998.x

FERNANDEZ-FIGUEROA, E. G.; BULEY, R. P.; BARROS, M. U. G.; GLADFELTER, M. F.; MCCLIMANS, W. D.; WILSON, A. E. Carlson's Trophic State Index is a poor predictor of cyanobacterial dominance in drinking water reservoirs. **AWWA Water Science**, v. 3, n. 2, 2021. Disponível em: https://doi.org/10.1002/aws2.1219

FERNÁNDEZ, C.; ESTRADA, V.; PARODI, E. R. Factors Triggering Cyanobacteria Dominance and Succession During Blooms in a Hypereutrophic Drinking Water Supply Reservoir. **Water, Air, & Soil Pollution**, v. 226, n. 3, p. 73, 2015. Disponível em: https://doi.org/10.1007/s11270-014-2290-5

FILBRUN, J. E.; CONROY, J. D.; CULVER, D. A. Understanding seasonal phosphorus dynamics to guide effective management of shallow, hypereutrophic Grand Lake St. Marys, Ohio. **Lake and Reservoir Management**, v. 29, n. 3, p. 165–178, 2013. Disponível em: https://doi.org/10.1080/10402381.2013.823469

FRANCY, D. S.; BRADY, A. M. G.; STELZER, E. A.; CICALE, J. R.; HACKNEY, C.; DALBY, H. D.; STRUFFOLINO, P.; DWYER, D. F. Predicting microcystin concentration action-level exceedances resulting from cyanobacterial blooms in selected lake sites in Ohio. **Environmental Monitoring and Assessment**, v. 192, n. 8, 2020. Disponível em: https://doi.org/10.1007/s10661-020-08407-x

GOES, M. C. A.; BARROS, M. U. G.; NETO, I. E. L. Prediction of total phosphorus in reservoir cascade systems. **Environmental Monitoring and Assessment**, v. 195, n. 12, p. 1–12, 2023. Disponível em: https://doi.org/10.1007/s10661-023-12155-z

GOLSHAN, A.; EVANS, C.; GEARY, P.; MORROW, A.; MAEDER, M.; TAULER, R. Patterns of cyanobacterial abundance in a major drinking water reservoir: what 3 years of comprehensive monitoring data reveals? **Environmental Monitoring and Assessment**, v. 192, n. 2, 2020. Disponível em: https://doi.org/10.1007/s10661-020-8090-z

GORNEY, R. M.; JUNE, S. G.; STAINBROOK, K. M.; SMITH, A. J. Detections of cyanobacteria harmful algal blooms (cyanoHABs) in New York State, United States (2012–2020). **Lake and Reservoir Management**, v. 39, n. 1, p. 21–36, 2023. Disponível em: https://doi.org/10.1080/10402381.2022.2161436

GOYENS, C.; LAVIGNE, H.; DILLE, A.; VERVAEREN, H. Using Hyperspectral Remote Sensing to Monitor Water Quality in Drinking Water Reservoirs. **Remote Sensing**, v. 14, n. 21, p. 5607, 2022. Disponível em: https://doi.org/10.3390/rs14215607

GUNKEL, G.; SOBRAL, M. Re-oligotrophication as a challenge for tropical reservoir management with reference to Itaparica Reservoir, São Francisco, Brazil. **Water Science and Technology**, v. 67, n. 4, p. 708–714, 2013. Disponível em: https://doi.org/10.2166/wst.2012.583

HANDLER, A. M.; COMPTON, J. E.; HILL, R. A.; LEIBOWITZ, S. G.; SCHAEFFER, B. A. Identifying lakes at risk of toxic cyanobacterial blooms using satellite imagery and field surveys across the United States. **Science of the Total Environment**, v. 869, n. October 2022, p. 161784, 2023. Disponível em: https://doi.org/10.1016/j.scitotenv.2023.161784

HARRIS, T. D.; GRAHAM, J. L. Predicting cyanobacterial abundance, microcystin, and geosmin in a eutrophic drinking-water reservoir using a 14-year dataset. **Lake and Reservoir Management**, v. 33, n. 1, p. 32–48, 2017. Disponível em: https://doi.org/10.1080/10402381.2016.1263694

HURTADO, I.; POUGET, L.; FERNÁNDEZ, S.; CASCALES, P. Monitoring and forecasting cyanobacteria risk for a drinking water plant in Spain. **Water Supply**, v. 22, n. 7, p. 6296–6307, 2022. Disponível em: https://doi.org/10.2166/ws.2022.246

HWAN, B. J.; KIM, H. N.; KANG, T. G.; KIM, B. H.; BYEON, M. S. Study of the cause of the generation of odor compounds (geosmin and 2-methylisoborneol) in the Han River system, the drinking water source, Republic of Korea. **Water Supply**, v. 23, n. 3, p. 1081–1093, 2023. Disponível em: https://doi.org/10.2166/ws.2023.037

HWANG, S. Y.; CHOI, B. W.; PARK, J. H.; SHIN, D. S.; CHUNG, H. S.; SON, M. S.; LIM, C. H.; CHAE, H. M.; HA, D. W.; JUNG, K. Y. Evaluating Statistical Machine Learning Algorithms for Classifying Dominant Algae in Juam Lake and Tamjin Lake, Republic of Korea. **Water (Switzerland)**, v. 15, n. 9, 2023. Disponível em: https://doi.org/10.3390/w15091738

JEONG, B.; CHAPETA, M. R.; KIM, M.; KIM, J.; SHIN, J.; CHA, Y. K. Machine learning-based prediction of harmful algal blooms in water supply reservoirs. **Water Quality Research Journal**, v. 57, n. 4, p. 304–318, 2022. Disponível em: https://doi.org/10.2166/wqrj.2022.019

KAKOUEI, K.; KRAEMER, B. M.; ADRIAN, R. Variation in the predictability of lake plankton metric types. **Limnology and Oceanography**, v. 67, n. 3, p. 608–620, 2022. Disponível em: https://doi.org/10.1002/lno.12021

KIM, M. E.; SHON, T. S.; SHIN, H. S. Forecasting algal bloom (chl-a) on the basis of coupled wavelet transform and artificial neural networks at a large lake. **Desalination and Water Treatment**, v. 51, n. 19–21, p. 4118–4128, 2013. Disponível em: https://doi.org/10.1080/19443994.2013.781108

KRÖGER, B.; SELMECZY, G. B.; CASPER, P.; SOININEN, J.; PADISÁK, J. Longterm phytoplankton community dynamics in Lake Stechlin (north-east Germany) under sudden and heavily accelerating eutrophication. **Freshwater Biology**, v. 68, n. 5, p. 737–751, 2023. Disponível em: https://doi.org/10.1111/fwb.14060

KUDELA, R. M.; PALACIOS, S. L.; AUSTERBERRY, D. C.; ACCORSI, E. K.; GUILD, L. S.; TORRES-PEREZ, J. Application of hyperspectral remote sensing to cyanobacterial blooms in inland waters. **Remote Sensing of Environment**, v. 167, p. 196–205, 2015. Disponível em: https://doi.org/10.1016/j.rse.2015.01.025

KWON, D. H.; HONG, S. M.; ABBAS, A.; PYO, J.; LEE, H. K.; BAEK, S. S.; CHO, K. H. Inland harmful algal blooms (HABs) modeling using internet of things (IoT) system and deep learning. **Environmental Engineering Research**, v. 28, n. 1, p. 0–1, 2023. Disponível em: https://doi.org/10.4491/eer.2021.280

LEI, L.; PENG, L.; HUANG, X.; HAN, B. P. Occurrence and dominance of Cylindrospermopsis raciborskii and dissolved cylindrospermopsin in urban reservoirs used for drinking water supply, South China. **Environmental Monitoring and Assessment**, v. 186, n. 5, p. 3079–3090, 2014. Disponível em: https://doi.org/10.1007/s10661-013-3602-8

LI, J.; HANSSON, L. A.; PERSSON, K. M. Nutrient control to prevent the occurrence of cyanobacterial blooms in a eutrophic lake in Southern Sweden, used for drinking water supply. **Water (Switzerland)**, v. 10, n. 7, 2018. Disponível em: https://doi.org/10.3390/w10070919

LI, J.; PERSSON, K. M.; PEKAR, H.; JANSSON, D. Evaluation of indicators for cyanobacterial risk in 108 temperate lakes using 23 years of environmental monitoring data. **Environmental Sciences Europe**, v. 33, n. 1, p. 1–13, 2021. Disponível em: https://doi.org/10.1186/s12302-021-00483-1

LI, L.; LI, L.; SONG, K. Remote sensing of freshwater cyanobacteria: An extended IOP Inversion Model of Inland Waters (IIMIW) for partitioning absorption coefficient and estimating phycocyanin. **Remote Sensing of Environment**, v. 157, p. 9–23, 2015. Disponível em: https://doi.org/10.1016/j.rse.2014.06.009

LIN, J. L.; KARANGAN, A.; HUANG, Y. M.; KANG, S. F. Eutrophication factor analysis using Carlson trophic state index (CTSI) towards non-algal impact reservoirs in Taiwan. **Sustainable Environment Research**, v. 32, n. 1, 2022. Disponível em: https://doi.org/10.1186/s42834-022-00134-x

LOAIZA-GONZÁLEZ, J. M.; RUBIO-CLEMENTE, A.; PEÑUELA, G. A. Cyanotoxin Monitoring and Detection Using Passive Sampling Application. **Water, Air, & Soil Pollution**, v. 235, n. 7, p. 423, 2024. Disponível em: https://doi.org/10.1007/s11270-024-07195-1

LORENZI, A. S.; CORDEIRO-ARAÚJO, M. K.; CHIA, M. A.; BITTENCOURTOLIVEIRA, M. do C. Cyanotoxin contamination of semiarid drinking water supply reservoirs. **Environmental Earth Sciences**, v. 77, n. 16, p. 0, 2018. Disponível em: https://doi.org/10.1007/s12665-018-7774-y

LOU, I.; XIE, Z.; UNG, W. K.; MOK, K. M. Integrating Support Vector Regression with Particle Swarm Optimization for numerical modeling for algal blooms of freshwater. **Applied Mathematical Modelling**, v. 39, n. 19, p. 5907–5916, 2015. Disponível em:

https://doi.org/10.1016/j.apm.2015.04.001

LUDOLF GOMES, L. N.; ALVES CORRÊA OLIVEIRA, S. M.; GIANI, A.; VON SPERLING, E. Association between biotic and abiotic parameters and the occurrence of cyanobacteria in a Brazilian reservoir. **Environmental Monitoring and Assessment**, v. 184, n. 8, p. 4635–4645, 2012. Disponível em: https://doi.org/10.1007/s10661-011-2291-4

LUNETTA, R. S.; SCHAEFFER, B. A.; STUMPF, R. P.; KEITH, D.; JACOBS, S. A.; MURPHY, M. S. Evaluation of cyanobacteria cell count detection derived from MERIS imagery across the eastern USA. **Remote Sensing of Environment**, v. 157, p. 24–34, 2015. Disponível em: https://doi.org/10.1016/j.rse.2014.06.008

LUONG, H. A.; ROHLFS, A. M.; FACEY, J. A.; COLVILLE, A.; MITROVIC, S. M. Long-term study of phytoplankton dynamics in a supply reservoir reveals signs of trophic state shift linked to changes in hydrodynamics associated with flow management and extreme events. **Water Research**, v. 256, n. November 2023, p. 121547, 2024. Disponível em: https://doi.org/10.1016/j.watres.2024.121547

MÂNICA, A. N.; DE LIMA ISAAC, R. Seasonal dynamics and diversity of cyanobacteria in a eutrophied Urban River in Brazil. **Water Supply**, v. 23, n. 9, p. 3868–3880, 2023. Disponível em: https://doi.org/10.2166/ws.2023.216

MEDINA-COBO, M.; DOMÍNGUEZ, J. A.; QUESADA, A.; DE HOYOS, C. Estimation of cyanobacteria biovolume in water reservoirs by MERIS sensor. **Water Research**, v. 63, p. 10–20, 2014. Disponível em: https://doi.org/10.1016/j.watres.2014.06.001

MELARAM, R.; LOPEZ-DUEÑAS, B. Detection and Occurrence of Microcystins and Nodularins in Lake Manatee and Lake Washington-Two Floridian Drinking Water Systems. **Frontiers in Water**, v. 4, 2022. Disponível em: https://doi.org/10.3389/frwa.2022.899572

MEZZANOTTE, V. *et al.* Monitoring, environmental emergencies management and water treatment improvement of freshwater lakes in China: The Chao Lake case study. **Water Science and Technology: Water Supply**, v. 11, n. 4, p. 490–496, 2011.

Disponível em: https://doi.org/10.2166/ws.2011.071

MOE, S. J.; HAANDE, S.; COUTURE, R.-M. Climate change, cyanobacteria blooms and ecological status of lakes: A Bayesian network approach. **Ecological Modelling**, v. 337, p. 330–347, 2016. Disponível em: https://doi.org/10.1016/j.ecolmodel.2016.07.004

MOHAMED, Z. A.; DEYAB, M. A.; ABOU-DOBARA, M. I.; EL-RAGHI, W. M. Occurrence of toxic cyanobacteria and microcystin toxin in domestic water storage reservoirs, Egypt. **Journal of Water Supply: Research and Technology-Aqua**, v. 65, n. 5, p. 431–440, 2016. Disponível em: https://doi.org/10.2166/aqua.2016.115

NI, J.; LIU, R.; LI, Y.; TANG, G.; SHI, P. An Improved Transfer Learning Model for Cyanobacterial Bloom Concentration Prediction. **Water**, v. 14, n. 8, p. 1300, 2022. Disponível em: https://doi.org/10.3390/w14081300

NIETCH, C. T.; GAINS-GERMAIN, L.; LAZORCHAK, J.; KEELY, S. P.; YOUNGSTROM, G.; URICHICH, E. M.; ASTIFAN, B.; DASILVA, A.; MAYFIELD, H. Development of a Risk Characterization Tool for Harmful Cyanobacteria Blooms on the Ohio River. **Water (Switzerland)**, v. 14, n. 4, 2022. Disponível em: https://doi.org/10.3390/w14040644

NIU, C.; TAN, K.; JIA, X.; WANG, X. Deep learning based regression for optically inactive inland water quality parameter estimation using airborne hyperspectral imagery. **Environmental Pollution**, v. 286, n. May, p. 117534, 2021. Disponível em: https://doi.org/10.1016/j.envpol.2021.117534

OMONDI, A. N.; OUMA, Y.; KOSGEI, J. R.; KONGO, V.; KEMBOI, E. J.; NJOROGE, S. M.; MECHA, A. C.; KIPKORIR, E. C. Estimation and mapping of water quality parameters using satellite images: a case study of Two Rivers Dam, Kenya. **Water Practice and Technology**, v. 18, n. 2, p. 428–443, 2023. Disponível em: https://doi.org/10.2166/wpt.2023.010

PATERSON, A. M.; RÜHLAND, K. M.; ANSTEY, C. V.; SMOL, J. P. Climate as a driver of increasing algal production in Lake of the Woods, Ontario, Canada. **Lake and Reservoir Management**, v. 33, n. 4, p. 403–414, 2017. Disponível em: https://doi.org/10.1080/10402381.2017.1379574

PAZO, M.; GERASSIS, S.; ARAÚJO, M.; MARGARIDA ANTUNES, I.; RIGUEIRA, X. Enhancing water quality prediction for fluctuating missing data scenarios: A dynamic Bayesian network-based processing system to monitor cyanobacteria proliferation. **Science of the Total Environment**, v. 927, n. January, p. 172340, 2024. Disponível em: https://doi.org/10.1016/j.scitotenv.2024.172340

PÉREZ-GONZÁLEZ, R.; SÒRIA-PERPINYÀ, X.; SORIA, J. M.; DELEGIDO, J.; URREGO, P.; SENDRA, M. D.; RUÍZ-VERDÚ, A.; VICENTE, E.; MORENO, J. Phycocyanin monitoring in some spanish water bodies with sentinel-2 imagery. **Water (Switzerland)**, v. 13, n. 20, p. 1–22, 2021. Disponível em: https://doi.org/10.3390/w13202866

PRESTIGIACOMO, A. R.; GORNEY, R. M.; HYDE, J. B.; DAVIS, C.; CLINKHAMMER, A. Patterns and impacts of cyanobacteria in a deep, thermally stratified, oligotrophic lake. **AWWA Water Science**, v. 5, n. 2, p. 1–18, 2023. Disponível em: https://doi.org/10.1002/aws2.1326

REYNOLDS, N.; SCHAEFFER, B. A.; GUERTAULT, L.; NELSON, N. G. Satellite and in situ cyanobacteria monitoring: Understanding the impact of monitoring frequency on management decisions. **Journal of Hydrology**, v. 619, n. February, p. 129278, 2023. Disponível em: https://doi.org/10.1016/j.jhydrol.2023.129278

SCHAEFFER, B. A.; IIAMES, J.; DWYER, J.; URQUHART, E.; SALLS, W.; ROVER, J.; SEEGERS, B. An initial validation of Landsat 5 and 7 derived surface water temperature for U.S. lakes, reservoirs, and estuaries. **International Journal of Remote Sensing**, v. 39, n. 22, p. 7789–7805, 2018. Disponível em: https://doi.org/10.1080/01431161.2018.1471545

SCHEIFHACKEN, N.; HORN, H.; PAUL, L. Comparing in situ particle monitoring to microscopic counts of plankton in a drinking water reservoir. **Water Research**, v. 44, n. 11, p. 3496–3510, 2010. Disponível em: https://doi.org/10.1016/j.watres.2010.03.027

SHIZUKA, K.; MAIE, N.; KAKINO, W.; TARUYA, H.; TANJI, H. Forecasting a 2methylisoborneol outbreak in a brackish lake. **Environmental Monitoring and Assessment**, v. 193, n. 7, 2021. Disponível em: https://doi.org/10.1007/s10661-021-09140-9

SMUCKER, N. J.; BEAULIEU, J. J.; NIETCH, C. T.; YOUNG, J. L. Increasingly severe cyanobacterial blooms and deep water hypoxia coincide with warming water temperatures in reservoirs. **Global Change Biology**, v. 27, n. 11, p. 2507–2519, 2021. Disponível em: https://doi.org/10.1111/gcb.15618

SU, J.; WANG, X.; ZHAO, S.; CHEN, B.; LI, C.; YANG, Z. A Structurally Simplified Hybrid Model of Genetic Algorithm and Support Vector Machine for Prediction of Chlorophyll a in Reservoirs. **Water**, v. 7, n. 4, p. 1610–1627, 2015. Disponível em: https://doi.org/10.3390/w7041610

SU, M.; ZHU, Y.; JIA, Z.; LIU, T.; YU, J.; BURCH, M.; YANG, M. Identification of MIB producers and odor risk assessment using routine data: A case study of an estuary drinking water reservoir. **Water Research**, v. 192, p. 116848, 2021. Disponível em: https://doi.org/10.1016/j.watres.2021.116848

SUAREZ, E. L.; DE VENTURA, L.; STÖCKLI, A.; ORDÓÑEZ, C.; THOMAS, M. K.; IBELINGS, B. W.; MCGINNIS, D. F. The emergence and dominance of Planktothrix rubescens as an hypolimnetic cyanobacterium in response to re-oligotrophication of a deep peri-alpine lake. **Limnology and Oceanography**, v. 68, n. 6, p. 1346–1359, 2023. Disponível em: https://doi.org/10.1002/lno.12351

WEI, J.; LI, Q.; LIU, W.; ZHANG, S.; XU, H.; PEI, H. Changes of phytoplankton and water environment in a highly urbanized subtropical lake during the past ten years. **Science of the Total Environment**, v. 879, n. March, p. 162985, 2023. Disponível em: https://doi.org/10.1016/j.scitotenv.2023.162985

WHEELER, S. M.; MORRISSEY, L. A.; LEVINE, S. N.; LIVINGSTON, G. P.; VINCENT, W. F. Mapping cyanobacterial blooms in Lake Champlain’s Missisquoi Bay using QuickBird and MERIS satellite data. **Journal of Great Lakes Research**, v. 38, n. SUPPL. 1, p. 68–75, 2012. Disponível em: https://doi.org/10.1016/j.jglr.2011.06.009

WILLIAMS, B. J.; COLE, B. Mining monitored data for decision-making with a Bayesian network model. **Ecological Modelling**, v. 249, p. 26–36, 2013. Disponível em: https://doi.org/10.1016/j.ecolmodel.2012.07.008

WU, T.; ZHU, G.; WANG, Z.; ZHU, M.; XU, H. Seasonal dynamics of odor compounds concentration driven by phytoplankton succession in a subtropical drinking water reservoir, southeast China. **Journal of Hazardous Materials**, v. 425, n. August 2021, p. 128056, 2022. Disponível em: https://doi.org/10.1016/j.jhazmat.2021.128056

XIAO, X.; HE, J.; HUANG, H.; MILLER, T. R.; CHRISTAKOS, G.; REICHWALDT, E. S.; GHADOUANI, A.; LIN, S.; XU, X.; SHI, J. A novel single-parameter approach for forecasting algal blooms. **Water Research**, v. 108, p. 222–231, 2017. Disponível em: https://doi.org/10.1016/j.watres.2016.10.076

XIE, Z.; LOU, I.; UNG, W. K.; MOK, K. M. Freshwater algal bloom prediction by support vector machine in Macau storage reservoirs. **Mathematical Problems in Engineering**, v. 2012, 2012. Disponível em: https://doi.org/10.1155/2012/397473

XU, H.; PAERL, H. W.; QIN, B.; ZHU, G.; HALL, N. S.; WU, Y. Determining critical nutrient thresholds needed to control harmful cyanobacterial blooms in eutrophic Lake Taihu, China. **Environmental Science and Technology**, v. 49, n. 2, p. 1051–1059, 2015. Disponível em: https://doi.org/10.1021/es503744q

ZHANG, J. *et al.* Ancient DNA reveals potentially toxic cyanobacteria increasing with climate change. **Water Research**, v. 229, n. September 2022, p. 119435, 2023. Disponível em: https://doi.org/10.1016/j.watres.2022.119435

ZHANG, S.; WANG, W.; CHANG, J. The contribution of cyanobacteria bloom decline to phosphorus in water column of Dianchi Lake, China. **Polish Journal of Environmental Studies**, v. 28, n. 5, p. 3513–3520, 2019. Disponível em: https://doi.org/10.15244/pjoes/94217

ZHANG, W.; LIU, Z. Daily water quality evaluation of reservoir and cyanobacteria pollution index calculation. **Water Science and Technology: Water Supply**, v. 21, n. 2, p. 836–847, 2021. Disponível em: https://doi.org/10.2166/ws.2020.350

ZHANG, X.; RECKNAGEL, F.; CHEN, Q.; CAO, H.; LI, R. Spatially-explicit modelling and forecasting of cyanobacteria growth in Lake Taihu by evolutionary computation. **Ecological Modelling**, v. 306, p. 216–225, 2015. Disponível em: https://doi.org/10.1016/j.ecolmodel.2014.05.013

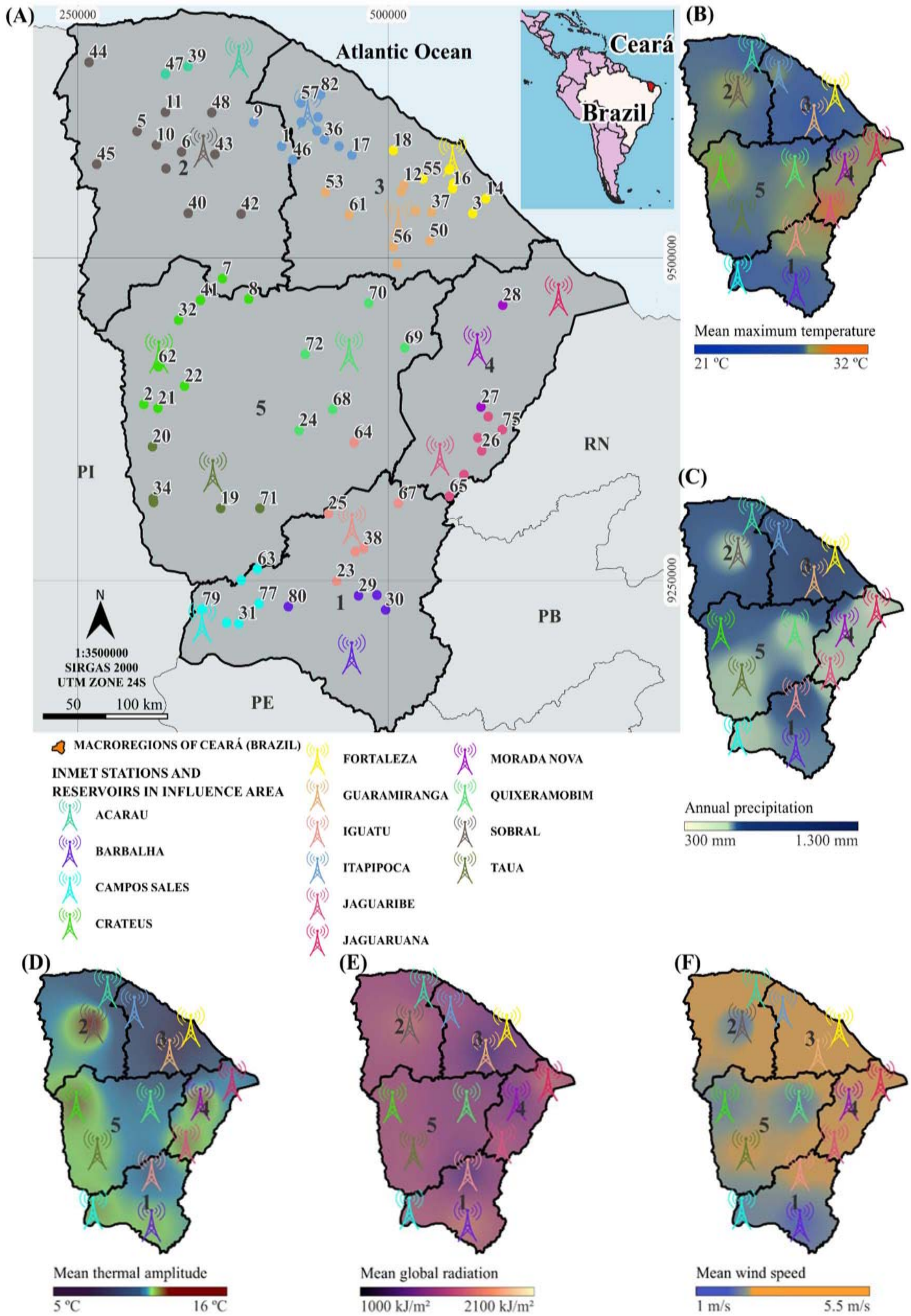
ZONG, J. M.; WANG, X. X.; ZHONG, Q. Y.; XIAO, X. M.; MA, J.; ZHAO, B. Increasing outbreak of cyanobacterial blooms in large lakes and reservoirs under pressures from climate change and anthropogenic interferences in the middle-lower Yangtze River basin. **Remote Sensing**, v. 11, n. 15, p. 1–21, 2019. Disponível em: https://doi.org/10.3390/rs11151754

APPENDIX B – SUPPLEMENTARY MATERIALS TO CHAPTER 3

SECTION 1: FIELD OF STUDY AND ENVIRONMENTAL GRADIENT

Based on the location of these stations, the 82 reservoirs analyzed were classified according to their geographic proximity, using the NNJoin plugin in QGIS (Figure 1 and 2). Climatic data were used to evaluate the environmental gradient based on variables such as Maximum Temperature (Figure 3), Thermal Amplitude (Figure 4), Global Radiation (Figure 5), Wind Speed (Figure 6), and Precipitation (Figure 7), obtained from the National Institute of Meteorology (INMET, 2023). Of the 16 meteorological stations in Ceará, 14 were selected, as only those with continuous data from 2009 to 2020 were considered. The Maximum Temperature, Global Radiation, and Wind Speed values were calculated as monthly and overall (annual) means. For Thermal Amplitude, the daily difference between Maximum Temperature and Minimum Temperature was considered, and these daily variations were then summarized into monthly and annual means. Precipitation was recorded as the total rainfall height for each period (monthly or annual).

Figure 21 - : Location of Ceará (Brazil). In (A) the macro-regions (Region 1 to 5) and the location of thPe 82 reservoirs studied in their respective areas of influence of the INMET meteorological posts are shown. Mean ranges of the gradients of maximum temperature



Source: prepared by author

- **DEFINITION OF THE INFLUENCE AREAS OF INMET STATIONS**

Table A – List of reservoir and definition of the influence areas of INMET stations

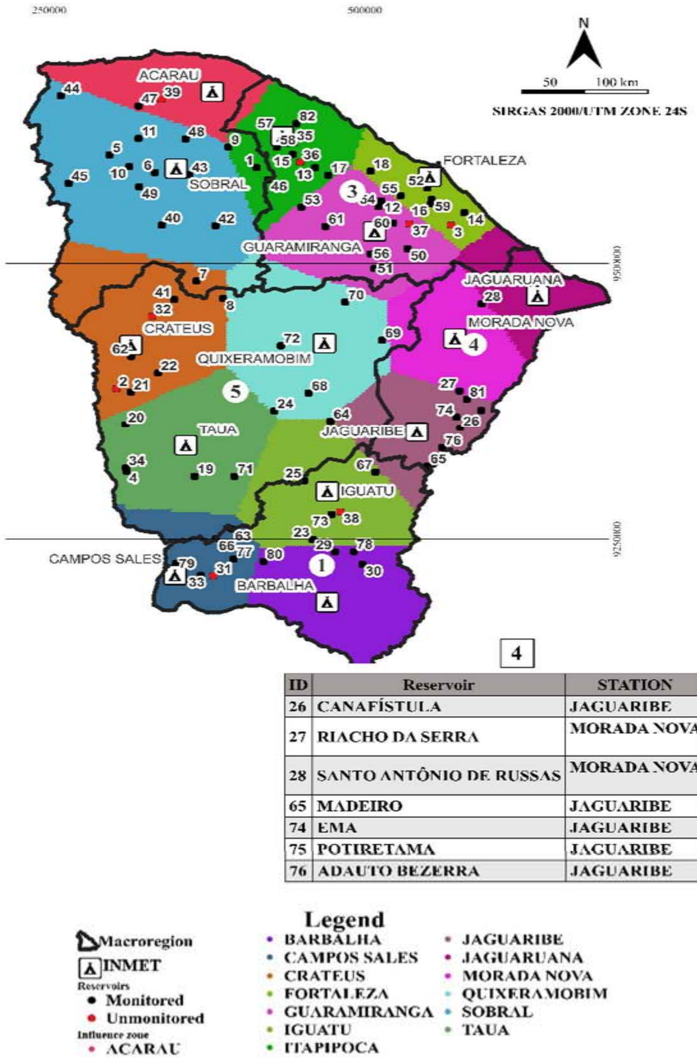
Code	Reservoir	INMET Staton	Macroregion	Macroregion Code	Code	Reservoir	INMET station	Macroregion	Macroregion Code
1	MISSI	ITAPIPOCA	Noroeste Cearense	2	42	EDSON QUEIROZ	SOBRAL	Noroeste Cearense	2
2	RODRIGO	CRATEUS	Sertões Cearenses	5	43	FORQUILHA	SOBRAL	Noroeste Cearense	2
3	CORIOLANO ALVES DE SOUSA BRITO	FORTALEZA	Metrololitana-Norte	3	44	ITAÚNA	SOBRAL	Noroeste Cearense	2
4	FACUNDO	TAUA	Sertões Cearenses	5	45	JABURU I	SOBRAL	Noroeste Cearense	2
5	ANGICOS	SOBRAL	Noroeste Cearense	2	46	JERIMUM	ITAPIPOCA	Metrololitana-Norte	3
6	AYRES DE SOUZA	SOBRAL	Noroeste Cearense	2	47	MARTINÓPOLE	ACARAU	Noroeste Cearense	2
7	CARMINA	CRATEUS	Noroeste Cearense	2	48	ACARAÚ MIRIM	SOBRAL	Noroeste Cearense	2
8	MONSENHOR TABOSA	CRATEUS	Sertões Cearenses	5	49	TAQUARA	SOBRAL	Noroeste Cearense	2
9	SÃO PEDRO TIMBAÚBA	ITAPIPOCA	Noroeste Cearense	2	50	ARACOIABA	GUARAMIRANGA	Metrololitana-Norte	3
10	TRAPIÁ III	SOBRAL	Noroeste Cearense	2	51	CASTRO	GUARAMIRANGA	Metrololitana-Norte	3
11	VÁRZEA DA VOLTA	SOBRAL	Noroeste Cearense	2	52	GAVIÃO	FORTALEZA	Metrololitana-Norte	3
12	AMANARY	GUARAMIRANGA	Metrololitana-Norte	3	53	GENERAL SAMPAIO	GUARAMIRANGA	Metrololitana-Norte	3
13	CAXITORÉ	ITAPIPOCA	Metrololitana-Norte	3	54	ITAPEBUSSU	GUARAMIRANGA	Metrololitana-Norte	3
14	MALCOZINHADO	FORTALEZA	Metrololitana-Norte	3	55	PENEDO	FORTALEZA	Metrololitana-Norte	3
15	MUNDAÚ	ITAPIPOCA	Metrololitana-Norte	3	56	PESQUEIRO	GUARAMIRANGA	Metrololitana-Norte	3
16	PACOTI	FORTALEZA	Metrololitana-Norte	3	57	POÇO VERDE	ITAPIPOCA	Metrololitana-Norte	3
17	PENTECOSTE	ITAPIPOCA	Metrololitana-Norte	3	58	QUANDÚ	ITAPIPOCA	Metrololitana-Norte	3
18	SITIOS NOVOS	FORTALEZA	Metrololitana-Norte	3	59	RIACHÃO	FORTALEZA	Metrololitana-Norte	3
19	ARNEIROZ II	TAUA	Sertões Cearenses	5	60	ACARAPE DO MEIO	GUARAMIRANGA	Metrololitana-Norte	3
20	COLINA	TAUA	Sertões Cearenses	5	61	SÃO DOMINGOS	GUARAMIRANGA	Metrololitana-Norte	3
21	FLOR DO CAMPO	CRATEUS	Sertões Cearenses	5	62	CARNAUBAL	CRATEUS	Sertões Cearenses	5
22	JABURU II	CRATEUS	Sertões Cearenses	5	63	DO CORONEL	CAMPOS SALES	Sertões Cearenses	5
23	OLHO D'ÁGUA	IGUATU	Centro-Sul Cearense	1	64	JENIPAPEIRO	IGUATU	Sertões Cearenses	5
24	SERAFIM DIAS	QUIXERAMOBIM	Sertões Cearenses	5	65	MADEIRO	JAGUARIBE	Jaguaribe	4
25	TRUSSU	IGUATU	Centro-Sul Cearense	1	66	MAMOEIRO	CAMPOS SALES	Sertões Cearenses	5

26	CANAFÍSTULA	JAGUARIBE	Jaguaribe	4	67	ORÓS	IGUATU	Centro-Sul Cearense	1
27	RIACHO DA SERRA	MORADA NOVA	Jaguaribe	4	68	PATU	QUIXERAMOBIM	Sertões Cearenses	5
28	SANTO ANTÔNIO DE RUSSAS	MORADA NOVA	Jaguaribe	4	69	PEDRAS BRANCAS	QUIXERAMOBIM	Sertões Cearenses	5
29	JUNCO	BARBALHA	Centro-Sul Cearense	1	70	POMPEU SOBRINHO	QUIXERAMOBIM	Sertões Cearenses	5
30	CACHOEIRA	BARBALHA	Centro-Sul Cearense	1	71	RIVALDO DE CARVALHO	TAUA	Sertões Cearenses	5
31	BELO HORIZONTE	CAMPOS SALES	Centro-Sul Cearense	1	72	SÃO JOSÉ I	QUIXERAMOBIM	Sertões Cearenses	5
32	JOÃO JOSÉ DE CASTRO	CRATEUS	Sertões Cearenses	5	73	UBALDINHO	IGUATU	Centro-Sul Cearense	1
33	JOÃO LUÍS	CAMPOS SALES	Centro-Sul Cearense	1	74	EMA	JAGUARIBE	Jaguaribe	4
34	PUIÚ	TAUA	Sertões Cearenses	5	75	POTIRETAMA	JAGUARIBE	Jaguaribe	4
35	IPÚ MAZAGÃO	ITAPIPOCA	Metrololitana-Norte	3	76	ADAUTO BEZERRA	JAGUARIBE	Jaguaribe	4
36	SÃO JOAQUIM	ITAPIPOCA	Metrololitana-Norte	3	77	CANOAS	CAMPOS SALES	Centro-Sul Cearense	1
37	BOQUEIRÃO	GUARAMIRANGA	Metrololitana-Norte	3	78	ROSÁRIO	BARBALHA	Centro-Sul Cearense	1
38	UBALDINHO (RIACHO SÃO MIGUEL)	IGUATU	Centro-Sul Cearense	1	79	POÇO DA PEDRA	CAMPOS SALES	Centro-Sul Cearense	1
39	SEBASTIÃO DE ABREU	ACARAU	Noroeste Cearense	2	80	VALÉRIO	BARBALHA	Centro-Sul Cearense	1
40	ARARAS	SOBRAL	Noroeste Cearense	2	81	FIGUEIREDO	JAGUARIBE	Jaguaribe	4
41	CARÃO	CRATEUS	Sertões Cearenses	5	82	GAMELEIRA	ITAPIPOCA	Metrololitana-Norte	3

Figure 22 - Location and definition of the areas of influence of the 14 official INMET meteorological stations and the list of water supply reservoirs (monitored and unmonitored by COGERH) within these areas.

2

ID	Reservoir	STATION
1	MISSI	ITAPIPOCA
5	ANGICOS	SOBRAL
6	AYRES DE SOUZA	SOBRAL
7	CARMINA	CRATEUS
9	SÃO PEDRO TIMBAÚBA	ITAPIPOCA
10	TRAPIÁ III	SOBRAL
11	VÁRZEA DA VOLTA	SOBRAL
39	SEBASTIÃO DE ABREU	ACARAU
40	ARARAS	SOBRAL
42	EDSON QUEIROZ	SOBRAL
43	FORQUILHA	SOBRAL
44	ITAÚNA	SOBRAL
45	JABURU I	SOBRAL
47	MARTINÓPOLE	ACARAU



3

ID	Reservoir	STATION
3	CORIOLANO ALVES DE SOUSA BRITO	FORTALEZA
12	AMANARY	GUARAMIRANGA
13	CAXITORÉ	ITAPIPOCA
14	MALCOZINHADO	FORTALEZA
15	MUNDAÚ	ITAPIPOCA
16	PACOTI	FORTALEZA
17	PENTECOSTE	ITAPIPOCA
18	SITIOS NOVOS	FORTALEZA
35	IPÚ MAZAGÃO	ITAPIPOCA
36	SÃO JOAQUIM	ITAPIPOCA
37	BOQUEIRÃO	GUARAMIRANGA
46	JERIMUM	ITAPIPOCA
	ARACOIABA	GUARAMIRANGA
51	CASTRO	GUARAMIRANGA
52	GAVIÃO	FORTALEZA
53	GENERAL SAMPAIO	GUARAMIRANGA
54	ITAPEBUSSU	GUARAMIRANGA
55	PENEDO	FORTALEZA
56	PESQUEIRO	GUARAMIRANGA

5

ID	Reservoir	STATION
2	RODRIGO	CRATEUS
4	FACUNDO	TAUA
8	MONSENHOR TABOSA	CRATEUS
19	ARNEIROZ II	TAUA
20	COLINA	TAUA
21	FLOR DO CAMPO	CRATEUS
22	JABURU II	CRATEUS
24	SERAFIM DIAS	QUIXERAMOBIM
32	JOÃO JOSÉ DE CASTRO	CRATEUS
34	PUIÚ	TAUA
41	CARÃO	CRATEUS
62	CARNAUBAL	CRATEUS
63	DO CORONEL	CAMPOS SALES
64	JENIPAPEIRO	IGUATU
66	MAMOEIRO	CAMPOS SALES
68	PATU	QUIXERAMOBIM
69	PEDRAS BRANCAS	QUIXERAMOBIM
70	POMPEU SOBRINHO	QUIXERAMOBIM
71	RIVALDO DE CARVALHO	TAUA
72	SÃO JOSÉ I	QUIXERAMOBIM

4

ID	Reservoir	STATION
26	CANAFÍSTULA	JAGUARIBE
27	RIACHO DA SERRA	MORADA NOVA
28	SANTO ANTÔNIO DE RUSSAS	MORADA NOVA
65	MADEIRO	JAGUARIBE
74	EMA	JAGUARIBE
75	POTIRETAMA	JAGUARIBE
76	ADAUTO BEZERRA	JAGUARIBE

1

ID	Reservoir	STATION
23	OLHO D'ÁGUA	IGUATU
25	TRUSSU	IGUATU
29	JUNCO	BARBALHA
30	CACHOEIRA	BARBALHA
31	BELO HORIZONTE	CAMPOS SALES
33	JOÃO LUÍS	CAMPOS SALES
38	UBALDINHO (RIACHO SÃO MIGUEL)	IGUATU
67	ORÓS	IGUATU
73	UBALDINHO	IGUATU
77	CANOAS	CAMPOS SALES
78	ROSÁRIO	BARBALHA
79	POÇO DA PEDRA	CAMPOS SALES
80	VALÉRIO	BARBALHA

Figure 23 - Mean Maximum Temperature Gradient: Annual and Monthly Mean of Maximum Temperature in Ceará (Brazil).

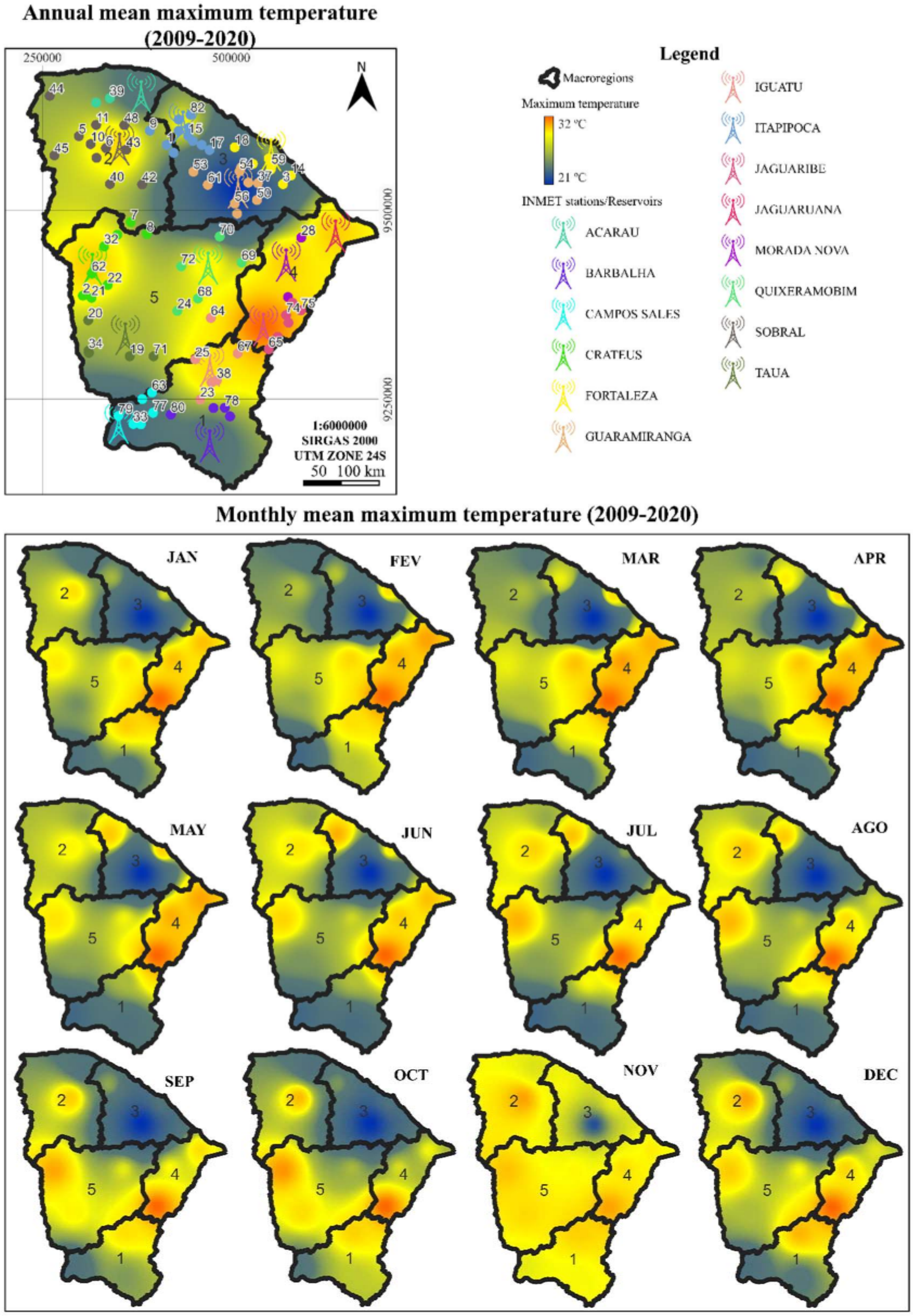


Figure 24 - Thermal Amplitude Gradient: Annual and Monthly Mean of Thermal Amplitude in Ceará (Brazil)

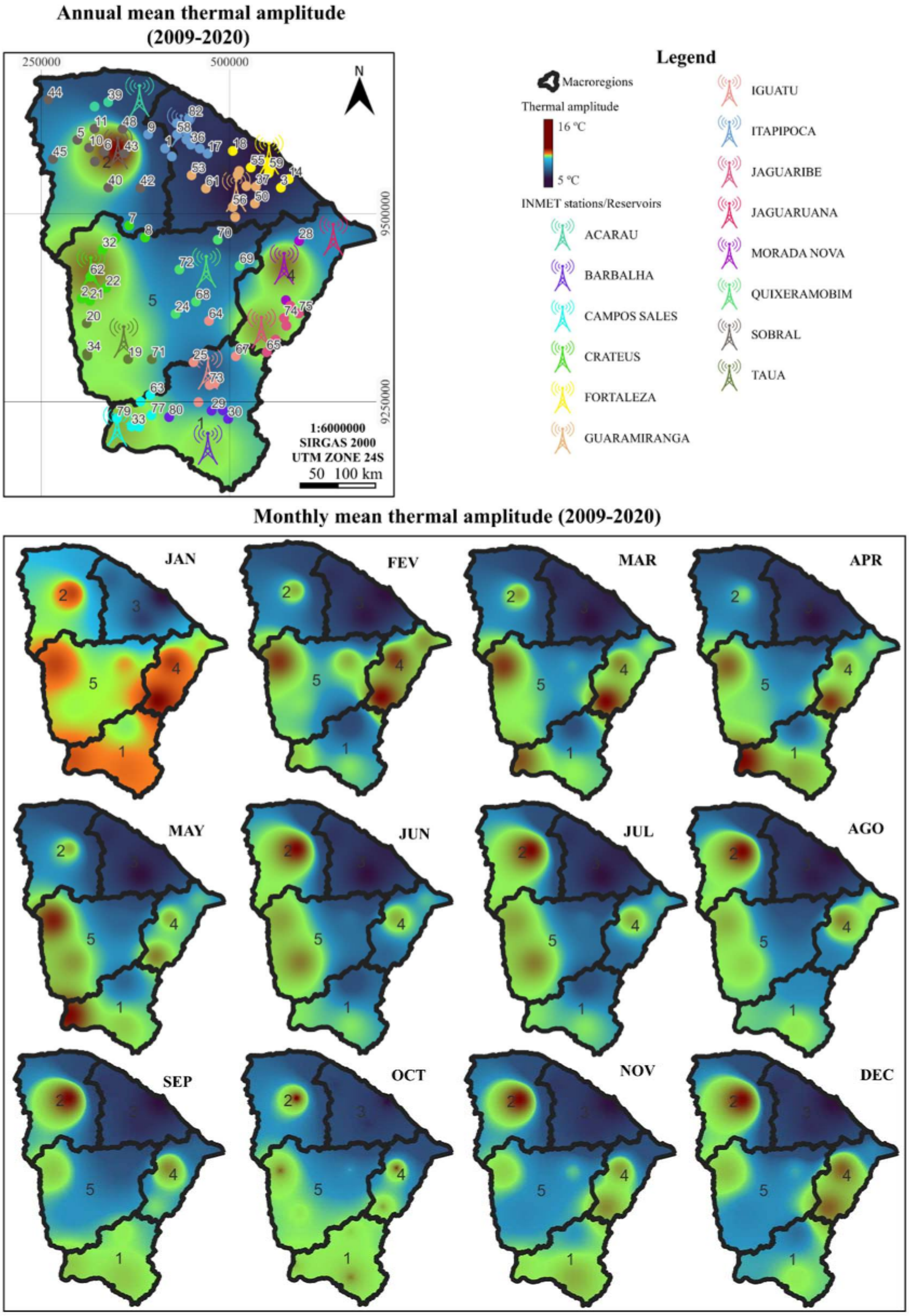


Figure 25 - Gradiente de Radiação Global: Média Anual e Mensal da Radiação Global no Ceará (Brasil)

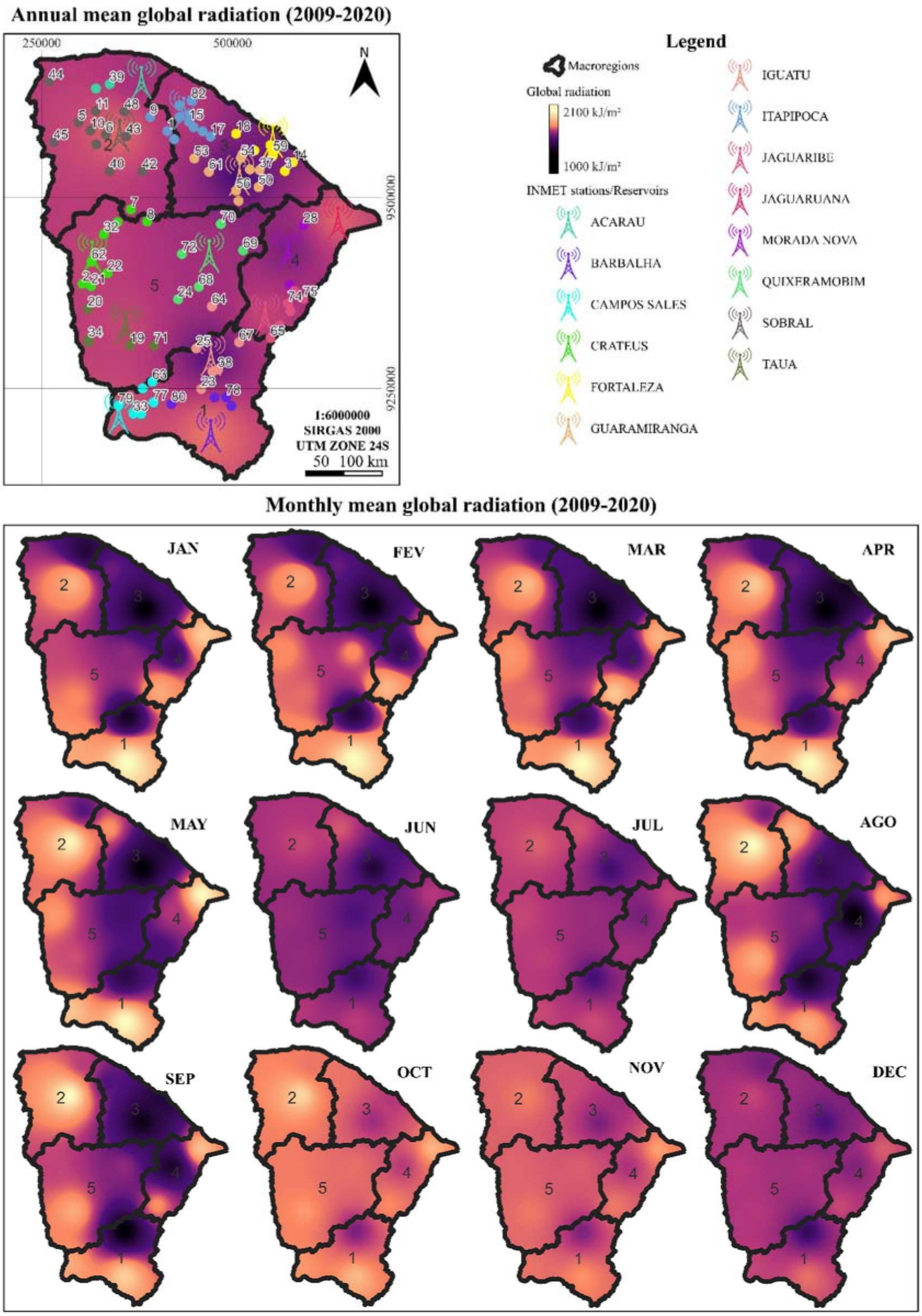


Figure 26 -Wind Speed Gradient: Annual and Monthly Mean of Wind Speed in Ceará (Brazil)

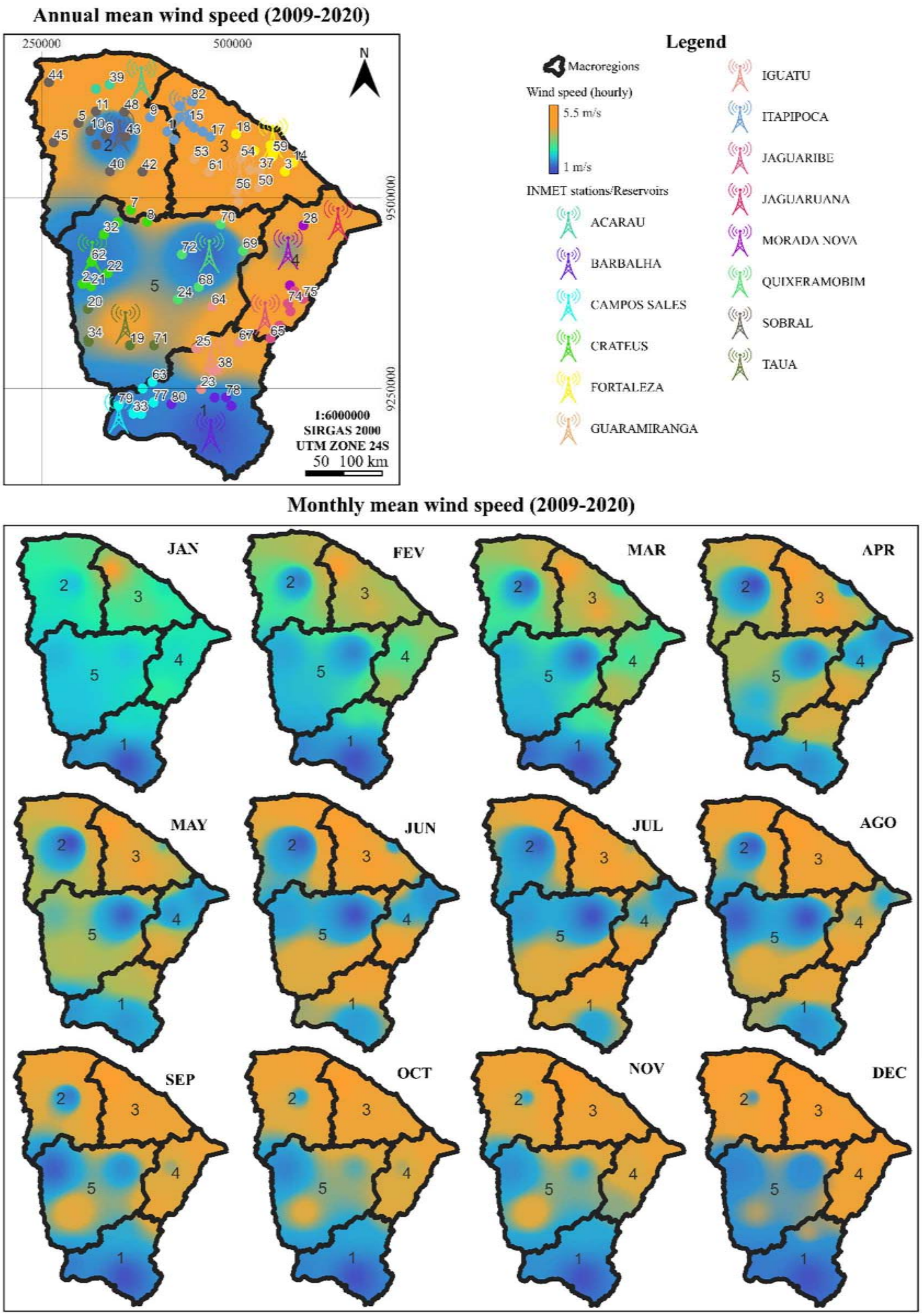
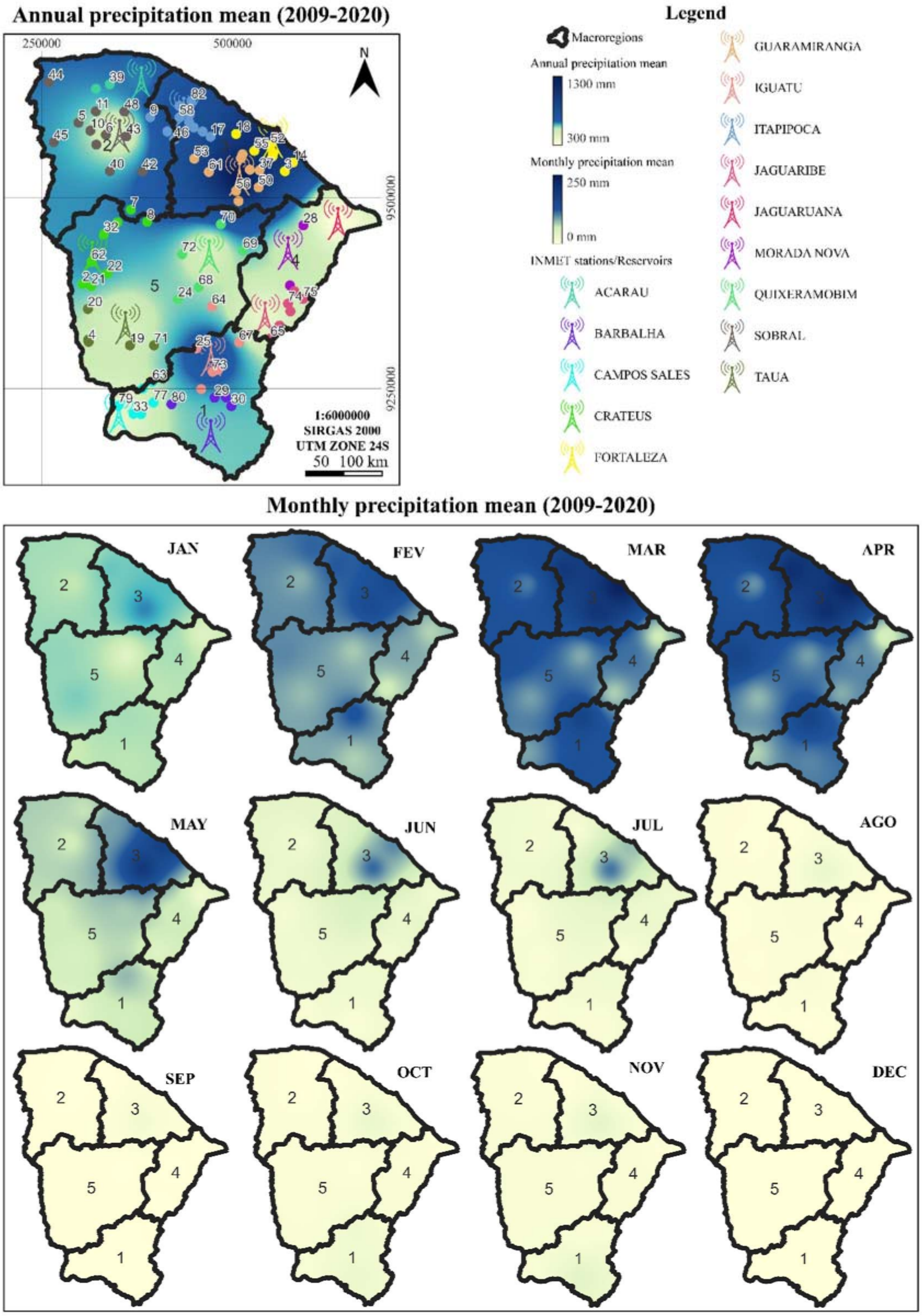


Figure 27 - Precipitation Gradient: Annual and Monthly Total of Precipitation in Ceará (Brazil)



SECTION 2: RESERVOIR DATA, TREATMENT SYSTEMS, SAMPLING PROTOCOL, AND PRIMARY ORGANISMS IDENTIFIED

RESERVOIR DATA AND TREATMENT SYSTEMS

Table B - Details of the water supply system and reservoir

Loc	Water supply system	Reservoir	Unidad	Techno.	consumer units	inhab/ consumer units	Dammed river	Vol (m^3)	Lat.	Long.	Flow (m^3/h)	Crest (m)	Crest width (m)	maximum depth(m)
ANTONINA DO NORTE	SI ANTONINA DO NORTE	MAMOEIRO	ETA 01	DF (Upflow)	2.583	3,30	Rio Conceição	20.490.000	6º46'16"	40º03'53"	-	348	6,5	27,5
ARNEIROZ	ARNEIROZ	ARNEIROZ II	ETA 01	DF (Upflow)	1248	3,5								
COREAÚ	UBAÚNA	TRAPIÁ III	ETA 01	DF (Upflow)	1.480	3,60	Riacho Cajazeiras	5.510.000	3º44'16"	40º41'07"	32	-	533,59	5,2
AURORA	SI AURORA	CACHOEIRA	ETA 01 - AURORA	DF (Upflow)	4.492,00	3,30	Rio Caiçara	3.433.000	6º56'29"	38º58'03"	-	273	8,4	25,5
MIRAÍMA	BROTAS	MISSÍ	ETA 01 - BROTAS	DF (Upflow)	1.139	3,80	Rio Quandú	-	3º30'57"	39º51'58"	-	706	6	17,3
CASCAVEL	CASCAVEL	MALCOZINHADO	ETA 01 - CASCAVEL	DF (Upflow)	13.278	3,50	Rio Malcozinhado	37.840.000	4º08'23"	38º14'04"	233	755	6	15,11
CEDRO	SI CEDRO	UBALDINHO	ETA 01 - CEDRO	DF (Upflow)	5.211	3,30	SALGADO	31.800	6º35'58"	39º03'22"	130	475	6	17,9
ITAPIPOCA	ITAPIPOCA	POÇO VERDE	ETA 01 - ITAPIPOCA	DF (Upflow)	24.720	3,70	Riacho Sororó	12.430.000	3º30'43"	39º35'03"	-	711,47	5	
QUITAIÚS	QUITAIÚS	JOÃO LUIS	ETA 01 - JOÃO LUIS	Conventional	729	3,4								
LAVRAS DA MANGABEIRA	SI QUITAIÚS	ROSÁRIO	ETA 01 - QUITAIÚS	Conventional	4.962	3,40	Rio Rosário	47.220.000	6º53'12"	39º05'24"	151	670	6	20,8
LAVRAS DA MANGABEIRA	SI QUITAIÚS	ROSÁRIO	ETA 02	DF (Upflow)	762	3,40	Rio Rosário	47.220.000	6º45'36"	38º58'05"	151	670	6	20,8
MIRAÍMA	BROTAS	MISSÍ	ETA 02	DF (Upflow)	1.139	3,80	Rio Quandú	-	3º34'12"	39º46'02"	73	706	6	17,3
AURORA	SI AURORA	CACHOEIRA	ETA 02 - AURORA	Conventional	4.492	3,30	Rio Caiçara	3.433.000	6º56'32.75"	38º58'30"	100	273	8,4	25,5
ITAPIPOCA	ITAPIPOCA	GAMELEIRA	ETA 02 - ITAPIPOCA	Conventional	24.720	3,6								

ITAPIPOCA	ITAPIPOCA	IPU MAZAÇÃO	ETA 02 - ITAPIPOCA	Conventional	24.720,01	3,6								
ITAPIPOCA	ITAPIPOCA	POÇO VERDE	ETA 02 - ITAPIPOCA	Conventional	24.720	3,70	Riacho Sororó	12.430.000	3º29'22"	39º36'25"	434	711,47	5	
ITAPIPOCA	ITAPIPOCA	QUANDU	ETA 02 - ITAPIPOCA	Conventional	24.720	3,6								
CASCAVEL	CASCAVEL	MALCOZINHADO	ETA 02 - MALCOZINHADO	DF (Double)	13.278	3,50	Rio Malcozinhado	37.840.000	4º06'25"	38º17'23"	432	755	6	15,11
ANTONINA DO NORTE	ANTONINA DO NORTE	DO CORONEL	ETA 02 - MAMOEIRO	DF (Double)	2.583	3,30								
ANTONINA DO NORTE	SI ANTONINA DO NORTE	MAMOEIRO	ETA 02 - MAMOEIRO	DF (Double)	2.583	3,30	Rio Conceição	20.490.000	6º46'16"	40º03'53"	-	348	6,5	27,5
CATARINA	SI CATARINA	ARNEIROZ II	ETA 02 - SÃO GONÇALO	DF (Upflow)	202	3,5								
CATARINA	SI CATARINA	RIVALDO DE CARVALHO	ETA 02 - SÃO GONÇALO	Conventional	2581	3,50	Riacho Condado	19.520.000	6º15'48"	39º55'39"	35	413,55	1,5	18,23
COREAÚ	UBAÚNA	TRAPIÁ III	ETA 02 - UBAÚNA	DF (Upflow)	-	-	Riacho Cajazeiras	-	-	-	-	-	533,59	5,2
CEDRO	SI CEDRO	UBALDINHO (RIACHO SÃO MIGUEL)	ETA 02 - VÁRZEA DA CONCEIÇÃO	DF (Upflow)	5.211	3,30	SALGADO	31.800	6º32'47"	39º02'35"	-	475	6	17,9
ACOPIARA	ACOPIARA	TRUSSU	ETA ACOPIARA	DF (Upflow)	9.115	3,20	ALTO JAGUARIBE	268.800.000	6º05'17"	39º27'17"	190	1247,9	8	36,3
ARARIPE	SI ALAGOINHA	JOÃO LUÍS	ETA ALAGOINHA	DF (Upflow)	441	3,60	Bela Cruz	5.500.000	7º05'08"	40º10'23"	-	0	0	0
JENIPAPO	ALCÂNTARAS	JENIPAPO	ETA ALCÂNTARAS	DF (Upflow)	1.566	3,50	Jenipapo	3.500.000	3º 33' 43"	40º 26' 55"	53,02	100	x	x
ALTANEIRA	ALTANEIRA	VALÉRIO	ETA ALTANEIRA	DF (Upflow)	2.382	3,20	ALTO JAGUARIBE	1.860.000	6º59'06"	38º43'13"	93	180	6	19,9
ALTO SANTO	ALTO SANTO	RIACHO DA SERRA	ETA ALTO SANTO	Conventional	1.807	3,40	Rch. da Serra	-	5º31'05"	38º16'18"	90	-	-	-
MARANGUAPE	SI AMANARI	AMANARI	ETA AMANARI	DF (Upflow)	1.492	3,80	RIACHO POCINHOS	11.011.000	4º01'23"	38º53'18"	49	435	3	19,1
ARACOIABA	SI ARACOIABA	ARACOIABA	ETA ARACOIABA	DF (Upflow)	3.835	3,50	Rio Aracoiaba	162.000.000	4º24'04"	38º43'01"	142	2	8	35
ASSARÉ	ASSARÉ	CANOAS	ETA ASSARÉ	DF (Upflow)	5.031	3,30	Riacho São Gonçalo	69.250.000	6º53'09"	39º52'05"	210	116,5	6	50,8
RUSSAS	BONHU	SANTO ANTÔNIO DE RUSSAS	ETA BONHU	DF (Upflow)	137	3,30	BAIXO JAGUARIBE	-	4º51'19"	38º09'51"	-	620	4	14
CAMPOS SALES	CAMPOS SALES	POÇO DA PEDRA	ETA CAMPOS SALES	DF (Downflow)	6.800	3,40	Rio Conceição	52.000.000	6º59'00"	40º21'16"	144	797	7	31,4

CAPISTRANO	CAPISTRANO	PESQUEIRO	ETA CAPISTRANO	DF (Upflow)	2.454	3,50	Riacho Lagoa Nova	-	4º28'00"	38º54'18"	80	331	6	24,02
CARIDADE	SI CARIDADE	GENERAL SAMPAIO	ETA CARIDADE	DF (Upflow)	2.013	3,80	Rio Curu	322.200.000	4º13'33"	39º16'59"	40	222	8	36,6
CARIDADE	SI CARIDADE	SÃO DOMINGOS	ETA CARIDADE	DF (Upflow)	2013	3,8								
CAUCAIA	SI CATUANA	SITIOS NOVOS	ETA CATUANA	DF (Upflow)	891	3,60	METROPOLIT ANA	-	3º40'57"	38º55'11"	144	1818	6,5	21,5
CATUNDA	CATUNDA	CARMINA	ETA CATUNDA	Conventional	2.354	3,50	Riacho dos Macacos	13.480.000	4º39'41"	40º11'57"	49	770	6	19
CHAVAL	SI CHAVAL	ITAÚNA	ETA CHAVAL	DF (Upflow)	2.962	3,90	Rio Timonha	72.580.000	3º08'56"	41º09'48"	-	436	6	18
CHORÓ	CHORÓ	POMPEU SOBRINHO	ETA CHORÓ	DF (Double)	1.266	3,70	Rio Choró	143.000.000	4º50'33"	39º08'29"	75	235	9,7	31
IRACEMA	EMA	EMA	ETA EMA	DF (Upflow)	563	3,40	Riacho Bonsucesso	1.039.000.000	5º46'30"	38º20'40"	80	341,5	4	15,2
FORQUILHA	FORQUILHA	FORQUILHA	ETA FORQUILHA	Conventional	6008	3,5								
FRECHEIRINHA	FRECHEIRINHA	ANGICOS	ETA FRECHEIRINHA	DF (Upflow)	4149	3,6								
PACATUBA	SI FORTALEZA	GAVIÃO	ETA GAVIÃO	DF (Downflow)	115.080	3,40	Rio Cocó	32.900	3º54'10"	38º33'28"	36.000	845,56	5,08	14,63
GENERAL SAMPAIO	GENERAL SAMPAIO	GENERAL SAMPAIO	ETA GENERAL SAMPAIO	DF (Upflow)	1.566	3,80	Rio Curu	322.200.000	4º03'12"	39º27'18"	70	222	8	36,6
GRANJEIRO	GRANJEIRO	JUNCO	ETA GRANJEIRO	DF (Upflow)	821	-	Rio Acarape	2.030.000	6º53'26"	39º13'05"	23	-	-	-
HORIZONTE	SI HORIZONTE	PACOTI	ETA HORIZONTE	DF (Upflow)	19.374	3,50	Rio Pacoti	-	4º04'32"	38º31'02"	-	278	6	20,8
INDEPENDÊNCIA	INDEPENDÊNCIA	JABURU II	ETA INDEPENDÊNCIA	DF (Upflow)	5.380	3,20	Rio Poti	106.000.000	5º22'55"	40º18'02"	110	1054	4	16,2
MASSAPÊ	SI IPAGUAÇU	ACARAPE DO MEIO	ETA IPAGUAÇU	DF (Upflow)	524	3,70	Riacho Acaraú Mirim	52.768.000	3º30'36"	40º16'44"	120	442	8	18
SOBRAL	SI IPUEIRAS	AYRES DE SOUSA	ETA IPUEIRAS	DF (Upflow)	158	3,70	Riacho Jaibaras	96.800.000	3º46'32"	40º30'48"	12	358	6	28,8
IRACEMA	IRACEMA	CANAFÍSTULA	ETA IRACEMA	DF (Upflow)	3.594	3,40	Foice	13.110.000	5º49'01"	38º18'17"	120	850	4	14,8
HIDROLÂNDIA	SI IRAJÁ	ARARAS	ETA IRAJÁ	DF (Double)	691	-	Rio Acaraú	859.530.000	4º 17' 25"	40º 28' 24"	120	2,6	8	38
IRAUÇUBA	IRAUÇUBA	JERIMUM	ETA IRAUÇUBA	DF (Upflow)	3.743	3,90								
IRAUÇUBA	IRAUÇUBA	MISSÍ	ETA IRAUÇUBA	DF (Upflow)	3.743	3,90	Rio Quandú	-	3º44'45"	39º46'27"	-	706	6	17,3
ITAITINGA	ITAITINGA	RIACHÃO	ETA ITAITINGA	DF (Upflow)	6712	3,50								
ITAITINGA	CARAPIÓ	RIACHÃO	ETA ITAITINGA	DF (Upflow)	513	3,50								

ITAITINGA	LAGOA DE DENTRO	RIACHÃO	ETA ITAITINGA	DF (Upflow)	335	3,50								

SAMPLING PROTOCOL

Collection points closest to the intake were used, following the specific routine of each treatment plant. Cell counts were conducted using an inverted microscope (Zeiss Axio, A1) with a Sedgewick-Rafter chamber. On average, fields were counted at multiple magnifications (200x and 400x, depending on the phytoplankton size) until at least 100 individuals or colonies/filaments (≥4 cells) were observed per sample, ensuring adequate phytoplankton diversity coverage. Phytoplankton was identified to the species level whenever possible; otherwise, identification was limited to the genus, family, etc. During the analyzed period (2009-2019), the protocol followed by the company was based on the (APHA-AWWA-WEF, 2005) and updated versions, as per legal requirements (BRAZIL, 2004, 2011, 2017, 2021). These regulations also recommend analysis frequency (weekly, monthly, or semiannual) depending on cyanobacteria density (Cell mL^{-1}) and mandate toxin analysis based on observed cell density. Enzyme-linked immunosorbent assays (ELISA) were used to quantify total microcystin (MC), saxitoxin (STX) and cylindrospermopsin (CYN) in raw water.

PRIMARY ORGANISMS IDENTIFIED

Table C - Lists of organisms from the phylum cyanobacteria (genus and species) identified in raw water collected in 82 reservoirs in Ceará between 2009-2020

Genus	Species	Genus	Species
Anabaena sp.	Anabaena spiroides	Lyngbya sp.	-
Anabaenopsis sp.	-	Merismopedia sp.	Merismopedia cf. glauca
Anagnostidinema sp.	Anagnostidinema amphibium		Merismopedia glauca
Anathece sp.	-		Merismopedia tenuissima
Aphanizomenon sp.	-	Microcystis sp.	Microcystis aeruginosa
Aphanocapsa sp.	Aphanocapsa cf. annulata		Microcystis botrys
	Aphanocapsa cf. holsatica		Microcystis brasiliense
	Aphanocapsa delicatissima		Microcystis cf. protocystis
	Aphanocapsa kordese		Microcystis panniformis
Aphanothece sp.	-		Microcystis protocystis
Arthrospira sp.	-		Microcystis weisenbergii
Bacularia sp.	-	Oscillatoria sp.	-
Cephalothrix sp.	-	Pannus sp.	Pannus cf. microcystiformisa
Chroococcus sp.	Chroococcus cf. turgidus	Phormidium sp.	-
Coelastrum sp.	Coelastrum cf. indicum	Planktolyngbya sp.	Planktolyngbya contorta
Coelomoron sp.	Coelomoron tropicale		Planktolyngbya limnetica
Coelosphaerium sp.	-	Planktothricoides sp.	Planktothricoides raciborskii
Cronbergia sp.	-	Planktothrix sp.	Planktothrix agardhii
Cuspidothrix sp.	Cuspidothrix issatschenkoi		Planktothrix isothrix
	Cuspidothrix tropicalis	Pseudanabaena sp.	Pseudanabaena cf. acicularis
Cyanodictyon sp.	-		Pseudanabaena cf. raphidioides

Raphidiopsis sp.	Raphidiopsis catemaco		Pseudanabaena galeata
	Raphidiopsis phillippinensis		Pseudanabaena mucicola
	Raphidiopsis raciborskii		Pseudanabaena catenata
Dolichospermum sp.	Dolichospermum circinale	Radiocystis sp.	-
	Dolichospermum spiroides	Romeria sp.	-
Epigloesphaera sp.	-	Snowella sp.	-
Eucapsis sp.	-	Sphaerocavum sp.	Sphaerocavum brasiliense
Geitlerinema sp.	Geitlerinema amphibium	Sphaerospermopsis sp.	Sphaerospermopsis aphanizomenoide
	Geitlerinema splendidum	Spirulina sp.	-
Jaagnema sp.	-	Synechocystis sp.	-
Gloeothece sp.	-	Trichodesmium sp.	Trichodesmium iwanoffianum
Komvophoron sp.	-	Trichodesmium sp.	-
Limnococcus sp.	-	Woronichinia sp.	-
Limnoraphis sp.	-		

The list of the main genera (species) belonging to other taxonomic levels identified in Brazil(SANT'ANNA, 2012):

- **Oscillatoriales**: *Geitlerinema sp. (Geitlerinema amphibium, Geitlerinema splendidum), Komvophoron sp. (Komvophoron cf. crassum, Komvophoron schmidlei) Leptolyngbya sp. (Leptolyngbya perelegans) Limnothrix sp., Oscillatoria sp. (Oscillatoria limosa, Oscillatoria perornata), Phormidium sp. (Phormidium aerugineo-caeruleum, Phormidium tergestinum), Planktolyngbya sp. (Planktolyngbya contorta, Planktolyngbya limnetica), Planktothrix sp. (Planktothrix agardhii, Planktothrix isothrix), Pseudanabaena sp. (Pseudanabaena catenata, Pseudanabaena galeata, Pseudanabaena mucicola), Romeria sp. (Romeria victoriae), Planktolyngbya sp. (Planktolyngbya contorta, Planktolyngbya limnetica) Spirulina sp. (Spirulina subsalsa);*
- **Chroococcales***: Microcystis (isolated) or (Microcystis aeruginosa, Microcystis wesenbergii, Microcystis botrys, Microcystis panniformis, Microcystis protocystis, Microcystis pulchella), Aphanocapsa sp. (Aphanocapsa delicatissima, Aphanocapsa elachista, Aphanocapsa incerta, Aphanocapsa holsatica, Aphanocapsa koordersii) Aphanothece sp. (Aphanothece zulanirae), Bacularia sp. (Bacularia gracilis), Chroococcus sp. (Chroococcus dispersus), Coelomoron sp. (Coelomoron tropicalis), Coelosphaerium sp. (Coelosphaerium evidentir-marginatum), Cyanodictyon sp. (Cyanodictyon plantonicum), Eucapsis sp. (Eucapsis dense), Epigloeosphaera sp. (Epigloeosphaera brasílica), Gomphosphaeria sp. (Gomphosphaeria aponina), Lemmermanniella sp. (Lemmermanniella obesa), Merismopedia sp. (Merismopedia convoluta,*

Merismopedia tenuissima, Merismopedia glauca), Radiocystis sp. (Radiocystis fernandoi), Rhabdoderma sp. (Rhabdoderma sancti-pauli), Rhabdoderma sp. (Rhabdoderma lineare), Snowella sp. (Snowella lacustrine), Sphaerocavum sp. (Sphaerocavum brasiliense), Synechococcus sp. (Synechococcus nidulans), Synechocystis sp. (Synechocystis aquatilis), Woronichinias sp. (Woronichinia naegeliana).

- ***Nostocaceae**: Aphanizomenon sp. (Aphanizomenon gracile), Cuspidothrix sp. (Cuspidothrix issatschenkoi, Cuspidothrix tropicalis), Raphidiopsis sp. (Raphidiopsis raciborskii, Nostoc poolale) Dolichospermum sp. (Dolichospermum circinalis, Dolichospermum crassum, Dolichospermum flos-aquae, Dolichospermum mendotae, Dolichospermum mucosum, Dolichospermum nygaardii, Dolichospermum planctonicum, Dolichospermum solitarium, Dolichospermum spiroides) Nostoc sp. (Nostoc swimming pool), Sphaerospermopsis sp. (Sphaerospermopsis torques-reginae)*

SECTION 3: RAREFACTION CURVES

PLATEAU FORMATION: 1 NEW ORGANISMS FOR EVERY 100 INDIVIDUALS OR OBSERVATIONS)

- 55 reservoirs (67,1 %)

Figure 1 - Reservoirs: Sítios Novos, Valério e Orós

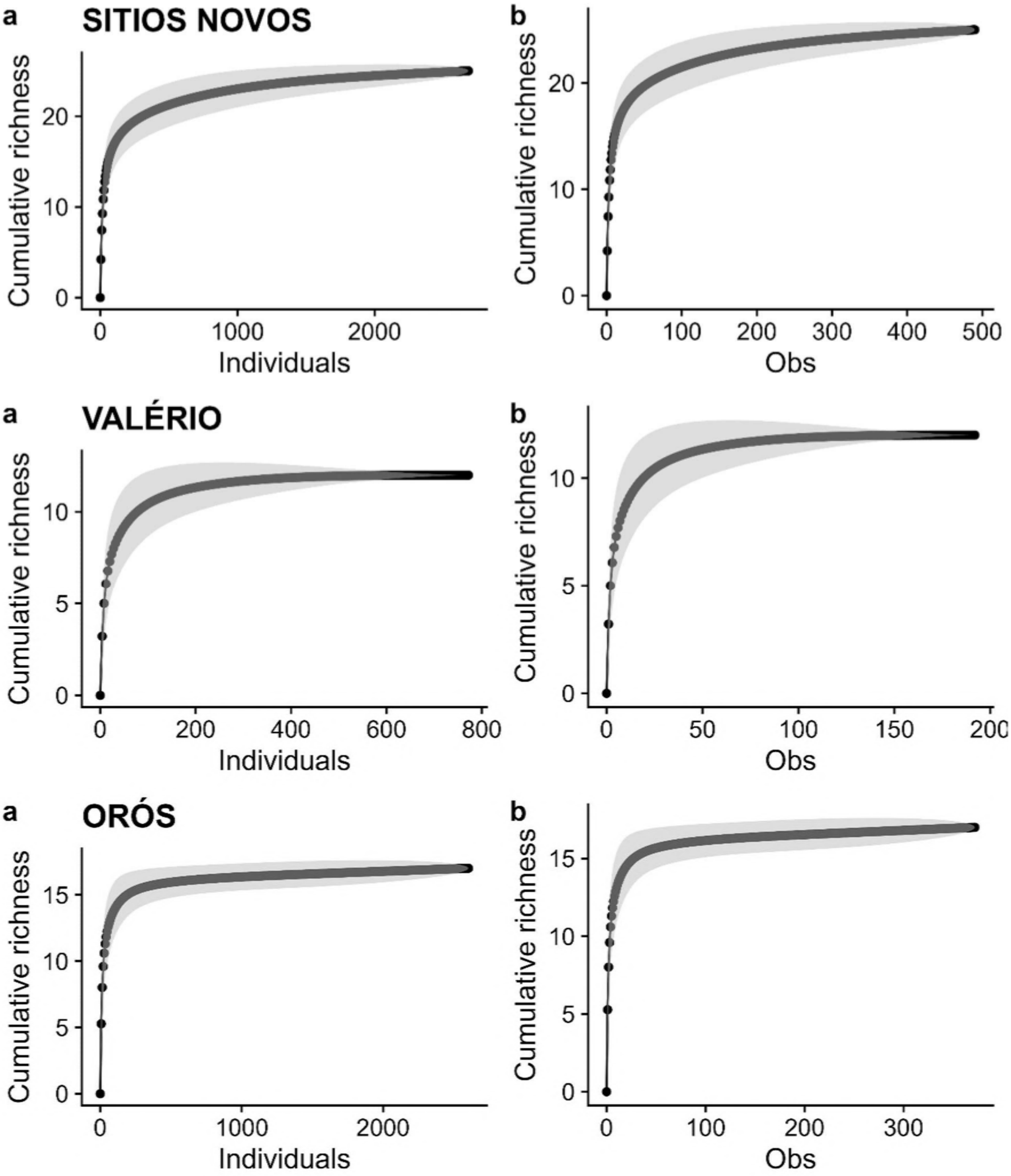


Figure 2 - Reservoirs: Acarape do Meio, Acaraú Mirim, Adauto Bezerra e Amanry

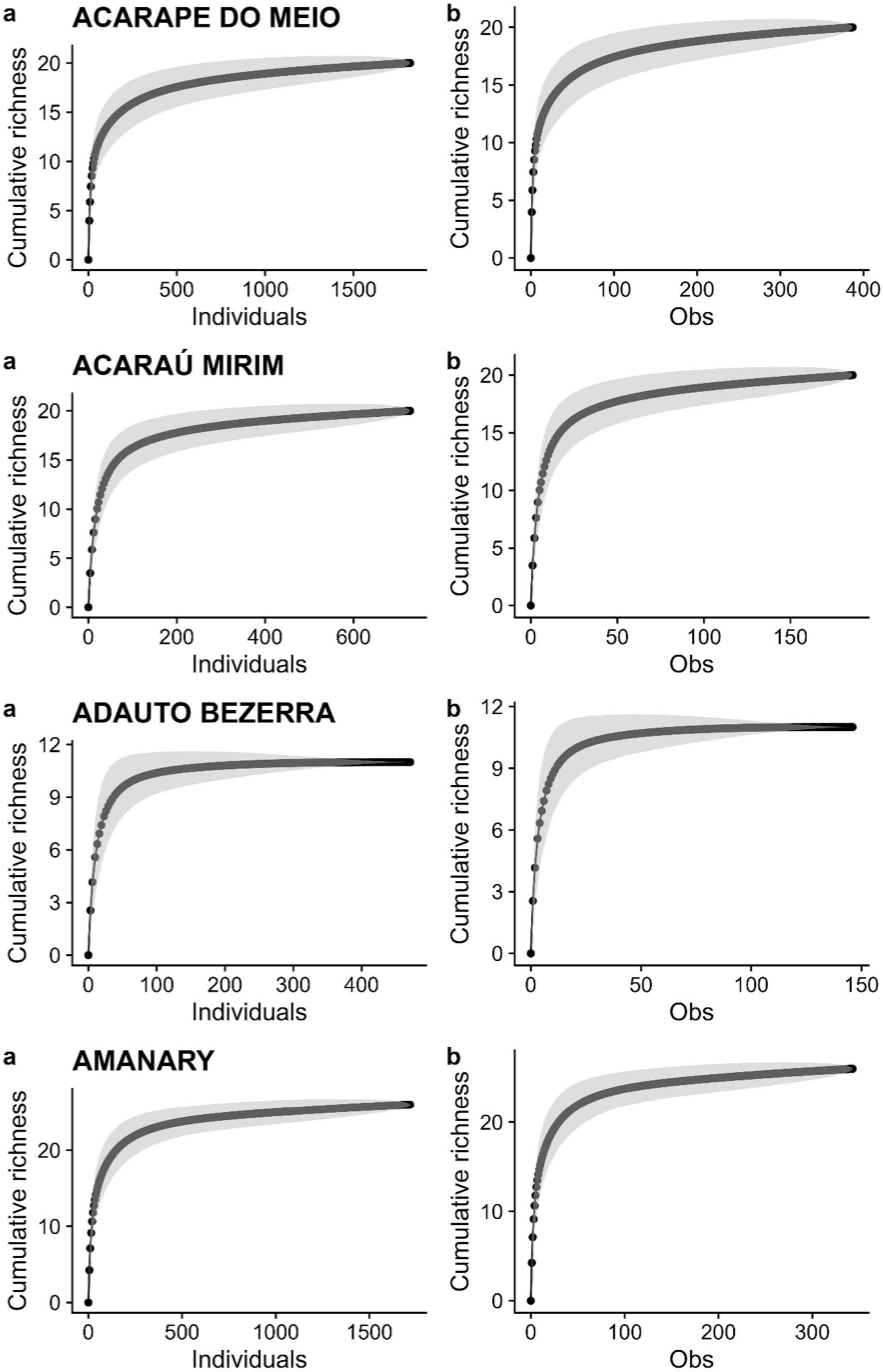


Figure 3 - Reservoirs: Angicos, Ayres de Souza, Belo Horizonte e Boqueirão

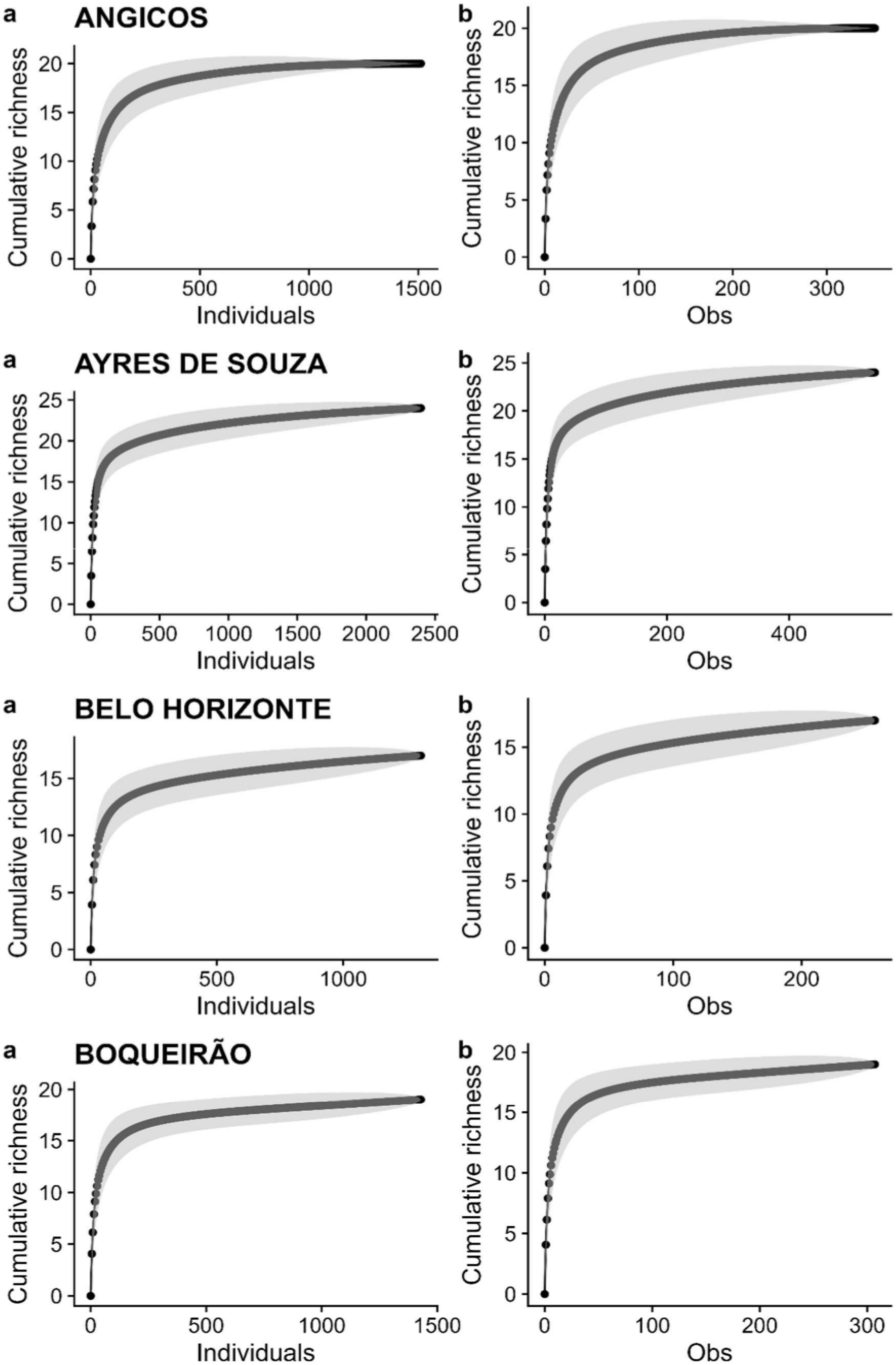


Figure 4 - Reservoirs: Canafístula, Canos, Carmina e Carnaubal

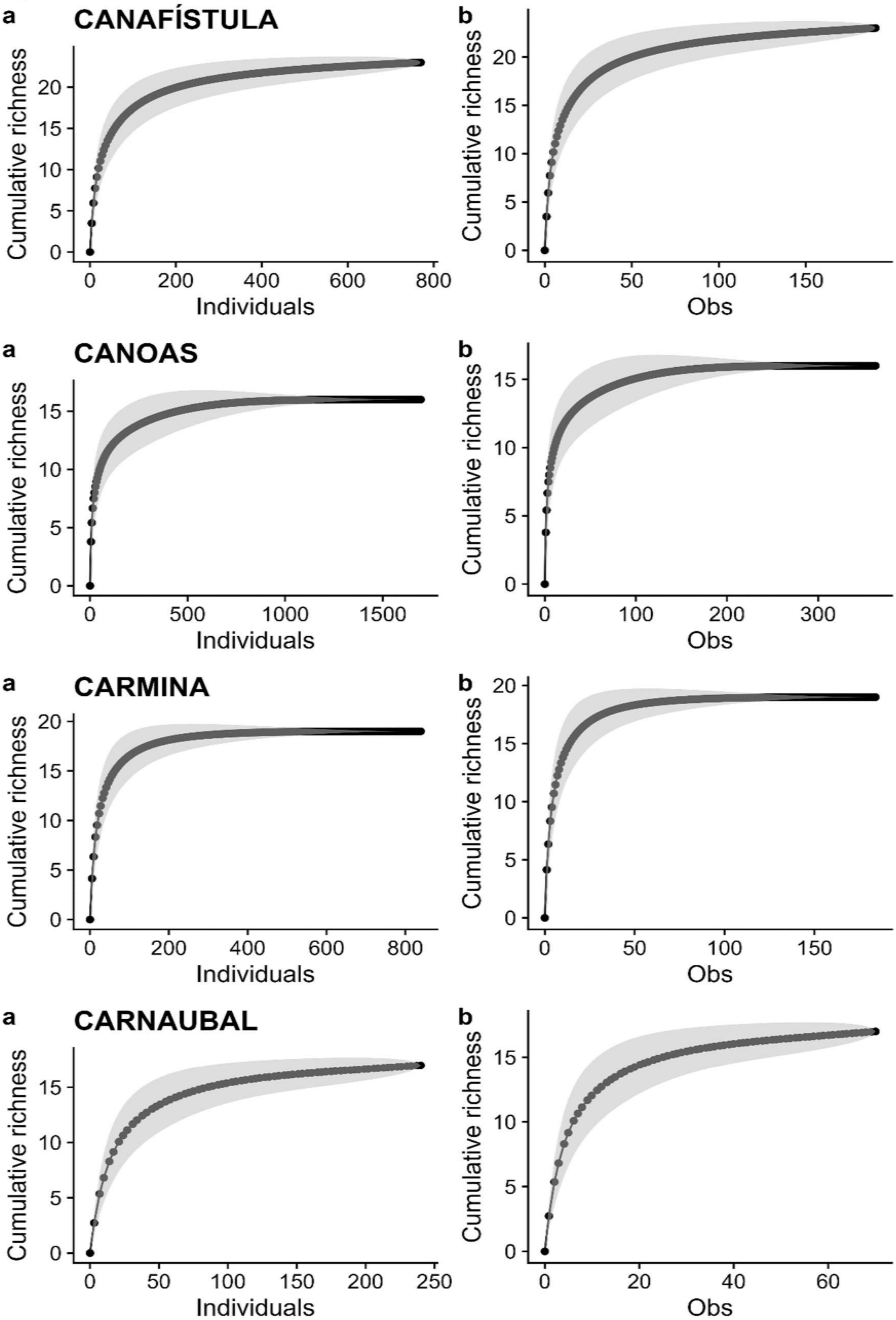


Figure 5 - Reservoirs: Carão, Castro, Caxitoré e Colina

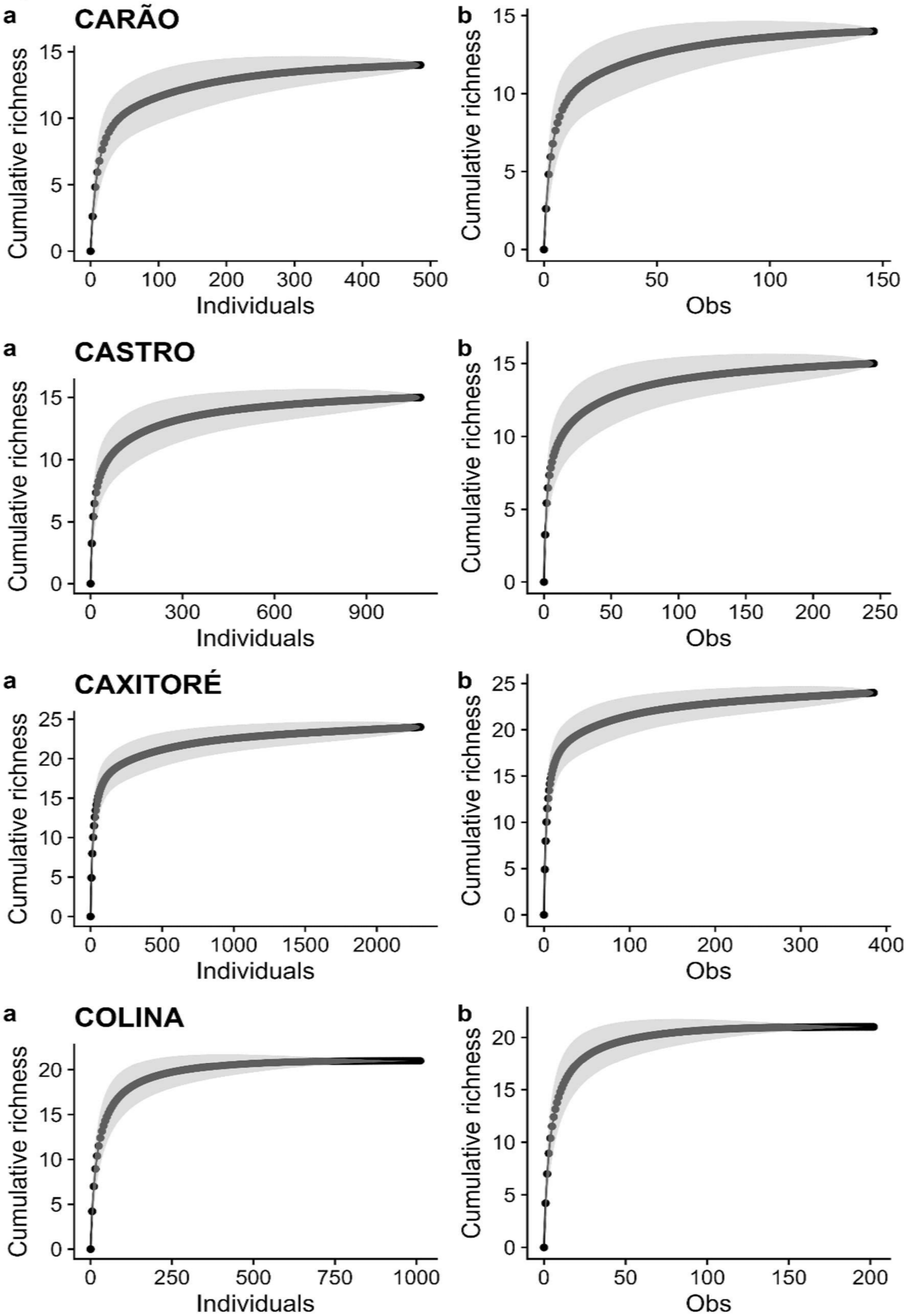


Figure 6 - Reservoirs: Coriolano Alves de Sousa Brito, Olho d'Água, Edson Queiroz e Ema

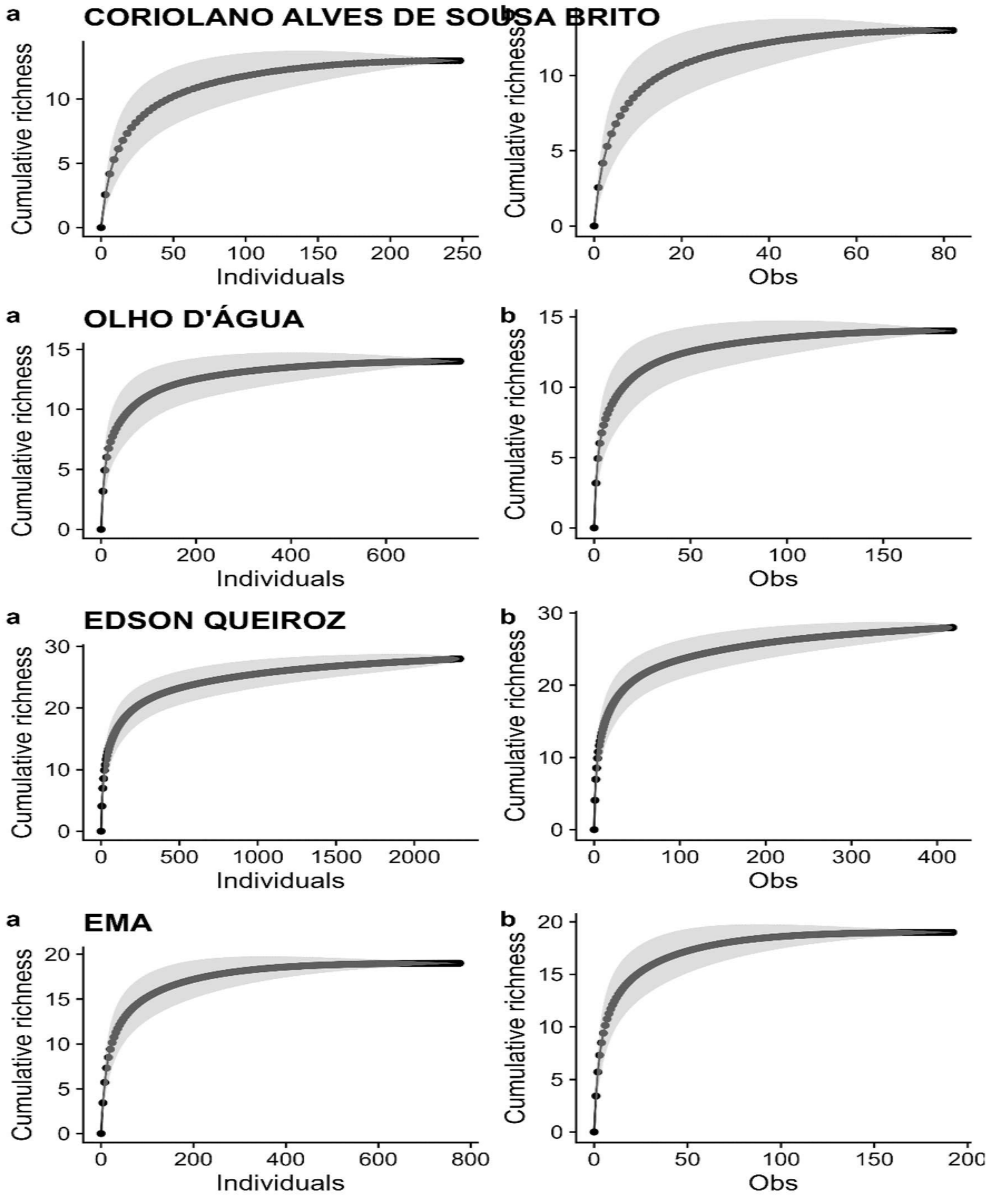
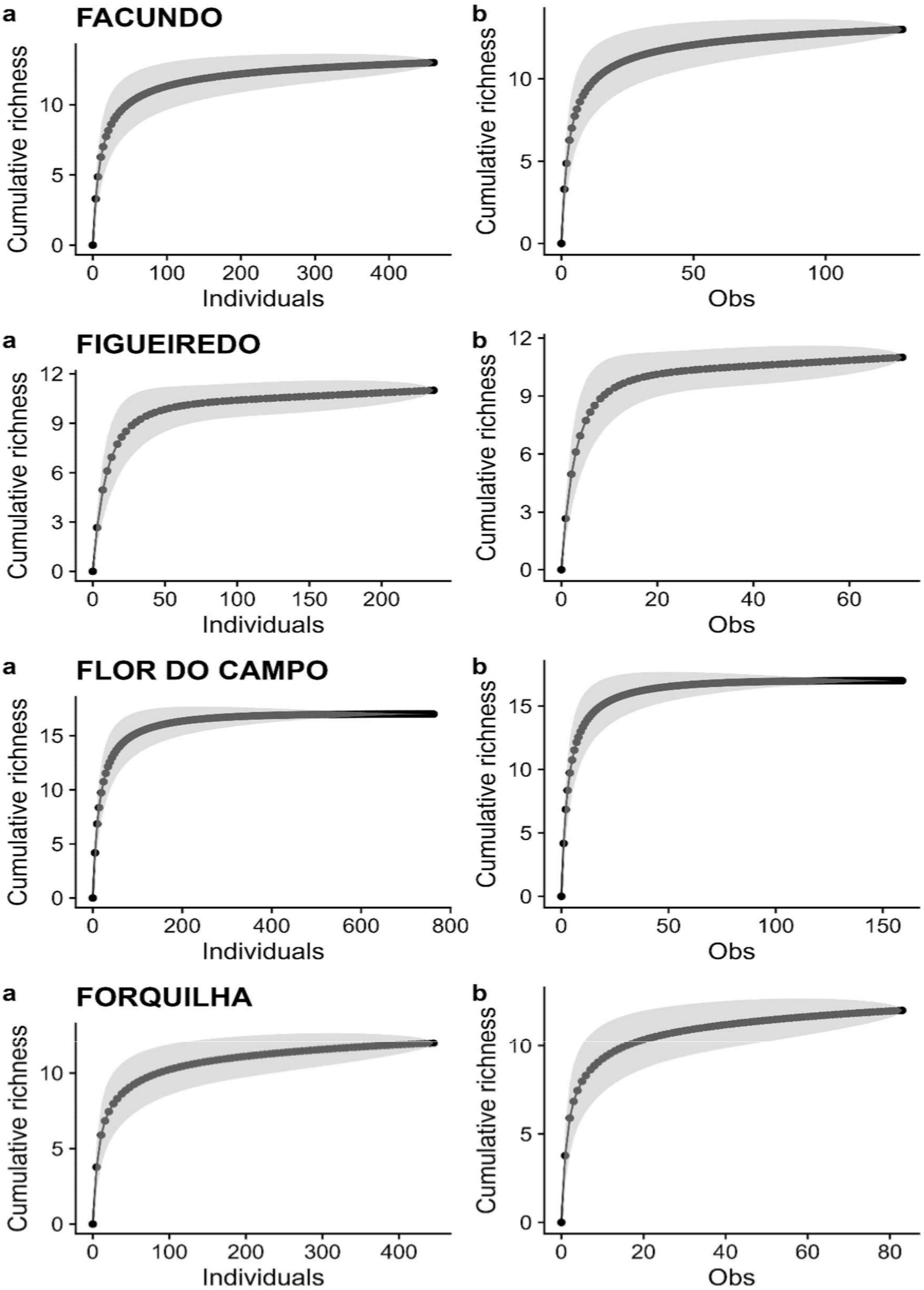


Figure 7 - Reservoirs: Facundo, Figueiredo, Flor do Campo e Forquilha



a
FACUNDO
Cumulative richness
0
5
10
0
100
200
300
400
Individuals
b
Cumulative richness
0
5
10
0
50
100
Obs
a
FIGUEIREDO
Cumulative richness
0
3
6
9
12
0
50
100
150
200
Individuals
b
Cumulative richness
0
3
6
9
12
0
20
40
60
Obs
a
FLOR DO CAMPO
Cumulative richness
0
5
10
15
0
200
400
600
800
Individuals
b
Cumulative richness
0
5
10
15
0
50
100
150
Obs
a
FORQUILHA
Cumulative richness
0
5
10
0
100
200
300
400
Individuals
b
Cumulative richness
0
5
10
0
20
40
60
80
Obs

Figure 8 - Reservoirs: Gameleira, General Sampaio, Itapebussu e Jaburu I

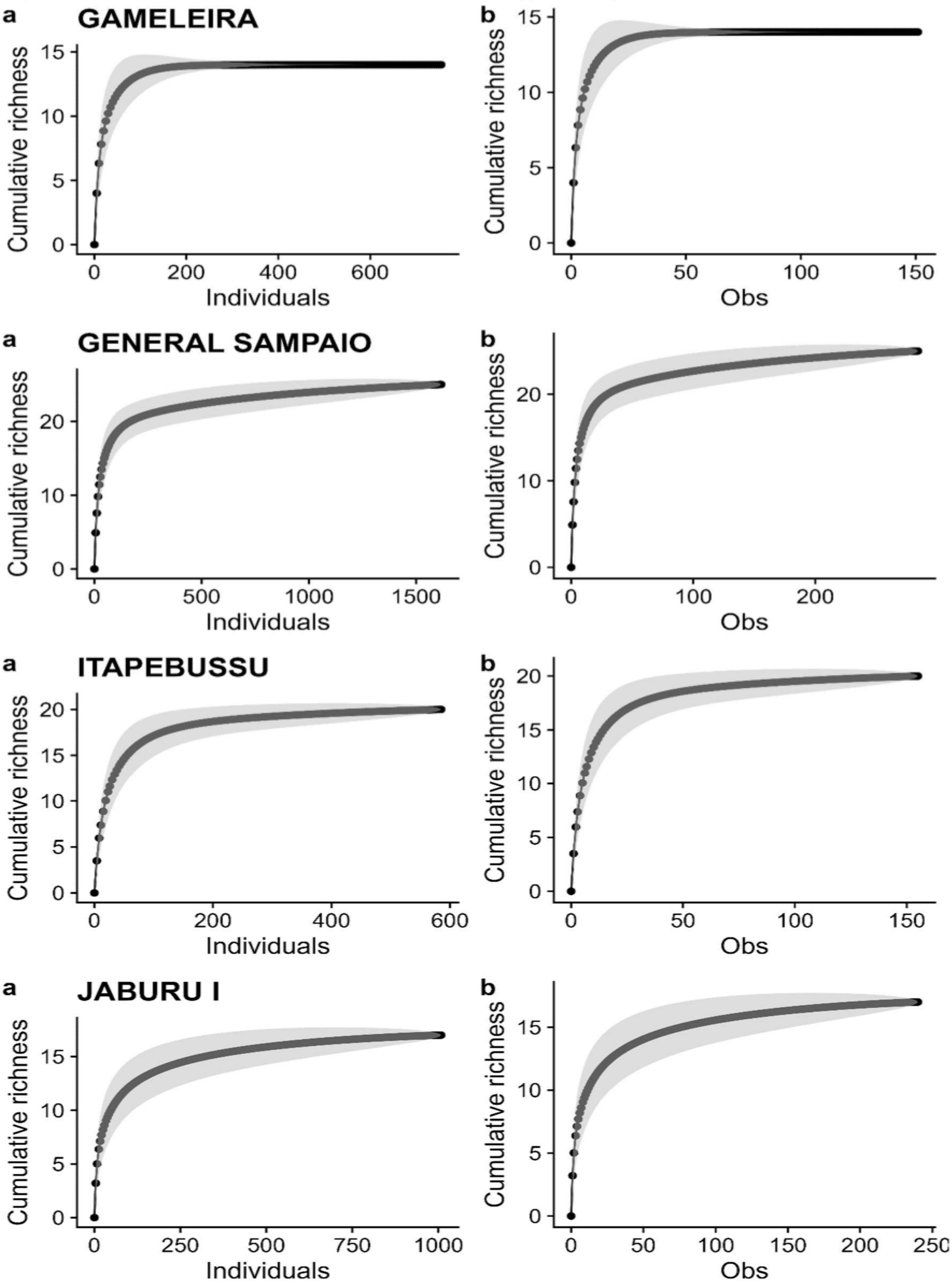


Figure 9 - Reservoirs: Jaburu II, Jerimum, João José de Castro e Junco

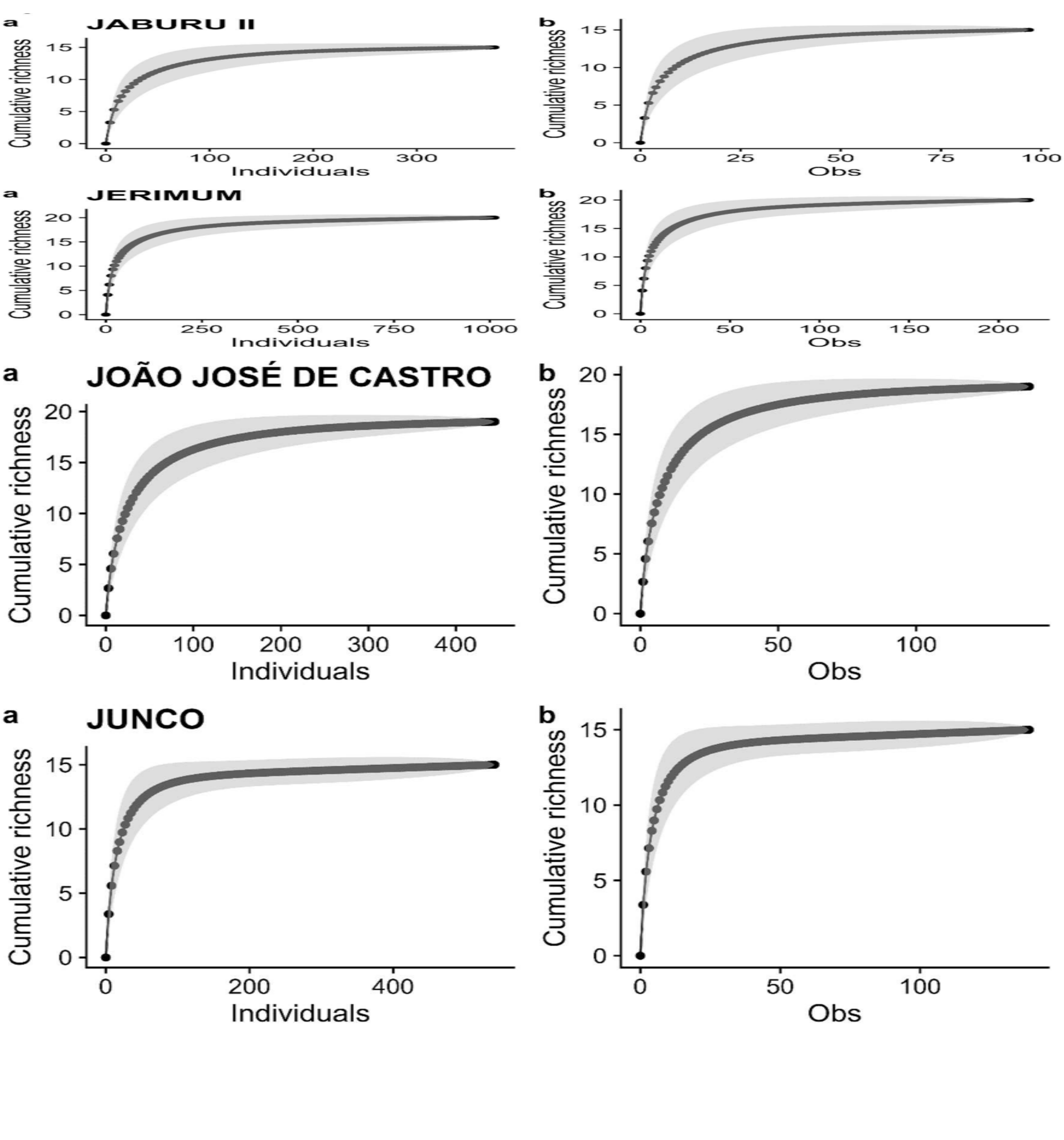


Figure 10 - Reservoirs: Malcozinhado, Mamoeiro, Martinópole e Missí

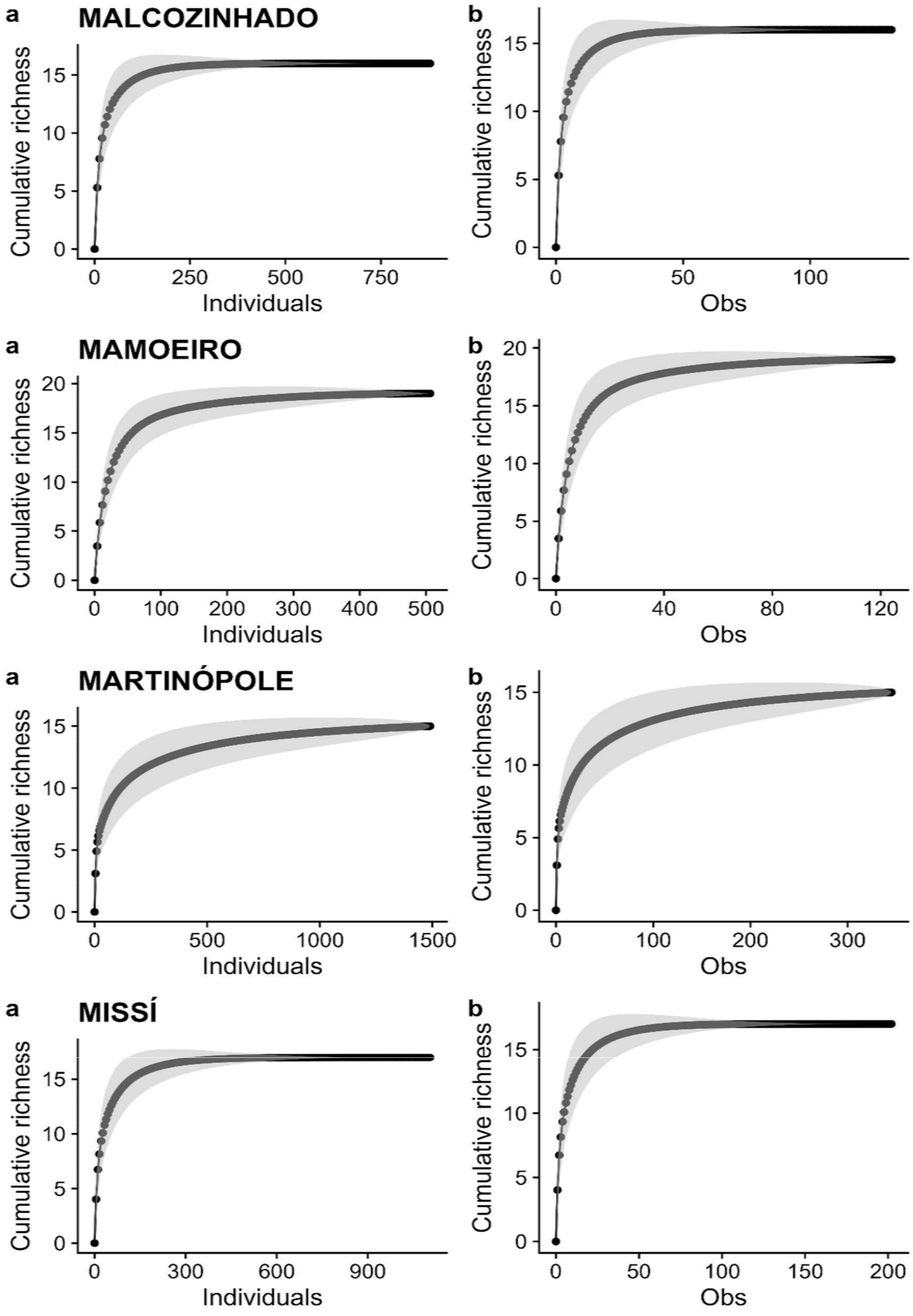

a
MALCOZINHADO
Cumulative richness
Individuals
b
Cumulative richness
Obs
a
MAMOEIRO
Cumulative richness
Individuals
b
Cumulative richness
Obs
a
MARTINÓPOLE
Cumulative richness
Individuals
b
Cumulative richness
Obs
a
MISSÍ
Cumulative richness
Individuals
b
Cumulative richness
Obs

Figure 11 - Reservoirs: Monsenhor Tabosa, Mundaú, Pacoti e Patu

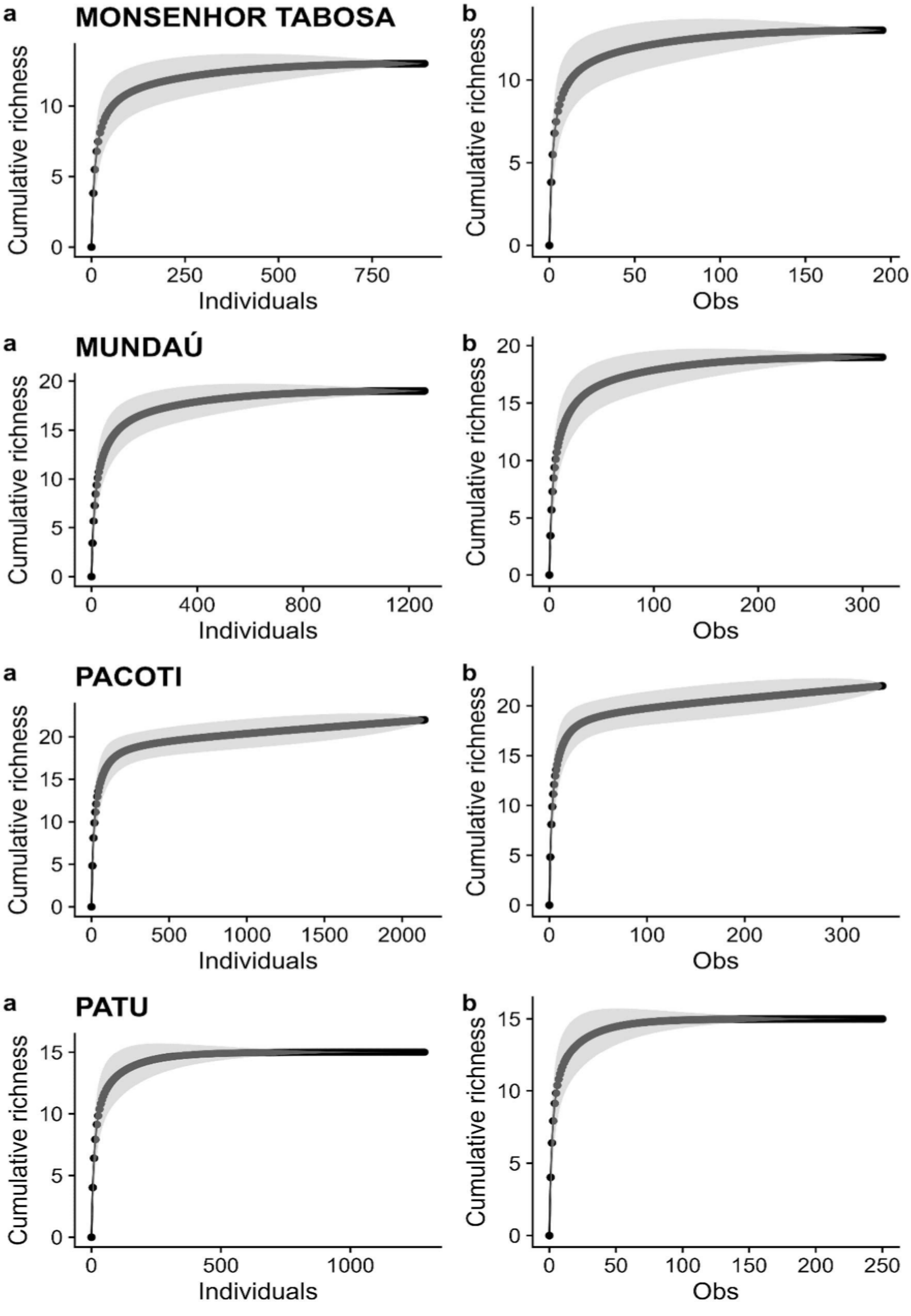
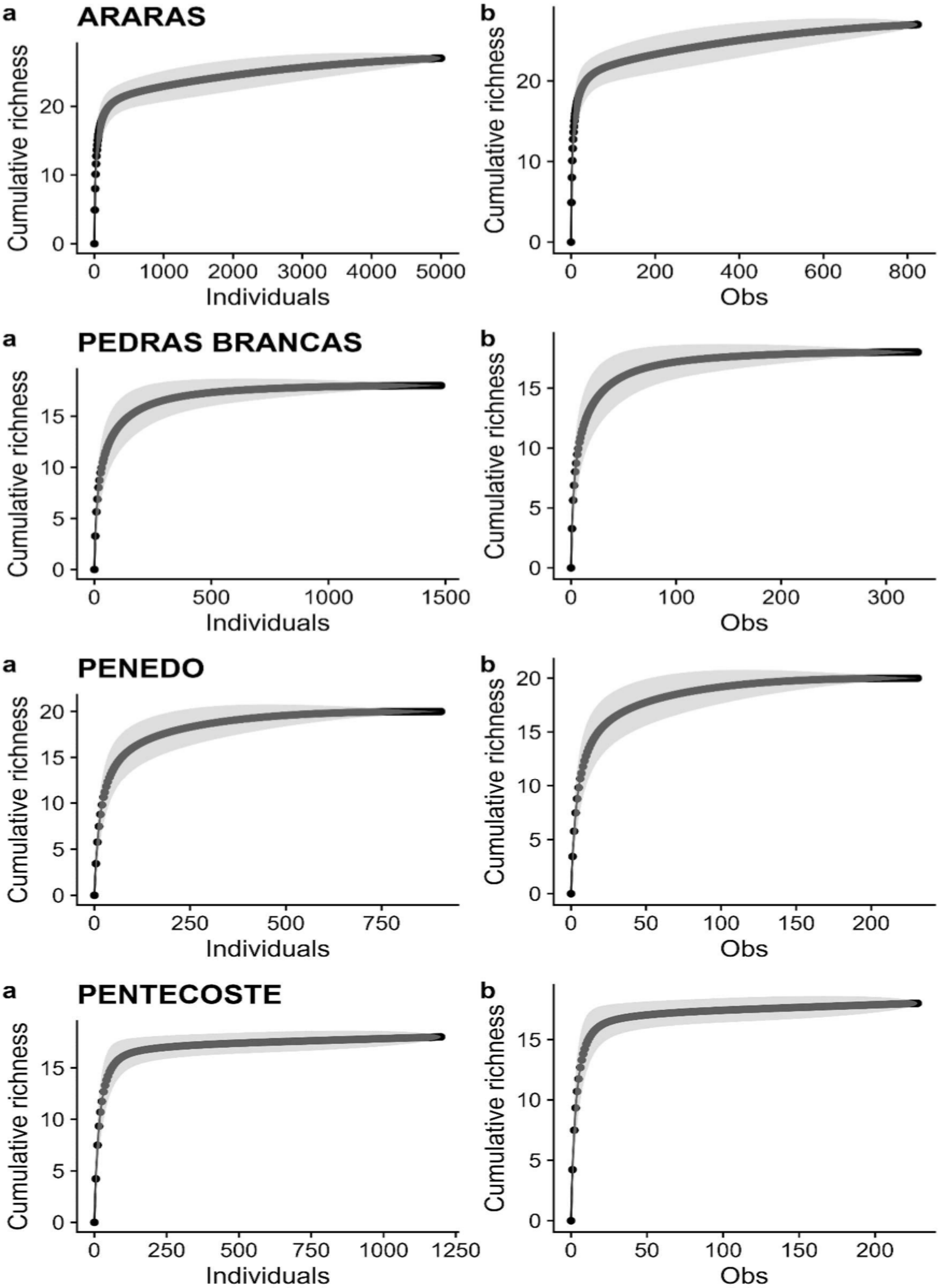


Figure 12 - Reservoirs: Araras, Pedras Brancas, Penedo e Pentecoste



a

Figure 13 - Reservoirs: Pesqueiro, Poço da Pedra, Quandú e Riacho da Serra

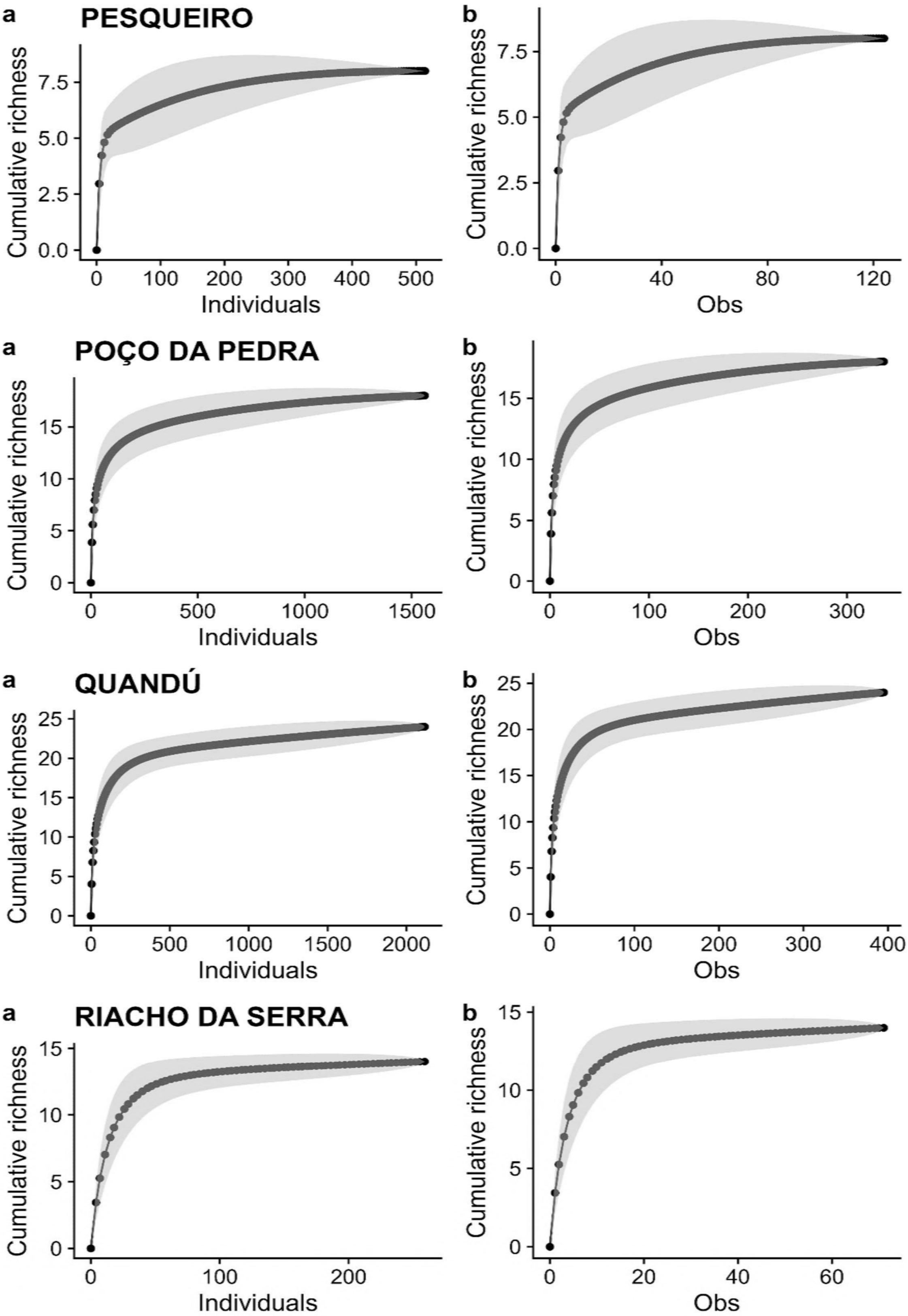
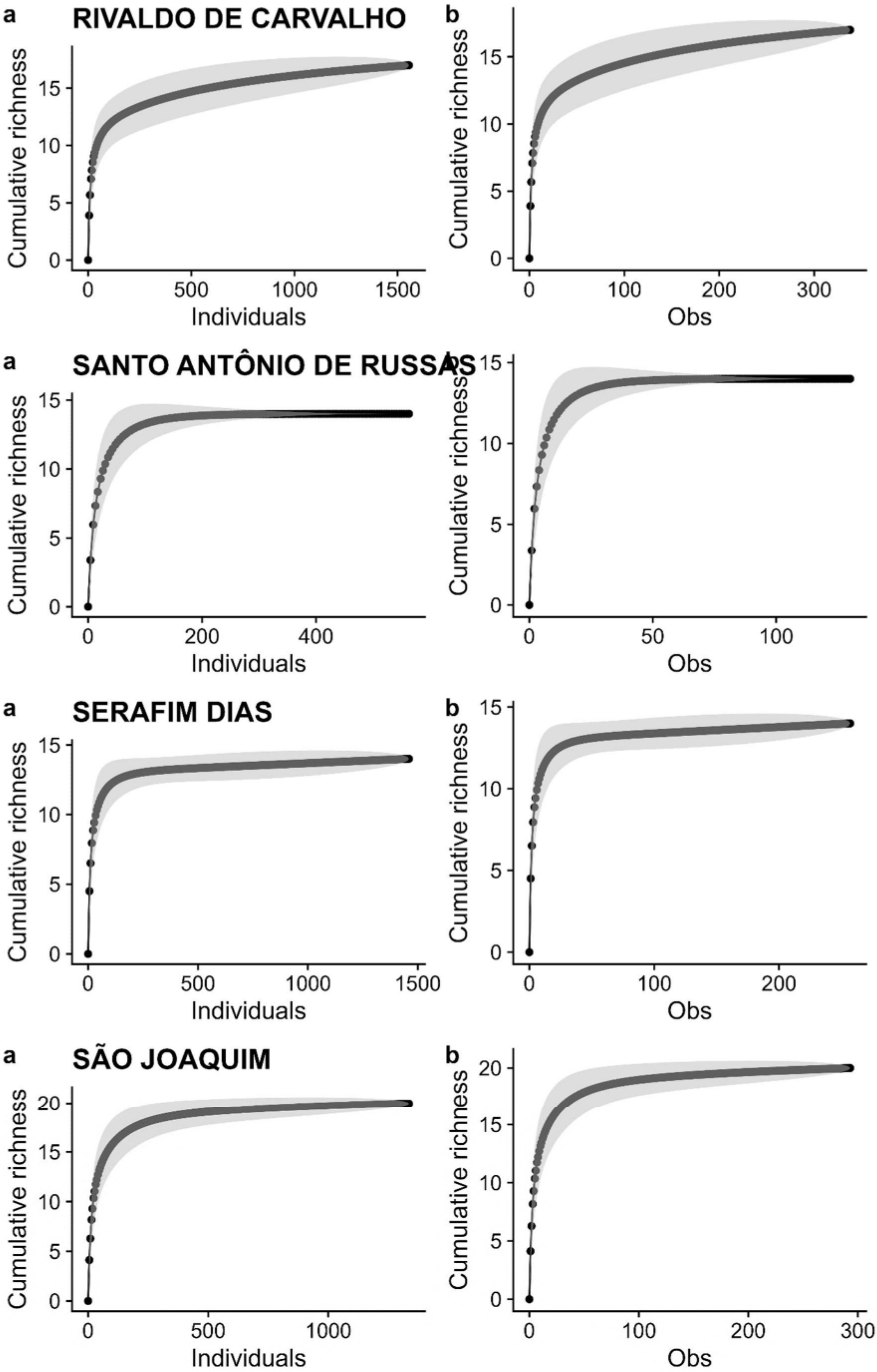


Figure 14 - Reservoirs: Rivaldo de Carvalho, Santo Antônio de Russas, Serafim Dias e São Joaquim


a
RIVALDO DE CARVALHO
b
Cumulative richness
Individuals
Obs
a
SANTO ANTÔNIO DE RUSSAS
Cumulative richness
Individuals
Obs
a
SERAFIM DIAS
b
Cumulative richness
Individuals
Obs
a
SÃO JOAQUIM
b
Cumulative richness
Individuals
Obs

DISCRETE GROWTH: UP TO 5 NEW ORGANISMS FOR EVERY 100 INDIVIDUALS OR OBSERVATIONS

- 27 reservoirs (32,9 %)

Figure 15 - Reservoirs: Ubaldinho (Riacho São Miguel), Ubaldinho e Várzea da Volta

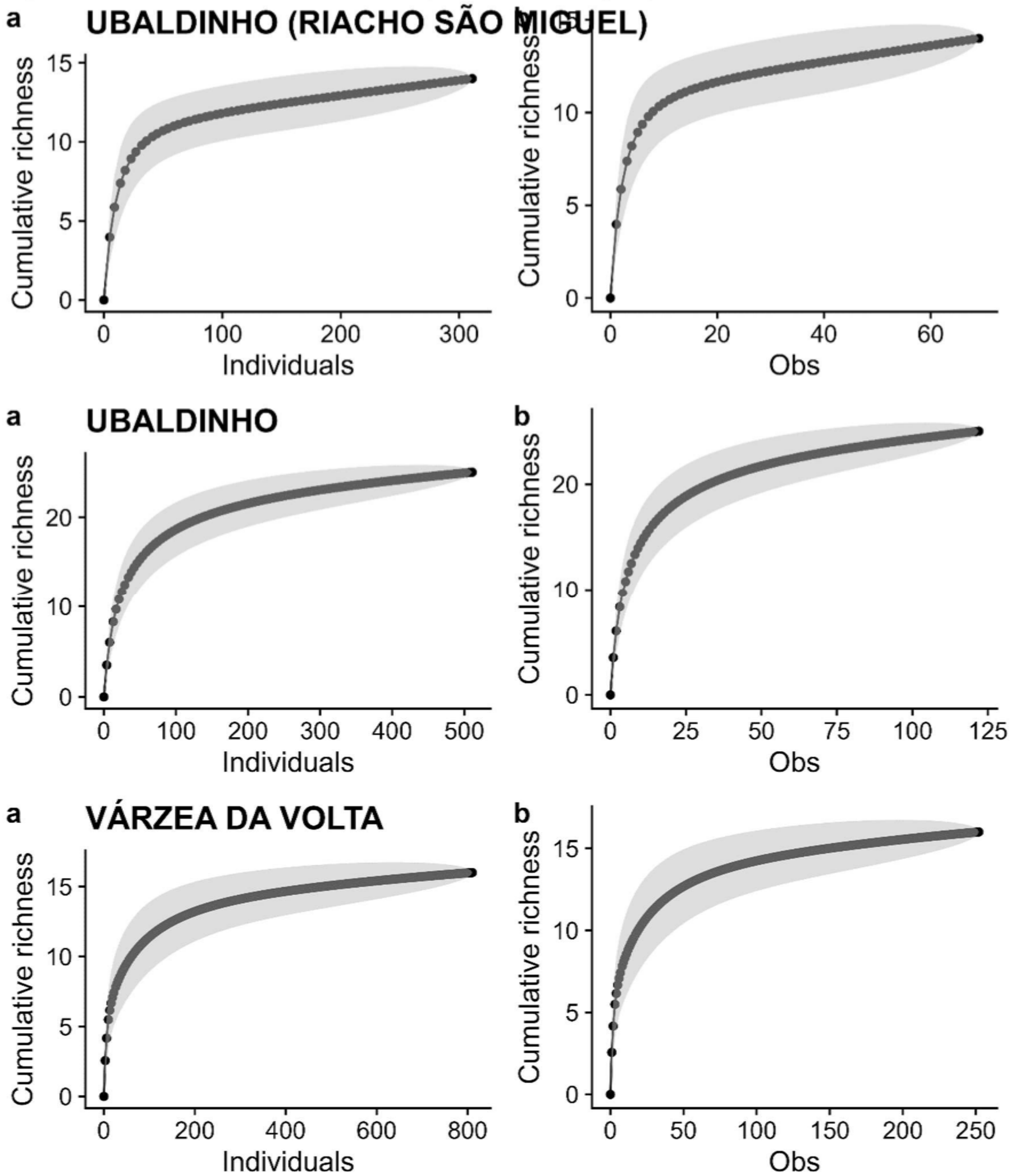


Figure 16 - Reservoirs: Aracoiaba, Arneiroz II, Cachoeira e Do Coronel

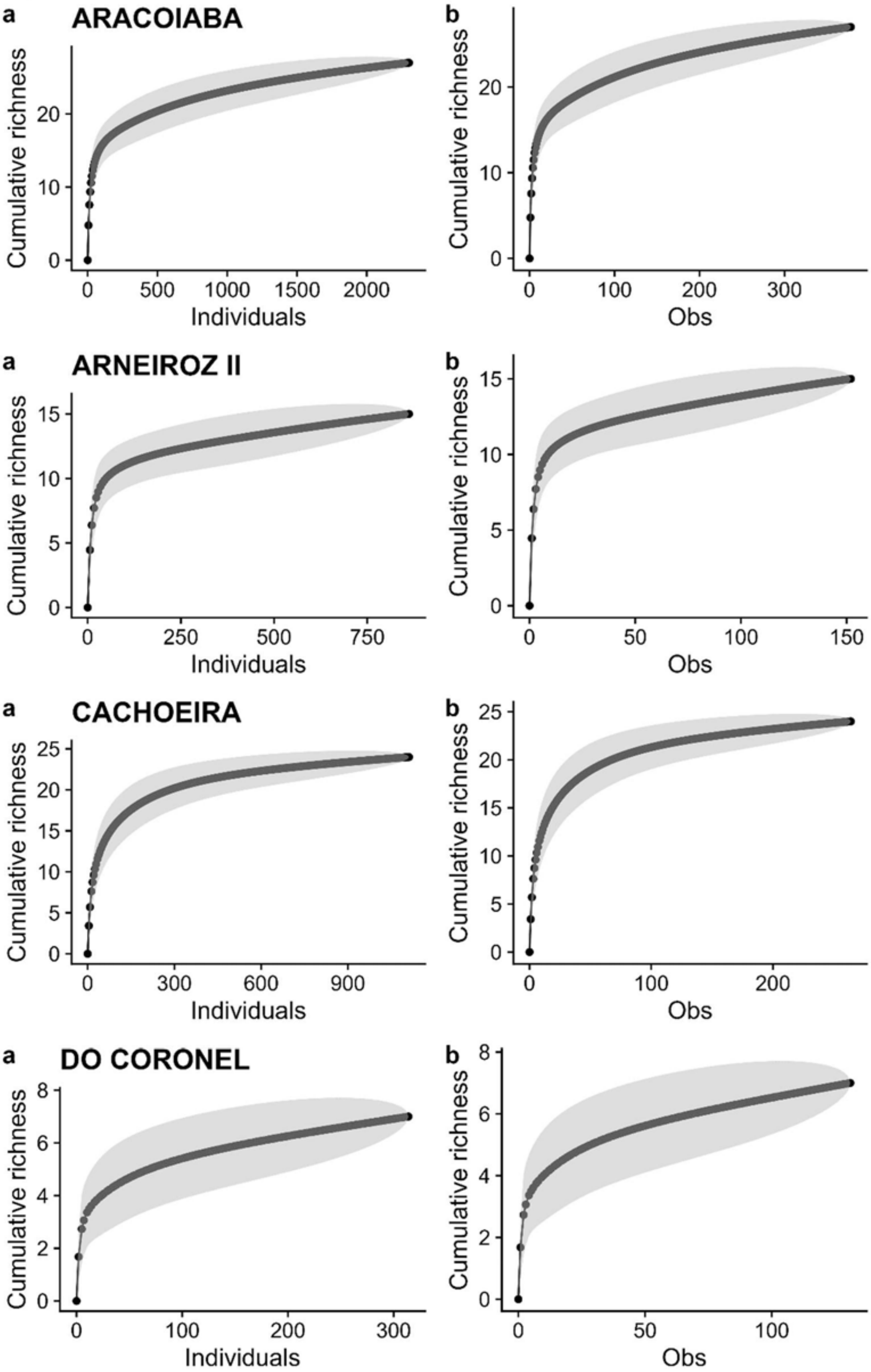


Figure 17 - Reservoirs: João Luís, Madeiro, Poço Verde e Pompeu Sobrinho

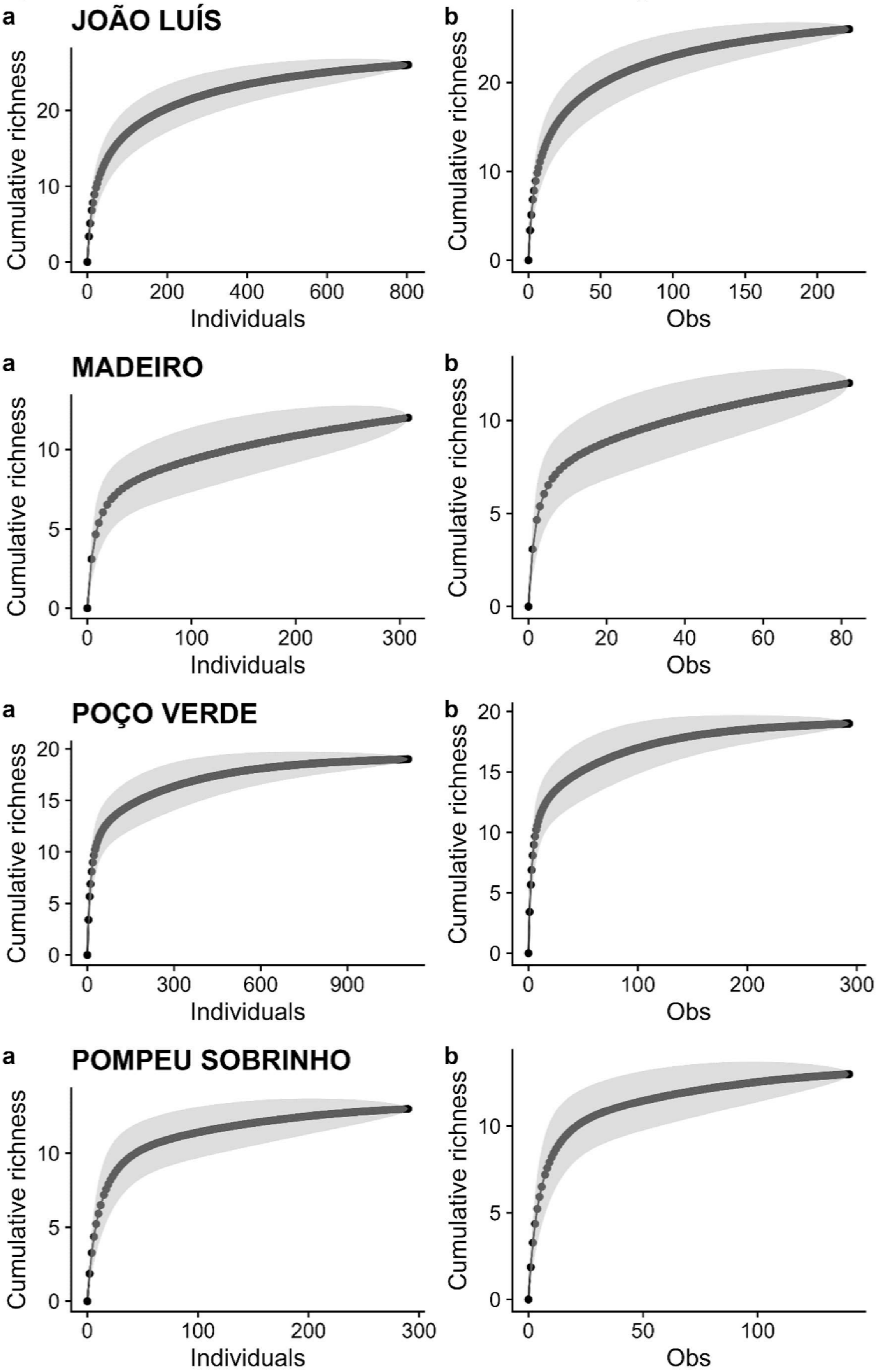


Figure 18 - Reservoirs: Potiretama, Puiú, Riachão e Rodrigo

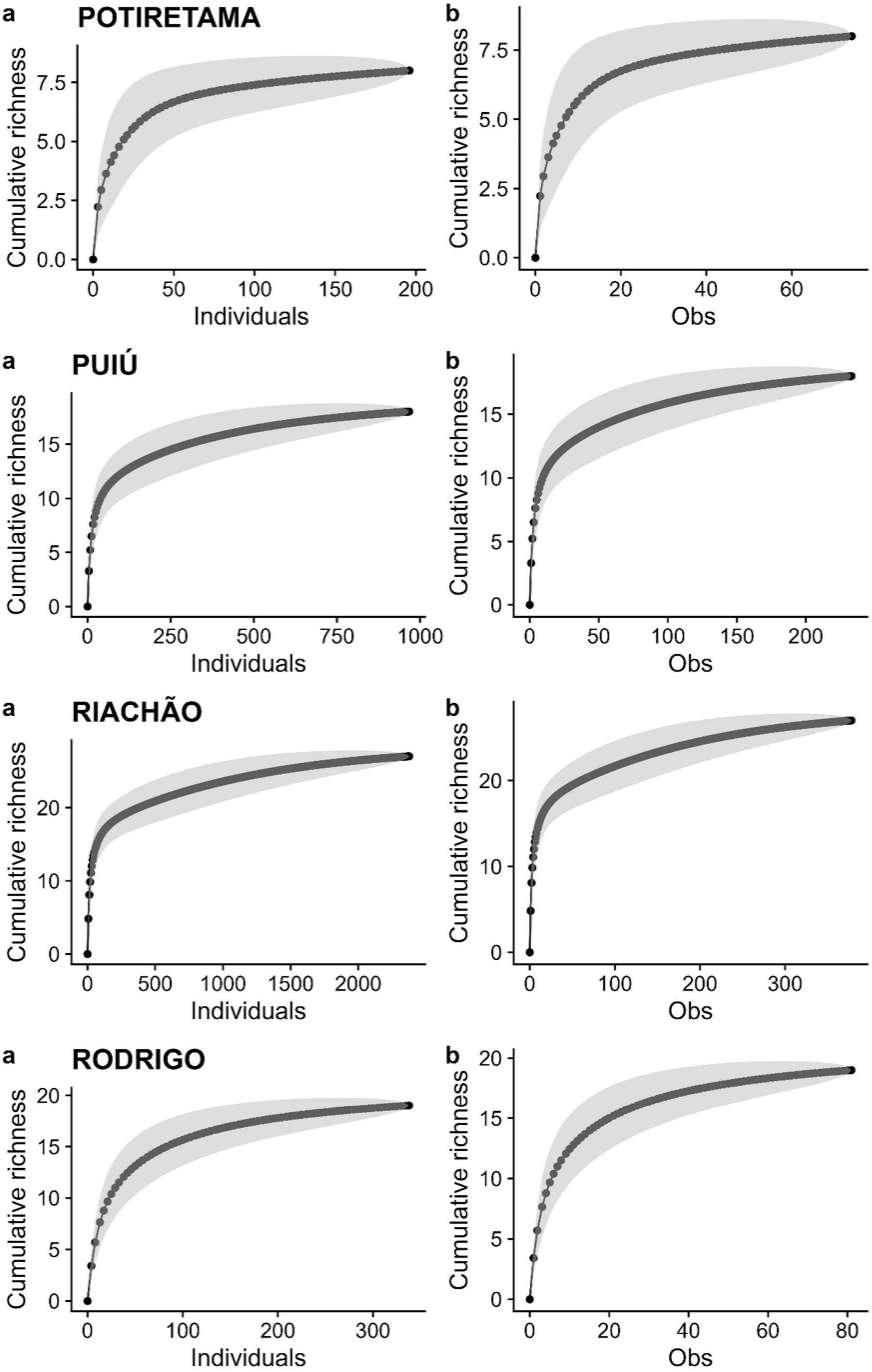


Figure 19 - Reservoirs: Rosário, São Domingos, São José I, São Pedro Timbaúba e Trussu

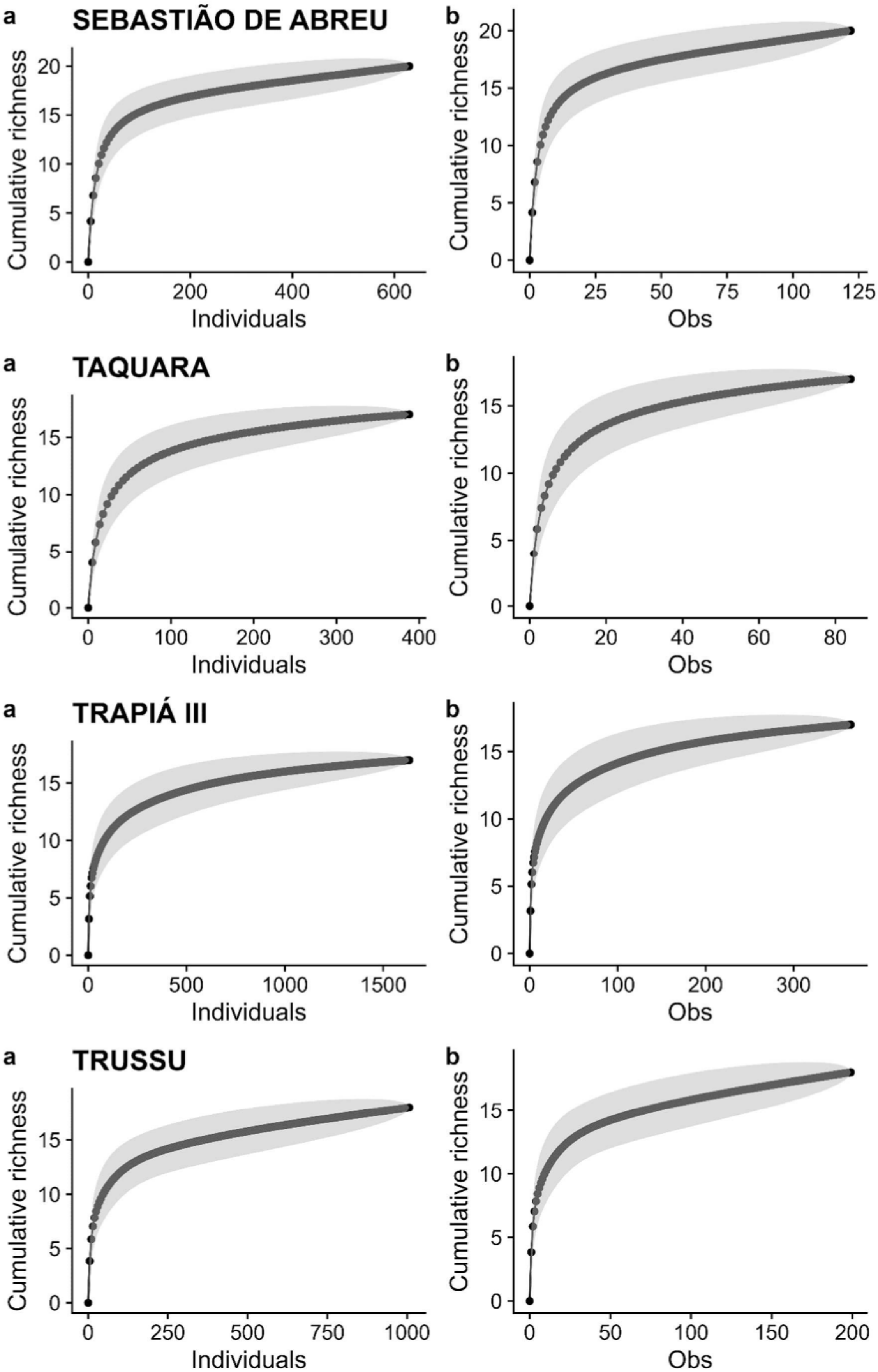
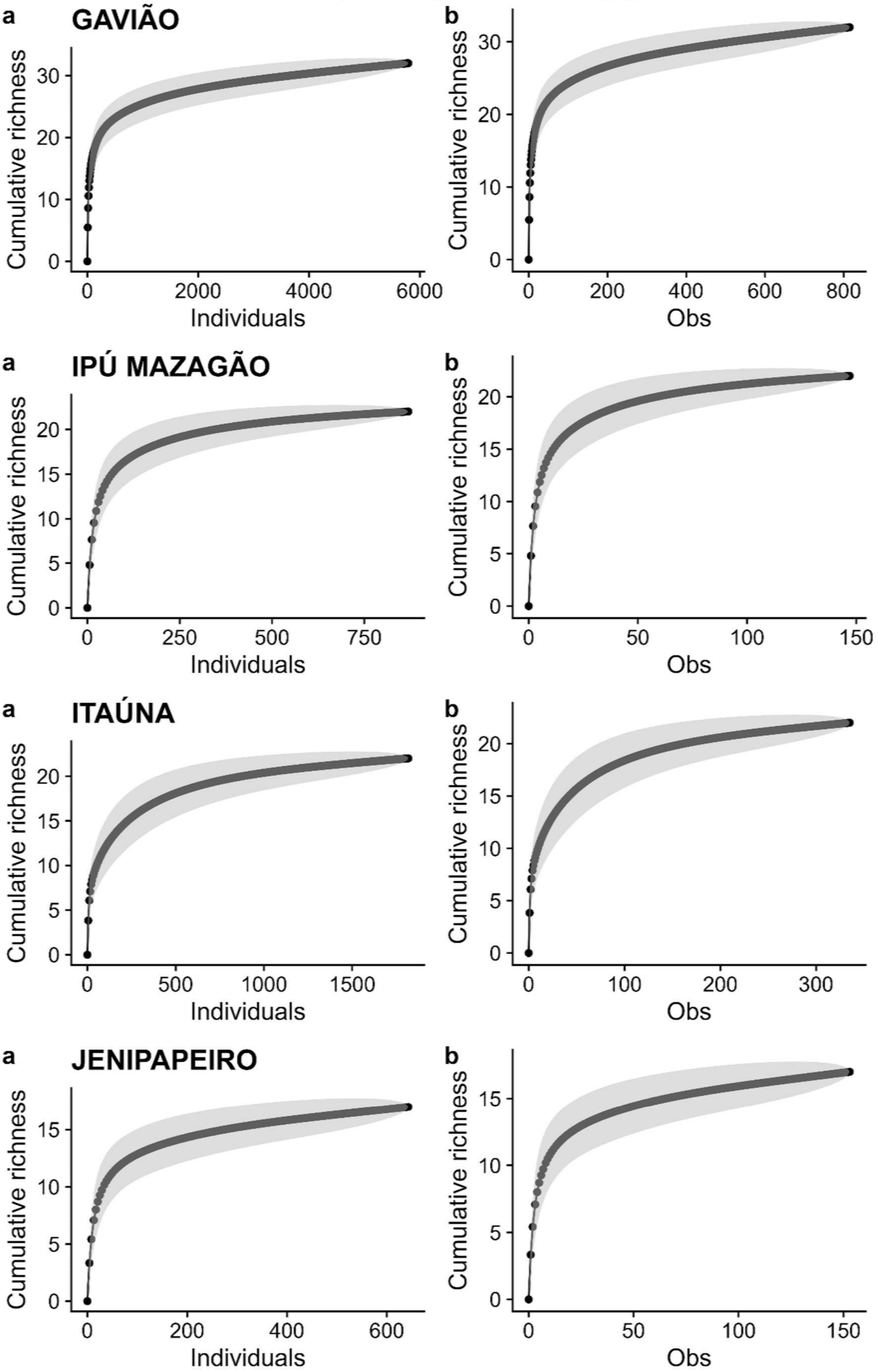


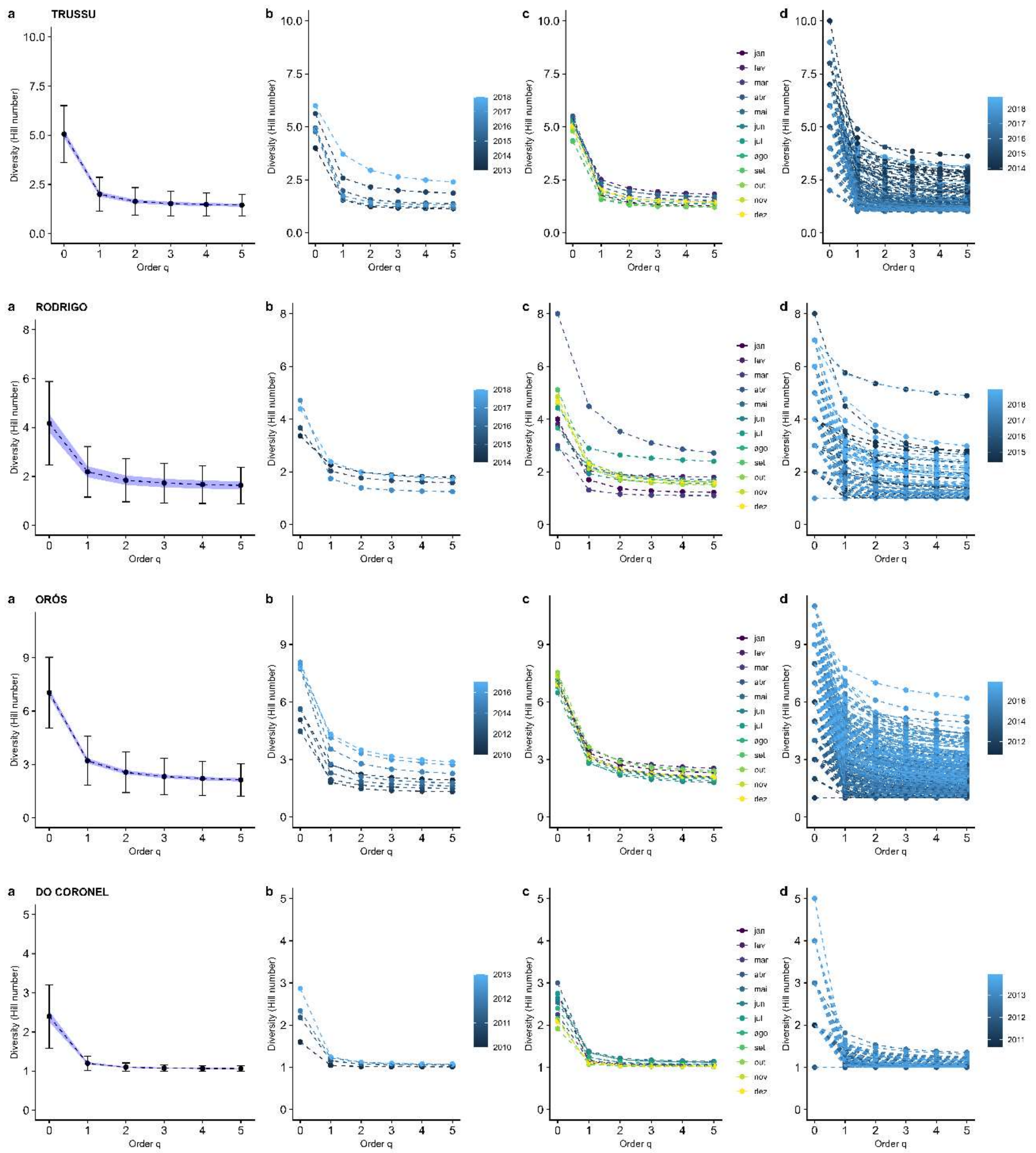
Figure 20 - Reservoirs: Gavião, Ipú Mazagão, Itaúna e Jenipapeiro

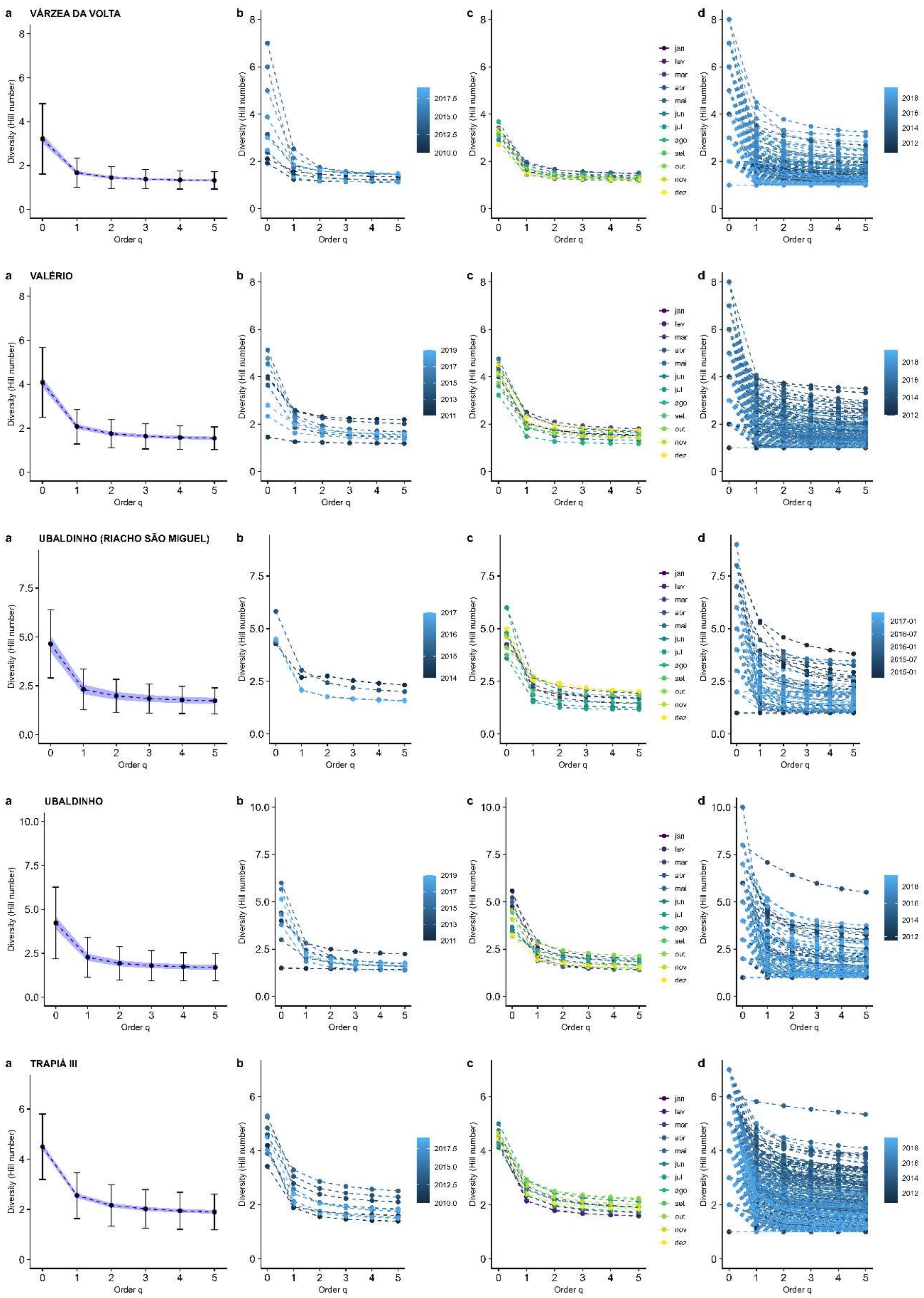


SECTION 4: HILL NUMBERS

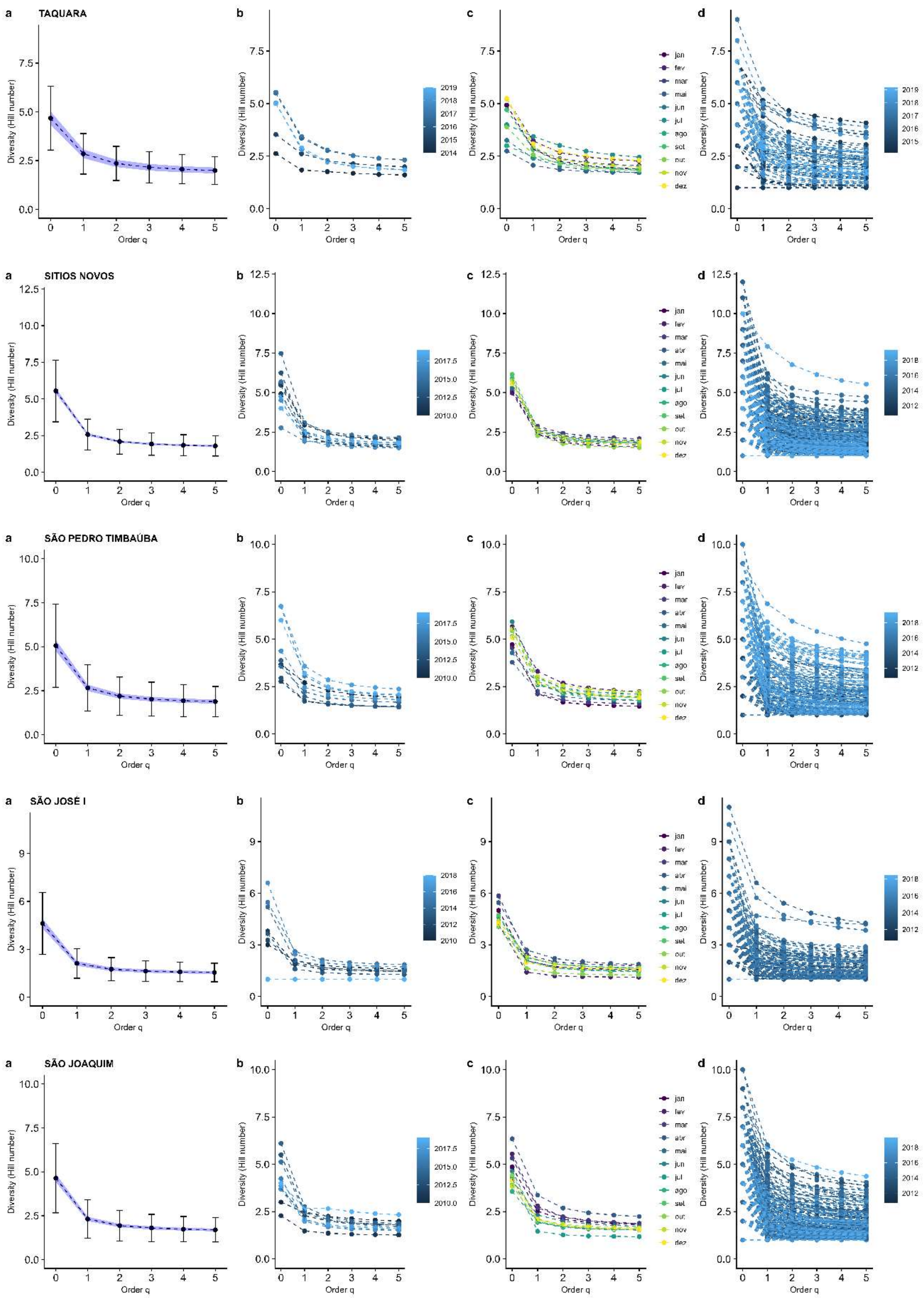
Legend:

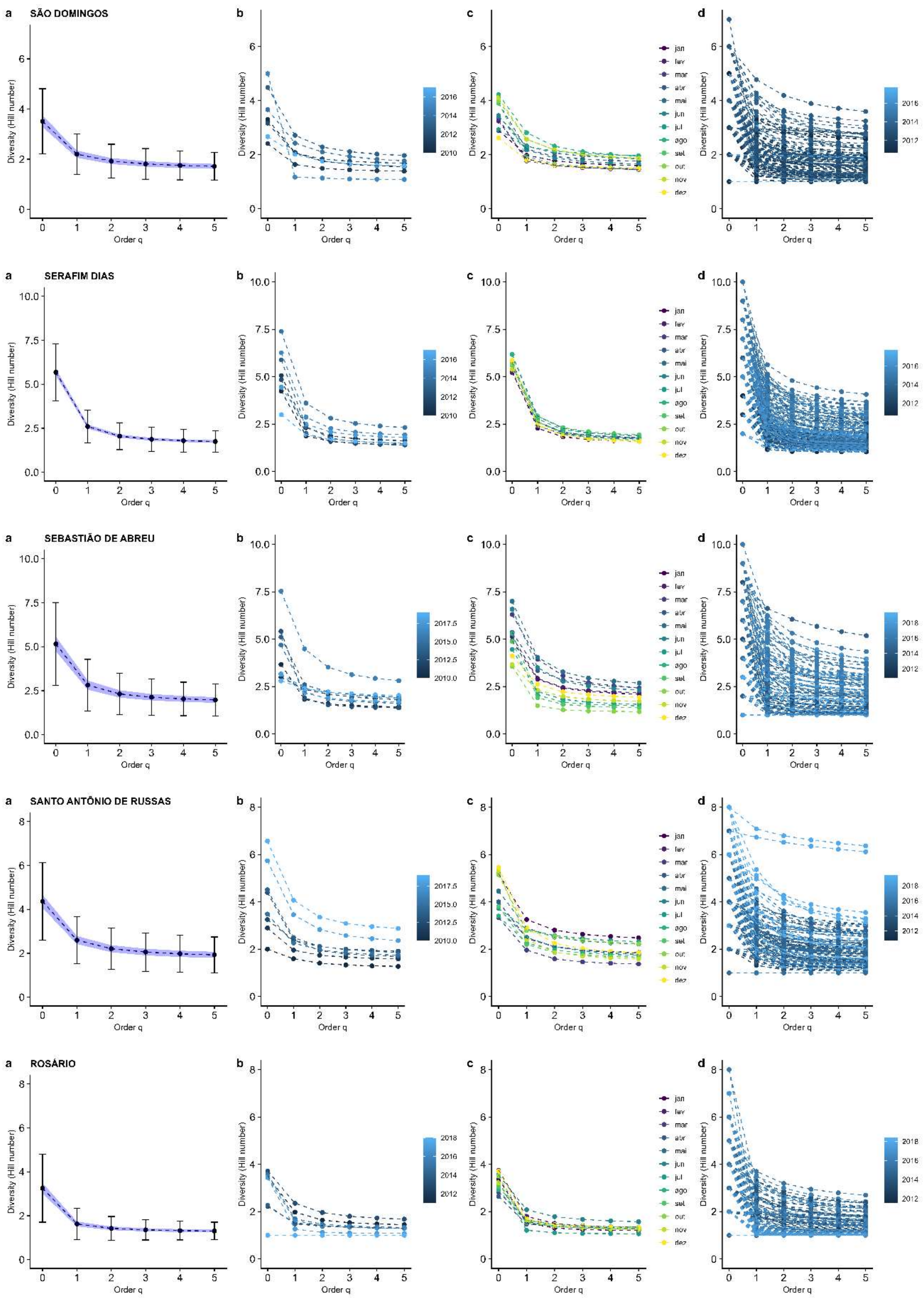
a) general behavior: average of all available observations.
b) annual behavior: average of all observations each year.
c) monthly behavior: average of a given month over all available years.
d) individual behavior: values of each observation evaluated

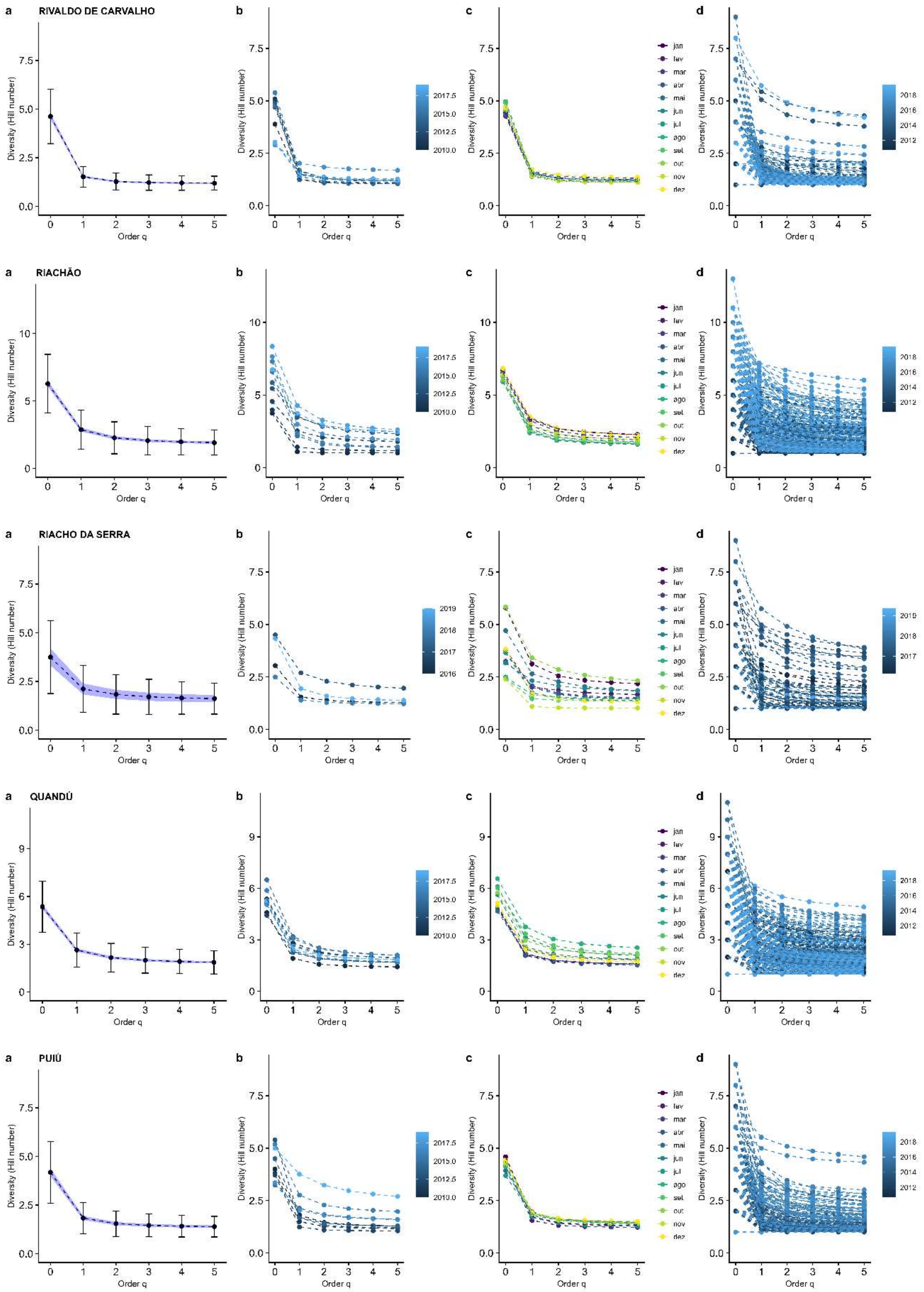


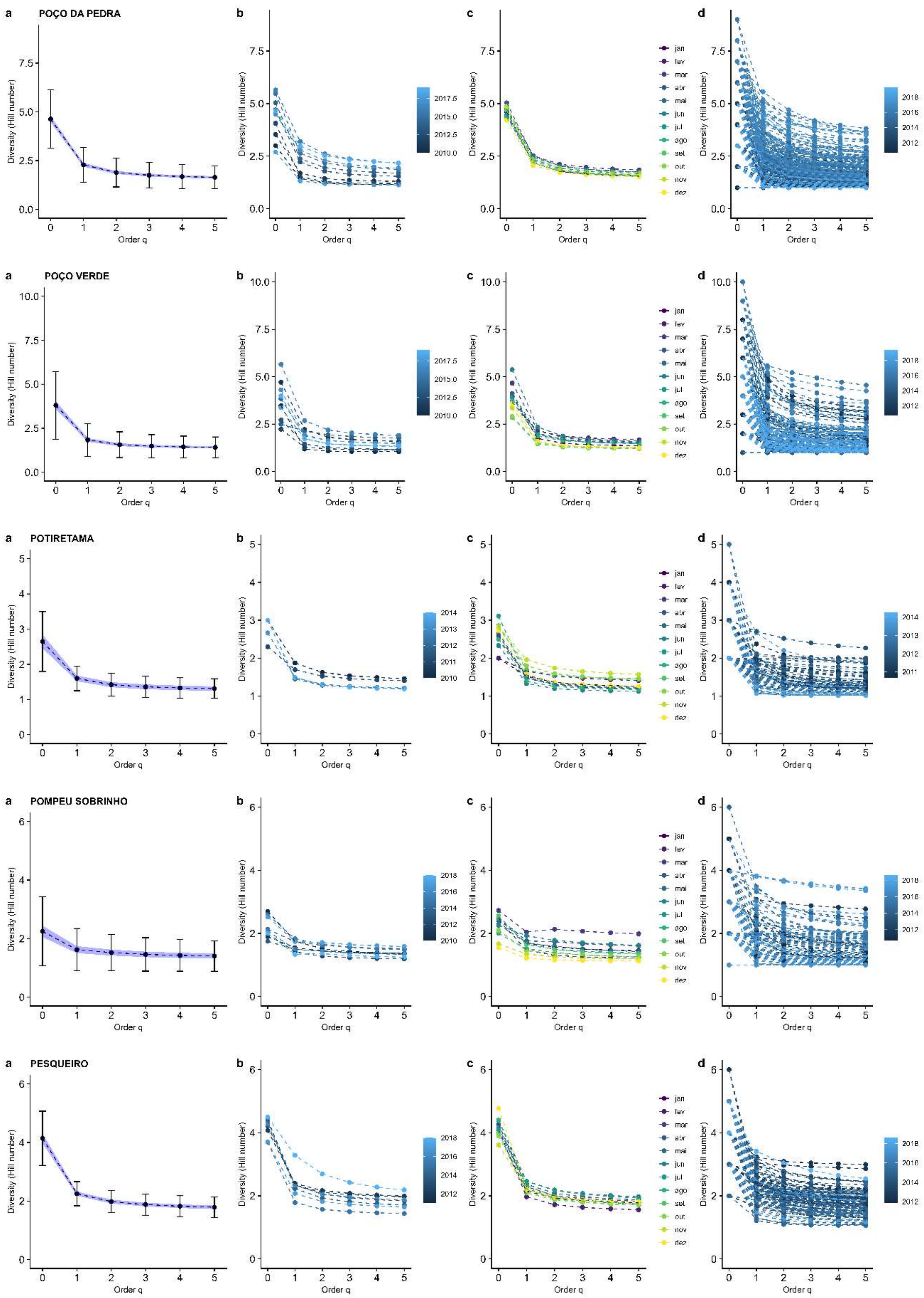


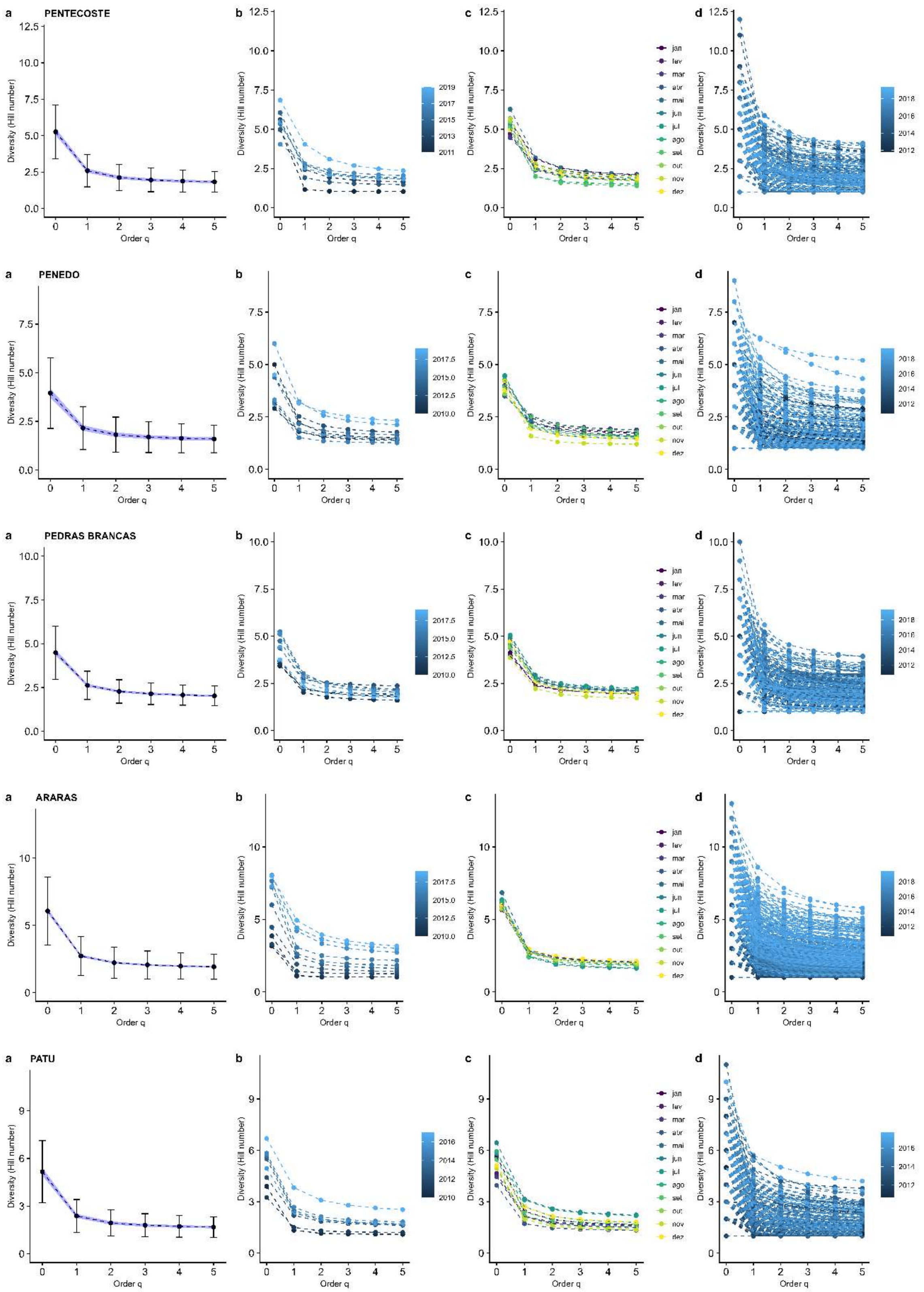
a VÁRZEA DA VOLTA
a VALÉRIO
a UBALDINHO (RIACHO SÃO MIGUEL)
a UBALDINHO
a TRAPIÁ III
b
c
d
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez


a
TAQUARA
b
c
d
a
SITIOS NOVOS
b
c
d
a
SÃO PEDRO TIMBAÚBA
b
c
d
a
SÃO JOSÉ I
b
c
d
a
SÃO JOAQUIM
b
c
d
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez

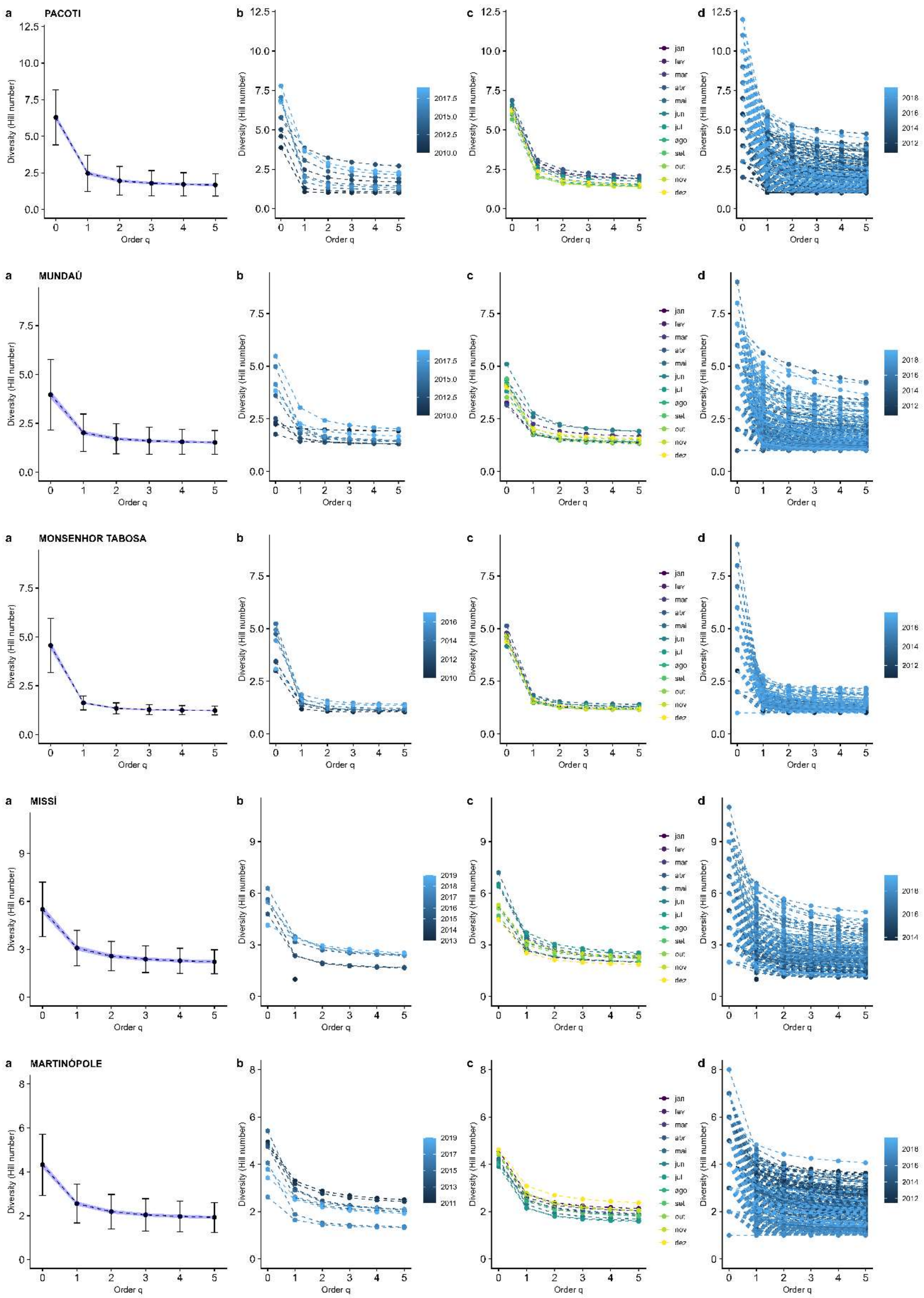

a
SÃO DOMINGOS
b
c
d
a
SERAFIM DIAS
b
c
d
a
SEBASTIÃO DE ABREU
b
c
d
a
SANTO ANTÔNIO DE RUSSAS
b
c
d
a
ROSÁRIO
b
c
d
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez

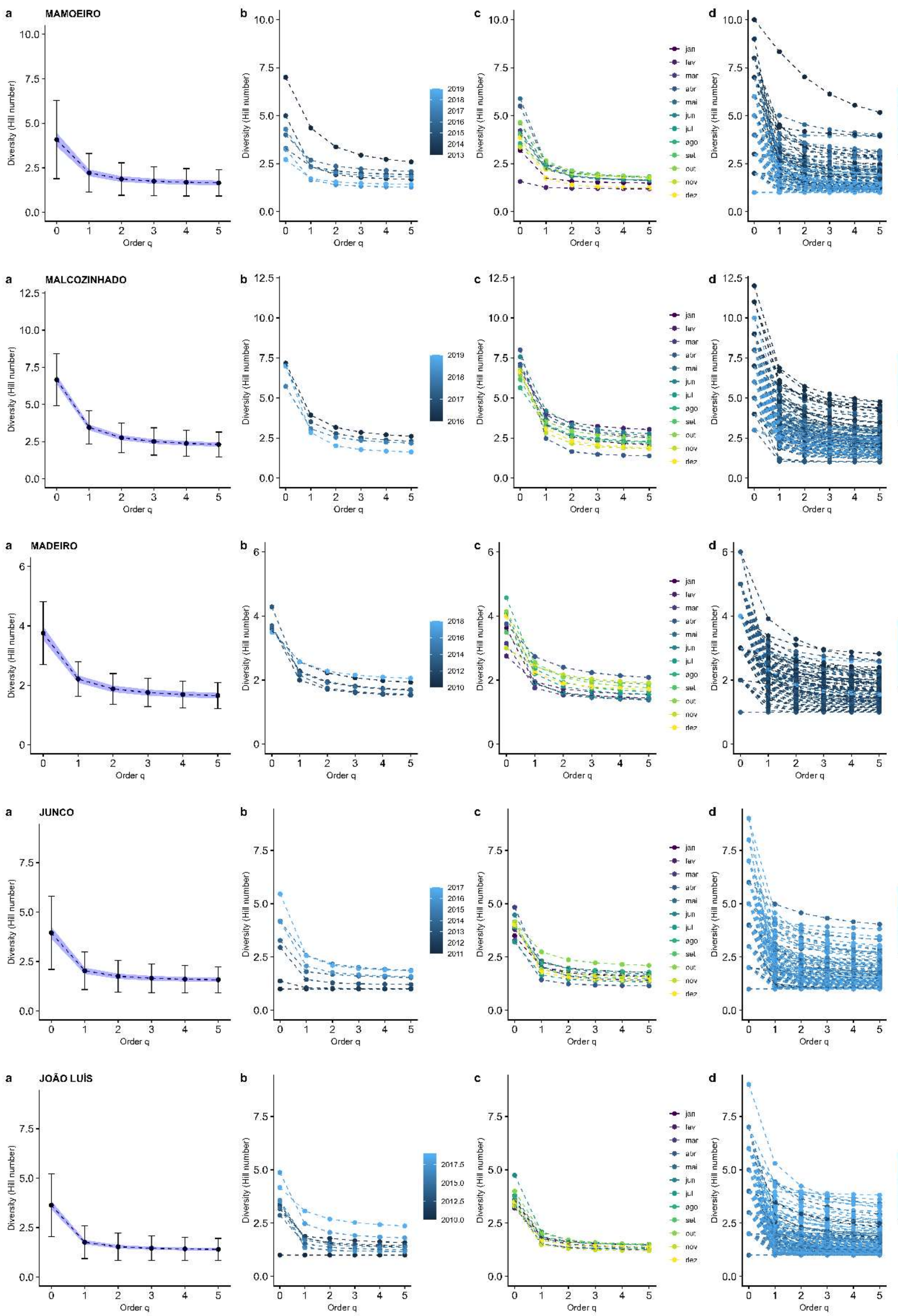

a RIVALDO DE CARVALHO
a RIACHÃO
a RIACHO DA SERRA
a QUANDÚ
a PUIÚ
b
c
d
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez

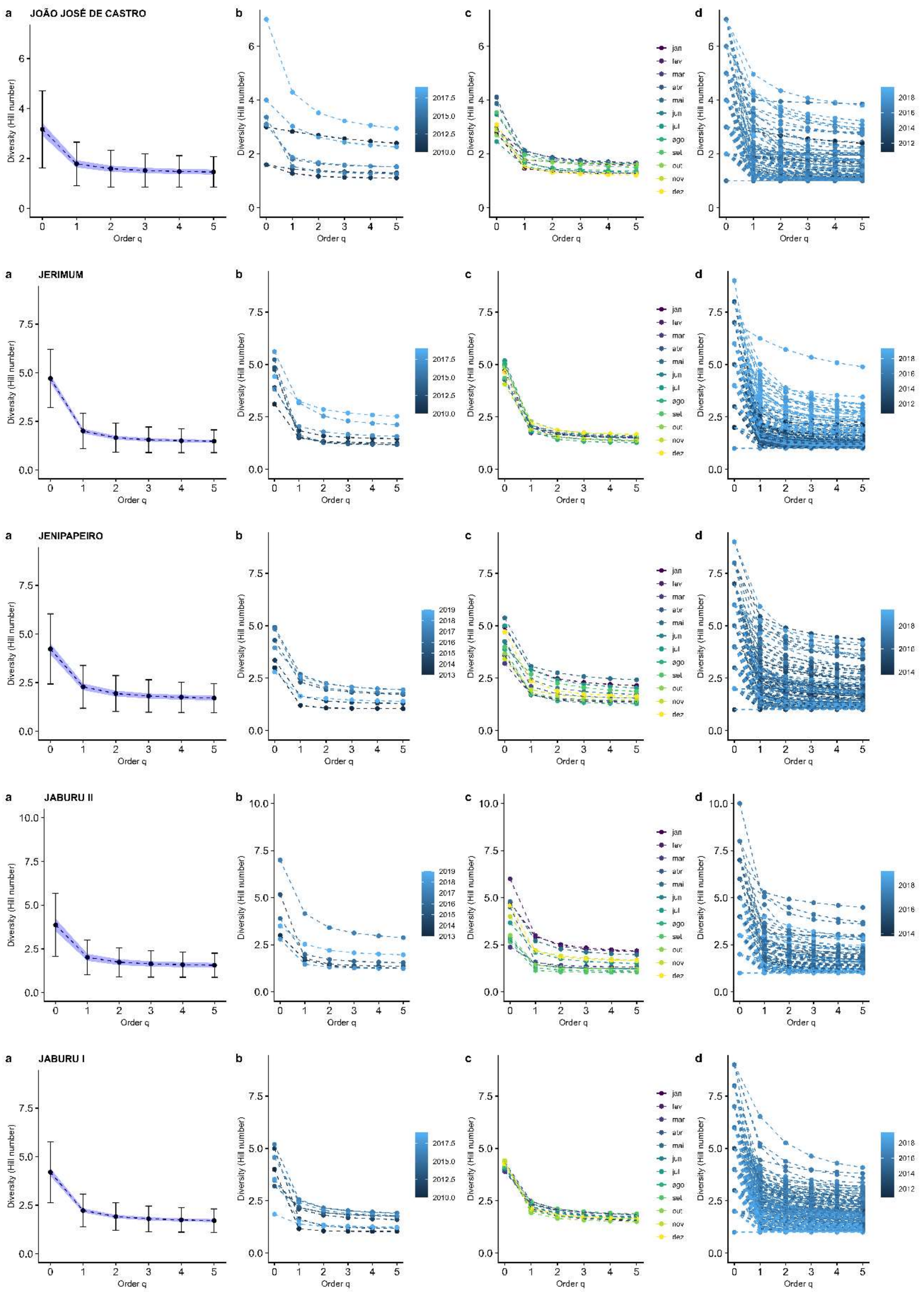

a POÇO DA PEDRA
b
c
d
a POÇO VERDE
b
c
d
a POTIRETAMA
b
c
d
a POMPEU SOBRINHO
b
c
d
a PESQUEIRO
b
c
d
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez

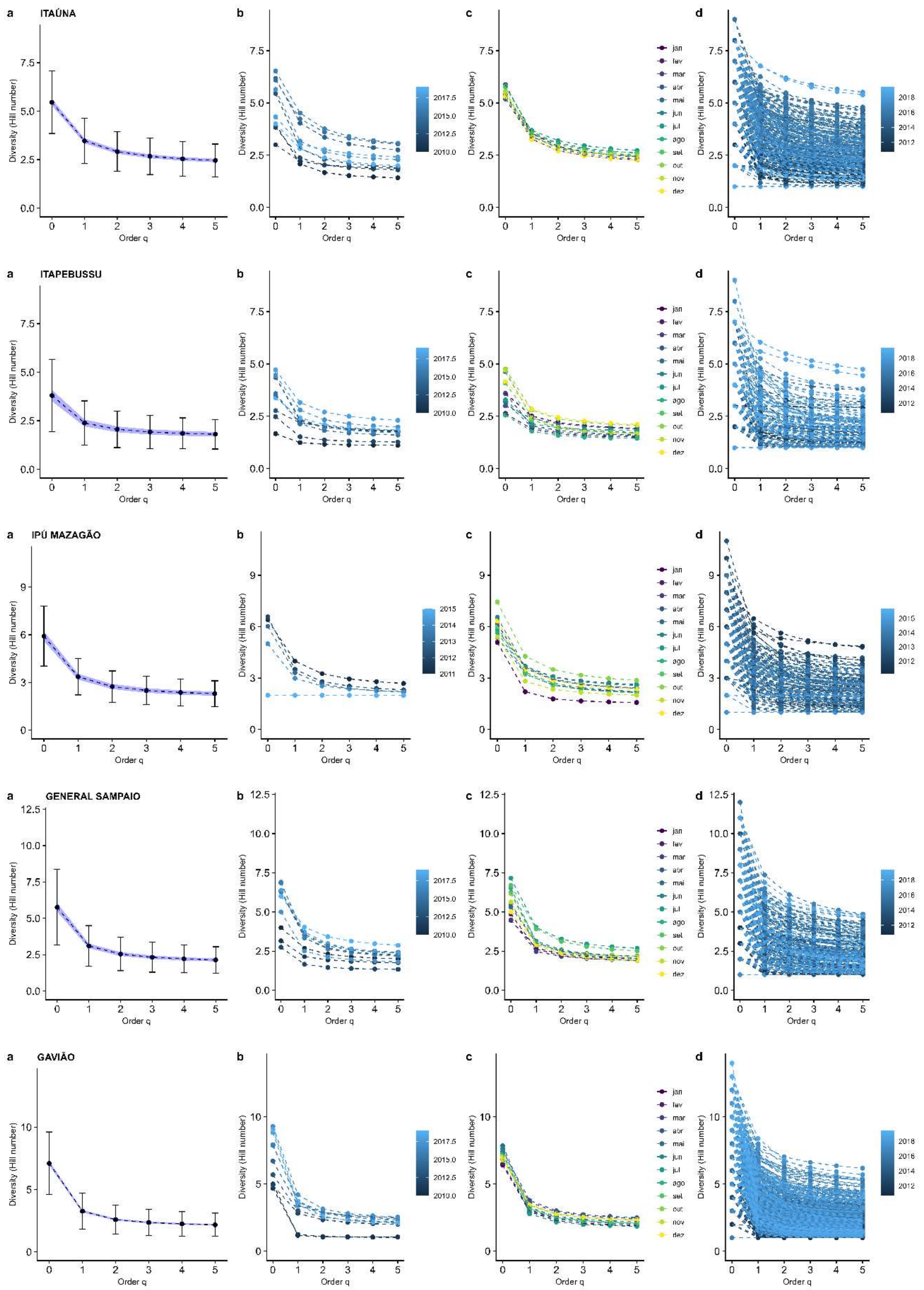


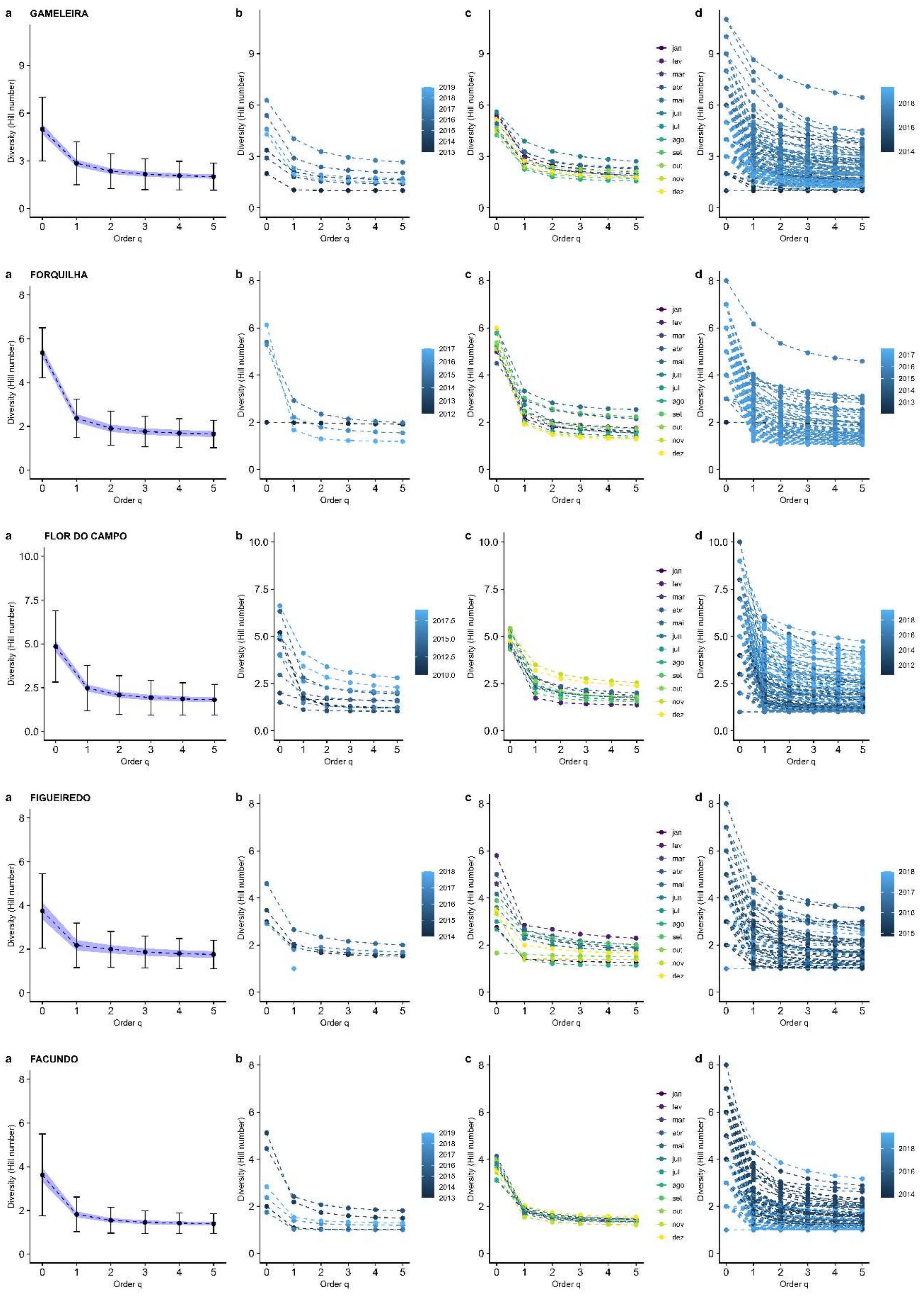
a PENTECOSTE
b
c
d
a PENEDO
b
c
d
a PEDRAS BRANCAS
b
c
d
a ARARAS
b
c
d
a PATU
b
c
d
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez


a PACOTI
b
c
d
a MUNDAÚ
a MONSENHOR TABOSA
a MISSÍ
a MARTINÓPOLE
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez

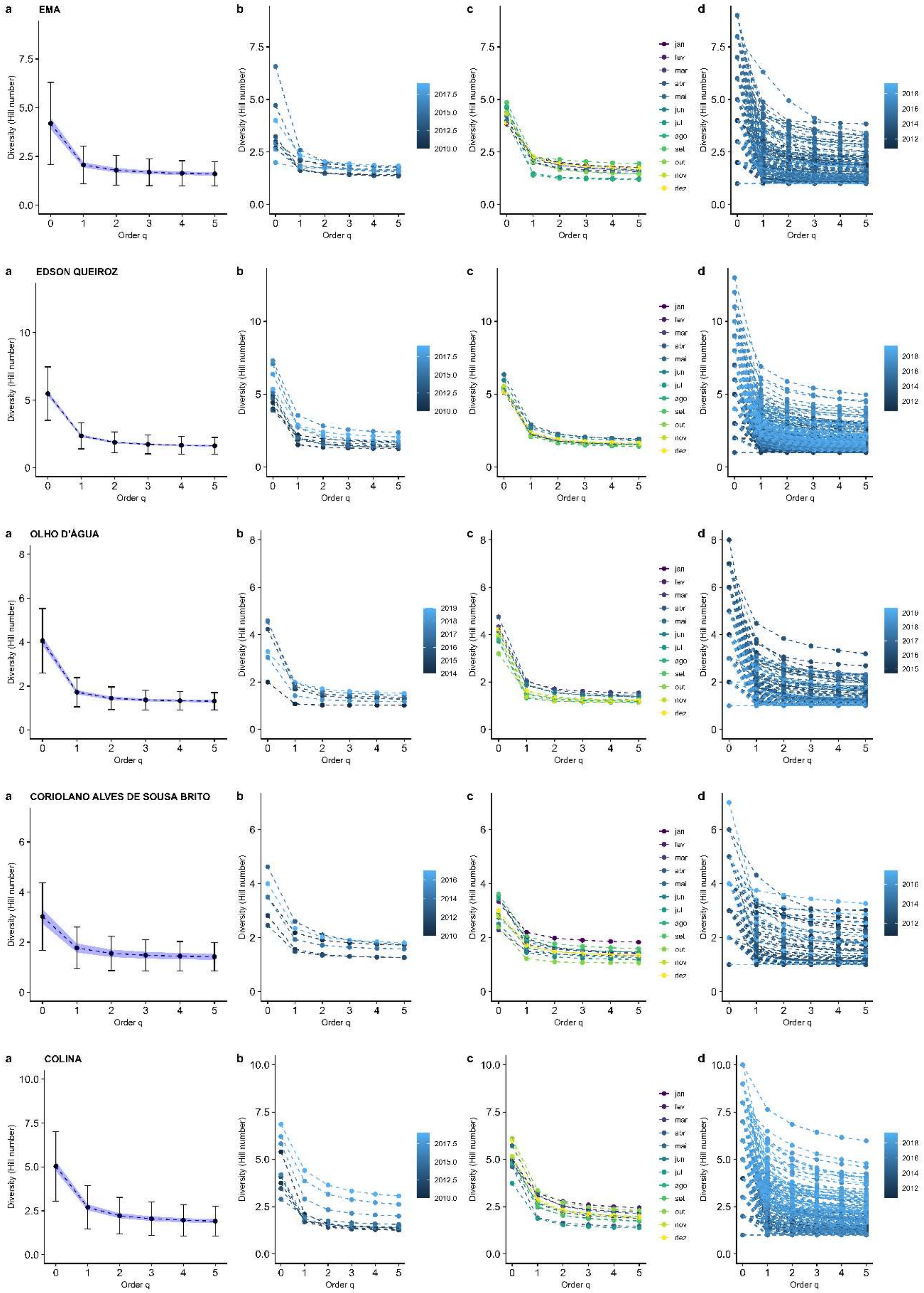

a
MAMOEIRO
b
c
d
MALCOZINHADO
MADEIRO
JUNCO
JOÃO LUÍS
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez
2019
2018
2017
2016
2015
2014
2013
2012
2011
2010
2017.5
2015.0
2012.5
2010.0

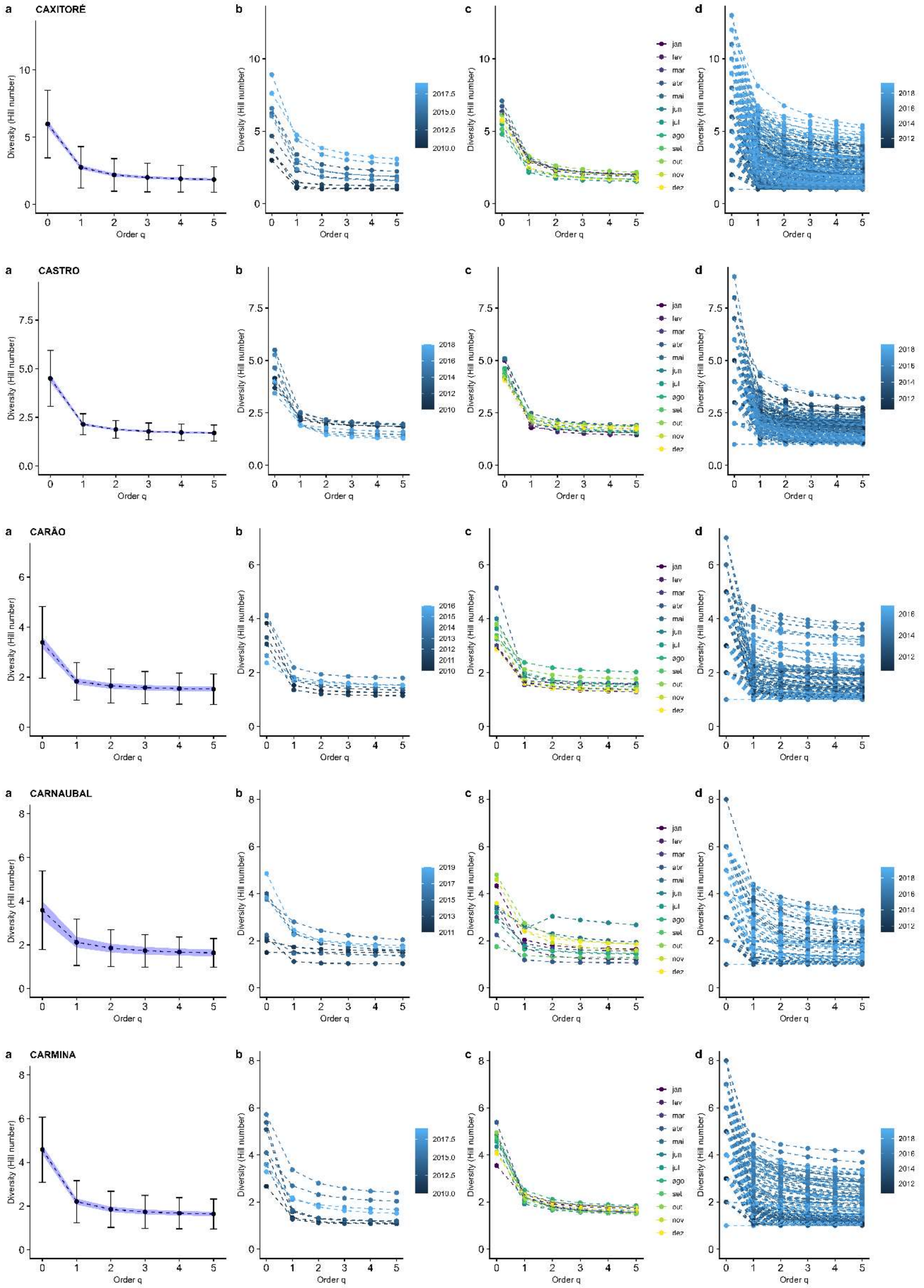

a JOÃO JOSÉ DE CASTRO
a JERIMUM
a JENIPAPEIRO
a JABURU II
a JABURU I
b
c
d
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez

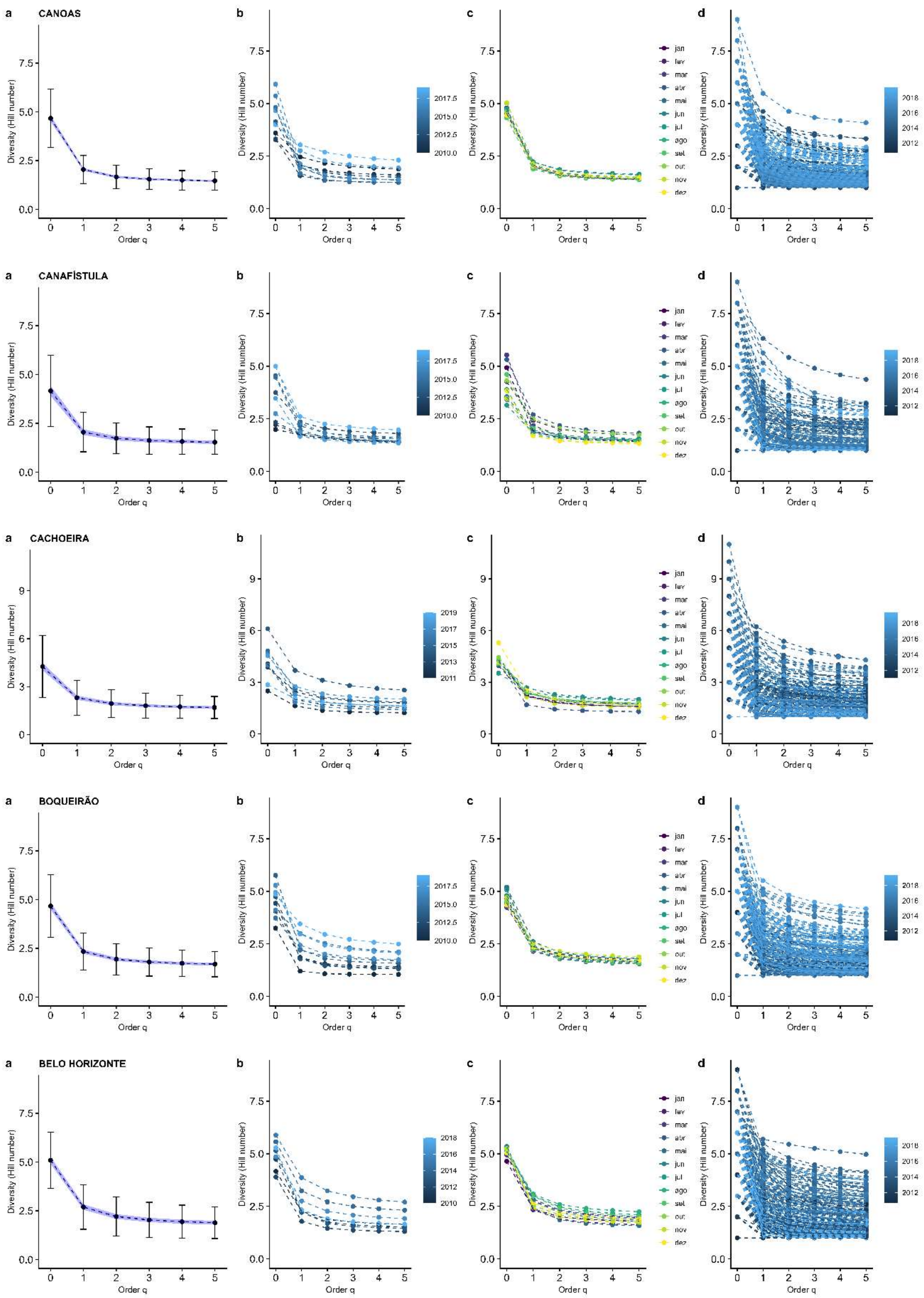

a ITAÚNA
a ITAPEBUSSU
a IPÚ MAZAGÃO
a GENERAL SAMPAIO
a GAVIÃO
b
c
d
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez
2010.0
2012.5
2015.0
2017.5
2011
2012
2013
2014
2015
2016
2018



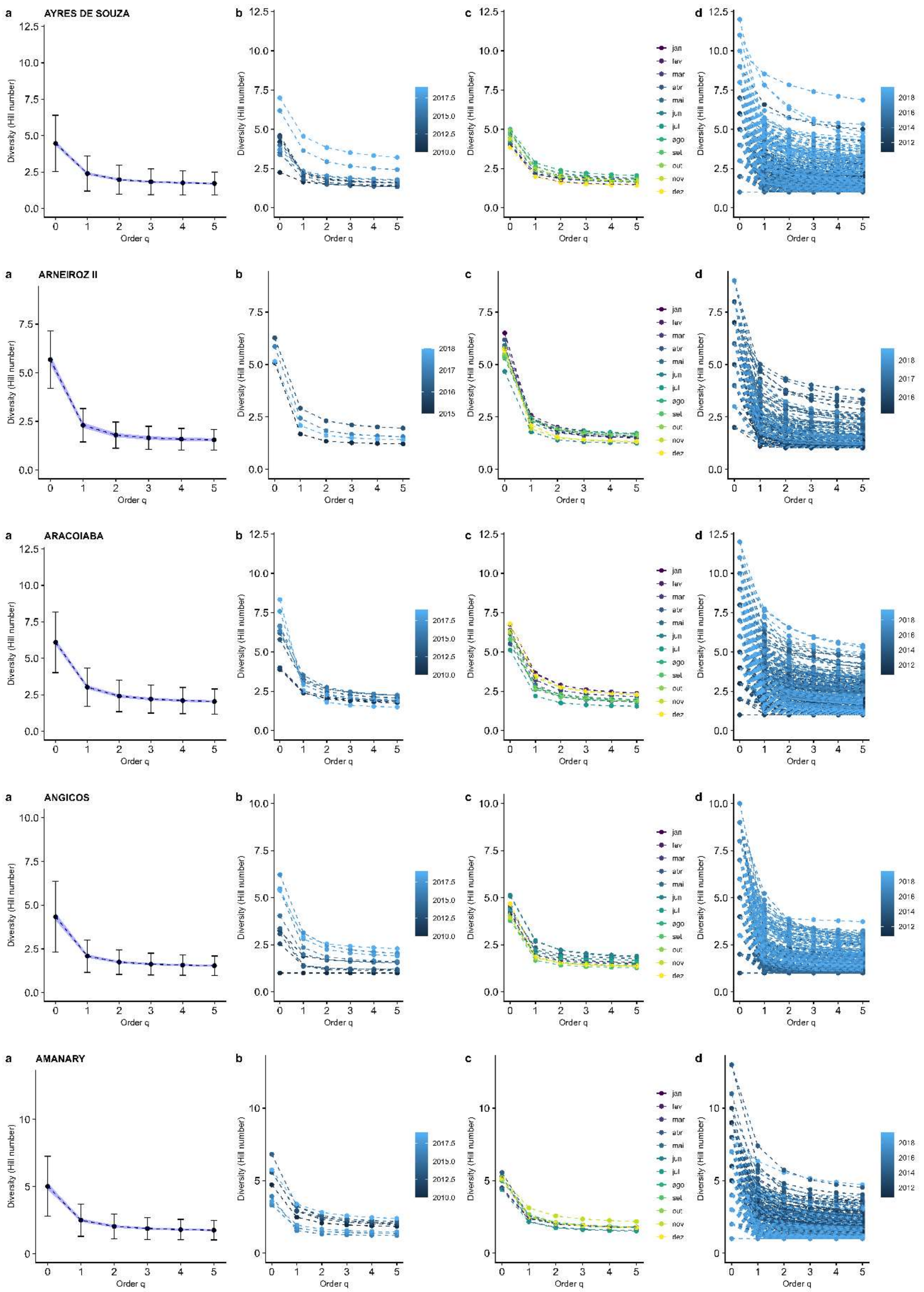
a GAMELEIRA
b
c
d
a FORQUILHA
a FLOR DO CAMPO
a FIGUEIREDO
a FACUNDO
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez

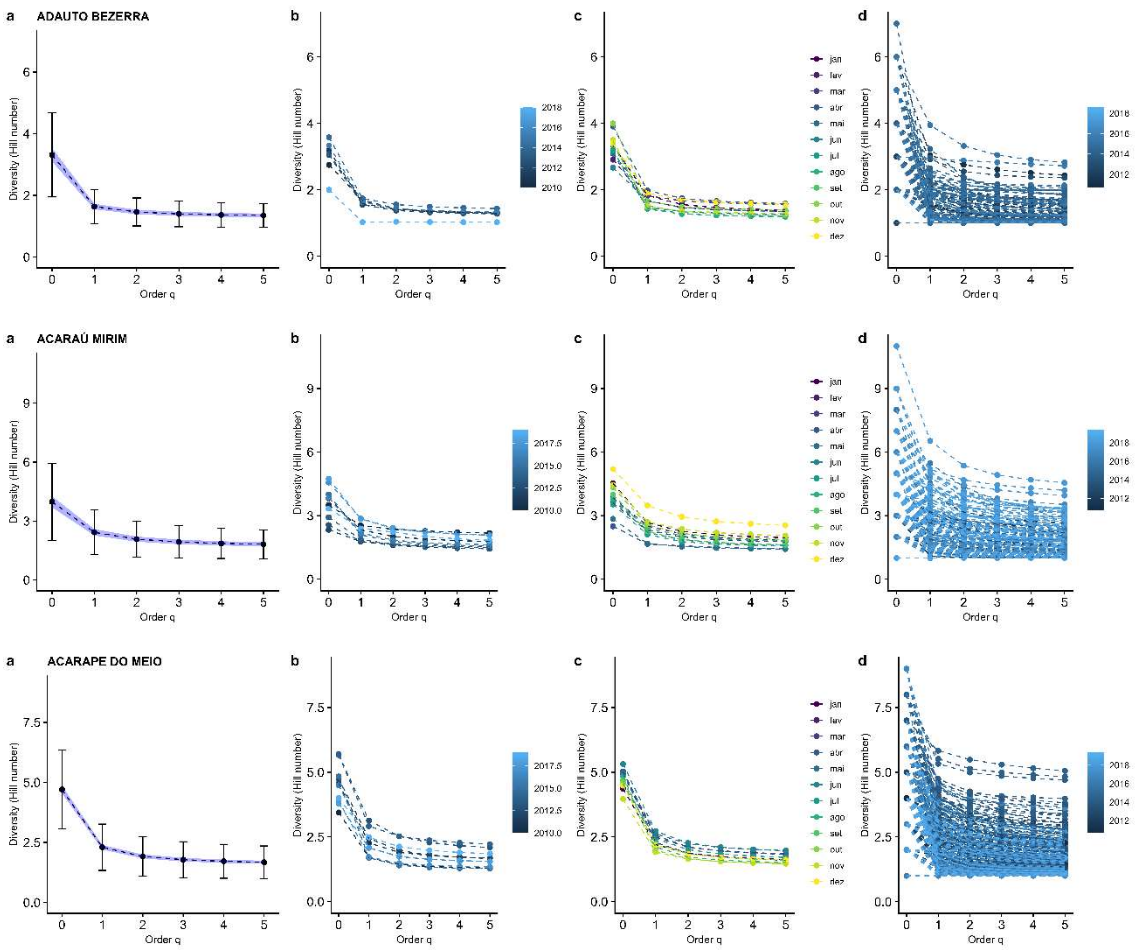

EMA
EDSON QUEIROZ
OLHO D'ÁGUA
CORIOLANO ALVES DE SOUSA BRITO
COLINA
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez


a CAXITORÉ
b
c
d
a CASTRO
b
c
d
a CARÃO
b
c
d
a CARNAUBAL
b
c
d
a CARMINA
b
c
d
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez

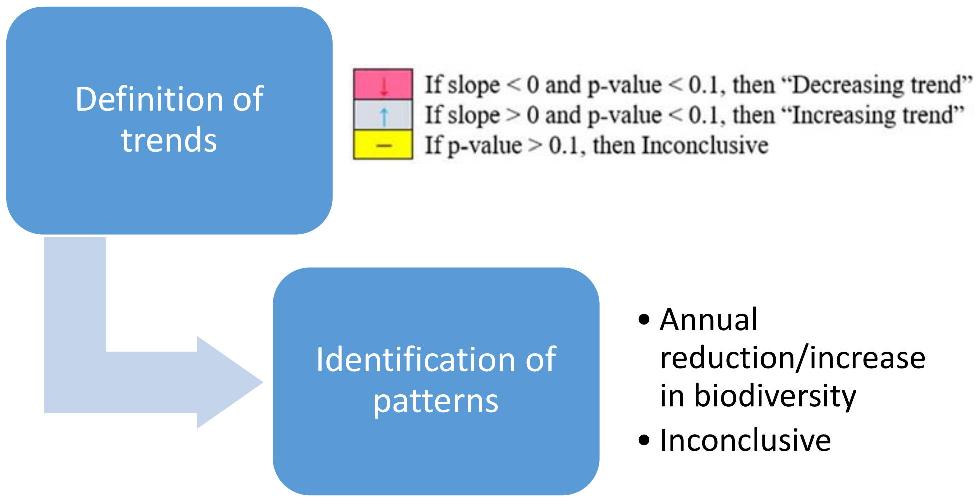


a
CANOAS
CANAFÍSTULA
CACHOEIRA
BOQUEIRÃO
BELO HORIZONTE
b
c
d
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez


a
AYRES DE SOUZA
b
c
d
ARNEIROZ II
ARACOIABA
ANGICOS
AMANARY
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez
2010.0
2012.5
2015.0
2017.5
2012
2014
2016
2018
2015
2017


a
ADAUTO BEZERRA
b
c
d
Diversity (Hill number)
Order q
jan
fev
mar
abr
mai
jun
jul
ago
set
out
nov
dez
2018
2016
2014
2012
2010
a
ACARAÚ MIRIM
b
c
d
Diversity (Hill number)
Order q
2017.5
2015.0
2012.5
2010.0
a
ACARAPE DO MEIO
b
c
d
Diversity (Hill number)
Order q

Annual Behavior of Hill numbers: long-term effect



Defining annual behavioral trends

Legend

↓ If slope < 0 and p-value < 0.1, then "Decreasing trend"

↑ If slope > 0 and p-value < 0.1, then "Increasing trend"

— If p-value > 0.1, then Inconclusive

Table D - Definition of the trend by linear regression of Hill numbers in the assessment of the annual behavior of the 82 reservoirs

Reservoir	Number	slope	Intercept	p_slope	p-value	Rsquared	Pattern
ACARAPE DO MEIO	0	-0.055	114.939	0.544	0.544	0	—
	1	-0.014	31.201	0.811	0.811	0	—
	2	-0.005	11.502	0.923	0.923	0	—
	3	-0.002	5.824	0.963	0.963	0	—
	4	-0.001	3.941	0.978	0.978	0	—
	5	-0.001	3.229	0.984	0.984	0	—
ACARAÚ MIRIM	0	0.127	-252.963	0.156	0.156	0.2	—
	1	0.056	-109.576	0.227	0.227	0.2	—
	2	0.038	-74.197	0.326	0.326	0.1	—
	3	0.025	-48.6	0.482	0.482	0.1	—
	4	0.019	-36.629	0.574	0.574	0	—
	5	0.016	-30.451	0.626	0.626	0	—
ADAUTO BEZERRA	0	-0.079	162.961	0.349	0.349	0.2	—
	1	-0.064	129.607	0.074	0.074	0.5	↓
	2	-0.038	77.483	0.131	0.131	0.4	—
	3	-0.031	62.784	0.171	0.171	0.3	—

	4	-0.027	55.521	0.198	0.198	0.3	—
	5	-0.025	51.376	0.215	0.215	0.3	—
AMANARY	0	-0.186	380.24	0.235	0.235	0.2	—
	1	-0.072	147.589	0.396	0.396	0.1	—
	2	-0.044	91.581	0.482	0.482	0.1	—
	3	-0.037	77.056	0.497	0.497	0.1	—
	4	-0.034	70.975	0.499	0.499	0.1	—
	5	-0.033	67.501	0.5	0.5	0.1	—
ANGICOS	0	0.492	-986.873	0	0	0.8	↑
	1	0.218	-437.029	0.001	0.001	0.8	↑
	2	0.16	-320.86	0.001	0.001	0.8	↑
	3	0.139	-279.053	0.001	0.001	0.7	↑
	4	0.129	-257.413	0.001	0.001	0.7	↑
	5	0.122	-243.986	0.002	0.002	0.7	↑
ARACOIABA	0	0.485	-972.081	0	0	0.9	↑
	1	0.047	-91.103	0.324	0.324	0.1	—
	2	0.002	-1.936	0.955	0.955	0	—
	3	-0.006	14.583	0.851	0.851	0	—
	4	-0.009	19.413	0.778	0.778	0	—
	5	-0.009	21.054	0.746	0.746	0	—
ARNEIROZ II	0	-0.024	54.374	0.946	0.946	0	—
	1	0.075	-149.429	0.814	0.814	0	—
	2	0.036	-69.923	0.888	0.888	0	—
	3	0.024	-46.708	0.914	0.914	0	—
	4	0.02	-37.938	0.924	0.924	0	—
	5	0.017	-33.646	0.929	0.929	0	—
AYRES DE SOUZA	0	0.299	-597.463	0.04	0.04	0.4	↑
	1	0.23	-460.266	0.009	0.009	0.6	↑
	2	0.189	-378.91	0.008	0.008	0.6	↑
	3	0.169	-337.521	0.009	0.009	0.6	↑
	4	0.157	-313.465	0.009	0.009	0.6	↑
	5	0.149	-297.702	0.009	0.009	0.6	↑
BELO HORIZONTE	0	0.155	-307.271	0.106	0.106	0.4	—
	1	0.1	-197.973	0.331	0.331	0.2	—
	2	0.086	-171.735	0.343	0.343	0.2	—
	3	0.078	-154.404	0.345	0.345	0.1	—
	4	0.072	-143.024	0.347	0.347	0.1	—
	5	0.068	-135.293	0.35	0.35	0.1	—
BOQUEIRÃO	0	0.127	-251.385	0.132	0.132	0.3	—
	1	0.21	-420.301	0	0	0.9	↑
	2	0.176	-352.236	0	0	0.9	↑
	3	0.155	-310.402	0	0	0.9	↑
	4	0.143	-285.735	0	0	0.9	↑
	5	0.135	-269.881	0	0	0.9	↑
CACHOEIRA	0	0.04	-76.909	0.792	0.792	0	—

	1	0.022	-42.783	0.798	0.798	0	—
	2	0.023	-44.148	0.754	0.754	0	—
	3	0.023	-45.426	0.71	0.71	0	—
	4	0.023	-44.069	0.694	0.694	0	—
	5	0.022	-42.368	0.688	0.688	0	—
CANAFÍSTULA	0	0.252	-503.421	0.042	0.042	0.4	↑
	1	0.044	-87.488	0.227	0.227	0.2	—
	2	0.028	-54.626	0.316	0.316	0.1	—
	3	0.023	-43.836	0.376	0.376	0.1	—
	4	0.02	-39.13	0.399	0.399	0.1	—
	5	0.019	-36.305	0.412	0.412	0.1	—
CANOAS	0	0.182	-362.281	0.052	0.052	0.4	↑
	1	0.073	-145.415	0.158	0.158	0.2	—
	2	0.052	-103.551	0.301	0.301	0.1	—
	3	0.045	-88.765	0.334	0.334	0.1	—
	4	0.041	-81.289	0.342	0.342	0.1	—
	5	0.039	-76.761	0.344	0.344	0.1	—
CARMINA	0	0.036	-68.364	0.782	0.782	0	—
	1	0.139	-278.048	0.075	0.075	0.4	↑
	2	0.119	-237.949	0.07	0.07	0.4	↑
	3	0.104	-208.327	0.079	0.079	0.4	↑
	4	0.096	-191.33	0.086	0.086	0.4	↑
	5	0.09	-180.71	0.09	0.09	0.4	↑
CARNAUBAL	0	0.369	-740.281	0.006	0.006	0.7	↑
	1	0.136	-271.142	0.078	0.078	0.4	↑
	2	0.092	-183.018	0.111	0.111	0.4	—
	3	0.071	-142.062	0.148	0.148	0.3	—
	4	0.061	-120.607	0.176	0.176	0.3	—
	5	0.054	-107.732	0.198	0.198	0.3	—
CARÃO	0	-0.158	321.737	0.274	0.274	0.2	—
	1	0.071	-141.407	0.158	0.158	0.4	—
	2	0.083	-165.031	0.052	0.052	0.6	↑
	3	0.079	-156.898	0.051	0.051	0.6	↑
	4	0.075	-150.294	0.055	0.055	0.6	↑
	5	0.073	-145.396	0.059	0.059	0.5	↑
CASTRO	0	0.021	-38.803	0.831	0.831	0	—
	1	-0.057	117.413	0.076	0.076	0.4	↓
	2	-0.074	151.343	0.024	0.024	0.5	↓
	3	-0.079	161.289	0.013	0.013	0.6	↓
	4	-0.08	161.933	0.01	0.01	0.6	↓
	5	-0.079	160.359	0.009	0.009	0.7	↓
CAXITORÉ	0	0.581	-1164.36	0	0	0.9	↑
	1	0.373	-748.711	0	0	0.8	↑
	2	0.273	-546.835	0.001	0.001	0.8	↑
	3	0.232	-465.59	0.001	0.001	0.8	↑

	4	0.211	-422.953	0.001	0.001	0.8	↑
	5	0.198	-396.764	0.001	0.001	0.8	↑
COLINA	0	0.24	-477.863	0.116	0.116	0.3	—
	1	0.252	-504.467	0.01	0.01	0.6	↑
	2	0.21	-420.752	0.007	0.007	0.7	↑
	3	0.189	-378.57	0.006	0.006	0.7	↑
	4	0.177	-354.36	0.006	0.006	0.7	↑
	5	0.169	-338.959	0.005	0.005	0.7	↑
CORIOLANO ALVES DE SOUSA BRITO	0	0.142	-283.235	0.369	0.369	0.2	—
	1	0.071	-140.462	0.432	0.432	0.2	—
	2	0.056	-111.122	0.398	0.398	0.2	—
	3	0.051	-101.191	0.367	0.367	0.2	—
	4	0.048	-95.893	0.352	0.352	0.2	—
	5	0.047	-92.497	0.343	0.343	0.2	—
OLHO D'ÁGUA	0	0.082	-160.757	0.778	0.778	0	—
	1	0.112	-223.629	0.217	0.217	0.3	—
	2	0.088	-175.149	0.163	0.163	0.4	—
	3	0.075	-150.178	0.156	0.156	0.4	—
	4	0.069	-137.522	0.154	0.154	0.4	—
	5	0.065	-130.414	0.152	0.152	0.4	—
EDSON QUEIROZ	0	0.29	-579.592	0.018	0.018	0.5	↑
	1	0.153	-305.941	0.009	0.009	0.6	↑
	2	0.102	-202.951	0.03	0.03	0.5	↑
	3	0.084	-167.402	0.046	0.046	0.4	↑
	4	0.076	-150.596	0.055	0.055	0.4	↑
	5	0.071	-141.016	0.059	0.059	0.4	↑
EMA	0	0.019	-35.24	0.904	0.904	0	—
	1	0.038	-74.615	0.322	0.322	0.1	—
	2	0.041	-81.826	0.077	0.077	0.3	↑
	3	0.036	-71.707	0.08	0.08	0.3	↑
	4	0.034	-66.045	0.086	0.086	0.3	↑
	5	0.032	-62.504	0.092	0.092	0.3	↑
FACUNDO	0	-0.203	412.959	0.499	0.499	0.1	—
	1	-0.058	117.738	0.649	0.649	0.1	—
	2	-0.034	70.754	0.71	0.71	0	—
	3	-0.026	54.522	0.74	0.74	0	—
	4	-0.023	47.711	0.753	0.753	0	—
	5	-0.021	44.216	0.759	0.759	0	—
FIGUEIREDO	0	0.088	-173.951	0.854	0.854	0	—
	1	-0.202	408.728	0.348	0.348	0.3	—
	2	0.13	-260.81	0.433	0.433	0.3	—
	3	0.114	-228.768	0.415	0.415	0.3	—
	4	0.104	-206.972	0.414	0.414	0.3	—
	5	0.096	-191.439	0.418	0.418	0.3	—
FLOR DO CAMPO	0	0.15	-297.833	0.458	0.458	0.1	—

	1	0.223	-447.724	0.021	0.021	0.5	↑
	2	0.187	-375.199	0.012	0.012	0.6	↑
	3	0.164	-328.97	0.012	0.012	0.6	↑
	4	0.15	-300.722	0.013	0.013	0.6	↑
	5	0.141	-282.478	0.013	0.013	0.6	↑
FORQUILHA	0	0.824	-1655.616	0.032	0.032	0.9	↑
	1	-0.028	57.975	0.887	0.887	0	—
	2	-0.108	218.635	0.468	0.468	0.3	—
	3	-0.123	250.471	0.33	0.33	0.4	—
	4	-0.128	260.424	0.268	0.268	0.5	—
	5	-0.13	263.355	0.236	0.236	0.6	—
GAMELEIRA	0	0.453	-909.659	0.104	0.104	0.4	—
	1	0.191	-382.528	0.327	0.327	0.2	—
	2	0.138	-275.915	0.355	0.355	0.2	—
	3	0.119	-237.338	0.353	0.353	0.2	—
	4	0.109	-217.846	0.349	0.349	0.2	—
	5	0.103	-206.106	0.345	0.345	0.2	—
GAVIÃO	0	0.577	-1156.099	0	0	0.9	↑
	1	0.288	-576.335	0.002	0.002	0.7	↑
	2	0.178	-356.708	0.015	0.015	0.5	↑
	3	0.144	-287.464	0.026	0.026	0.5	↑
	4	0.128	-256.494	0.031	0.031	0.5	↑
	5	0.12	-239.321	0.035	0.035	0.4	↑
GENERAL SAMPAIO	0	0.327	-653.152	0.047	0.047	0.4	↑
	1	0.164	-328.243	0.04	0.04	0.4	↑
	2	0.123	-246.065	0.042	0.042	0.4	↑
	3	0.103	-205.694	0.05	0.05	0.4	↑
	4	0.093	-184.298	0.057	0.057	0.4	↑
	5	0.086	-171.456	0.061	0.061	0.4	↑
IPÚ MAZAGÃO	0	-1.038	2093.687	0.056	0.056	0.8	↓
	1	-0.452	912.107	0.01	0.01	0.9	↓
	2	-0.278	562.962	0.008	0.008	0.9	↓
	3	-0.208	421.747	0.012	0.012	0.9	↓
	4	-0.172	348.519	0.019	0.019	0.9	↓
	5	-0.15	304.724	0.028	0.028	0.8	↓
ITAPEBUSSU	0	0.198	-395.625	0.065	0.065	0.4	↑
	1	0.143	-285.099	0.007	0.007	0.6	↑
	2	0.11	-220.013	0.01	0.01	0.6	↑
	3	0.095	-190.538	0.013	0.013	0.6	↑
	4	0.087	-173.935	0.015	0.015	0.5	↑
	5	0.082	-163.25	0.016	0.016	0.5	↑
ITAÚNA	0	0.131	-258.07	0.358	0.358	0.1	—
	1	0.131	-260.545	0.181	0.181	0.2	—
	2	0.117	-232.472	0.145	0.145	0.2	—
	3	0.104	-206.928	0.144	0.144	0.2	—

	4	0.096	-191.705	0.142	0.142	0.2	—
	5	0.091	-181.804	0.141	0.141	0.2	—
JABURU I	0	-0.104	213.8	0.373	0.373	0.1	—
	1	0.027	-53.402	0.642	0.642	0	—
	2	0.027	-52.756	0.584	0.584	0	—
	3	0.024	-47.017	0.588	0.588	0	—
	4	0.022	-43.002	0.598	0.598	0	—
	5	0.021	-40.316	0.606	0.606	0	—
JABURU II	0	0.021	-37.18	0.958	0.958	0	—
	1	0.139	-277.184	0.56	0.56	0.1	—
	2	0.134	-268.856	0.477	0.477	0.1	—
	3	0.126	-253.103	0.452	0.452	0.1	—
	4	0.121	-241.553	0.442	0.442	0.2	—
	5	0.117	-233.249	0.437	0.437	0.2	—
JENIPAPEIRO	0	0.04	-76.368	0.831	0.831	0	—
	1	0.12	-239.617	0.309	0.309	0.2	—
	2	0.11	-219.56	0.22	0.22	0.3	—
	3	0.097	-194.101	0.219	0.219	0.3	—
	4	0.09	-179.606	0.219	0.219	0.3	—
	5	0.085	-170.328	0.219	0.219	0.3	—
JERIMUM	0	0.109	-215.322	0.265	0.265	0.2	—
	1	0.184	-368.991	0.007	0.007	0.7	↑
	2	0.155	-311.266	0.008	0.008	0.7	↑
	3	0.139	-277.792	0.01	0.01	0.7	↑
	4	0.129	-259.046	0.012	0.012	0.7	↑
	5	0.124	-247.25	0.012	0.012	0.7	↑
JOÃO JOSÉ DE CASTRO	0	0.378	-757.947	0.032	0.032	0.6	↑
	1	0.157	-313.32	0.2	0.2	0.3	—
	2	0.107	-212.788	0.291	0.291	0.2	—
	3	0.089	-178.249	0.325	0.325	0.2	—
	4	0.082	-162.598	0.335	0.335	0.2	—
	5	0.077	-153.933	0.336	0.336	0.2	—
JOÃO LUÍS	0	0.259	-517.596	0.007	0.007	0.6	↑
	1	0.122	-242.99	0.056	0.056	0.4	↑
	2	0.091	-182.302	0.089	0.089	0.3	↑
	3	0.083	-164.806	0.087	0.087	0.3	↑
	4	0.078	-155.247	0.085	0.085	0.3	↑
	5	0.075	-148.786	0.084	0.084	0.3	↑
JUNCO	0	0.721	-1449.589	0	0	1	↑
	1	0.281	-564.264	0.003	0.003	0.9	↑
	2	0.202	-406.179	0.005	0.005	0.8	↑
	3	0.176	-352.833	0.005	0.005	0.8	↑
	4	0.163	-326.493	0.004	0.004	0.8	↑
	5	0.155	-310.751	0.004	0.004	0.8	↑
MADEIRO	0	-0.012	27.661	0.824	0.824	0	—

	1	0.021	-39.928	0.622	0.622	0.1	—
	2	0.026	-50.185	0.547	0.547	0.1	—
	3	0.029	-55.65	0.483	0.483	0.1	—
	4	0.03	-58.974	0.438	0.438	0.2	—
	5	0.031	-61.206	0.407	0.407	0.2	—
MALCOZINHADO	0	-0.179	367.713	0.656	0.656	0.1	—
	1	-0.369	748.153	0.007	0.007	1	↓
	2	-0.368	745.211	0.011	0.011	1	↓
	3	-0.341	689.577	0.024	0.024	1	↓
	4	-0.32	647.559	0.031	0.031	0.9	↓
	5	-0.305	618.292	0.035	0.035	0.9	↓
MAMOEIRO	0	-0.549	1110.36	0.023	0.023	0.7	↓
	1	-0.334	675.408	0.031	0.031	0.6	↓
	2	-0.237	480.113	0.035	0.035	0.6	↓
	3	-0.194	393.218	0.037	0.037	0.6	↓
	4	-0.172	347.913	0.038	0.038	0.6	↓
	5	-0.158	321.245	0.039	0.039	0.6	↓
MARTINÓPOLE	0	-0.24	488.887	0.023	0.023	0.5	↓
	1	-0.135	275.527	0.049	0.049	0.4	↓
	2	-0.112	227.409	0.069	0.069	0.4	↓
	3	-0.101	205.883	0.075	0.075	0.4	↓
	4	-0.095	193.551	0.078	0.078	0.4	↓
	5	-0.091	185.419	0.08	0.08	0.4	↓
MISSÍ	0	-0.116	239.226	0.574	0.574	0.1	—
	1	0.367	-737.999	0.009	0.009	0.8	↑
	2	0.232	-464.937	0.013	0.013	0.8	↑
	3	0.211	-423.291	0.015	0.015	0.8	↑
	4	0.197	-394.025	0.018	0.018	0.8	↑
	5	0.187	-374.146	0.02	0.02	0.8	↑
MONSENHOR TABOSA	0	0.126	-248.699	0.406	0.406	0.1	—
	1	0.068	-136.071	0.018	0.018	0.6	↑
	2	0.055	-110.017	0.009	0.009	0.7	↑
	3	0.049	-98.023	0.008	0.008	0.7	↑
	4	0.046	-91.439	0.008	0.008	0.7	↑
	5	0.044	-87.337	0.008	0.008	0.7	↑
MUNDAÚ	0	0.34	-680.539	0.003	0.003	0.7	↑
	1	0.1	-200.041	0.034	0.034	0.4	↑
	2	0.053	-104.934	0.146	0.146	0.2	—
	3	0.037	-73.391	0.257	0.257	0.2	—
	4	0.03	-58.069	0.341	0.341	0.1	—
	5	0.025	-49.194	0.402	0.402	0.1	—
PACOTI	0	0.378	-754.738	0	0	0.8	↑
	1	0.187	-373.437	0.101	0.101	0.3	—
	2	0.116	-231.27	0.189	0.189	0.2	—
	3	0.092	-183.596	0.225	0.225	0.2	—

	4	0.081	-161.532	0.242	0.242	0.2	—
	5	0.075	-149.15	0.251	0.251	0.2	—
PATU	0	0.387	-773.284	0.009	0.009	0.7	↑
	1	0.276	-552.482	0.003	0.003	0.8	↑
	2	0.21	-421.666	0.004	0.004	0.8	↑
	3	0.181	-362.877	0.004	0.004	0.8	↑
	4	0.166	-332.054	0.004	0.004	0.8	↑
	5	0.156	-313.266	0.004	0.004	0.8	↑
ARARAS	0	0.619	-1240.584	0	0	0.9	↑
	1	0.456	-915.061	0	0	1	↑
	2	0.348	-698.164	0	0	1	↑
	3	0.301	-603.504	0	0	1	↑
	4	0.276	-553.083	0	0	1	↑
	5	0.26	-522.344	0	0	1	↑
PEDRAS BRANCAS	0	0.112	-221.812	0.158	0.158	0.2	—
	1	0.053	-104.745	0.181	0.181	0.2	—
	2	0.024	-46.549	0.437	0.437	0.1	—
	3	0.013	-24.843	0.643	0.643	0	—
	4	0.008	-14.649	0.765	0.765	0	—
	5	0.006	-9.114	0.839	0.839	0	—
PENEDO	0	0.086	-168.718	0.489	0.489	0.1	—
	1	0.096	-191.186	0.164	0.164	0.3	—
	2	0.087	-173.888	0.095	0.095	0.3	↑
	3	0.078	-155.752	0.083	0.083	0.4	↑
	4	0.072	-143.726	0.08	0.08	0.4	↑
	5	0.068	-135.802	0.078	0.078	0.4	↑
PENTECOSTE	0	0.097	-190.201	0.449	0.449	0.1	—
	1	0.261	-523.038	0.008	0.008	0.7	↑
	2	0.195	-390.867	0.004	0.004	0.8	↑
	3	0.162	-325.158	0.003	0.003	0.8	↑
	4	0.145	-290.406	0.002	0.002	0.8	↑
	5	0.135	-269.987	0.002	0.002	0.8	↑
PESQUEIRO	0	-0.019	42.516	0.719	0.719	0	—
	1	0.054	-107.24	0.452	0.452	0.1	—
	2	0.024	-46.574	0.665	0.665	0	—
	3	0.008	-14.797	0.861	0.861	0	—
	4	0	1.291	0.995	0.995	0	—
	5	-0.004	10.15	0.918	0.918	0	—
POMPEU SOBRINHO	0	0.001	0.093	0.983	0.983	0	—
	1	-0.01	20.809	0.699	0.699	0	—
	2	0.011	-21.633	0.568	0.568	0	—
	3	0.011	-20.964	0.556	0.556	0.1	—
	4	0.011	-20.138	0.558	0.558	0.1	—
	5	0.01	-19.529	0.559	0.559	0.1	—
POTIRETAMA	0	0.037	-70.98	0.764	0.764	0	—

	1	-0.097	197.159	0.067	0.067	0.7	↓
	2	-0.088	178.623	0.048	0.048	0.8	↓
	3	-0.081	164.778	0.041	0.041	0.8	↓
	4	-0.076	154.828	0.039	0.039	0.8	↓
	5	-0.073	147.702	0.039	0.039	0.8	↓
POÇO VERDE	0	0.096	-190.709	0.439	0.439	0.1	—
	1	0.049	-97.874	0.353	0.353	0.1	—
	2	0.035	-68.467	0.398	0.398	0.1	—
	3	0.03	-59.271	0.407	0.407	0.1	—
	4	0.028	-54.729	0.409	0.409	0.1	—
	5	0.027	-51.968	0.411	0.411	0.1	—
POÇO DA PEDRA	0	0.088	-173.696	0.455	0.455	0.1	—
	1	0.126	-250.948	0.096	0.096	0.3	↑
	2	0.107	-214.475	0.067	0.067	0.4	↑
	3	0.096	-190.802	0.061	0.061	0.4	↑
	4	0.088	-176.105	0.059	0.059	0.4	↑
	5	0.083	-166.419	0.058	0.058	0.4	↑
PUIÚ	0	0.061	-117.749	0.567	0.567	0	—
	1	0.212	-425.164	0.008	0.008	0.7	↑
	2	0.19	-380.986	0.005	0.005	0.7	↑
	3	0.171	-342.913	0.004	0.004	0.7	↑
	4	0.158	-317.552	0.004	0.004	0.7	↑
	5	0.15	-299.914	0.004	0.004	0.7	↑
QUANDÚ	0	0.072	-138.875	0.296	0.296	0.1	—
	1	0.036	-70.662	0.428	0.428	0.1	—
	2	0.032	-62.05	0.342	0.342	0.1	—
	3	0.029	-55.771	0.313	0.313	0.1	—
	4	0.027	-51.666	0.302	0.302	0.1	—
	5	0.025	-48.932	0.297	0.297	0.1	—
RIACHO DA SERRA	0	0.188	-374.688	0.752	0.752	0.1	—
	1	-0.011	23.1	0.977	0.977	0	—
	2	-0.041	83.761	0.884	0.884	0	—
	3	-0.043	88.753	0.858	0.858	0	—
	4	-0.043	87.632	0.847	0.847	0	—
	5	-0.042	85.336	0.843	0.843	0	—
RIACHÃO	0	0.481	-963.771	0	0	0.9	↑
	1	0.251	-502.246	0.015	0.015	0.5	↑
	2	0.168	-335.375	0.042	0.042	0.4	↑
	3	0.14	-280.938	0.051	0.051	0.4	↑
	4	0.128	-255.646	0.055	0.055	0.4	↑
	5	0.121	-241.046	0.056	0.056	0.4	↑
RIVALDO DE CARVALHO	0	-0.187	381.368	0.075	0.075	0.3	↓
	1	0.024	-47.642	0.361	0.361	0.1	—
	2	0.03	-59.551	0.207	0.207	0.2	—
	3	0.029	-56.42	0.199	0.199	0.2	—

	4	0.027	-53.529	0.2	0.2	0.2	—
	5	0.026	-51.381	0.202	0.202	0.2	—
ROSÁRIO	0	-0.129	262.308	0.432	0.432	0.1	—
	1	-0.12	243.711	0.05	0.05	0.5	↓
	2	-0.09	182.286	0.034	0.034	0.6	↓
	3	-0.075	153.098	0.034	0.034	0.6	↓
	4	-0.068	138.052	0.035	0.035	0.6	↓
	5	-0.063	129.213	0.035	0.035	0.5	↓
SANTO ANTÔNIO DE RUSSAS	0	0.44	-882.288	0.001	0.001	0.9	↑
	1	0.237	-474.753	0	0	0.9	↑
	2	0.181	-362.259	0.001	0.001	0.9	↑
	3	0.159	-318.824	0.001	0.001	0.9	↑
	4	0.148	-296.654	0.001	0.001	0.8	↑
	5	0.142	-283.168	0.001	0.001	0.8	↑
SEBASTIÃO DE ABREU	0	-0.043	90.225	0.83	0.83	0	—
	1	0.071	-141.313	0.471	0.471	0.1	—
	2	0.066	-130.835	0.362	0.362	0.1	—
	3	0.062	-122.191	0.313	0.313	0.1	—
	4	0.059	-117.53	0.282	0.282	0.2	—
	5	0.058	-114.551	0.263	0.263	0.2	—
SERAFIM DIAS	0	-0.072	150.361	0.758	0.758	0	—
	1	0.062	-121.644	0.53	0.53	0.1	—
	2	0.058	-114.652	0.411	0.411	0.1	—
	3	0.053	-104.964	0.387	0.387	0.1	—
	4	0.049	-97.937	0.383	0.383	0.1	—
	5	0.047	-92.977	0.383	0.383	0.1	—
SÃO DOMINGOS	0	0.084	-164.808	0.644	0.644	0	—
	1	-0.021	44.617	0.829	0.829	0	—
	2	-0.014	29.143	0.854	0.854	0	—
	3	-0.013	28.599	0.834	0.834	0	—
	4	-0.013	27.688	0.824	0.824	0	—
	5	-0.013	26.788	0.82	0.82	0	—
SÃO JOAQUIM	0	0.092	-181.026	0.501	0.501	0.1	—
	1	0.017	-31.524	0.716	0.716	0	—
	2	0.025	-48	0.561	0.561	0	—
	3	0.021	-40.557	0.591	0.591	0	—
	4	0.02	-37.705	0.594	0.594	0	—
	5	0.019	-36.391	0.589	0.589	0	—
SÃO JOSÉ I	0	0.022	-39.554	0.938	0.938	0	—
	1	-0.041	84.364	0.607	0.607	0	—
	2	-0.035	71.397	0.526	0.526	0.1	—
	3	-0.031	63.747	0.511	0.511	0.1	—
	4	-0.029	59.035	0.51	0.51	0.1	—
	5	-0.027	55.853	0.513	0.513	0.1	—
SÃO PEDRO TIMBAÚBA	0	0.411	-823.403	0.007	0.007	0.6	↑

	1	0.152	-303.487	0.029	0.029	0.5	↑
	2	0.097	-193.39	0.045	0.045	0.4	↑
	3	0.081	-162.039	0.048	0.048	0.4	↑
	4	0.075	-148.301	0.048	0.048	0.4	↑
	5	0.071	-140.524	0.047	0.047	0.4	↑
SITIOS NOVOS	0	-0.195	397.495	0.181	0.181	0.2	—
	1	-0.036	75.877	0.461	0.461	0.1	—
	2	-0.012	26.883	0.721	0.721	0	—
	3	-0.008	17.177	0.801	0.801	0	—
	4	-0.006	14.034	0.828	0.828	0	—
	5	-0.006	12.865	0.835	0.835	0	—
TAQUARA	0	0.469	-941.112	0.094	0.094	0.5	↑
	1	0.17	-339.484	0.259	0.259	0.3	—
	2	0.065	-129.407	0.541	0.541	0.1	—
	3	0.037	-73.258	0.685	0.685	0	—
	4	0.027	-51.966	0.752	0.752	0	—
	5	0.022	-42.61	0.783	0.783	0	—
TRAPIÁ III	0	0.02	-35.313	0.786	0.786	0	—
	1	0	2.402	1	1	0	—
	2	-0.01	21.489	0.85	0.85	0	—
	3	-0.012	25.479	0.806	0.806	0	—
	4	-0.012	26.512	0.787	0.787	0	—
	5	-0.012	26.631	0.778	0.778	0	—
UBALDINHO	0	0.282	-564.528	0.117	0.117	0.3	—
	1	0.007	-12.81	0.904	0.904	0	—
	2	-0.018	37.867	0.69	0.69	0	—
	3	-0.022	45.667	0.58	0.58	0	—
	4	-0.023	47.754	0.535	0.535	0.1	—
	5	-0.023	48.326	0.513	0.513	0.1	—
UBALDINHO (RIACHO SÃO MIGUEL)	0	-0.078	161.018	0.86	0.86	0	—
	1	-0.27	546.77	0.246	0.246	0.6	—
	2	-0.36	728.456	0.053	0.053	0.9	↓
	3	-0.316	639.58	0.047	0.047	0.9	↓
	4	-0.288	581.672	0.045	0.045	0.9	↓
	5	-0.268	542.897	0.043	0.043	0.9	↓
VALÉRIO	0	0.007	-9.675	0.97	0.97	0	—
	1	-0.028	57.535	0.657	0.657	0	—
	2	-0.038	77.448	0.434	0.434	0.1	—
	3	-0.044	89.502	0.336	0.336	0.1	—
	4	-0.047	95.414	0.291	0.291	0.2	—
	5	-0.048	98.744	0.266	0.266	0.2	—
VÁRZEA DA VOLTA	0	0.145	-287.525	0.486	0.486	0.1	—
	1	0.031	-60.984	0.5	0.5	0.1	—
	2	0.017	-33.004	0.497	0.497	0.1	—
	3	0.014	-26.642	0.491	0.491	0.1	—

	4	0.013	-24.379	0.479	0.479	0.1	—
	5	0.012	-23.335	0.468	0.468	0.1	—
DO CORONEL	0	0.396	-794.999	0.02	0.02	1	↑
	1	0.066	-132.337	0.042	0.042	0.9	↑
	2	0.035	-68.485	0.047	0.047	0.9	↑
	3	0.027	-52.821	0.047	0.047	0.9	↑
	4	0.024	-46.642	0.046	0.046	0.9	↑
	5	0.022	-43.469	0.046	0.046	0.9	↑
ORÓS	0	0.504	-1008.434	0.009	0.009	0.7	↑
	1	0.335	-670.61	0.007	0.007	0.7	↑
	2	0.259	-519.814	0.009	0.009	0.7	↑
	3	0.223	-447.569	0.011	0.011	0.7	↑
	4	0.204	-407.834	0.011	0.011	0.7	↑
	5	0.191	-383.06	0.012	0.012	0.7	↑
RODRIGO	0	0.31	-621.094	0.095	0.095	0.8	↑
	1	-0.005	12.625	0.966	0.966	0	—
	2	-0.037	77.371	0.758	0.758	0.1	—
	3	-0.042	87.354	0.708	0.708	0.1	—
	4	-0.044	89.586	0.685	0.685	0.1	—
	5	-0.044	89.907	0.671	0.671	0.1	—
TRUSSU	0	0.217	-432.47	0.228	0.228	0.3	—
	1	0.226	-452.458	0.307	0.307	0.3	—
	2	0.176	-353.326	0.315	0.315	0.2	—
	3	0.149	-298.463	0.326	0.326	0.2	—
	4	0.134	-269.453	0.332	0.332	0.2	—
	5	0.126	-252.07	0.335	0.335	0.2	—

Source: prepared by the author

Identification of annual patterns

Three patterns were identified:

1. Annual reduction in biodiversity: 3 reservoirs (3,7 %)
2. Annual increase in biodiversity: 16 reservoirs (19,5 %)
3. Non-significant annual effect: 63 reservoirs (76,8 %)

Annual reduction in biodiversity: 3 reservoirs (3,7 %)

Legend

↓	Pattern 1: If decreasing trend, then decreasing of Biodiversity.
↑	Pattern 2: If increasing trend, increasing of Biodiversity

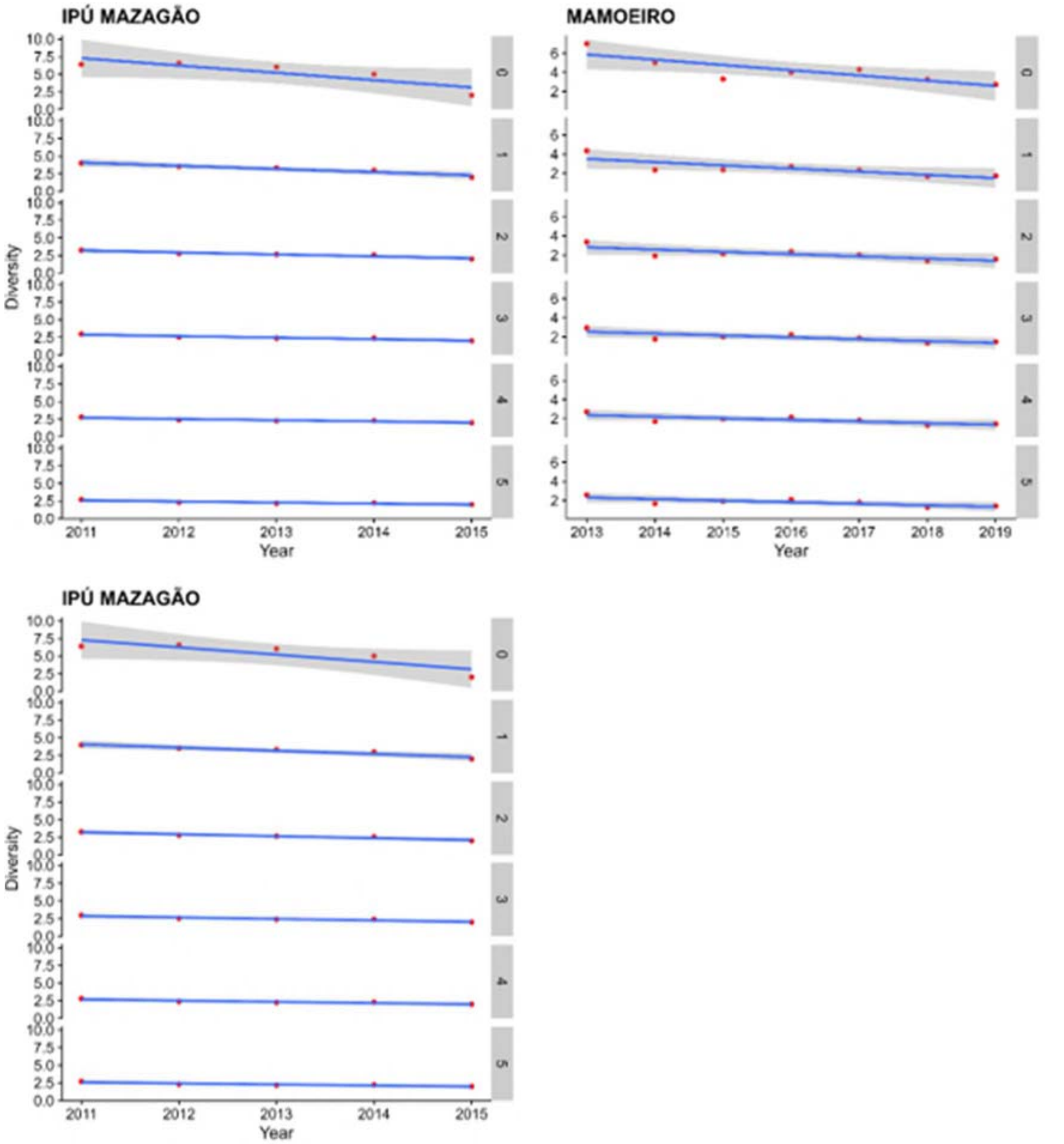
– Pattern 3: if trend is inconclusive then pattern is undetermined

Table E – Annual Behavior: reservoirs that showed a significant trend ($p < 0.10$) of decreasing biodiversity according to Hill numbers.

Reservoir	Hill number						Richness	Evenness	Class
	0	1	2	3	4	5			
IPÚ MAZAGÃO	↓	↓	↓	↓	↓	↓	↓	↓	↓
MAMOEIRO	↓	↓	↓	↓	↓	↓	↓	↓	↓
MARTINÓPOLIS	↓	↓	↓	↓	↓	↓	↓	↓	↓

Source: prepared by the author

Figure 21 – Hill numbers | Decreasing | IPÚ MAZAGÃO, MAMOEIRO and MARTINÓPOLE



Annual increase in biodiversity: 16 reservoirs (19,5 %)

Table F - Annual Behavior: reservoirs that showed a significant trend (p < 0.10) of increased biodiversity according to Hill numbers.

Reservoir	Hill number						Richness	Evenness	Class
	0	1	2	3	4	5			
ANGICOS	↑	↑	↑	↑	↑	↑	↑	↑	↑
AYRES DE SOUZA	↑	↑	↑	↑	↑	↑	↑	↑	↑
CAXITORÉ	↑	↑	↑	↑	↑	↑	↑	↑	↑
EDSON QUEIROZ	↑	↑	↑	↑	↑	↑	↑	↑	↑
GAVIÃO	↑	↑	↑	↑	↑	↑	↑	↑	↑
GENERAL SAMPAIO	↑	↑	↑	↑	↑	↑	↑	↑	↑
ITAPEBUSSU	↑	↑	↑	↑	↑	↑	↑	↑	↑
JOÃO LUÍS	↑	↑	↑	↑	↑	↑	↑	↑	↑
JUNCO	↑	↑	↑	↑	↑	↑	↑	↑	↑
PATU	↑	↑	↑	↑	↑	↑	↑	↑	↑
ARARAS	↑	↑	↑	↑	↑	↑	↑	↑	↑
RIACHÃO	↑	↑	↑	↑	↑	↑	↑	↑	↑
SANTO ANTÔNIO DE RUSSAS	↑	↑	↑	↑	↑	↑	↑	↑	↑
SÃO PEDRO TIMBAÚBA	↑	↑	↑	↑	↑	↑	↑	↑	↑
DO CORONEL	↑	↑	↑	↑	↑	↑	↑	↑	↑
ORÓS	↑	↑	↑	↑	↑	↑	↑	↑	↑

Source: prepared by the author

Figure 22 - Hill numbers | Increasing | Reservoirs: ANGICOS, ARARAS, AYRES DE SOUZA and CAXITORÉ

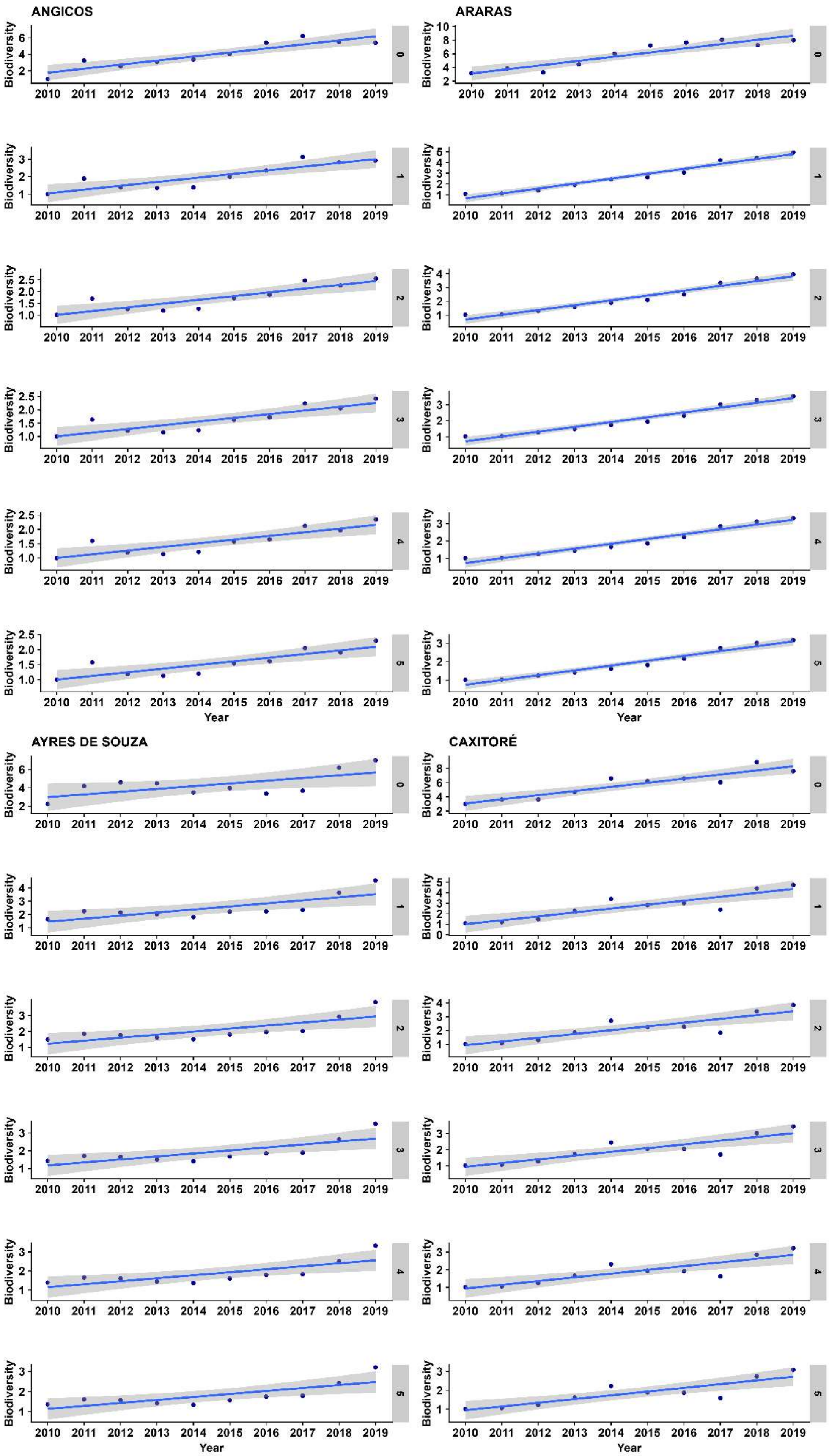

ANGICOS
ARARAS
AYRES DE SOUZA
CAXITORÉ
Biodiversity
2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year
0
1
2
3
4
5

Figure 23 - Hill numbers | Increasing | Reservoirs: DO CORONEL, EDSON QUEIROZ, GAVIÃO and GENERAL SAMPAIO

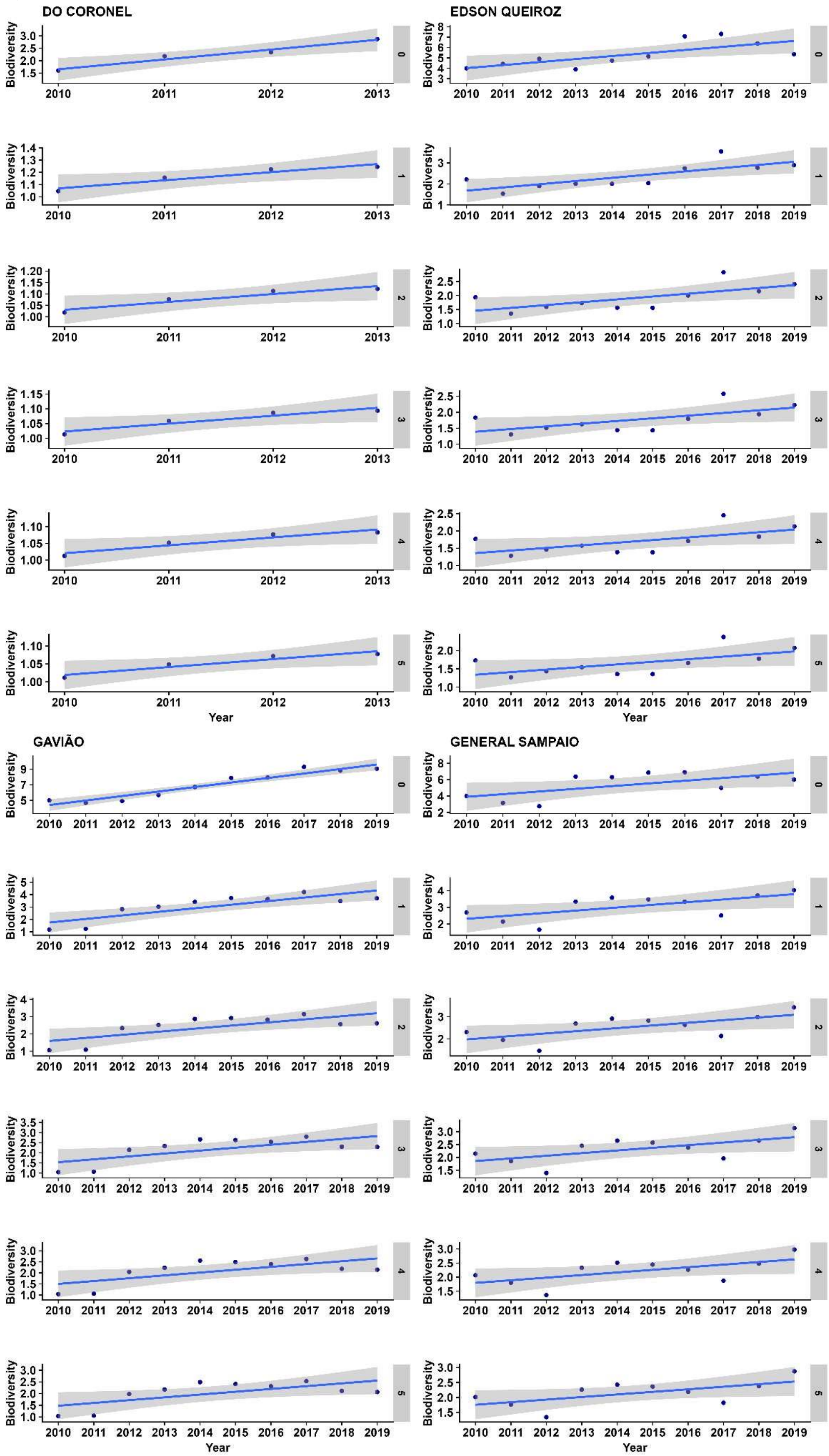


Figure 24 - Hill numbers | Increasing | Reservoirs: ITAPEBUSSU, JOÃO LUÍS, JUNCO and ORÓS

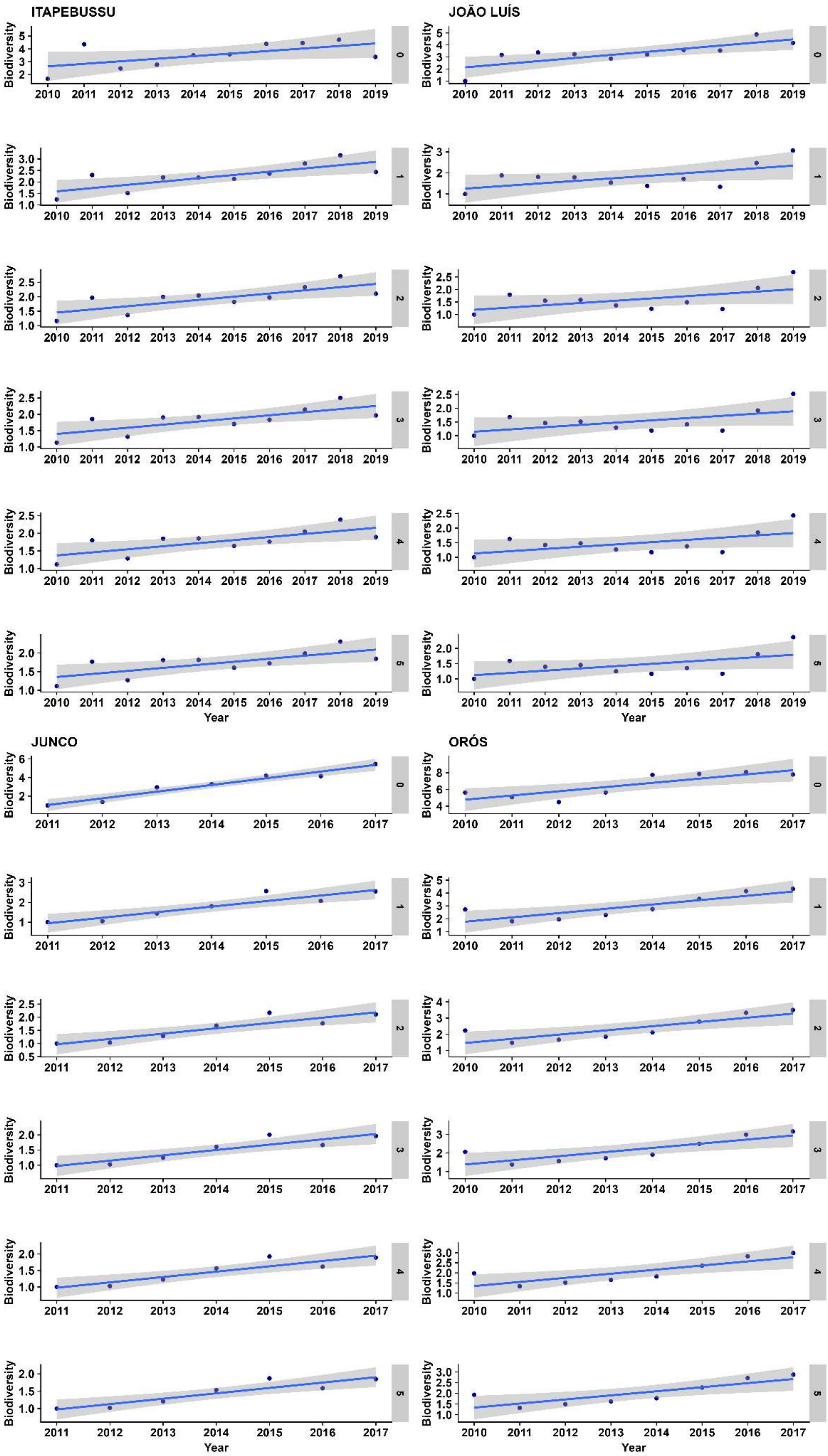

ITAPEBUSSU
JOÃO LUÍS
JUNCO
ORÓS
Biodiversity
Year
0
1
2
3
4
5

Figure 25 - Hill numbers | Increasing | Reservoirs: PATU, RIACHÃO, SANTO ANTÔNIO DE RUSSAS and SÃO PEDRO TIMBAÚBA

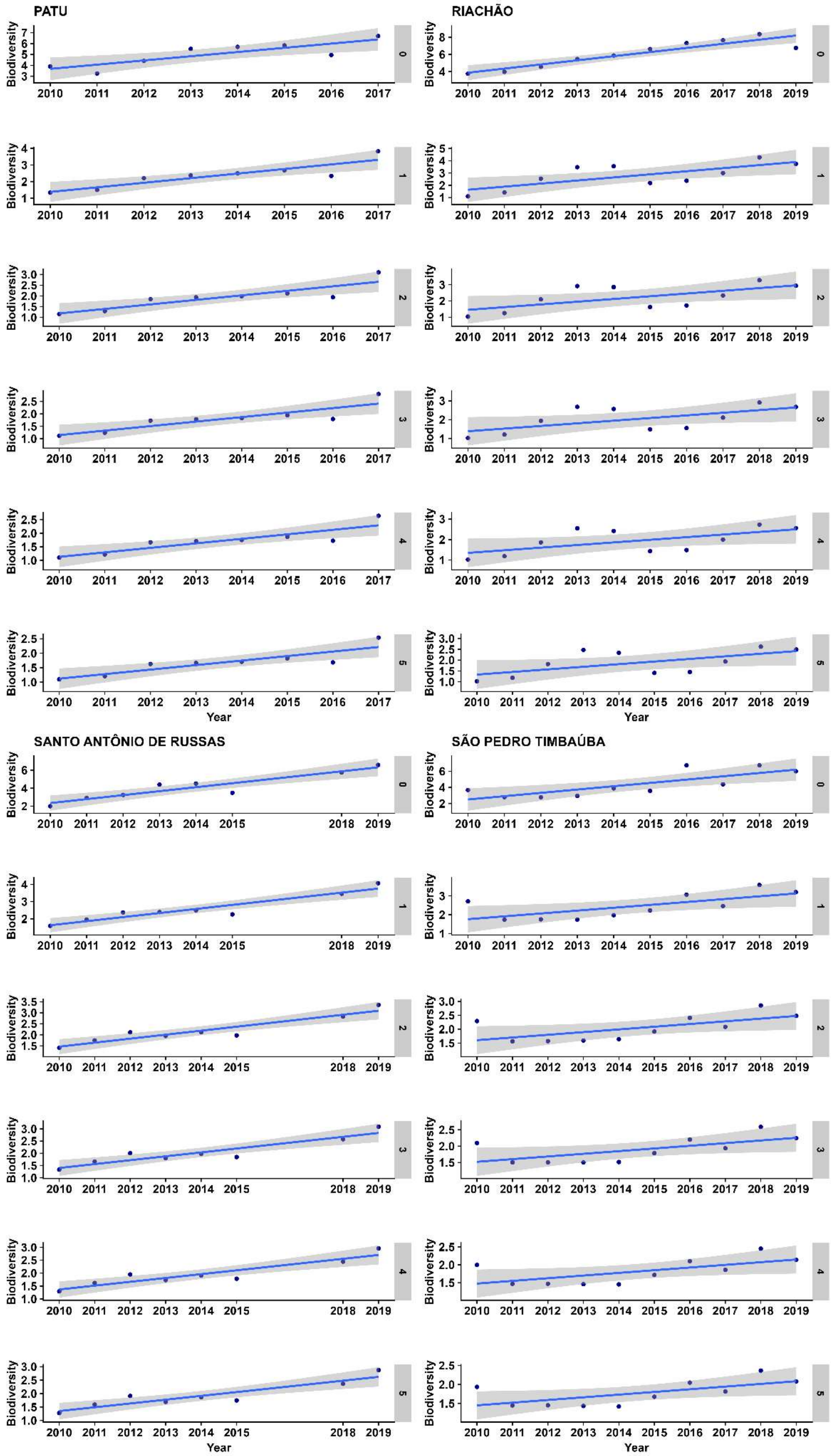
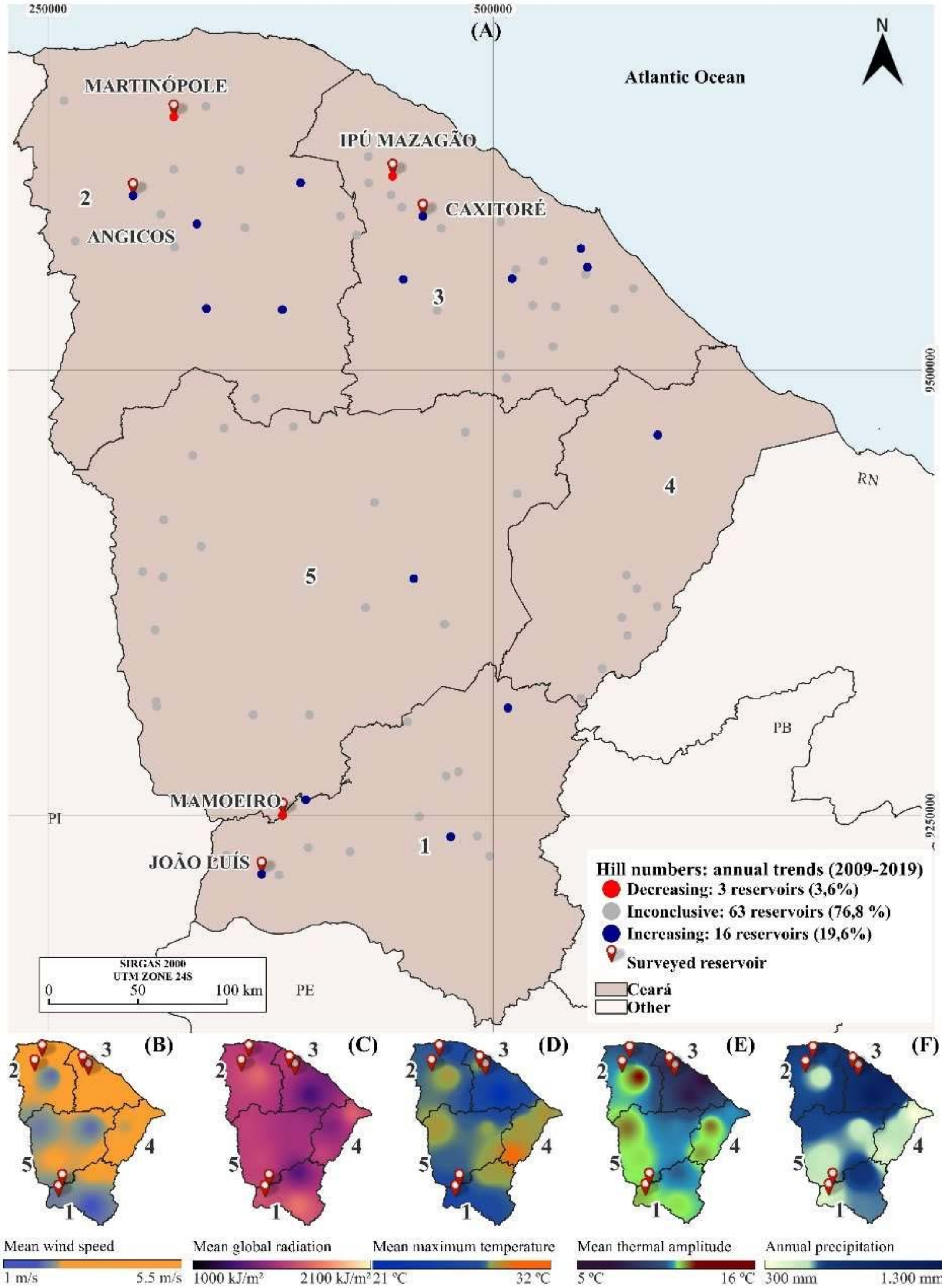
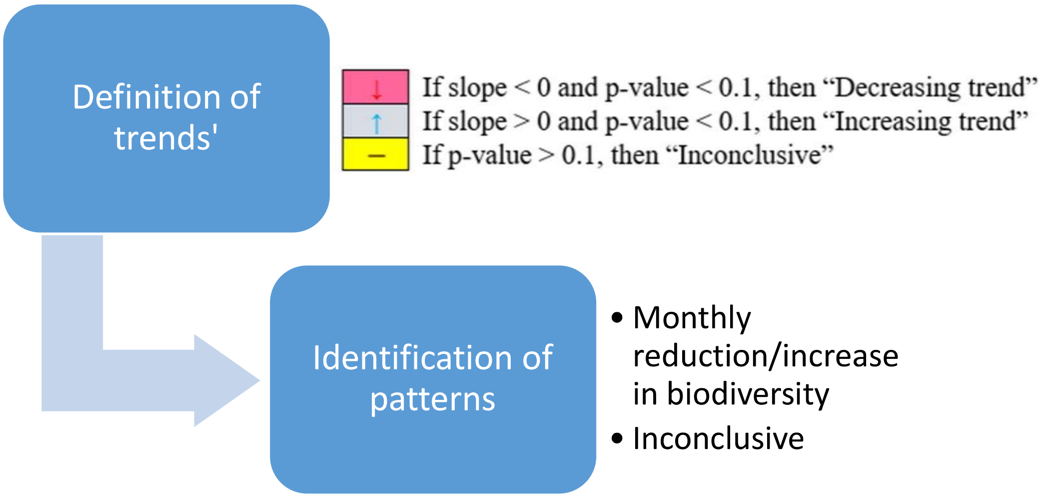


Figure 28 - In (A), location and categorization of reservoirs intended for human consumption in the 5 macro-regions of Ceará according to the trend of the annual average of cyanobacteria biodiversity in raw water. Red dots are reservoirs with a significant reduction, while blue dots represent those with an increase. Gray dots represent reservoirs without a significant trend. For example, 6 reservoirs with these trends were highlighted from the others. From (B) to (F) the gradients of the environmental variables and the location of the highlighted reservoirs are shown.



Source: prepared by the author

Monthly Behavior: The Effect of Seasonality on Biodiversity



Defining monthly behavioral trends

Legend

↓ If slope < 0 and p-value < 0.1, then "Decreasing trend"

↑ If slope > 0 and p-value < 0.1, then "Increasing trend"

— If p-value > 0.1, then "Inconclusive"

Table G - Tendências dos números de Hill durante o comportamento mensal

Reservoir	Number	slope	Intercept	p_slope	p-value	Rsquared	Pattern
ACARAPE DO MEIO	0	-0.047	5.037	0.128	0.128	0.2	—
	1	-0.037	2.556	0.081	0.081	0.3	↓
	2	-0.03	2.126	0.094	0.094	0.3	↓
	3	-0.027	1.969	0.105	0.105	0.2	—
	4	-0.025	1.889	0.113	0.113	0.2	—
	5	-0.024	1.841	0.118	0.118	0.2	—
ACARAÚ MIRIM	0	0.098	3.225	0.11	0.11	0.2	—
	1	0.061	1.986	0.141	0.141	0.2	—
	2	0.048	1.734	0.141	0.141	0.2	—
	3	0.043	1.625	0.138	0.138	0.2	—
	4	0.041	1.568	0.134	0.134	0.2	—
	5	0.039	1.535	0.132	0.132	0.2	—
ADAUTO BEZERRA	0	0.039	3.068	0.233	0.233	0.1	—
	1	-0.016	1.759	0.338	0.338	0.1	—
	2	-0.011	1.547	0.456	0.456	0.1	—
	3	-0.008	1.469	0.531	0.531	0	—

	4	-0.007	1.431	0.567	0.567	0	—
	5	-0.007	1.409	0.584	0.584	0	—
AMANARY	0	0.02	4.895	0.545	0.545	0	—
	1	0.035	2.28	0.094	0.094	0.3	↑
	2	0.023	1.9	0.174	0.174	0.2	—
	3	0.018	1.777	0.258	0.258	0.1	—
	4	0.015	1.715	0.307	0.307	0.1	—
	5	0.014	1.678	0.334	0.334	0.1	—
ANGICOS	0	-0.023	4.49	0.56	0.56	0	—
	1	-0.033	2.308	0.272	0.272	0.1	—
	2	-0.026	1.92	0.272	0.272	0.1	—
	3	-0.023	1.787	0.257	0.257	0.1	—
	4	-0.022	1.722	0.249	0.249	0.1	—
	5	-0.021	1.683	0.245	0.245	0.1	—
ARACOIABA	0	-0.001	6.09	0.982	0.982	0	—
	1	-0.029	3.199	0.469	0.469	0.1	—
	2	-0.019	2.533	0.551	0.551	0	—
	3	-0.015	2.298	0.582	0.582	0	—
	4	-0.013	2.184	0.595	0.595	0	—
	5	-0.012	2.117	0.6	0.6	0	—
ARNEIROZ II	0	-0.055	6.007	0.165	0.165	0.2	—
	1	-0.031	2.487	0.122	0.122	0.2	—
	2	-0.02	1.908	0.242	0.242	0.1	—
	3	-0.016	1.74	0.302	0.302	0.1	—
	4	-0.014	1.666	0.326	0.326	0.1	—
	5	-0.013	1.625	0.336	0.336	0.1	—
AYRES DE SOUZA	0	0.027	4.266	0.444	0.444	0.1	—
	1	0.018	2.265	0.462	0.462	0.1	—
	2	0.006	1.926	0.745	0.745	0	—
	3	0.003	1.801	0.863	0.863	0	—
	4	0.002	1.738	0.916	0.916	0	—
	5	0.001	1.7	0.943	0.943	0	—
BELO HORIZONTE	0	0.026	4.918	0.107	0.107	0.2	—
	1	0.019	2.561	0.424	0.424	0.1	—
	2	0.016	2.093	0.44	0.44	0.1	—
	3	0.014	1.932	0.458	0.458	0.1	—
	4	0.013	1.852	0.47	0.47	0.1	—
	5	0.012	1.805	0.477	0.477	0.1	—
BOQUEIRÃO	0	-0.015	4.762	0.604	0.604	0	—
	1	0.009	2.278	0.478	0.478	0.1	—
	2	0.01	1.878	0.353	0.353	0.1	—
	3	0.009	1.741	0.336	0.336	0.1	—
	4	0.009	1.674	0.326	0.326	0.1	—

	5	0.009	1.634	0.317	0.317	0.1	—
CACHOEIRA	0	0.042	3.974	0.237	0.237	0.1	—
	1	0.029	2.126	0.188	0.188	0.2	—
	2	0.016	1.848	0.412	0.412	0.1	—
	3	0.012	1.733	0.492	0.492	0	—
	4	0.011	1.675	0.535	0.535	0	—
	5	0.01	1.64	0.56	0.56	0	—
CANAFÍSTULA	0	-0.106	4.881	0.104	0.104	0.2	—
	1	-0.035	2.285	0.219	0.219	0.1	—
	2	-0.024	1.886	0.23	0.23	0.1	—
	3	-0.02	1.748	0.227	0.227	0.1	—
	4	-0.018	1.679	0.229	0.229	0.1	—
	5	-0.016	1.638	0.234	0.234	0.1	—
CANOAS	0	-0.001	4.683	0.959	0.959	0	—
	1	-0	2.049	0.977	0.977	0	—
	2	0.004	1.64	0.65	0.65	0	—
	3	0.005	1.516	0.529	0.529	0	—
	4	0.005	1.461	0.488	0.488	0	—
	5	0.005	1.43	0.47	0.47	0.1	—
CARMINA	0	-0.015	4.672	0.739	0.739	0	—
	1	0	2.21	0.99	0.99	0	—
	2	-0.003	1.878	0.832	0.832	0	—
	3	-0.003	1.762	0.785	0.785	0	—
	4	-0.003	1.704	0.765	0.765	0	—
	5	-0.003	1.669	0.756	0.756	0	—
CARNAUBAL	0	0.035	3.168	0.664	0.664	0	—
	1	0.05	1.704	0.263	0.263	0.1	—
	2	0.034	1.61	0.473	0.473	0.1	—
	3	0.027	1.545	0.524	0.524	0	—
	4	0.024	1.509	0.546	0.546	0	—
	5	0.022	1.487	0.555	0.555	0	—
CARÃO	0	-0.006	3.508	0.919	0.919	0	—
	1	0.024	1.681	0.27	0.27	0.1	—
	2	0.013	1.566	0.512	0.512	0	—
	3	0.011	1.513	0.587	0.587	0	—
	4	0.009	1.487	0.634	0.634	0	—
	5	0.008	1.472	0.665	0.665	0	—
CASTRO	0	-0.059	4.88	0.059	0.059	0.3	↓
	1	0.015	2.047	0.44	0.44	0.1	—
	2	0.016	1.77	0.274	0.274	0.1	—
	3	0.017	1.668	0.229	0.229	0.1	—
	4	0.017	1.617	0.213	0.213	0.2	—
	5	0.017	1.587	0.205	0.205	0.2	—

CAXITORÉ	0	-0.096	6.616	0.073	0.073	0.3	↓
	1	-0.033	2.977	0.316	0.316	0.1	—
	2	-0.025	2.361	0.304	0.304	0.1	—
	3	-0.022	2.151	0.276	0.276	0.1	—
	4	-0.021	2.049	0.255	0.255	0.1	—
	5	-0.02	1.989	0.241	0.241	0.1	—
COLINA	0	0.08	4.524	0.154	0.154	0.2	—
	1	0.001	2.709	0.979	0.979	0	—
	2	-0.009	2.299	0.79	0.79	0	—
	3	-0.012	2.144	0.704	0.704	0	—
	4	-0.013	2.065	0.657	0.657	0	—
	5	-0.014	2.017	0.63	0.63	0	—
CORIOLANO ALVES DE SOUSA BRITO	0	-0.02	3.123	0.638	0.638	0	—
	1	-0.027	1.925	0.248	0.248	0.1	—
	2	-0.026	1.701	0.188	0.188	0.2	—
	3	-0.025	1.621	0.164	0.164	0.2	—
	4	-0.025	1.582	0.149	0.149	0.2	—
	5	-0.025	1.558	0.14	0.14	0.2	—
OLHO D'ÁGUA	0	-0.049	4.374	0.125	0.125	0.2	—
	1	-0.052	2.052	0.005	0.005	0.6	↓
	2	-0.039	1.691	0.003	0.003	0.6	↓
	3	-0.033	1.576	0.002	0.002	0.6	↓
	4	-0.03	1.521	0.002	0.002	0.6	↓
	5	-0.028	1.49	0.002	0.002	0.6	↓
EDSON QUEIROZ	0	-0.015	5.587	0.641	0.641	0	—
	1	-0.017	2.478	0.461	0.461	0.1	—
	2	-0.011	1.956	0.577	0.577	0	—
	3	-0.008	1.789	0.641	0.641	0	—
	4	-0.007	1.712	0.667	0.667	0	—
	5	-0.006	1.669	0.679	0.679	0	—
EMA	0	0.039	3.969	0.192	0.192	0.2	—
	1	-0.008	2.101	0.786	0.786	0	—
	2	-0.006	1.821	0.803	0.803	0	—
	3	-0.005	1.711	0.822	0.822	0	—
	4	-0.004	1.657	0.833	0.833	0	—
	5	-0.004	1.624	0.837	0.837	0	—
FACUNDO	0	-0.006	3.641	0.839	0.839	0	—
	1	0.003	1.806	0.828	0.828	0	—
	2	0.002	1.53	0.808	0.808	0	—
	3	0.001	1.449	0.902	0.902	0	—
	4	0	1.412	0.967	0.967	0	—
	5	-0	1.391	0.993	0.993	0	—
FIGUEIREDO	0	-0.158	4.69	0.093	0.093	0.3	↓

	1	-0.058	2.507	0.221	0.221	0.1	—
	2	-0.048	2.258	0.236	0.236	0.1	—
	3	-0.038	2.069	0.282	0.282	0.1	—
	4	-0.032	1.969	0.315	0.315	0.1	—
	5	-0.029	1.908	0.339	0.339	0.1	—
FLOR DO CAMPO	0	0.044	4.553	0.212	0.212	0.2	—
	1	0.075	2.047	0.062	0.062	0.3	↑
	2	0.063	1.729	0.072	0.072	0.3	↑
	3	0.058	1.604	0.061	0.061	0.3	↑
	4	0.055	1.543	0.054	0.054	0.3	↑
	5	0.053	1.506	0.049	0.049	0.3	↑
FORQUILHA	0	0.047	5.088	0.242	0.242	0.1	—
	1	-0.032	2.684	0.423	0.423	0.1	—
	2	-0.036	2.244	0.333	0.333	0.1	—
	3	-0.034	2.074	0.338	0.338	0.1	—
	4	-0.032	1.985	0.352	0.352	0.1	—
	5	-0.031	1.932	0.365	0.365	0.1	—
GAMELEIRA	0	-0.06	5.345	0.142	0.142	0.2	—
	1	-0.054	3.193	0.173	0.173	0.2	—
	2	-0.044	2.641	0.204	0.204	0.2	—
	3	-0.039	2.42	0.212	0.212	0.2	—
	4	-0.036	2.305	0.213	0.213	0.2	—
	5	-0.034	2.236	0.211	0.211	0.2	—
GAVIÃO	0	0.019	6.968	0.611	0.611	0	—
	1	0.016	3.161	0.548	0.548	0	—
	2	0.016	2.482	0.478	0.478	0.1	—
	3	0.016	2.257	0.451	0.451	0.1	—
	4	0.015	2.147	0.439	0.439	0.1	—
	5	0.014	2.082	0.432	0.432	0.1	—
GENERAL SAMPAIO	0	0.123	4.877	0.099	0.099	0.2	↑
	1	0.046	2.766	0.258	0.258	0.1	—
	2	0.025	2.355	0.412	0.412	0.1	—
	3	0.018	2.187	0.497	0.497	0	—
	4	0.015	2.095	0.535	0.535	0	—
	5	0.013	2.038	0.552	0.552	0	—
IPÚ MAZAGÃO	0	0.064	5.549	0.234	0.234	0.1	—
	1	0.047	3.061	0.287	0.287	0.1	—
	2	0.041	2.483	0.275	0.275	0.1	—
	3	0.035	2.276	0.292	0.292	0.1	—
	4	0.031	2.173	0.313	0.313	0.1	—
	5	0.028	2.112	0.331	0.331	0.1	—
ITAPEBUSSU	0	0.122	2.887	0.056	0.056	0.3	↑
	1	0.044	2.019	0.128	0.128	0.2	—

	2	0.027	1.813	0.253	0.253	0.1	—
	3	0.022	1.716	0.297	0.297	0.1	—
	4	0.02	1.664	0.314	0.314	0.1	—
	5	0.019	1.631	0.321	0.321	0.1	—
ITAÚNA	0	0.007	5.413	0.769	0.769	0	—
	1	-0.004	3.496	0.736	0.736	0	—
	2	-0	2.92	0.994	0.994	0	—
	3	0.003	2.659	0.843	0.843	0	—
	4	0.005	2.513	0.717	0.717	0	—
	5	0.006	2.422	0.628	0.628	0	—
JABURU I	0	0.027	4.021	0.059	0.059	0.3	↑
	1	-0.023	2.374	0.116	0.116	0.2	—
	2	-0.017	2.027	0.17	0.17	0.2	—
	3	-0.014	1.894	0.222	0.222	0.1	—
	4	-0.012	1.826	0.259	0.259	0.1	—
	5	-0.011	1.784	0.284	0.284	0.1	—
JABURU II	0	-0.103	4.501	0.292	0.292	0.1	—
	1	-0.078	2.468	0.159	0.159	0.2	—
	2	-0.06	2.079	0.166	0.166	0.2	—
	3	-0.054	1.943	0.165	0.165	0.2	—
	4	-0.05	1.875	0.164	0.164	0.2	—
	5	-0.048	1.833	0.163	0.163	0.2	—
JENIPAPEIRO	0	-0.02	4.316	0.749	0.749	0	—
	1	-0.014	2.372	0.761	0.761	0	—
	2	-0.018	2.05	0.661	0.661	0	—
	3	-0.018	1.921	0.62	0.62	0	—
	4	-0.018	1.853	0.596	0.596	0	—
	5	-0.017	1.811	0.581	0.581	0	—
JERIMUM	0	-0.016	4.809	0.639	0.639	0	—
	1	0.009	1.948	0.52	0.52	0	—
	2	0.004	1.635	0.76	0.76	0	—
	3	0.002	1.537	0.834	0.834	0	—
	4	0.002	1.491	0.871	0.871	0	—
	5	0.001	1.465	0.895	0.895	0	—
JOÃO JOSÉ DE CASTRO	0	-0.02	3.282	0.645	0.645	0	—
	1	-0.016	1.87	0.438	0.438	0.1	—
	2	-0.017	1.686	0.364	0.364	0.1	—
	3	-0.015	1.604	0.38	0.38	0.1	—
	4	-0.014	1.56	0.393	0.393	0.1	—
	5	-0.013	1.533	0.401	0.401	0.1	—
JOÃO LUÍS	0	0.007	3.583	0.862	0.862	0	—
	1	-0.003	1.782	0.855	0.855	0	—
	2	-0.004	1.561	0.733	0.733	0	—

	3	-0.004	1.486	0.691	0.691	0	—
	4	-0.004	1.447	0.693	0.693	0	—
	5	-0.004	1.423	0.706	0.706	0	—
JUNCO	0	0.023	3.79	0.577	0.577	0	—
	1	0.006	1.989	0.854	0.854	0	—
	2	0.009	1.692	0.719	0.719	0	—
	3	0.01	1.592	0.695	0.695	0	—
	4	0.009	1.544	0.685	0.685	0	—
	5	0.009	1.515	0.679	0.679	0	—
MADEIRO	0	0.049	3.398	0.296	0.296	0.1	—
	1	0.048	1.886	0.061	0.061	0.3	↑
	2	0.038	1.619	0.101	0.101	0.2	—
	3	0.034	1.524	0.114	0.114	0.2	—
	4	0.032	1.476	0.116	0.116	0.2	—
	5	0.03	1.447	0.116	0.116	0.2	—
MALCOZINHADO	0	-0.066	7.195	0.213	0.213	0.2	—
	1	-0.052	3.747	0.238	0.238	0.1	—
	2	-0.047	3.026	0.306	0.306	0.1	—
	3	-0.044	2.759	0.311	0.311	0.1	—
	4	-0.041	2.62	0.315	0.315	0.1	—
	5	-0.04	2.534	0.318	0.318	0.1	—
MAMOEIRO	0	0.041	3.775	0.682	0.682	0	—
	1	0.04	1.915	0.247	0.247	0.1	—
	2	0.023	1.683	0.357	0.357	0.1	—
	3	0.017	1.609	0.442	0.442	0.1	—
	4	0.014	1.572	0.493	0.493	0	—
	5	0.012	1.549	0.524	0.524	0	—
MARTINÓPOLE	0	0.008	4.264	0.717	0.717	0	—
	1	0.027	2.365	0.246	0.246	0.1	—
	2	0.021	2.034	0.356	0.356	0.1	—
	3	0.017	1.922	0.441	0.441	0.1	—
	4	0.014	1.865	0.49	0.49	0	—
	5	0.013	1.829	0.518	0.518	0	—
MISSI	0	-0.056	5.89	0.484	0.484	0.1	—
	1	-0.015	3.187	0.654	0.654	0	—
	2	-0.009	2.635	0.716	0.716	0	—
	3	-0.008	2.434	0.704	0.704	0	—
	4	-0.008	2.33	0.692	0.692	0	—
	5	-0.007	2.265	0.684	0.684	0	—
MONSENHOR TABOSA	0	-0.025	4.741	0.254	0.254	0.1	—
	1	-0.012	1.705	0.249	0.249	0.1	—
	2	-0.009	1.405	0.287	0.287	0.1	—
	3	-0.008	1.331	0.295	0.295	0.1	—

	4	-0.007	1.299	0.294	0.294	0.1	—
	5	-0.007	1.282	0.291	0.291	0.1	—
MUNDAÚ	0	0.033	3.776	0.487	0.487	0	—
	1	-0.007	2.067	0.829	0.829	0	—
	2	-0.005	1.746	0.832	0.832	0	—
	3	-0.005	1.637	0.825	0.825	0	—
	4	-0.004	1.584	0.817	0.817	0	—
	5	-0.004	1.553	0.811	0.811	0	—
PACOTI	0	-0.044	6.583	0.131	0.131	0.2	—
	1	-0.085	3.056	0.002	0.002	0.6	↓
	2	-0.073	2.455	0.001	0.001	0.7	↓
	3	-0.065	2.241	0.001	0.001	0.7	↓
	4	-0.061	2.133	0	0	0.7	↓
	5	-0.058	2.069	0	0	0.7	↓
PATU	0	0.037	4.962	0.548	0.548	0	—
	1	0.021	2.259	0.596	0.596	0	—
	2	0.01	1.881	0.744	0.744	0	—
	3	0.008	1.747	0.772	0.772	0	—
	4	0.007	1.681	0.783	0.783	0	—
	5	0.007	1.643	0.789	0.789	0	—
ARARAS	0	-0.007	6.111	0.811	0.811	0	—
	1	-0.009	2.772	0.597	0.597	0	—
	2	-0.004	2.248	0.778	0.778	0	—
	3	-0.003	2.069	0.843	0.843	0	—
	4	-0.002	1.981	0.869	0.869	0	—
	5	-0.002	1.928	0.879	0.879	0	—
PEDRAS BRANCAS	0	0.029	4.283	0.44	0.44	0.1	—
	1	-0.005	2.648	0.786	0.786	0	—
	2	-0.012	2.348	0.413	0.413	0.1	—
	3	-0.013	2.223	0.323	0.323	0.1	—
	4	-0.013	2.152	0.295	0.295	0.1	—
	5	-0.013	2.105	0.284	0.284	0.1	—
PENEDO	0	0.013	3.871	0.679	0.679	0	—
	1	-0.047	2.481	0.031	0.031	0.4	↓
	2	-0.04	2.093	0.032	0.032	0.4	↓
	3	-0.037	1.942	0.026	0.026	0.4	↓
	4	-0.035	1.862	0.023	0.023	0.4	↓
	5	-0.033	1.814	0.023	0.023	0.4	↓
PENTECOSTE	0	0.021	5.13	0.63	0.63	0	—
	1	-0.035	2.83	0.28	0.28	0.1	—
	2	-0.033	2.346	0.208	0.208	0.2	—
	3	-0.029	2.157	0.192	0.192	0.2	—
	4	-0.027	2.058	0.191	0.191	0.2	—

	5	-0.025	1.997	0.193	0.193	0.2	—
PESQUEIRO	0	0.004	4.129	0.867	0.867	0	—
	1	0.002	2.236	0.903	0.903	0	—
	2	0.001	1.974	0.958	0.958	0	—
	3	0	1.875	0.978	0.978	0	—
	4	0	1.821	0.996	0.996	0	—
	5	-0	1.787	0.988	0.988	0	—
POMPEU SOBRINHO	0	-0.076	2.737	0.004	0.004	0.6	↓
	1	-0.042	1.892	0.025	0.025	0.4	↓
	2	-0.043	1.806	0.046	0.046	0.3	↓
	3	-0.041	1.728	0.05	0.05	0.3	↓
	4	-0.039	1.684	0.051	0.051	0.3	↓
	5	-0.038	1.655	0.051	0.051	0.3	↓
POTIRETAMA	0	0.065	2.154	0.012	0.012	0.5	↑
	1	0.022	1.437	0.149	0.149	0.2	—
	2	0.017	1.293	0.216	0.216	0.1	—
	3	0.016	1.239	0.215	0.215	0.1	—
	4	0.016	1.212	0.206	0.206	0.2	—
	5	0.015	1.197	0.199	0.199	0.2	—
POÇO VERDE	0	-0.097	4.472	0.097	0.097	0.3	↓
	1	-0.053	2.204	0.021	0.021	0.4	↓
	2	-0.04	1.845	0.013	0.013	0.5	↓
	3	-0.035	1.726	0.011	0.011	0.5	↓
	4	-0.033	1.669	0.011	0.011	0.5	↓
	5	-0.031	1.635	0.01	0.01	0.5	↓
POÇO DA PEDRA	0	-0.01	4.694	0.64	0.64	0	—
	1	-0.017	2.397	0.195	0.195	0.2	—
	2	-0.012	1.968	0.266	0.266	0.1	—
	3	-0.01	1.816	0.319	0.319	0.1	—
	4	-0.009	1.74	0.349	0.349	0.1	—
	5	-0.008	1.695	0.365	0.365	0.1	—
PUIÚ	0	0.002	4.176	0.918	0.918	0	—
	1	0.018	1.717	0.066	0.066	0.3	↑
	2	0.017	1.436	0.035	0.035	0.4	↑
	3	0.016	1.354	0.029	0.029	0.4	↑
	4	0.015	1.317	0.027	0.027	0.4	↑
	5	0.015	1.296	0.025	0.025	0.4	↑
QUANDÚ	0	0.074	4.886	0.161	0.161	0.2	—
	1	0.079	2.14	0.078	0.078	0.3	↑
	2	0.058	1.799	0.093	0.093	0.3	↑
	3	0.049	1.688	0.098	0.098	0.2	↑
	4	0.045	1.633	0.1	0.1	0.2	—
	5	0.043	1.6	0.1	0.1	0.2	—

RIACHO DA SERRA	0	-0.103	4.421	0.32	0.32	0.1	—
	1	-0.073	2.58	0.214	0.214	0.1	—
	2	-0.055	2.191	0.221	0.221	0.1	—
	3	-0.046	2.015	0.228	0.228	0.1	—
	4	-0.043	1.928	0.222	0.222	0.1	—
	5	-0.04	1.878	0.213	0.213	0.2	—
RIACHÃO	0	0	6.283	0.99	0.99	0	—
	1	-0.031	3.079	0.405	0.405	0.1	—
	2	-0.027	2.448	0.358	0.358	0.1	—
	3	-0.026	2.239	0.306	0.306	0.1	—
	4	-0.025	2.138	0.272	0.272	0.1	—
	5	-0.025	2.077	0.251	0.251	0.1	—
RIVALDO DE CARVALHO	0	0.043	4.339	0.026	0.026	0.4	↑
	1	0.003	1.5	0.703	0.703	0	—
	2	-0.001	1.282	0.933	0.933	0	—
	3	-0.001	1.234	0.857	0.857	0	—
	4	-0.002	1.215	0.828	0.828	0	—
	5	-0.002	1.204	0.813	0.813	0	—
ROSÁRIO	0	0.015	3.158	0.63	0.63	0	—
	1	-0.002	1.627	0.933	0.933	0	—
	2	-0.002	1.427	0.905	0.905	0	—
	3	-0.002	1.358	0.897	0.897	0	—
	4	-0.001	1.324	0.897	0.897	0	—
	5	-0.001	1.304	0.899	0.899	0	—
SANTO ANTÔNIO DE RUSSAS	0	0.082	3.764	0.227	0.227	0.1	—
	1	-0.005	2.613	0.879	0.879	0	—
	2	-0.014	2.278	0.667	0.667	0	—
	3	-0.016	2.141	0.602	0.602	0	—
	4	-0.017	2.067	0.574	0.574	0	—
	5	-0.017	2.022	0.558	0.558	0	—
SEBASTIÃO DE ABREU	0	-0.177	6.262	0.046	0.046	0.3	↓
	1	-0.14	3.681	0.031	0.031	0.4	↓
	2	-0.109	2.994	0.031	0.031	0.4	↓
	3	-0.096	2.723	0.029	0.029	0.4	↓
	4	-0.089	2.584	0.027	0.027	0.4	↓
	5	-0.085	2.501	0.026	0.026	0.4	↓
SERAFIM DIAS	0	0.021	5.547	0.429	0.429	0.1	—
	1	0.013	2.509	0.473	0.473	0.1	—
	2	0.003	2.027	0.834	0.834	0	—
	3	-0.001	1.877	0.938	0.938	0	—
	4	-0.002	1.805	0.828	0.828	0	—
	5	-0.003	1.763	0.771	0.771	0	—
SÃO DOMINGOS	0	0.039	3.247	0.395	0.395	0.1	—

	1	0.036	1.955	0.219	0.219	0.1	—
	2	0.024	1.753	0.282	0.282	0.1	—
	3	0.02	1.669	0.321	0.321	0.1	—
	4	0.017	1.623	0.342	0.342	0.1	—
	5	0.016	1.594	0.354	0.354	0.1	—
SÃO JOAQUIM	0	-0.153	5.585	0.014	0.014	0.5	↓
	1	-0.084	2.823	0.039	0.039	0.4	↓
	2	-0.057	2.283	0.051	0.051	0.3	↓
	3	-0.047	2.085	0.059	0.059	0.3	↓
	4	-0.042	1.988	0.064	0.064	0.3	↓
	5	-0.039	1.932	0.066	0.066	0.3	↓
SÃO JOSÉ I	0	-0.076	5.123	0.107	0.107	0.2	—
	1	-0.004	2.135	0.896	0.896	0	—
	2	-0.003	1.764	0.897	0.897	0	—
	3	-0.003	1.651	0.879	0.879	0	—
	4	-0.003	1.596	0.872	0.872	0	—
	5	-0.003	1.563	0.873	0.873	0	—
SÃO PEDRO TIMBAÚBA	0	0.068	4.615	0.203	0.203	0.2	—
	1	0.041	2.417	0.198	0.198	0.2	—
	2	0.036	1.975	0.144	0.144	0.2	—
	3	0.031	1.834	0.153	0.153	0.2	—
	4	0.028	1.768	0.164	0.164	0.2	—
	5	0.026	1.73	0.175	0.175	0.2	—
SITIOS NOVOS	0	0.068	5.102	0.007	0.007	0.5	↑
	1	0.001	2.571	0.929	0.929	0	—
	2	-0.006	2.134	0.711	0.711	0	—
	3	-0.006	1.966	0.708	0.708	0	—
	4	-0.005	1.883	0.719	0.719	0	—
	5	-0.005	1.834	0.728	0.728	0	—
TAQUARA	0	0.026	4.114	0.791	0.791	0	—
	1	0.007	2.75	0.864	0.864	0	—
	2	0.01	2.287	0.771	0.771	0	—
	3	0.008	2.109	0.8	0.8	0	—
	4	0.006	2.015	0.814	0.814	0	—
	5	0.006	1.957	0.82	0.82	0	—
TRAPIÁ III	0	0.013	4.412	0.556	0.556	0	—
	1	0.041	2.291	0.074	0.074	0.3	↑
	2	0.041	1.896	0.042	0.042	0.4	↑
	3	0.04	1.758	0.032	0.032	0.4	↑
	4	0.039	1.692	0.029	0.029	0.4	↑
	5	0.038	1.655	0.027	0.027	0.4	↑
UBALDINHO	0	-0.131	5.076	0.05	0.05	0.3	↓
	1	-0.018	2.42	0.543	0.543	0	—

	2	-0.002	1.959	0.945	0.945	0	—
	3	0.001	1.808	0.951	0.951	0	—
	4	0.003	1.736	0.91	0.91	0	—
	5	0.003	1.695	0.889	0.889	0	—
UBALDINHO (RIACHO SÃO MIGUEL)	0	-0.022	4.835	0.746	0.746	0	—
	1	0.018	2.129	0.636	0.636	0	—
	2	0.033	1.704	0.311	0.311	0.1	—
	3	0.032	1.578	0.26	0.26	0.1	—
	4	0.032	1.517	0.235	0.235	0.1	—
	5	0.031	1.483	0.222	0.222	0.1	—
VALÉRIO	0	-0.005	4.12	0.899	0.899	0	—
	1	-0.017	2.179	0.509	0.509	0	—
	2	-0.012	1.819	0.571	0.571	0	—
	3	-0.009	1.684	0.635	0.635	0	—
	4	-0.007	1.619	0.668	0.668	0	—
	5	-0.007	1.581	0.686	0.686	0	—
VÁRZEA DA VOLTA	0	-0.017	3.32	0.448	0.448	0.1	—
	1	-0.022	1.814	0.13	0.13	0.2	—
	2	-0.017	1.558	0.128	0.128	0.2	—
	3	-0.014	1.469	0.137	0.137	0.2	—
	4	-0.013	1.427	0.145	0.145	0.2	—
	5	-0.012	1.403	0.151	0.151	0.2	—
DO CORONEL	0	-0.049	2.712	0.081	0.081	0.3	↓
	1	-0.007	1.243	0.494	0.494	0	—
	2	-0.004	1.125	0.48	0.48	0.1	—
	3	-0.003	1.097	0.476	0.476	0.1	—
	4	-0.003	1.086	0.475	0.475	0.1	—
	5	-0.003	1.08	0.475	0.475	0.1	—
ORÓS	0	0.011	6.962	0.683	0.683	0	—
	1	-0.002	3.21	0.945	0.945	0	—
	2	-0.008	2.605	0.728	0.728	0	—
	3	-0.011	2.389	0.616	0.616	0	—
	4	-0.012	2.28	0.555	0.555	0	—
	5	-0.012	2.213	0.521	0.521	0	—
RODRIGO	0	0.059	4.006	0.619	0.619	0	—
	1	0.012	2.227	0.863	0.863	0	—
	2	0.001	1.926	0.98	0.98	0	—
	3	-0.001	1.806	0.988	0.988	0	—
	4	-0.001	1.742	0.975	0.975	0	—
	5	-0.002	1.703	0.967	0.967	0	—
TRUSSU	0	-0.062	5.462	0.051	0.051	0.3	↓
	1	-0.042	2.272	0.108	0.108	0.2	—
	2	-0.034	1.857	0.106	0.106	0.2	—

	3	-0.03	1.72	0.113	0.113	0.2	—
	4	-0.027	1.653	0.12	0.12	0.2	—
	5	-0.025	1.613	0.125	0.125	0.2	—

Source: prepared by the author

Identification of monthly patterns

Two patterns were identified:

1. Annual decreasing in biodiversity: 4 reservoirs (4,9 %)
2. Non-significant annual effect: 78 reservoirs (95,1 %)

Monthly reduction in biodiversity: 4 reservoirs (4,9 %)

Table H - Monthly behavior: reservoirs that showed a reduction in biodiversity according to Hill numbers

Reservoir	Hill number						Richness	Evenness	Class
	0	1	2	3	4	5			
POÇO VERDE	↓	↓	↓	↓	↓	↓	↓	↓	↓
POMPEU SOBRINHO	↓	↓	↓	↓	↓	↓	↓	↓	↓
SÃO JOAQUIM	↓	↓	↓	↓	↓	↓	↓	↓	↓
SEBASTIÃO DE ABREU	↓	↓	↓	↓	↓	↓	↓	↓	↓

Source: prepared by the author

Figure 26 - Hill numbers | Decreasing | POÇO VERDE, SÃO JOAQUIM and SEBASTIÃO DE ABREU

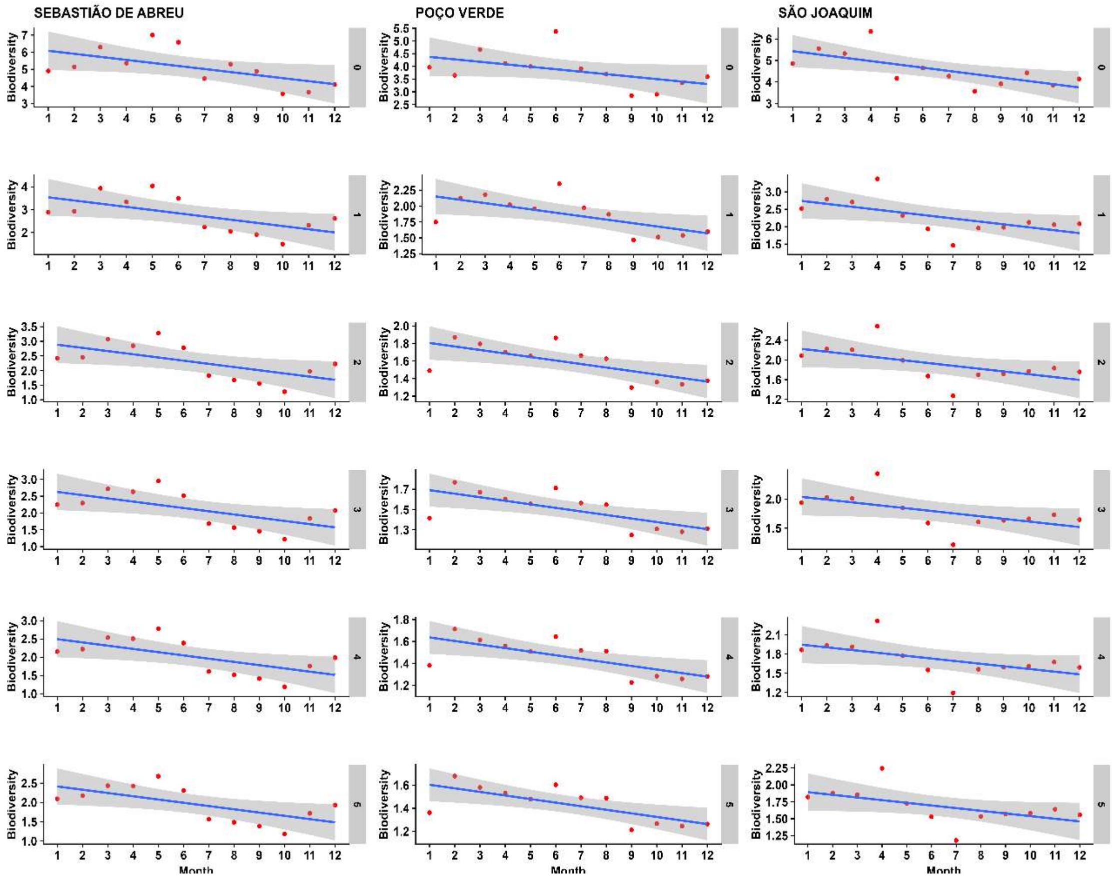


Figure 27 - Hill numbers | Decreasing | POMPEU SOBRINHO

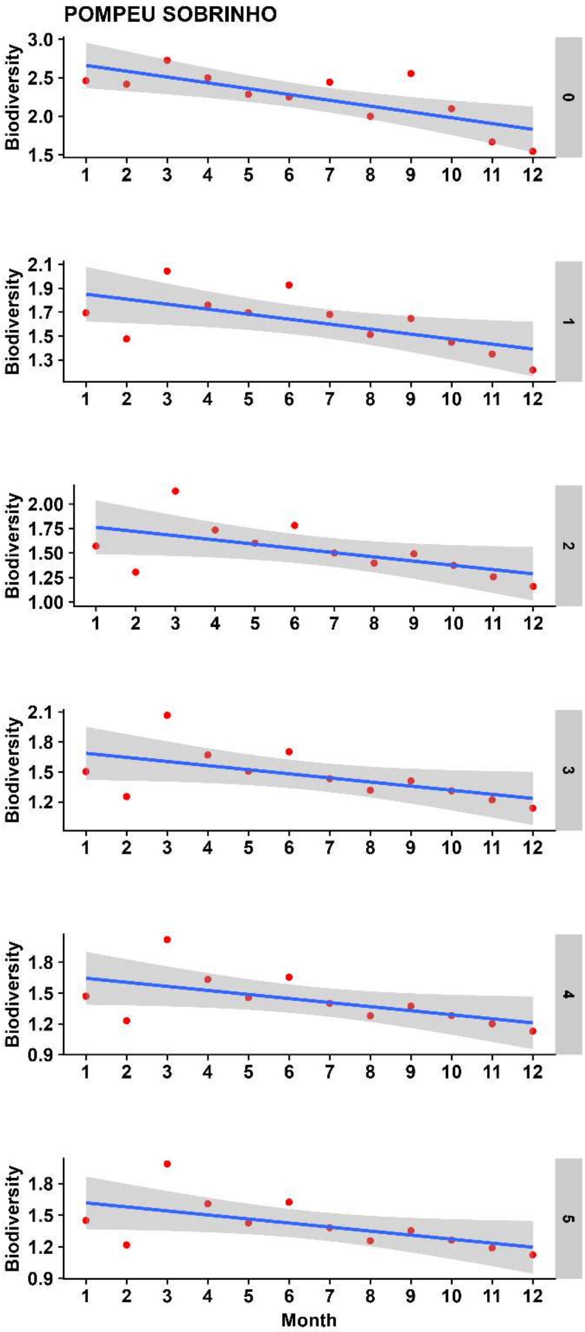
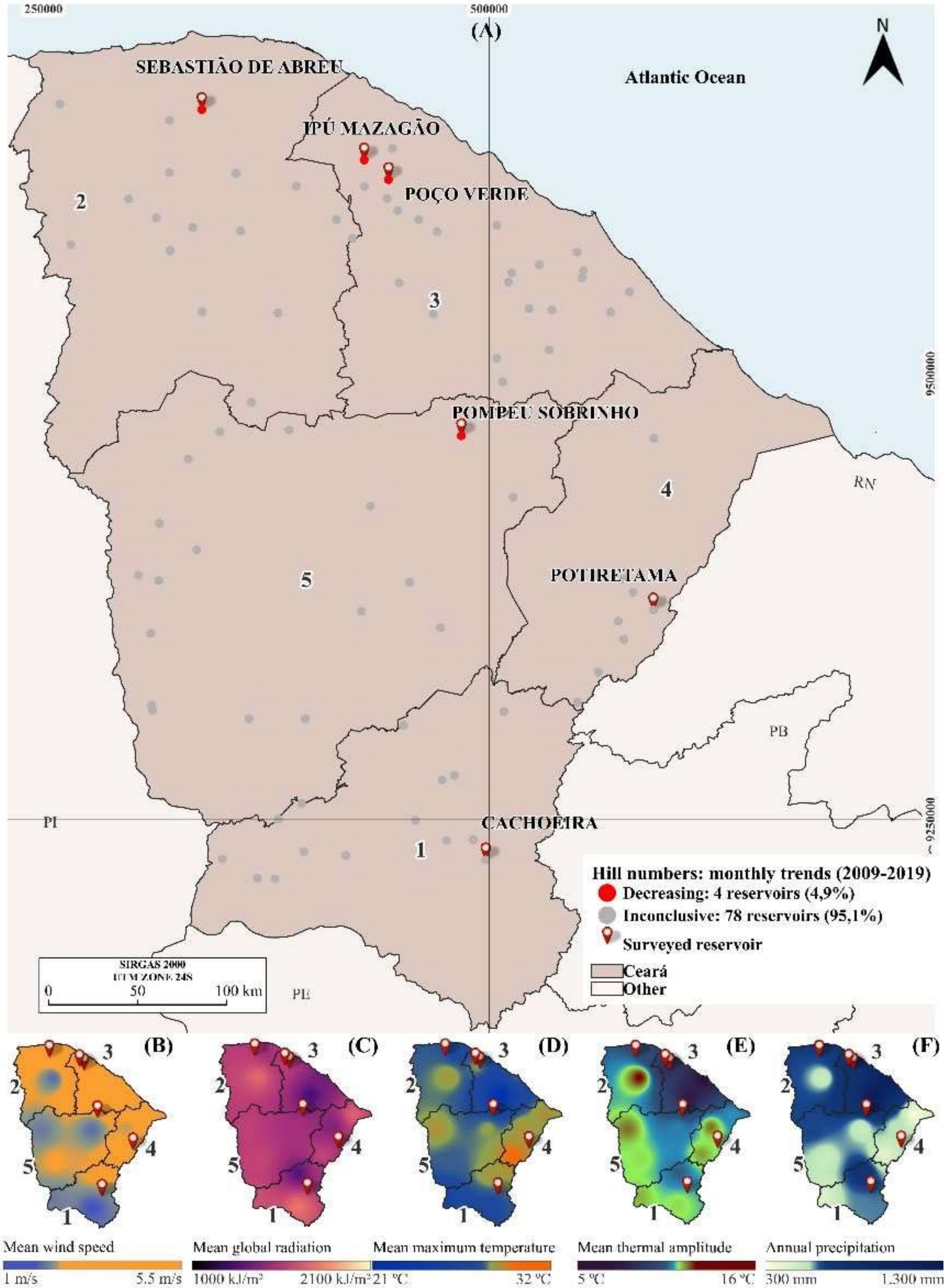

POMPEU SOBRINHO
Biodiversity
Month
0
1
2
3
4
5

Figura 28 – Em (A), localização e categorização dos reservatórios destinados ao consumo humano nas 5 macrorregiões cearenses de acordo com a tendência da média mensal da biodiversidade de cianobactérias na água bruta. Pontos vermelhos são reservatórias com redução significativa, enquanto os cinzas representam aqueles sem tendência significativa. Para exemplificação, 6 reservatórios com essas tendencias foram destacados dos demais. De (B) a (F) são mostrados os gradientes das variáveis ambientais e a localização dos reservatórios destacados



Section 5: Zivot-Andrews Test

To investigate whether the prolonged dry season had a significant impact on the behavior of the biodiversity of the cyanobacteria community in the dams intended for human supply, the Zivot-Andrew test (ZA) was performed (ZIVOT; ANDREWS, 1992). From this test, it was possible to evaluate whether the time series that describes biodiversity had a unit root with structural change, similar to other findings (LIN, B.; ULLAH, 2024;QUAN et al., 2024). This means that the test evaluated whether there was a significant change in the structure of the series during the specific period of the drought. According to the test, time series can be modeled:

$$y_t = \alpha + \beta t + \gamma D_U + \delta t D_T + \phi y_{t-1} + \sum_{i=1}^{p} \psi_i \Delta y_{t-i} + \varepsilon_t \qquad \text{(Eq. 1)}$$

Where:

- y_t is the time series value of the components describing biodiversity at time t;
- α: it is the intercept. In the context of the analysis, it represents the baseline of the series, that is the midpoint before considering any variation (with all other coefficients equal to zero);
- βD_U: β is the coefficient that captures the change in the base level of the series at the breakpoint and a D_U dummy variable that activates at the structural breakpoint, i.e. after the breakpoint $D_U = 1$, otherwise $D_U = 0$. If it is significantly different from zero ($p < 0.10$), this indicates that there was a significant change in the baseline of the sets after the breakpoint.
- γt: This term represents the linear trend over time. If significant ($p < 0.10$), the coefficient indicates the magnitude ($|\gamma|$) as the trend direction of the component series. If $\gamma > 0$ there was an upward trend, on the contrary, $\gamma < 0$ the trend was downward.
- $\delta t D_T$: The coefficient δ represents the change in the trend series after the breakout point and a D_T dummy variable that also activates at the breakout point. If significant ($p<0.10$), when, $\delta > 0$ there was a positive influence on the linear trend (increase in the magnitude of the trend) after the breakpoint, on the contrary $\delta < 0$ there was a negative influence (reduction in the magnitude of the trend). The magnitude of the change is defined by the $|\delta|$.
- ϕy_{t-1}: is associated with the previous value of the time series. If it is significant ($p < 0.1$) and equal to 1, this suggests that the time series is non-stationary (has a unit root) and that future values are a direct continuation of past values, without returning to a long-term average behavior. Alternatively, when $\phi < 1$, there is an indication of stationarity in the presence of a breakpoint. Stationary series were those with mean and constant variance over time (second-order or weakly stationary stationarity) (HAMILTON, J. D., 2020;MACHIWAL; JHA, 2012);

- $\sum_{i=1}^{p} \psi_i \Delta y_{t-i}$: This is the sum of the autoregressive terms that capture the relationship between the changes in the series and its previous values. Coefficients measure the impact that past changes have on current changes. Given the general characteristics of the graphs ACF (Autocorrelation Function), which measures the relationship between an observation and its previous ones in different time intervals, and PACF (Partial Autocorrelation Function), which evaluates this same relationship, but disregards the intermediate influences, it was decided to use Equation 2 instead of Equation 1 and not to consider this term in the analyses performed. ψ_i

$$y_t = \alpha + \beta D_U + \gamma t + \delta t D_T + \phi y_{t-1} + \varepsilon_t \qquad \text{(Eq. 2)}$$

- ε_t: An error or residue term that represents random variation that is not explained by the other components of the model. This term is expected to be independently and identically distributed, with zero mean and constant variance. The better these features are, the better the fit of the model. These analyses were performed graphically.

Table 3: List of the sub-grouped tendency for biodiversity measurement. NS means "No significant" and α, β, δ, γ are the coefficients of the ZA test

Trend	**ZA test**	**Breakpoint effect on trend or baseline level**	**Effect**
Increasing α or $\gamma > 0$	$\beta > 0$ or $\delta > 0$	Elevation in the magnitude of an increasing trend or baseline	Elevation
	$\beta < 0$ or $\delta < 0$	Reduction in the magnitude of an increasing trend or baseline	Reduction
	β or $\delta \rightarrow$ NS	Retention of a similar increasing trend or baseline level as before the breakpoint	Retention
	α or $\gamma \rightarrow$ NS $\beta > 0$ or $\delta > 0$	Initialization of an increasing trend or baseline level after the breakpoint	Initialization
Decreasing α or $\gamma < 0$	$\beta < 0$ or $\delta < 0$	Elevation in the magnitude of a decreasing trend or baseline	Elevation
	$\beta > 0$ or $\delta > 0$	Reduction in the magnitude of a decreasing trend or baseline	Reduction

	β or δ → NS	Retention of a similar decreasing trend or baseline level as before the breakpoint	Retention
	α or γ → NS β > 0 or δ > 0	Initialization of a decreasing trend or baseline level after the breakpoint	Initialization
Random	α, β, γ or δ →NS	No significant effect	Non-significant

Source: prepared by author

Table I - Coefficients of the Zivot-Andrews test for all reservoirs in which it was possible to evaluate the series that describe biodiversity. In blue, the positive values that were interpreted as elevation. In red, the negative values that were associated with reductions. BP represents the year-month in which the breakpoint occurred. RU indicates the presence (non-stationary) or absence (stationary) of a unit root of the series. The symbols "*", "**", "***" represent the significance levels of 10, 5, 1%, respectively. In parentheses, the estimated confidence interval with 95% certainty.

RESERVOIR	Components	Breakpoint effect on baseline level		Breakpoint effect on trend		Dependence on previous values	RU	BP
		Before	After	Before	After			
		α	β	γ	δ	ϕ		
ACARAPE DO MEIO	Temporal α diversity	**1.015 (0.615,1.416)*** **	-1.475 (-1.996,-0.955)***	0.03 (0.016,0.044)***	-0.022 (-0.037,-0.006)***	**0.379 (0.212,0.547)*** **	**Stationary** **	2014 - 06
ACARAPE DO MEIO	Temporal β diversity	**0.182 (0.12,0.243)*** **	**0.12 (0.044,0.196)*** **	**0.003 (0.002,0.004)*** **	**-0.004 (-0.006,-0.001)** **	**0.158 (-0.039,0.356)**	**Stationary** **	2016 - 02
ACARAPE DO MEIO	Turnover	0.022 (-0.034,0.078)	**-0.14 (-0.242,-0.039)*** **	**0.005 (0.003,0.006)*** **	0.001 (-0.003,0.006)	0.108 (-0.095,0.311)	**Stationary** **	2016 - 08
ACARAPE DO MEIO	Nestedness	**0.123 (0.067,0.18)*** **	**0.158 (0.071,0.245)*** **	0 (-0.002,0.001)	**-0.004 (-0.008,-0.001)** **	**0.275 (0.084,0.466)*** **	**Stationary** **	2016 - 07
ACARAÚ MIRIM	Temporal α diversity	**1.681 (0.988,2.374)*** **	-1.679 (-3.14,-0.217)**	0.039 (0.017,0.062)***	0.227 (-0.011,0.465)*	-0.136 (-0.415,0.143)	**Stationary** **	2018 - 04
ADAUTO BEZERRA	Temporal α diversity	1.372 (0.822,1.921)***	0.582 (0.104,1.06)**	-0.001 (-0.024,0.021)	-0.037 (-0.073,-0.001)**	0.162 (-0.126,0.45)	**Stationary** **	2013 - 11
AMANARY	Temporal β diversity	**0.213 (0.122,0.303)*** **	0.07 (-0.034,0.175)	**0.006 (0.003,0.009)*** **	**-0.008 (-0.012,-0.004)*** **	**0.055 (-0.154,0.264)**	**Stationary** **	2014 - 12
AMANARY	Nestedness	**0.096 (0.046,0.146)*** **	**-0.16 (-0.288,-0.032)** **	**0.004 (0.002,0.005)*** **	0.002 (-0.011,0.015)	**-0.338 (-0.528,-0.148)*** **	**Stationary** **	2018 - 01
AMANARY	Temporal α diversity	1.685 (1.024,2.345)***	-1.383 (-2.044,-0.721)***	0.011 (-0.004,0.025)	0.001 (-0.018,0.02)	0.372 (0.179,0.566)***	**Stationary** **	2014 - 12
AMANARY	Turnover	**0.305 (0.014,0.597)** **	**0.349 (0.121,0.577)*** **	-0.034 (-0.078,0.01)	0.034 (-0.01,0.078)	-0.106 (-0.31,0.099)	**Stationary** **	2011 - 08

ANGICOS	Temporal β diversity	**0.307 (0.207,0.408)*** **	**-0.278 (-0.393,-0.163)*** **	**0.012 (0.007,0.017)*** **	**-0.014 (-0.019,-0.008)*** **	**0.005 (-0.22,0.23)**	**Stationary** **	01/01/2015
ANGICOS	Nestedness	**0.135 (0.038,0.232)*** **	**-0.122 (-0.218,-0.026)** **	**0.009 (0.003,0.015)*** **	**-0.011 (-0.017,-0.004)*** **	0.05 (-0.179,0.278)	**Stationary** **	2014 - 01
ANGICOS	Temporal α diversity	**0.552 (0.265,0.838)*** **	0.955 (0.501,1.408)***	0.009 (0.002,0.016)***	-0.059 (-0.088,-0.03)***	**0.513 (0.339,0.687)*** **	**Stationary*** **	2017 - 02
ANGICOS	Turnover	0.063 (-0.035,0.162)	**-0.174 (-0.293,-0.054)*** **	**0.013 (0.007,0.019)*** **	**-0.015 (-0.021,-0.008)*** **	-0.067 (-0.292,0.157)	**Stationary** **	2014 - 09
ARACOIABA	Nestedness	**0.376 (0.242,0.511)*** **	**0.222 (0.121,0.323)*** **	**-0.038 (-0.056,-0.02)*** **	**0.038 (0.02,0.056)*** **	**-0.262 (-0.454,-0.07)*** **	**Stationary** **	2011 - 09
ARACOIABA	Temporal α diversity	**1.168 (0.628,1.708)*** **	-1.658 (-2.839,-0.476)***	0.008 (0,0.015)**	0.118 (-0.044,0.28)	**0.512 (0.345,0.678)*** **	**Stationary*** **	2018 - 05
ARACOIABA	Turnover	**0.166 (0.031,0.302)** **	**0.331 (0.197,0.465)*** **	-0.009 (-0.021,0.003)	0.007 (-0.005,0.019)	-0.122 (-0.318,0.073)	**Stationary** **	2012 - 04
ARACOIABA	Temporal β diversity	**0.242 (0.108,0.376)*** **	**0.317 (0.186,0.447)*** **	**-0.01 (-0.021,0.001)* **	0.008 (-0.003,0.019)	**0.168 (-0.024,0.359)* **	**Stationary** **	01/04/2012
ARARAS	Temporal β diversity	**0.218 (0.112,0.324)*** **	0.056 (-0.03,0.142)	**0.011 (0.004,0.017)*** **	**-0.012 (-0.019,-0.006)*** **	**0.191 (-0.017,0.399)* **	**Stationary** **	01/07/2012
ARARAS	Temporal α diversity	**0.477 (0.239,0.715)*** **	0.362 (0.003,0.72)**	0.017 (0.011,0.024)***	0.004 (-0.013,0.022)	**0.441 (0.265,0.617)*** **	**Stationary*** **	2016 - 01
ARARAS	Turnover	**0.143 (0.03,0.256)** **	**0.16 (0.049,0.271)*** **	0.004 (-0.003,0.012)	-0.005 (-0.013,0.003)	-0.024 (-0.227,0.178)	**Stationary** **	2012 - 07
ARARAS	Nestedness	**0.28 (0.099,0.461)*** **	**0.211 (0.092,0.33)*** **	-0.029 (-0.066,0.008)	0.028 (-0.009,0.064)	-0.024 (-0.217,0.169)	**Stationary** **	2011 - 03
ARNEIROZ II	Turnover	**0.26 (0.139,0.38)*** **	**-0.255 (-0.431,-0.08)*** **	**0.005 (-0.001,0.011)* **	-0.001 (-0.024,0.022)	-0.224 (-0.55,0.102)	**Stationary** **	2017 - 04
ARNEIROZ II	Temporal α diversity	1.59 (0.645,2.536)***	1.516 (0.707,2.325)***	-0.057 (-0.145,0.031)	0.02 (-0.069,0.108)	0.276 (-0.016,0.568)*	**Stationary** **	2016 - 11
ARNEIROZ II	Temporal β diversity	**0.463 (0.285,0.64)*** **	**0.139 (0.023,0.254)** **	-0.007 (-0.018,0.004)	0.001 (-0.012,0.013)	-0.072 (-0.411,0.268)	**Stationary** **	01/04/2016

ARNEIROZ II	Nestedness	**0.215 (0.135,0.295)*** **	**0.197 (0.073,0.322)*** **	-0.002 (-0.006,0.001)	-0.007 (-0.025,0.012)	**-0.35 (-0.68,-0.021)** **	**Stationary** **	2017 - 06
AYRES DE SOUZA	Temporal β diversity	**0.615 (0.356,0.874)*** **	**0.375 (0.199,0.552)*** **	**-0.064 (-0.1,-0.029)*** **	**0.064 (0.029,0.1)*** **	**0.377 (0.203,0.551)*** **	**Stationary** **	01/06/20 11
AYRES DE SOUZA	Turnover	**0.226 (0.131,0.321)*** **	**-0.237 (-0.378,-0.097)*** **	**0.003 (0.001,0.005)*** **	0.002 (-0.004,0.007)	0.12 (-0.075,0.316)	**Stationary** **	2016 - 06
AYRES DE SOUZA	Temporal α diversity	**1.467 (0.991,1.942)*** **	-0.75 (-1.465,-0.035)**	0.008 (0.001,0.014)**	0.14 (0.082,0.197)***	**0.172 (-0.025,0.369)* **	**Stationary*** **	2016 - 01
AYRES DE SOUZA	Nestedness	**0.169 (0.101,0.236)*** **	**0.121 (0.031,0.212)*** **	0 (-0.002,0.002)	**-0.003 (-0.006,0)* **	-0.038 (-0.235,0.159)	**Stationary** **	2015 - 08
BOQUEIRÃO	Nestedness	**0.223 (0.166,0.28)*** **	**0.358 (0.205,0.51)*** **	**-0.001 (-0.002,0)* **	**-0.064 (-0.093,-0.036)*** **	-0.099 (-0.286,0.088)	**Stationary** **	2018 - 05
BOQUEIRÃO	Temporal α diversity	1.36 (0.858,1.862)***	0.692 (0.184,1.2)***	0.001 (-0.009,0.011)	-0.006 (-0.024,0.013)	0.302 (0.1,0.503)***	**Stationary** **	2015 - 12
BOQUEIRÃO	Temporal β diversity	**0.254 (0.145,0.362)*** **	**0.17 (0.066,0.273)*** **	0 (-0.003,0.003)	-0.002 (-0.006,0.001)	**0.323 (0.128,0.517)*** **	**Stationary** **	2014 - 09
BOQUEIRÃO	Turnover	**0.151 (0.065,0.238)*** **	**0.218 (0.102,0.335)*** **	0.001 (-0.002,0.004)	**-0.006 (-0.01,-0.002)*** **	0.077 (-0.133,0.287)	**Stationary** **	2015 - 01
CACHOEIRA	Temporal α diversity	**1.546 (0.904,2.189)*** **	-1.463 (-2.242,-0.684)***	0.027 (0.008,0.046)***	-0.008 (-0.037,0.022)	**0.199 (-0.012,0.409)* **	**Stationary*** **	2017 - 08
CANOAS	Nestedness	**0.168 (-0.022,0.357)* **	**-0.443 (-0.611,-0.276)*** **	**0.053 (0.018,0.088)*** **	**-0.054 (-0.089,-0.019)*** **	-0.09 (-0.295,0.115)	**Stationary** **	2011 - 05
CANOAS	Temporal β diversity	0.068 (-0.156,0.291)	**-0.295 (-0.458,-0.131)*** **	**0.073 (0.029,0.118)*** **	**-0.074 (-0.118,-0.029)*** **	**-0.022 (-0.221,0.178)**	**Stationary** **	2011 - 04
CANOAS	Temporal α diversity	**0.982 (0.3,1.664)*** **	-0.889 (-1.472,-0.306)***	0.06 (-0.002,0.122)*	-0.052 (-0.114,0.01)	**0.307 (0.117,0.497)*** **	**Stationary*** **	2015 - 09
CANOAS	Turnover	-0.035 (-0.192,0.123)	**0.133 (0.002,0.265)** **	0.008 (-0.012,0.029)	-0.008 (-0.029,0.012)	-0.127 (-0.327,0.073)	**Stationary** **	2011 - 08
CARÃO	Nestedness	**0.388 (0.242,0.533)*** **	**0.098 (0,0.197)* **	**-0.035 (-0.051,-0.018)*** **	**0.036 (0.02,0.053)*** **	0.077 (-0.162,0.316)	**Stationary** **	2011 - 08

CARÃO	Temporal α diversity	**1.087 (0.673,1.5)*** **	-0.724 (-1.293,-0.155)**	0.02 (0.008,0.032)***	0.011 (-0.028,0.05)	0.077 (-0.182,0.335)	**Stationary*** **	2015 - 10
CARÃO	Temporal β diversity	**0.234 (-0.019,0.486)***	**0.239 (0.022,0.455)****	-0.02 (-0.052,0.011)	0.021 (-0.011,0.053)	**0.439 (0.212,0.666)*****	**Stationary****	2011 - 08
CARÃO	Turnover	0.081 (-0.025,0.188)	**0.253 (0.043,0.462)****	0.002 (-0.002,0.005)	**-0.023 (-0.04,-0.006)****	**0.424 (0.2,0.647)*****	**Stationary****	2014 - 11
CASTRO	Temporal α diversity	**1.879 (1.384,2.374)*****	-0.535 (-0.837,-0.234)***	0.009 (0.001,0.017)**	-0.012 (-0.025,0)*	0.12 (-0.091,0.331)	**Stationary*** **	2011 - 12
CASTRO	Turnover	0.027 (-0.029,0.083)	**0.535 (0.364,0.705)*****	**0.002 (0.001,0.004)*****	**-0.056 (-0.077,-0.036)*****	0.12 (-0.076,0.316)	**Stationary****	2016 - 09
CASTRO	Temporal β diversity	**0.233 (0.166,0.3)*****	**0.624 (0.45,0.799)*****	**0.002 (0.001,0.003)*****	**-0.067 (-0.088,-0.046)*****	**-0.077 (-0.298,0.143)**	**Stationary****	2016 - 10
CASTRO	Nestedness	**0.186 (0.112,0.259)*****	**-0.149 (-0.233,-0.064)*****	0.002 (-0.001,0.005)	0 (-0.004,0.003)	-0.15 (-0.371,0.071)	**Stationary****	2013 - 08
CAXITORÉ	Temporal β diversity	**0.204 (0.09,0.318)*****	**-0.197 (-0.314,-0.08)*****	**0.013 (0.007,0.019)*****	**-0.014 (-0.02,-0.008)*****	**-0.008 (-0.203,0.188)**	**Stationary****	2013 - 03
CAXITORÉ	Turnover	**0.11 (-0.013,0.233)***	**-0.245 (-0.38,-0.109)*****	**0.014 (0.007,0.021)*****	**-0.015 (-0.022,-0.008)*****	-0.13 (-0.328,0.069)	**Stationary****	2013 - 03
CAXITORÉ	Temporal α diversity	**0.411 (-0.022,0.844)***	-1.188 (-2.006,-0.37)***	0.023 (0.01,0.035)***	0.027 (-0.007,0.06)	**0.491 (0.315,0.667)*****	**Stationary*** **	2014 - 09
CAXITORÉ	Nestedness	**0.136 (0.072,0.2)*****	**0.097 (0.021,0.173)****	0 (-0.003,0.002)	-0.001 (-0.004,0.002)	-0.151 (-0.35,0.047)	**Stationary****	2014 - 05
DO CORONEL	Temporal α diversity	**0.609 (0.273,0.945)*****	-0.211 (-0.377,-0.045)**	0.034 (0.006,0.061)**	-0.03 (-0.057,-0.002)**	0.307 (-0.03,0.645)*	**Non-stationary**	2015 - 04
DO CORONEL	Nestedness	**0.518 (0.273,0.763)*****	**0.206 (0.02,0.393)****	**-0.036 (-0.062,-0.009)****	**0.034 (0.007,0.061)****	-0.154 (-0.511,0.203)	**Stationary****	2011 - 08
DO CORONEL	Turnover	-0.069 (-0.2,0.061)	**-0.142 (-0.295,0.011)***	**0.013 (0,0.025)****	-0.005 (-0.021,0.01)	-0.096 (-0.456,0.264)	**Stationary****	2012 - 01
DO CORONEL	Temporal β diversity	**0.256 (0.117,0.394)*****	**-0.304 (-0.555,-0.054)****	0.003 (-0.003,0.009)	**0.064 (0.013,0.116)****	**-0.02 (-0.364,0.325)**	**Stationary****	2013 - 02
EDSON QUEIROZ	Temporal β diversity	**0.169 (0.089,0.25)*****	**-0.363 (-0.474,-0.252)*****	**0.012 (0.008,0.016)*****	**-0.008 (-0.012,-0.004)*****	**-0.023 (-0.221,0.176)**	**Stationary****	2013 - 11

EDSON QUEIROZ	Temporal α diversity	**0.846 (0.482,1.211)*****	0.836 (0.314,1.357)***	0.007 (0.001,0.012)**	-0.044 (-0.074,-0.014)***	**0.469 (0.294,0.645)*****	**Stationary*****	2016 - 06
EDSON QUEIROZ	Turnover	0.011 (-0.073,0.096)	**-0.248 (-0.36,-0.136)*****	**0.008 (0.004,0.012)*****	**-0.005 (-0.009,0)****	0.105 (-0.092,0.302)	**Stationary****	2013 - 11
EDSON QUEIROZ	Nestedness	**0.314 (-0.017,0.645)***	**0.201 (0.029,0.372)****	-0.059 (-0.148,0.03)	0.058 (-0.031,0.147)	-0.005 (-0.201,0.192)	**Stationary****	2011 - 01
EMA	Temporal α diversity	1.48 (0.962,1.997)***	0.939 (0.426,1.451)***	-0.001 (-0.013,0.012)	-0.037 (-0.06,-0.014)***	0.24 (0.024,0.455)**	**Stationary****	2011 - 07
FIGUEIREDO	Temporal α diversity	1.988 (0.853,3.122)***	1.152 (0.091,2.213)**	0.029 (-0.082,0.139)	-0.097 (-0.223,0.03)	-0.171 (-0.53,0.188)	**Stationary****	2017 - 03
FLOR DO CAMPO	Temporal α diversity	**0.927 (0.443,1.41)*****	-1.094 (-1.779,-0.409)***	0.012 (0,0.024)*	0.046 (0.02,0.071)***	**0.352 (0.163,0.541)*****	**Stationary*****	2014 - 08
FORQUILHA	Turnover	-0.081 (-0.285,0.123)	**-0.209 (-0.418,0)***	**0.034 (0,0.068)***	-0.017 (-0.057,0.022)	0.211 (-0.228,0.65)	**Non-stationary**	2016 - 03
FORQUILHA	Temporal α diversity	2.922 (1.376,4.469)***	-0.689 (-1.558,0.18)	-0.037 (-0.121,0.047)	-0.019 (-0.196,0.158)	0.088 (-0.407,0.583)	**Non-stationary**	-
FORQUILHA	Temporal β diversity	0.124 (-0.132,0.38)	**-0.199 (-0.431,0.032)***	0.026 (-0.011,0.063)	-0.018 (-0.06,0.025)	0.159 (-0.348,0.665)	**Non-stationary**	2016 - 10
FORQUILHA	Nestedness	**0.223 (0.066,0.38)*****	0.124 (-0.028,0.277)	**-0.013 (-0.02,-0.006)*****	NA (NA,NA)	-0.401 (-0.891,0.088)	**Stationary****	-
GAMELEIRA	Temporal α diversity	2.2 (1.388,3.012)***	2.436 (1.393,3.479)***	0.023 (-0.009,0.054)	-0.163 (-0.227,-0.099)***	-0.157 (-0.426,0.112)	**Stationary****	2015 - 07
GAVIÃO	Nestedness	0.029 (-0.138,0.197)	**-0.105 (-0.222,0.012)***	**0.032 (0.001,0.063)****	**-0.032 (-0.063,0)****	0.047 (-0.15,0.245)	**Stationary****	2011 - 04
GAVIÃO	Temporal α diversity	**0.741 (0.329,1.153)*****	-0.941 (-1.793,-0.088)**	0.017 (0.008,0.026)***	0.01 (-0.065,0.085)	**0.524 (0.358,0.69)*****	**Stationary*****	2017 - 12
GAVIÃO	Temporal β diversity	**0.29 (0.18,0.4)*****	**0.121 (0.027,0.214)****	0.002 (-0.003,0.006)	-0.003 (-0.008,0.001)	**0.205 (0.011,0.398)****	**Stationary****	2013 - 03
GAVIÃO	Turnover	**0.18 (0.077,0.283)*****	**0.147 (0.038,0.256)*****	0 (-0.004,0.005)	-0.002 (-0.007,0.003)	0.129 (-0.067,0.325)	**Stationary****	2013 - 05
GENERAL SAMPAIO	Temporal α diversity	1.53 (0.656,2.405)***	1.113 (0.287,1.938)***	-0.024 (-0.071,0.023)	0.022 (-0.025,0.07)	0.384 (0.195,0.572)***	**Stationary****	2016 - 12

IPÚ MAZAGÃO	Turnover	**0.055 (-0.01,0.12)***	0.154 (-0.157,0.466)	**0.008 (0.004,0.011)*****	**-0.206 (-0.354,-0.057)*****	-0.097 (-0.416,0.223)	**Stationary****	2014 - 08
IPÚ MAZAGÃO	Nestedness	-0.026 (-0.124,0.072)	**-0.198 (-0.308,-0.088)*****	**0.026 (0.014,0.037)*****	**-0.027 (-0.039,-0.014)*****	-0.224 (-0.531,0.084)	**Stationary****	2012 - 06
IPÚ MAZAGÃO	Temporal α diversity	3.873 (2.516,5.231)***	-0.886 (-3.092,1.32)	-0.012 (-0.036,0.012)	-0.046 (-0.822,0.73)	-0.012 (-0.337,0.313)	**Stationary****	2014 - 10
IPÚ MAZAGÃO	Temporal β diversity	0.056 (-0.169,0.282)	**0.238 (0.053,0.422)****	0.012 (-0.03,0.053)	-0.012 (-0.054,0.029)	0.053 (-0.269,0.375)	**Stationary****	2011 - 11
ITAPEBUSSU	Temporal α diversity	**1.832 (1.298,2.367)*****	1.938 (0.604,3.272)***	0.01 (0.002,0.017)**	-0.389 (-0.644,-0.135)***	-0.006 (-0.219,0.206)	**Stationary*****	2017 - 12
ITAÚNA	Temporal α diversity	**1.809 (1.175,2.442)*****	-1.125 (-1.906,-0.345)***	0.046 (0.028,0.063)***	-0.065 (-0.108,-0.021)***	0.117 (-0.107,0.34)	**Stationary*****	2012 - 12
JABURU I	Temporal α diversity	**1.694 (1.219,2.168)*****	-1.077 (-1.553,-0.602)***	0.016 (0.008,0.024)***	-0.028 (-0.064,0.007)	0.096 (-0.126,0.319)	**Stationary*****	2017 - 05
JENIPAPEIRO	Temporal α diversity	**2.657 (1.906,3.409)*****	1.454 (0.012,2.896)**	0.016 (-0.002,0.035)*	-0.296 (-0.538,-0.054)**	**-0.296 (-0.56,-0.033)****	**Stationary*****	2018 - 08
JOÃO JOSÉ DE CASTRO	Temporal α diversity	1.447 (0.767,2.128)***	-0.443 (-1.729,0.842)	0.007 (-0.007,0.021)	0.721 (0.138,1.304)**	0.097 (-0.258,0.452)	**Stationary****	2018 - 02
JOÃO LUÍS	Temporal α diversity	2.188 (0.667,3.708)***	-1.198 (-2.183,-0.213)**	-0.033 (-0.372,0.306)	0.054 (-0.285,0.393)	0.23 (-0.025,0.485)*	**Stationary****	2016 - 08
- 06CO	Temporal α diversity	**1.324 (0.626,2.022)*****	-1.344 (-2.185,-0.504)***	0.072 (0.031,0.112)***	-0.03 (-0.086,0.026)	-0.09 (-0.375,0.196)	**Stationary*****	2017 - 05
MADEIRO	Temporal α diversity	**3.768 (2.3,5.235)*****	0.852 (0.144,1.56)**	-0.318 (-0.54,-0.097)***	0.316 (0.094,0.538)***	-0.107 (-0.426,0.212)	**Stationary*****	2018 - 04
MADEIRO	Temporal β diversity	0.033 (-0.055,0.121)	**-0.167 (-0.303,-0.031)****	**0.013 (0.006,0.019)*****	-0.005 (-0.017,0.007)	0.034 (-0.284,0.353)	**Stationary****	2012 - 11
MADEIRO	Nestedness	0.015 (-0.055,0.085)	**-0.217 (-0.354,-0.081)*****	**0.006 (0.001,0.01)*****	**0.018 (0.003,0.032)****	0.24 (-0.05,0.53)	**Stationary****	2013 - 03
MADEIRO	Turnover	0.077 (-0.021,0.175)	**0.183 (0.053,0.314)*****	-0.003 (-0.01,0.004)	-0.004 (-0.014,0.005)	**0.333 (0.046,0.619)****	**Non-stationary**	2012 - 06

MALCOZINHADO	Turnover	**0.122 (0.004,0.24)****	**-0.144 (-0.281,-0.006)****	**0.015 (0,0.03)****	**-0.018 (-0.034,-0.001)****	-0.148 (-0.561,0.265)	**Stationary****	2017 - 06
MALCOZINHADO	Temporal α diversity	5.146 (2.625,7.667)***	-1.948 (-3.685,-0.211)**	-0.019 (-0.051,0.013)		0.114 (-0.227,0.455)	**Stationary****	2018 - 03
MALCOZINHADO	Nestedness	**0.182 (0.094,0.269)*****	**0.169 (0.065,0.273)*****	-0.003 (-0.008,0.002)	-0.003 (-0.014,0.009)	-0.157 (-0.525,0.21)	**Stationary****	2018 - 01
MALCOZINHADO	Temporal β diversity	**0.342 (0.207,0.477)*****	0.233 (-0.055,0.52)	0.002 (-0.003,0.006)	**-0.15 (-0.279,-0.02)****	**-0.093 (-0.47,0.284)**	**Stationary****	2018 - 09
MAMOEIRO	Temporal α diversity	**4.328 (2.932,5.723)*****	2.124 (1.177,3.07)***	-0.195 (-0.283,-0.106)***	0.165 (0.078,0.253)***	0.001 (-0.249,0.251)	**Stationary*****	2014 - 08
- 03TINÓPOLE	Turnover	0.023 (-0.024,0.07)	**0.259 (0.13,0.387)*****	**0.001 (0,0.002)****	**-0.014 (-0.029,0.001)***	-0.144 (-0.363,0.075)	**Stationary****	2017 - 12
- 03TINÓPOLE	Temporal β diversity	**0.11 (0.043,0.177)*****	**0.164 (0.057,0.27)*****	**0.002 (0,0.004)****	-0.001 (-0.005,0.004)	-0.025 (-0.245,0.195)	**Stationary****	2016 - 05
- 03TINÓPOLE	Temporal α diversity	**1.278 (0.701,1.856)*****	0.899 (0.388,1.41)***	-0.011 (-0.017,-0.004)***	-0.037 (-0.083,0.009)	0.647 (0.496,0.797)***	**Non-stationary**	2015 - 09
- 03TINÓPOLE	Nestedness	**0.116 (0.058,0.175)*****	**0.233 (0.133,0.332)*****	0.001 (-0.001,0.002)	**-0.006 (-0.011,-0.002)*****	**-0.211 (-0.425,0.002)***	**Stationary****	2016 - 06
MISSI	Temporal α diversity	**1.855 (1.17,2.539)*****	-1.413 (-2.182,-0.644)***	0.066 (0.037,0.094)***	0.004 (-0.055,0.062)	-0.006 (-0.291,0.279)	**Stationary*****	2011 - 03
MISSI	Nestedness	**0.095 (0.034,0.156)*****	**-0.104 (-0.193,-0.016)****	0.002 (-0.001,0.005)	**0.012 (0.004,0.02)*****	-0.07 (-0.352,0.212)	**Stationary****	2017 - 07
MISSI	Temporal β diversity	**0.429 (0.315,0.543)*****	**0.145 (0.037,0.253)****	-0.001 (-0.004,0.003)	-0.005 (-0.013,0.004)	-0.126 (-0.39,0.139)	**Stationary****	2017 - 06
MISSI	Turnover	**0.315 (0.209,0.421)*****	**0.223 (0.085,0.361)*****	-0.002 (-0.006,0.002)	**-0.018 (-0.029,-0.006)*****	-0.116 (-0.385,0.153)	**Stationary****	2017 - 07
MONSENHOR TABOSA	Temporal α diversity	**0.837 (0.545,1.129)*****	0.344 (0.12,0.568)***	0.003 (0,0.006)*	-0.023 (-0.041,-0.005)**	**0.369 (0.161,0.577)*****	**Stationary*****	2016 - 06
MUNDAÚ	Nestedness	**0.378 (0.161,0.595)*****	**0.184 (0.036,0.332)****	**-0.038 (-0.072,-0.004)****	**0.038 (0.004,0.072)****	-0.155 (-0.371,0.062)	**Stationary****	2012 - 11

MUNDAÚ	Temporal α diversity	1.477 (1.007,1.947)***	0.996 (0.169,1.822)**	0.006 (-0.002,0.013)	-0.016 (-0.122,0.091)	0.083 (-0.146,0.311)	**Stationary****	2015 - 03
MUNDAÚ	Temporal β diversity	**0.358 (0.055,0.66)****	**0.343 (0.121,0.565)*****	-0.029 (-0.078,0.019)	0.028 (-0.02,0.077)	0.117 (-0.099,0.332)	**Stationary****	2012 - 11
MUNDAÚ	Turnover	0.067 (-0.209,0.342)	0.164 (-0.056,0.384)	0.005 (-0.037,0.048)	-0.006 (-0.048,0.036)	0.071 (-0.157,0.298)	**Stationary****	-
OLHO D'ÁGUA	Temporal α diversity	**1.328 (0.778,1.878)*****	-0.771 (-1.28,-0.261)***	0.017 (-0.003,0.037)*	-0.005 (-0.042,0.032)	0.173 (-0.114,0.461)	**Stationary*****	2014 - 06
OLHO D'ÁGUA	Nestedness	**0.239 (0.142,0.335)*****	**0.421 (0.197,0.645)*****	**-0.003 (-0.005,0)***	**-0.052 (-0.101,-0.002)****	-0.05 (-0.352,0.251)	**Stationary****	2018 - 07
OLHO D'ÁGUA	Temporal β diversity	**0.502 (0.337,0.667)*****	**0.322 (0.13,0.514)*****	**-0.007 (-0.011,-0.002)*****	0 (-0.025,0.025)	-0.053 (-0.358,0.251)	**Stationary****	2018 - 02
OLHO D'ÁGUA	Turnover	**0.268 (0.155,0.381)*****	**0.296 (0.093,0.498)*****	-0.003 (-0.007,0.001)	**-0.026 (-0.052,0.001)***	-0.179 (-0.473,0.114)	**Stationary****	2018 - 02
ORÓS	Nestedness	0.026 (-0.051,0.102)	**-0.157 (-0.24,-0.074)*****	**0.011 (0.005,0.017)*****	**-0.01 (-0.016,-0.003)*****	-0.034 (-0.265,0.198)	**Stationary****	2012 - 05
ORÓS	Temporal β diversity	**0.195 (0.119,0.27)*****	**-0.194 (-0.348,-0.041)****	**0.005 (0.003,0.007)*****	-0.008 (-0.022,0.007)	-0.046 (-0.268,0.175)	**Stationary****	2016 - 07
ORÓS	Turnover	**0.072 (0.009,0.135)****	**-0.219 (-0.379,-0.06)*****	**0.004 (0.002,0.006)*****	-0.006 (-0.023,0.011)	0.014 (-0.207,0.235)	**Stationary****	2016 - 08
ORÓS	Temporal α diversity	1.546 (0.631,2.462)***	0.408 (-0.274,1.09)	-0.042 (-0.101,0.017)	0.069 (0.007,0.132)**	0.385 (0.182,0.588)***	**Stationary****	2012 - 09
PACOTI	Temporal α diversity	**0.681 (0.205,1.158)*****	-1.62 (-2.358,-0.883)***	0.026 (0.011,0.041)***	0.005 (-0.016,0.026)	**0.463 (0.299,0.628)*****	**Stationary*****	2017 - 09
PATU	Temporal β diversity	**0.332 (0.23,0.435)*****	-0.052 (-0.155,0.051)	**0.003 (0.001,0.006)*****	**-0.01 (-0.015,-0.004)*****	**0.041 (-0.185,0.267)**	**Stationary****	2014 - 09
PATU	Turnover	**0.13 (0.047,0.213)*****	-0.089 (-0.206,0.028)	**0.005 (0.002,0.008)*****	**-0.011 (-0.017,-0.005)*****	0.031 (-0.197,0.258)	**Stationary****	2014 - 09
PATU	Nestedness	0.036 (-0.095,0.167)	**-0.239 (-0.361,-0.117)*****	**0.027 (0.01,0.044)*****	**-0.027 (-0.044,-0.01)*****	-0.024 (-0.245,0.198)	**Stationary****	2011 - 09

PATU	Temporal α diversity	**1.235 (0.76,1.71)*** **	-0.947 (-1.738,-0.156)**	0.019 (0.009,0.03)***	0.096 (0.024,0.167)***	0.148 (-0.075,0.37)	**Stationary*****	2017 - 07
PEDRAS BRANCAS	Nestedness	**-0.162 (-0.427,0.103)**	**-0.279 (-0.419,-0.14)*** **	**0.126 (0.053,0.199)*** **	**-0.126 (-0.199,-0.053)*** **	0.01 (-0.184,0.204)	**Stationary****	2011 - 03
PEDRAS BRANCAS	Temporal α diversity	**1.518 (1.086,1.95)*** **	-1.392 (-1.995,-0.788)***	0.017 (0.011,0.023)***	0.038 (-0.01,0.087)	**0.199 (0.004,0.394)****	**Stationary*****	2018 - 01
PEDRAS BRANCAS	Turnover	**0.063 (0.008,0.118)****	**-0.276 (-0.465,-0.087)*** **	**0.003 (0.002,0.004)*** **	**0.042 (0.01,0.074)****	0.074 (-0.131,0.279)	**Stationary****	2018 - 02
PEDRAS BRANCAS	Temporal β diversity	**0.305 (0.215,0.394)*** **	**0.141 (0.041,0.242)*** **	0.001 (-0.001,0.003)	-0.002 (-0.006,0.001)	-0.006 (-0.214,0.202)	**Stationary****	2015 - 05
PENTECOSTE	Temporal β diversity	**0.293 (0.116,0.47)*** **	**-0.166 (-0.312,-0.02)****	**0.019 (0.001,0.037)****	**-0.018 (-0.037,0)***	**0.009 (-0.238,0.256)**	**Stationary****	2014 - 07
PENTECOSTE	Turnover	0.061 (-0.117,0.238)	**-0.252 (-0.414,-0.09)*** **	**0.026 (0.006,0.046)****	**-0.024 (-0.044,-0.003)****	0.036 (-0.204,0.276)	**Stationary****	2014 - 07
PENTECOSTE	Temporal α diversity	**2.009 (1.095,2.922)*** **	1.483 (0.687,2.279)***	-0.051 (-0.099,-0.003)**	0.042 (-0.008,0.091)*	**0.237 (0.001,0.473)****	**Stationary*****	2017 - 05
PENTECOSTE	Nestedness	**0.126 (0.027,0.225)****	**-0.128 (-0.233,-0.022)****	**0.006 (-0.001,0.013)***	-0.006 (-0.013,0.002)	0.08 (-0.164,0.325)	**Stationary****	2015 - 04
POMPEU SOBRINHO	Temporal α diversity	1.029 (0.571,1.487)***	0.663 (0.082,1.244)**	0.001 (-0.008,0.01)	-0.041 (-0.076,-0.006)**	0.319 (0.103,0.534)***	**Stationary****	2012 - 04
POTIRETAMA	Temporal α diversity	**1.235 (0.668,1.803)*** **	-0.444 (-0.768,-0.12)***	0.028 (0.001,0.056)**	-0.026 (-0.057,0.006)	0.11 (-0.218,0.438)	**Stationary*****	2015 - 04
POTIRETAMA	Turnover	**-0.046 (-0.099,0.008)***	**-0.116 (-0.194,-0.038)*** **	**0.007 (0.003,0.011)*** **	-0.002 (-0.009,0.006)	0.01 (-0.315,0.335)	**Stationary****	2012 - 09
POTIRETAMA	Temporal β diversity	0.025 (-0.053,0.102)	**-0.294 (-0.523,-0.066)****	**0.006 (0.002,0.01)*** **	NA (NA,NA)	0.197 (-0.123,0.518)	**Stationary****	2013 - 11
POTIRETAMA	Nestedness	0.034 (-0.047,0.115)	**0.227 (0.062,0.392)*** **	0.003 (-0.001,0.008)	**-0.04 (-0.065,-0.014)*** **	0.222 (-0.086,0.529)	**Stationary****	2013 - 03
RIACHÃO	Temporal β diversity	**0.177 (0.103,0.252)*** **	**-0.092 (-0.166,-0.018)****	**0.002 (0,0.004)****	**-0.002 (-0.005,0)***	**0.408 (0.231,0.586)*** **	**Stationary****	2015 - 05

RIACHÃO	Nestedness	0.02 (-0.133,0.172)	**-0.246 (-0.36,-0.133)*** **	**0.045 (0.019,0.07)*** **	**-0.045 (-0.071,-0.02)*** **	-0.026 (-0.213,0.16)	**Stationary** **	2011 - 05
RIACHÃO	Temporal α diversity	0.241 (-0.233,0.715)	-1.002 (-1.646,-0.357)***	0.029 (0.009,0.049)***	-0.015 (-0.037,0.006)	0.66 (0.521,0.8)***	**Non-stationary**	2017 - 06
RIACHÃO	Turnover	**0.138 (0.045,0.231)*** **	**0.138 (0.038,0.239)*** **	-0.001 (-0.006,0.004)	-0.001 (-0.006,0.004)	0.174 (-0.018,0.365)	**Stationary** **	2013 - 03
RIACHO DA SERRA	Temporal α diversity	0.468 (-0.336,1.273)	-2.461 (-3.606,-1.317)***	0.176 (0.083,0.27)***	-0.186 (-0.394,0.021)*	0.211 (-0.125,0.548)	**Stationary*** **	2016 - 02
ROSÁRIO	Temporal α diversity	**2.227 (1.282,3.172)*** **	1.089 (0.425,1.753)***	-0.073 (-0.125,-0.022)***	0.037 (-0.015,0.089)	0.125 (-0.168,0.418)	**Stationary*** **	2015 - 04
SÃO DOMINGOS	Turnover	**-0.266 (-0.528,-0.005)** **	**-0.256 (-0.475,-0.038)** **	**0.069 (0.026,0.112)*** **	**-0.069 (-0.112,-0.025)*** **	-0.049 (-0.392,0.293)	**Stationary** **	2011 - 06
SÃO DOMINGOS	Nestedness	0.138 (-0.032,0.308)	-0.133 (-0.297,0.03)	**0.017 (-0.001,0.035)* **	**-0.02 (-0.04,-0.001)** **	-0.236 (-0.554,0.083)	**Stationary** **	2011 - 10
SÃO DOMINGOS	Temporal β diversity	-0.043 (-0.245,0.158)	**-0.283 (-0.477,-0.089)*** **	**0.063 (0.034,0.092)*** **	**-0.067 (-0.098,-0.036)*** **	**-0.126 (-0.439,0.186)**	**Stationary** **	2011 - 08
SÃO DOMINGOS	Temporal α diversity	3.042 (-0.43,6.514)*	-0.781 (-1.825,0.264)	-0.314 (-1.63,1.002)	0.362 (-0.955,1.678)	-0.019 (-0.373,0.335)	**Stationary** **	-
SÃO JOAQUIM	Nestedness	**0.135 (-0.025,0.295)* **	**-0.261 (-0.378,-0.143)*** **	**0.029 (-0.001,0.059)* **	**-0.028 (-0.057,0.002)* **	-0.111 (-0.319,0.097)	**Stationary** **	2012 - 08
SÃO JOAQUIM	Temporal β diversity	0.16 (-0.046,0.367)	**-0.328 (-0.509,-0.146)*** **	**0.033 (0.006,0.06)** **	**-0.029 (-0.056,-0.002)** **	**0.204 (-0.004,0.412)* **	**Stationary** **	2012 - 12
SÃO JOAQUIM	Turnover	-0.107 (-0.337,0.124)	**-0.376 (-0.578,-0.175)*** **	**0.052 (0.021,0.084)*** **	**-0.048 (-0.08,-0.017)*** **	0.181 (-0.027,0.388)	**Stationary** **	2012 - 12
SÃO JOAQUIM	Temporal α diversity	**1.388 (0.766,2.009)*** **	-1.05 (-1.725,-0.376)***	0.024 (0.003,0.046)**	-0.021 (-0.047,0.006)	**0.293 (0.081,0.504)*** **	**Stationary*** **	2015 - 05
SÃO JOSÉ I	Turnover	0.143 (-0.057,0.344)	**-0.252 (-0.43,-0.075)*** **	**0.033 (0.004,0.062)** **	**-0.033 (-0.063,-0.003)** **	-0.123 (-0.419,0.173)	**Stationary** **	2013 - 03
SÃO JOSÉ I	Temporal α diversity	1.153 (-0.952,3.258)	-1.661 (-2.851,-0.47)***	0.372 (-0.209,0.953)	-0.34 (-0.92,0.239)	-0.004 (-0.303,0.296)	**Stationary** **	2012 - 02

SÃO JOSÉ I	Temporal β diversity	**0.618 (0.439,0.796)*****	**0.162 (0.042,0.282)*****	**-0.006 (-0.012,0)***	-0.005 (-0.014,0.004)	-0.239 (-0.554,0.077)	**Stationary****	2014 - 04
SÃO JOSÉ I	Nestedness	**0.186 (0.098,0.273)*****	**0.144 (-0.021,0.308)***	0.001 (-0.002,0.004)	**-0.047 (-0.078,-0.016)*****	-0.117 (-0.413,0.179)	**Stationary****	2015 - 05
SÃO PEDRO TIMBAÚBA	Temporal α diversity	1.73 (0.999,2.461)***	1.223 (0.368,2.079)***	0 (-0.026,0.026)	0.004 (-0.031,0.04)	0.024 (-0.203,0.252)	**Stationary****	2014 - 06
SITIOS - 11OS	Temporal α diversity	**1.478 (0.965,1.991)*****	-1.056 (-1.607,-0.505)***	0.009 (0,0.018)**	0.01 (-0.008,0.028)	**0.339 (0.152,0.525)*****	**Stationary*****	2018 - 01
SITIOS - 11OS	Temporal β diversity	**0.168 (0.098,0.238)*****	**-0.272 (-0.39,-0.153)*****	**0.004 (0.003,0.006)*****	0.005 (-0.001,0.011)	**0.357 (0.181,0.532)*****	**Stationary****	01/03/2017
SITIOS - 11OS	Turnover	**0.09 (0.013,0.168)****	**-0.258 (-0.419,-0.096)*****	**0.004 (0.002,0.006)*****	0 (-0.009,0.009)	**0.225 (0.028,0.423)****	**Stationary****	2017 - 03
SITIOS - 11OS	Nestedness	**0.141 (0.09,0.192)*****	**-0.289 (-0.483,-0.095)*****	0 (0,0.001)	**0.105 (0.056,0.154)*****	0.051 (-0.151,0.252)	**Stationary****	2018 - 10
TRUSSU	Temporal β diversity	0.166 (-0.034,0.366)	**-0.238 (-0.408,-0.068)*****	**0.036 (0,0.072)****	**-0.032 (-0.068,0.003)***	**0.151 (-0.15,0.453)**	**Stationary****	2014 - 10
TRUSSU	Turnover	-0.004 (-0.21,0.202)	**-0.378 (-0.562,-0.195)*****	**0.053 (0.017,0.09)*****	**-0.046 (-0.083,-0.01)****	-0.146 (-0.442,0.15)	**Stationary****	2014 - 10
TRUSSU	Temporal α diversity	**3.833 (2.098,5.569)*****	0.801 (0.104,1.498)**	-0.31 (-0.491,-0.13)***	0.296 (0.119,0.473)***	0.132 (-0.199,0.463)	**Stationary*****	2015 - 08
TRUSSU	Nestedness	**0.211 (0.124,0.299)*****	0.079 (-0.021,0.179)	0 (-0.005,0.005)	**-0.008 (-0.016,0)****	-0.218 (-0.515,0.08)	**Stationary****	2016 - 01
UBALDINHO	Temporal α diversity	1.853 (1.121,2.586)***	1.286 (0.224,2.347)**	-0.008 (-0.02,0.005)	-0.15 (-0.29,-0.009)**	0.258 (0.01,0.505)**	**Stationary****	2015 - 03
UBALDINHO (RIACHO SÃO MIGUEL)	Temporal α diversity	1.274 (-0.755,3.302)	-1.764 (-3.238,-0.29)**	0.401 (0.052,0.75)**	-0.43 (-0.787,-0.073)**	-0.02 (-0.389,0.35)	**Stationary*****	2013 - 05
VALÉRIO	Temporal α diversity	1.078 (0.435,1.722)***	0.711 (0.135,1.288)**	0.016 (-0.041,0.073)	-0.025 (-0.083,0.032)	0.166 (-0.053,0.384)	**Stationary****	2013 - 05

Section 6 TEMPORAL ALPHA DIVERSITY

- In (a) the time series of alpha diversity and %Volume is presented. The values were rescaled to improve the visualization of the series.
- In (b) the time series of alpha diversity (dotted black line), the trend component obtained by decomposition using LOESS (red line), the breakpoint (vertical black line) and the ACF and PACF plots of the raw data are shown.
- In (c) the residuals obtained by LOESS decomposition are presented and in (d) the residuals after the ZA test adjustment, both with their respective ACF and PACF plots.

Figure 29 - Temporal α diversity | Acarape do meio

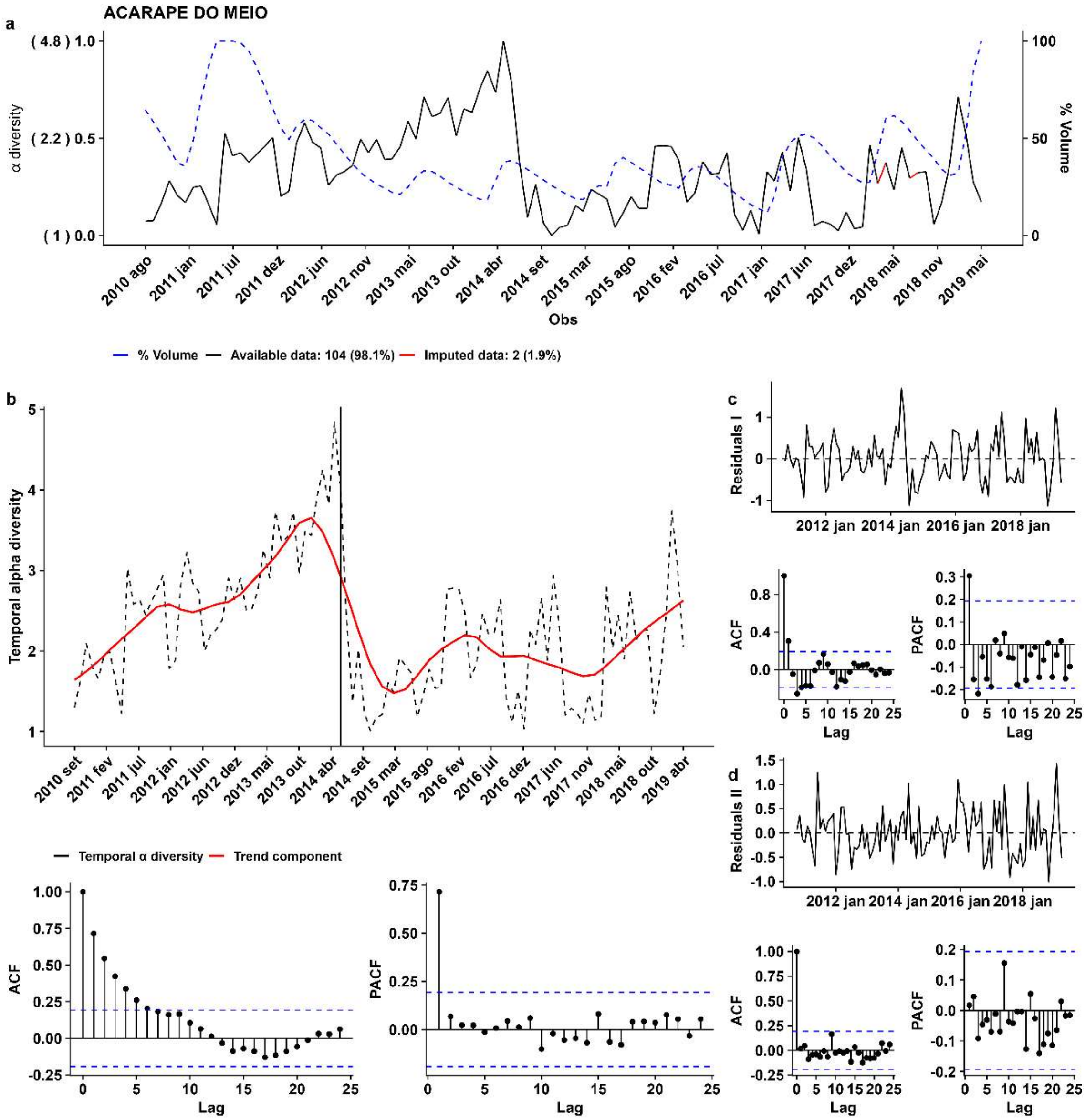


Figure 30 – Temporal α diversity | ACARAÚ MIRIM Residuals

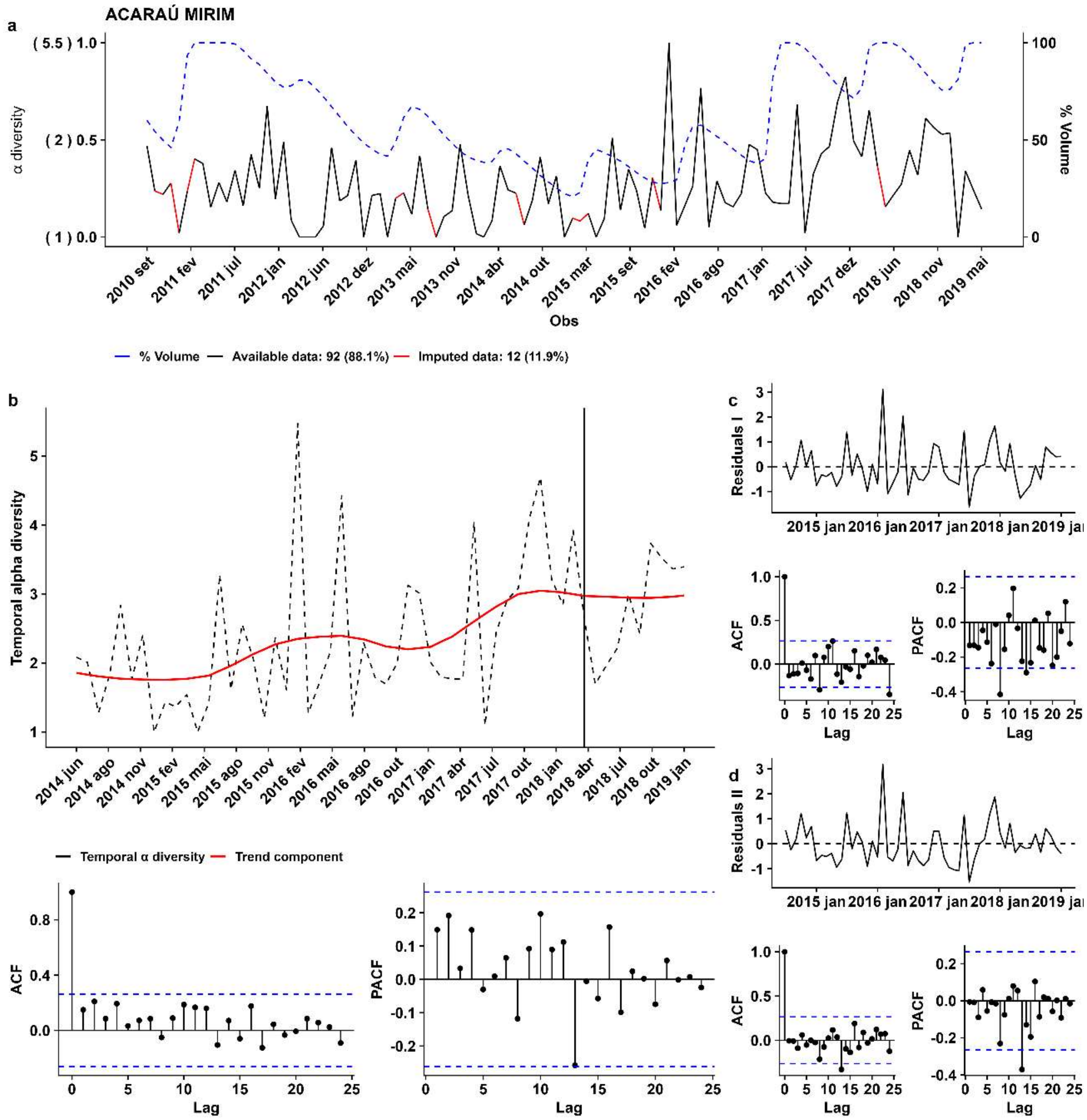


Figure 31 – Temporal α diversity | ADAUTO BEZERRA Residuals

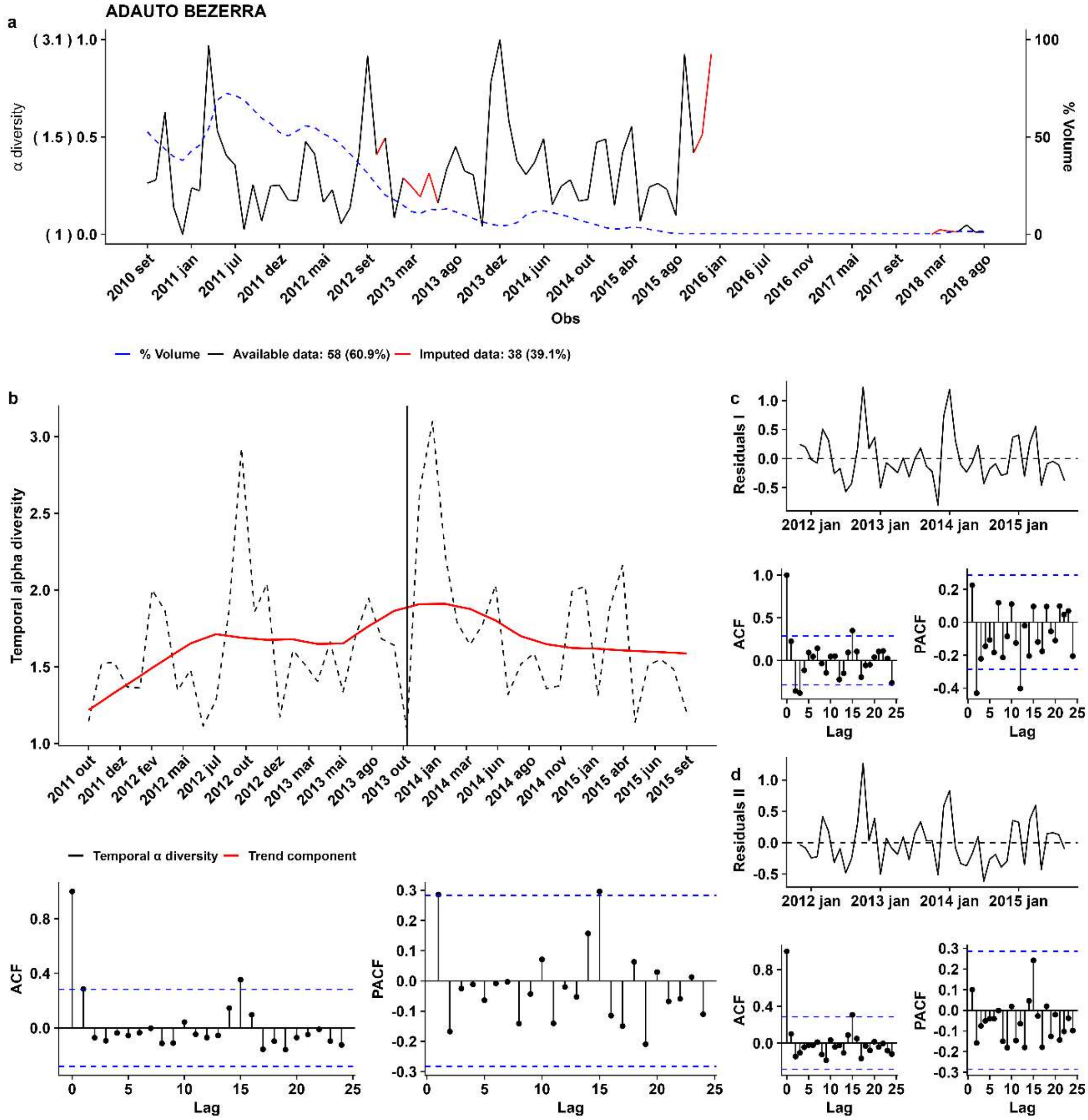


Figure 32 – Temporal α diversity | AMANARY Residuals

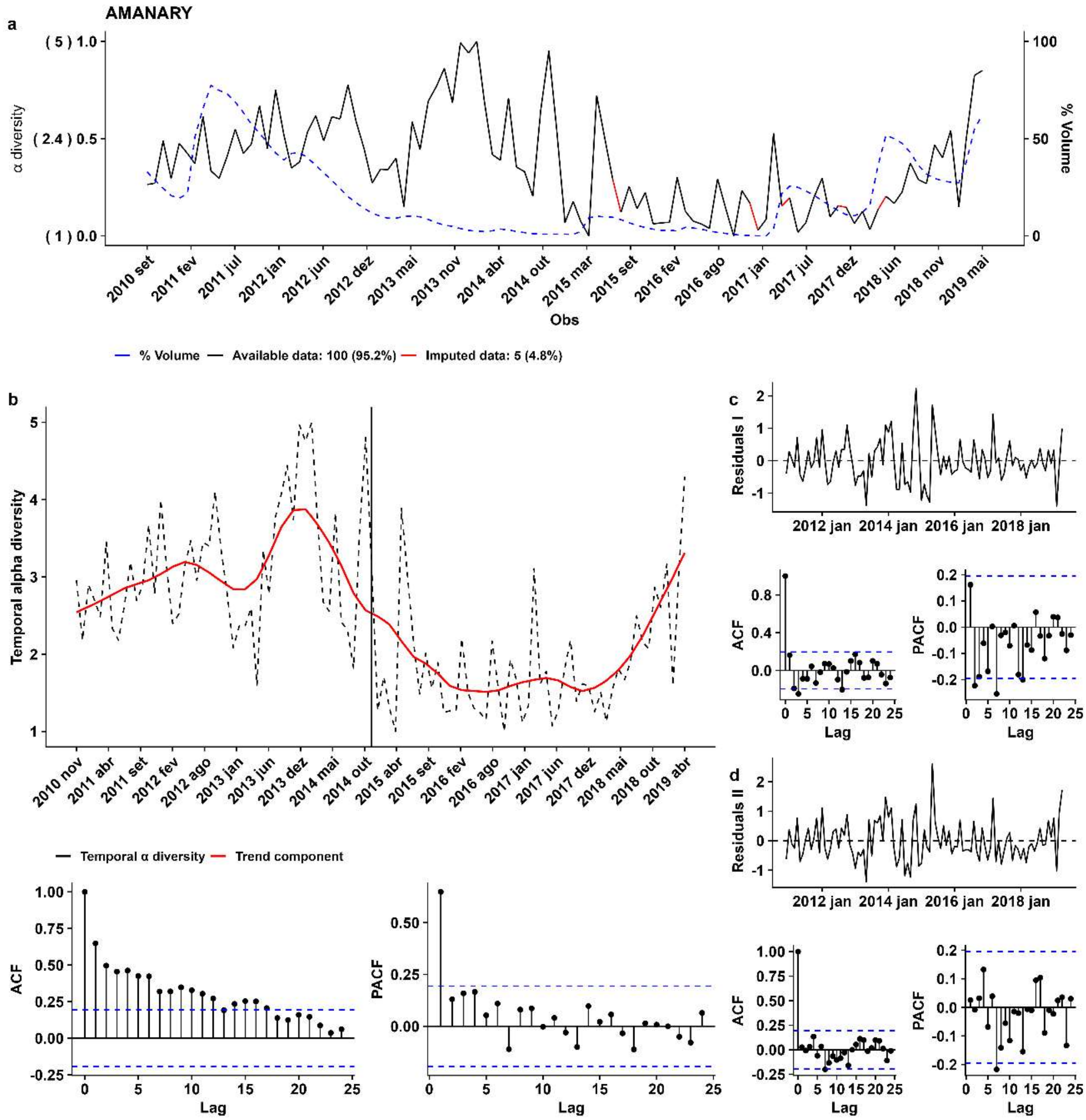


Figure 33 – Temporal α diversity | ANGICOS Residuals

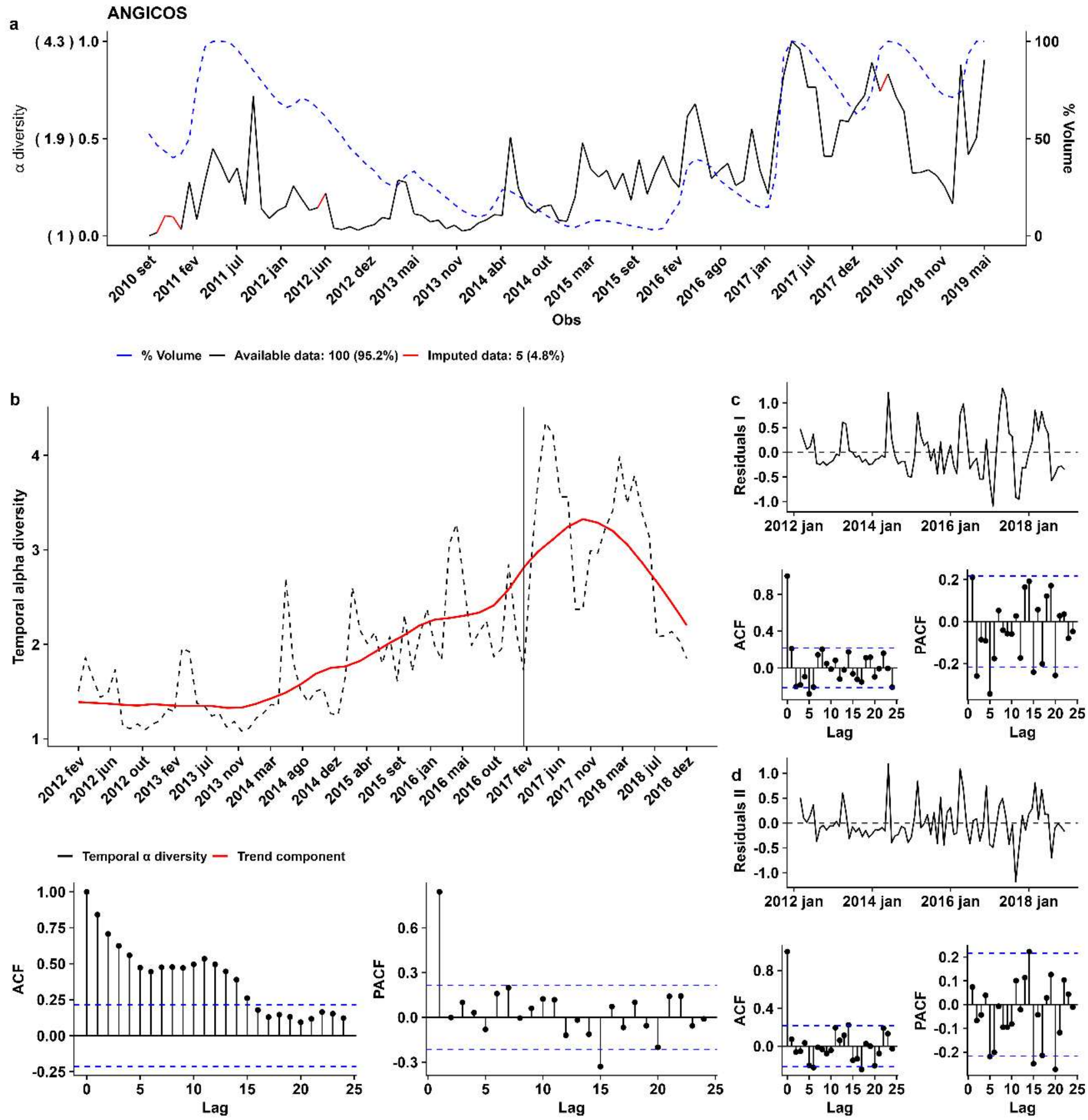


Figure 34 – Temporal α diversity | ARACOIABA Residuals

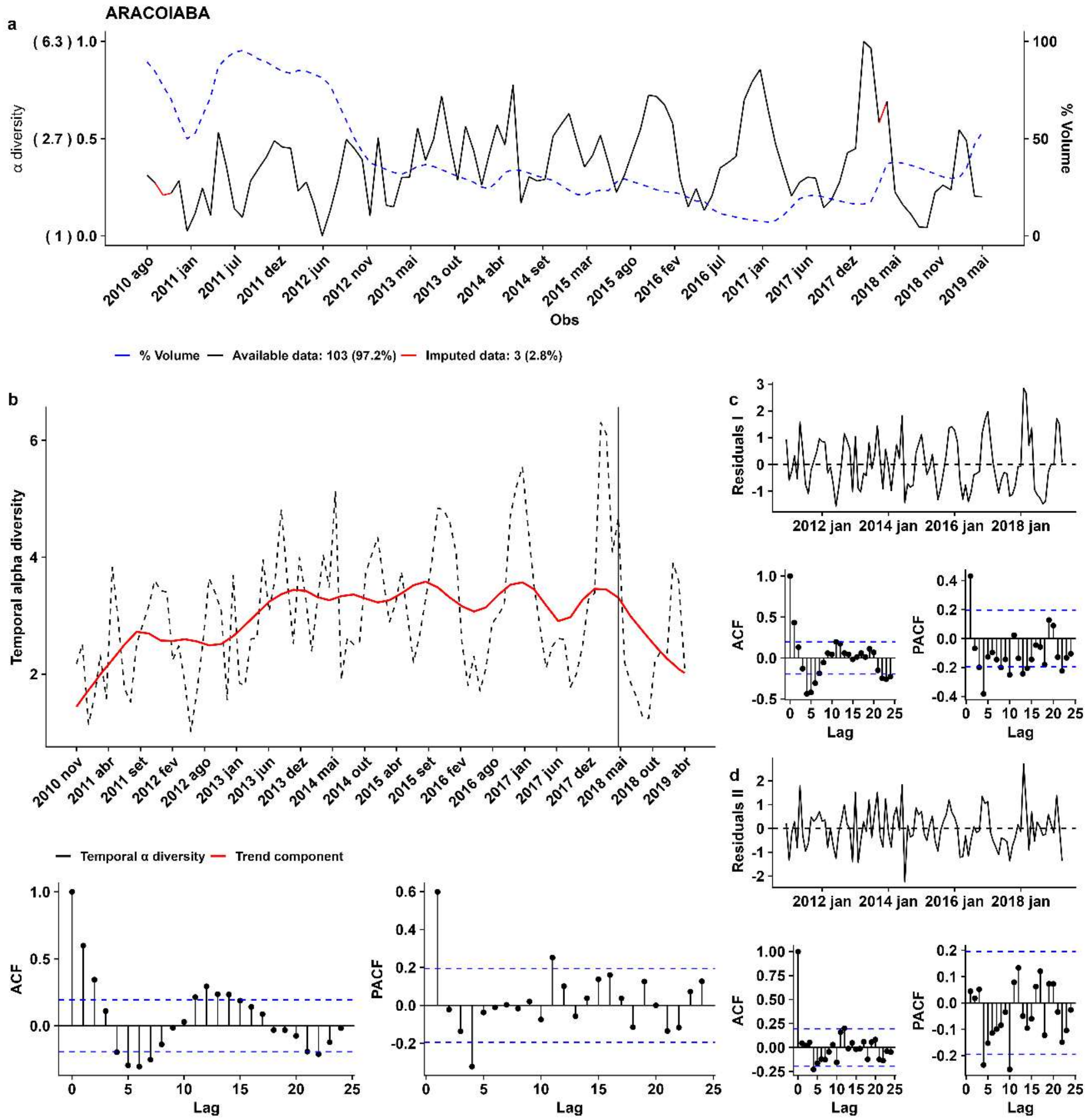


Figure 35 – Temporal α diversity | ARARAS Residuals

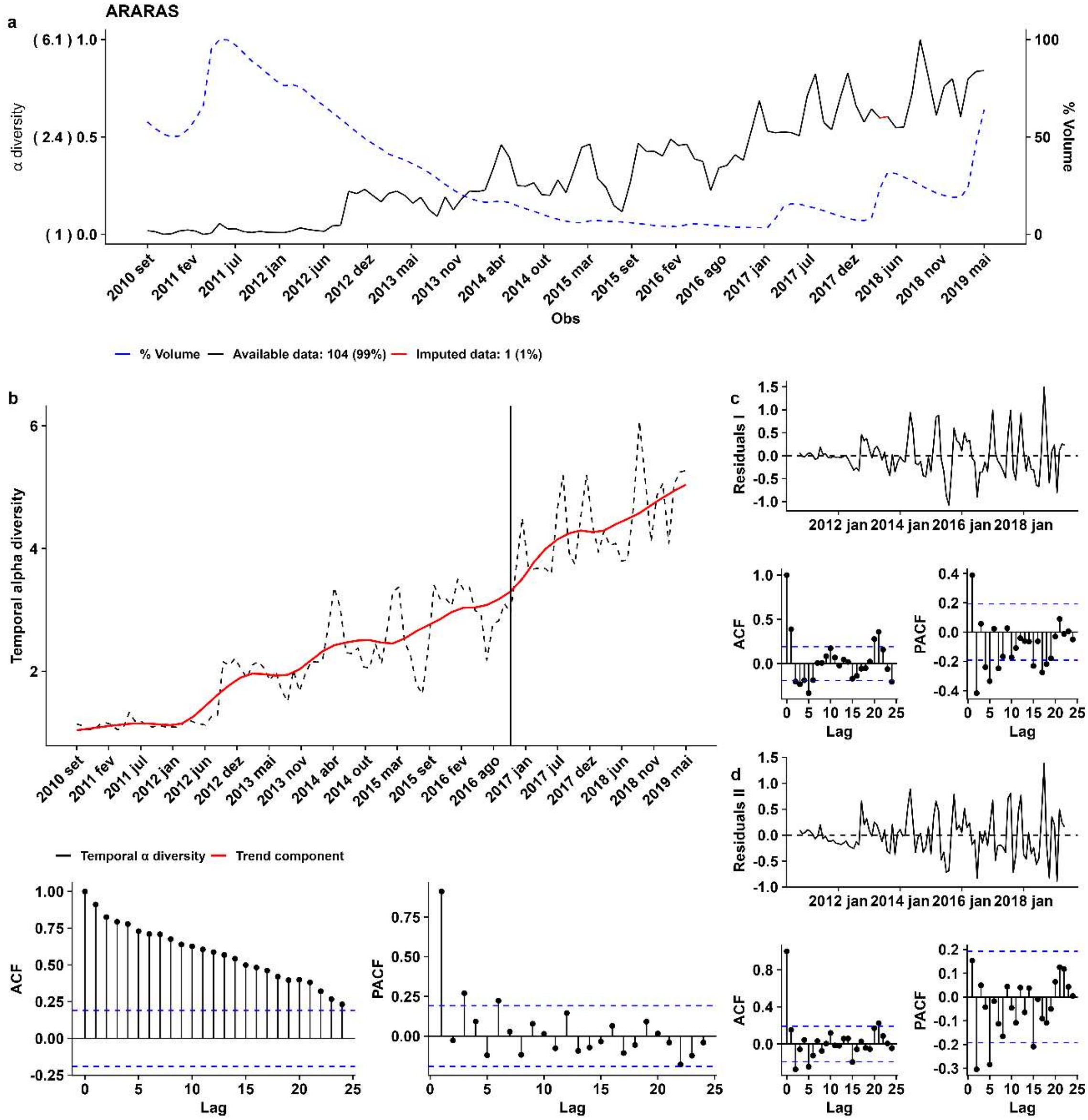


Figure 36 – Temporal α diversity | ARNEIROZ II Residuals

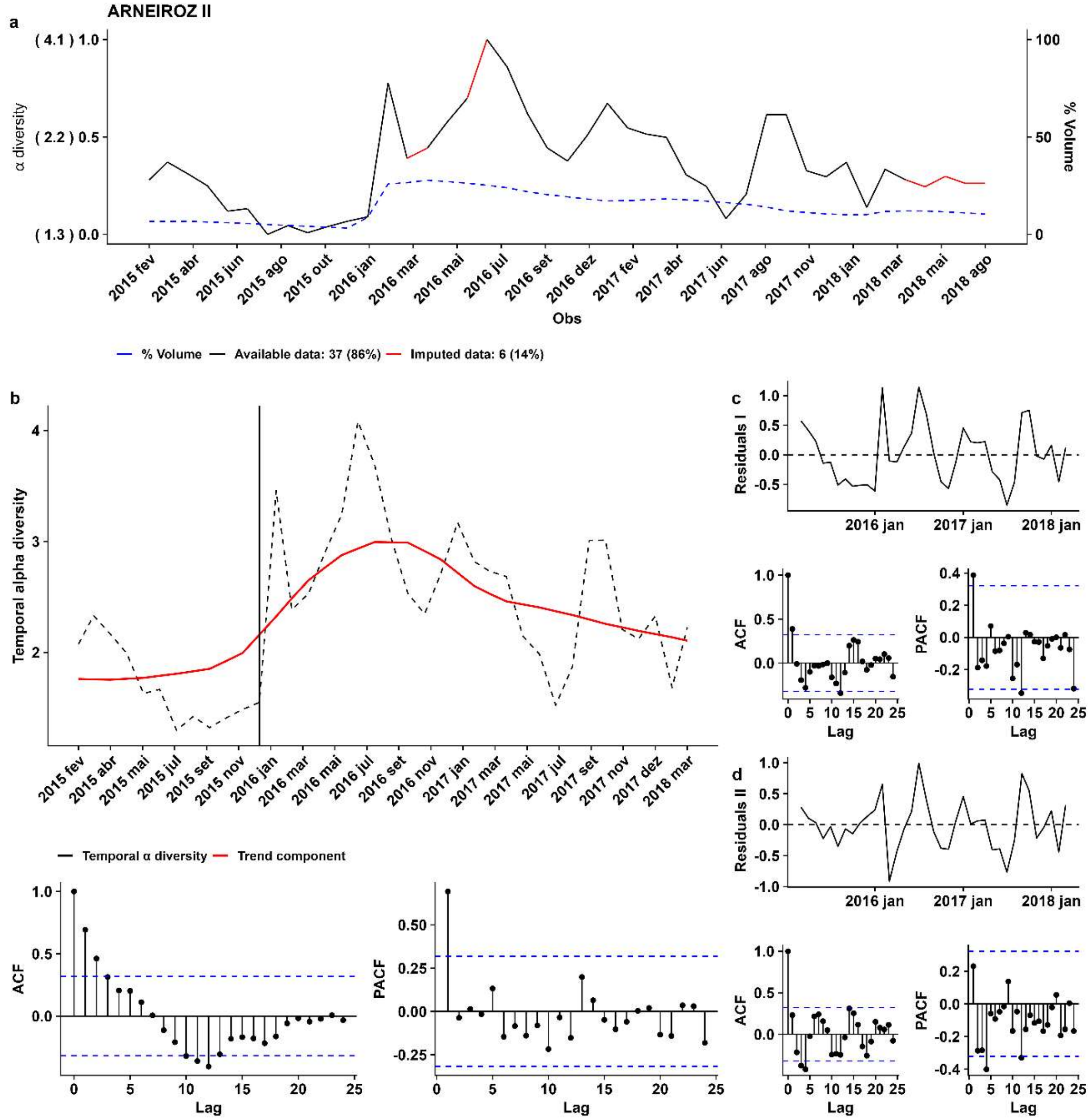


Figure 37 – Temporal α diversity | AYRES DE SOUZA Residuals

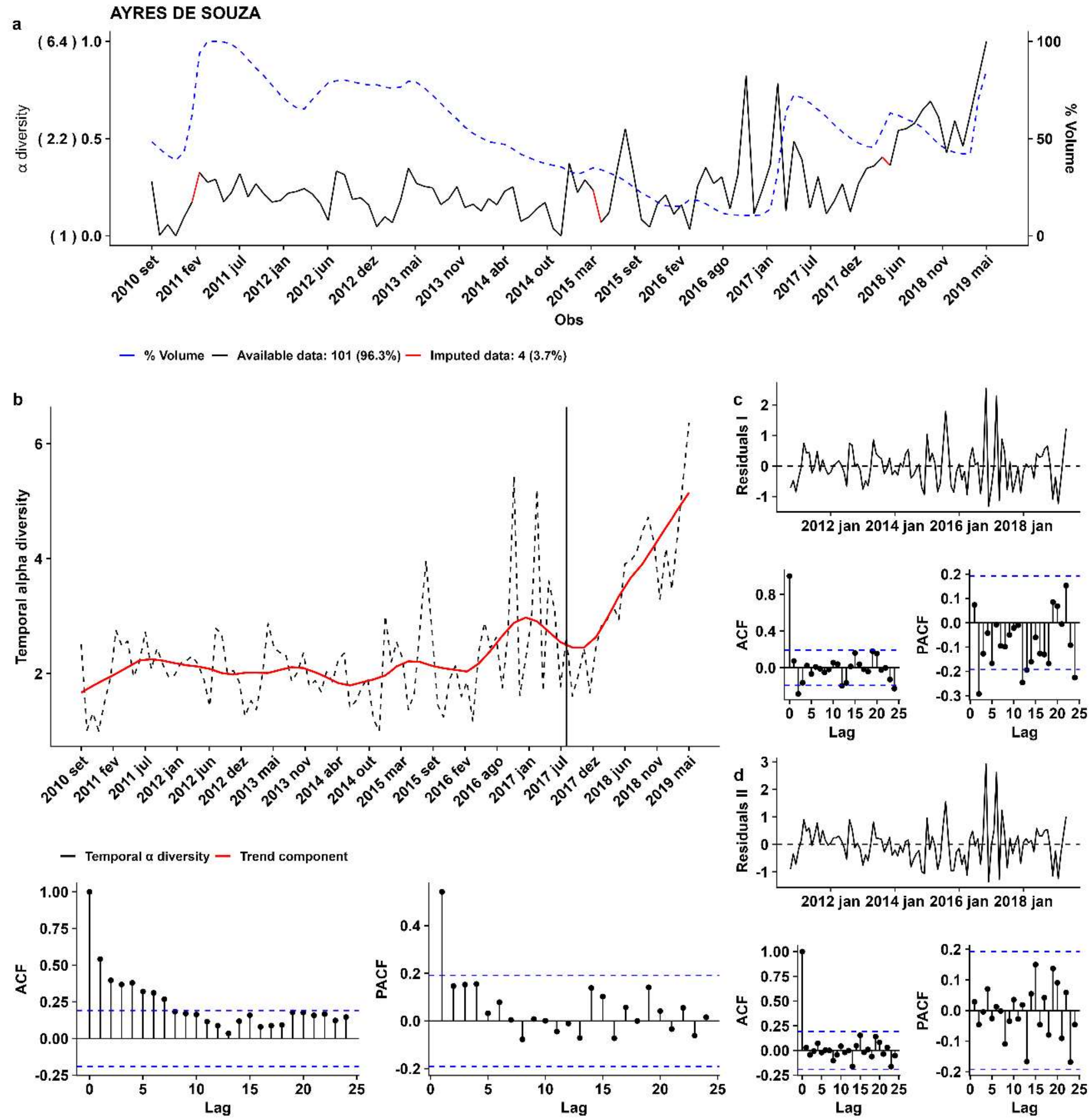


Figure 38 – Temporal α diversity | CACHOEIRA Residuals

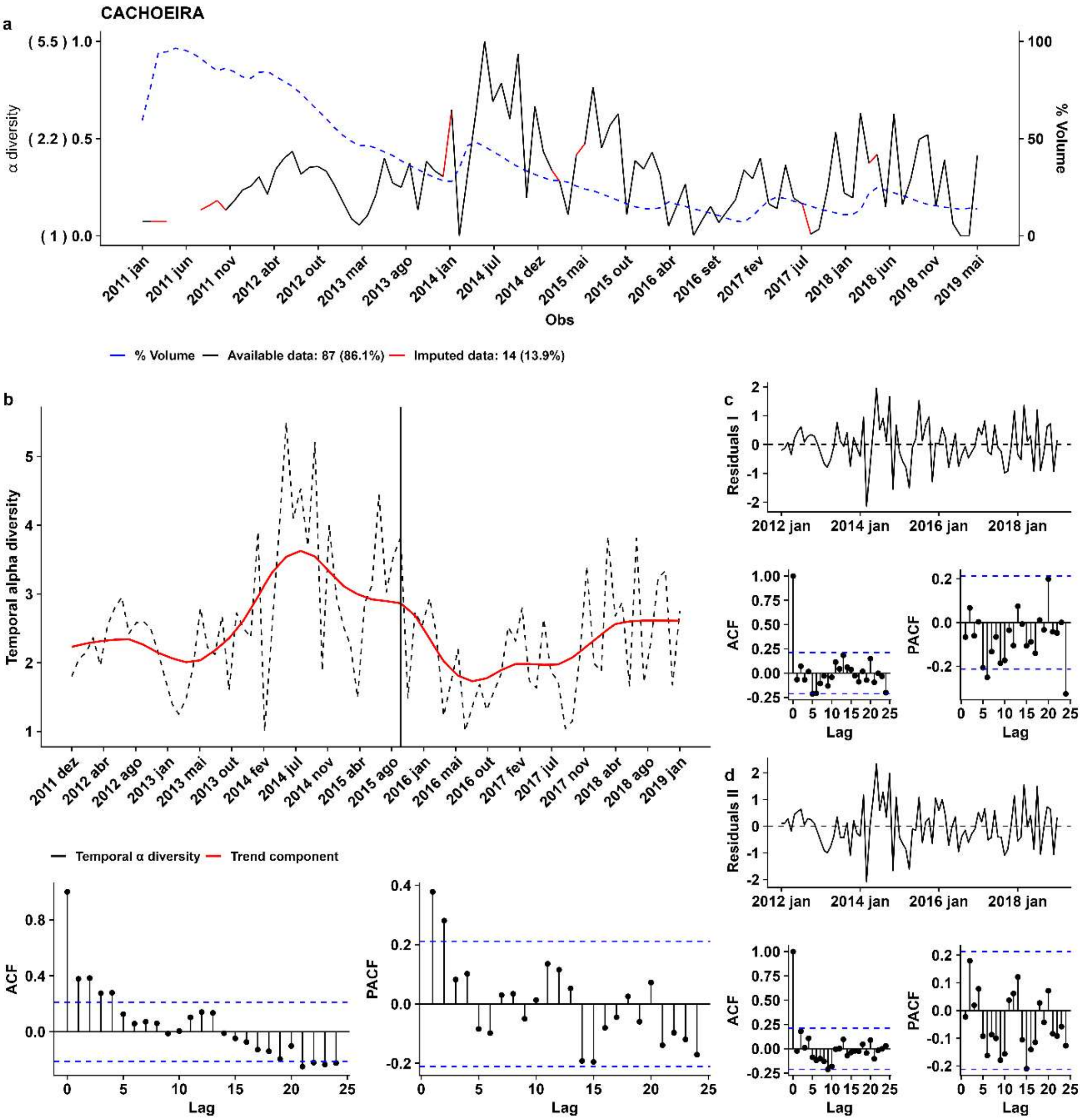


Figure 39 – Temporal α diversity | CANOAS Residuals

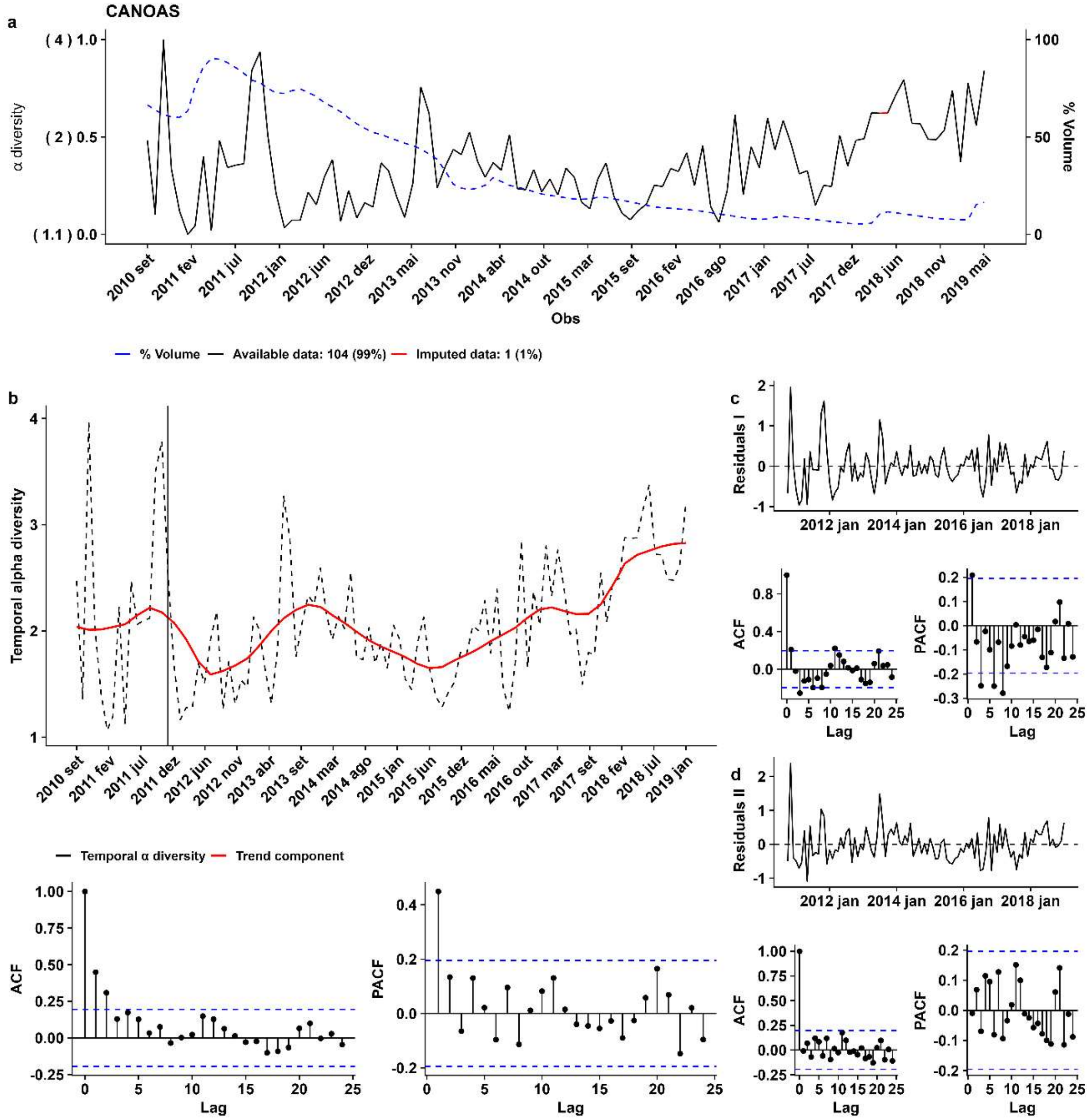


Figure 40 – Temporal α diversity | CARÃO Residuals

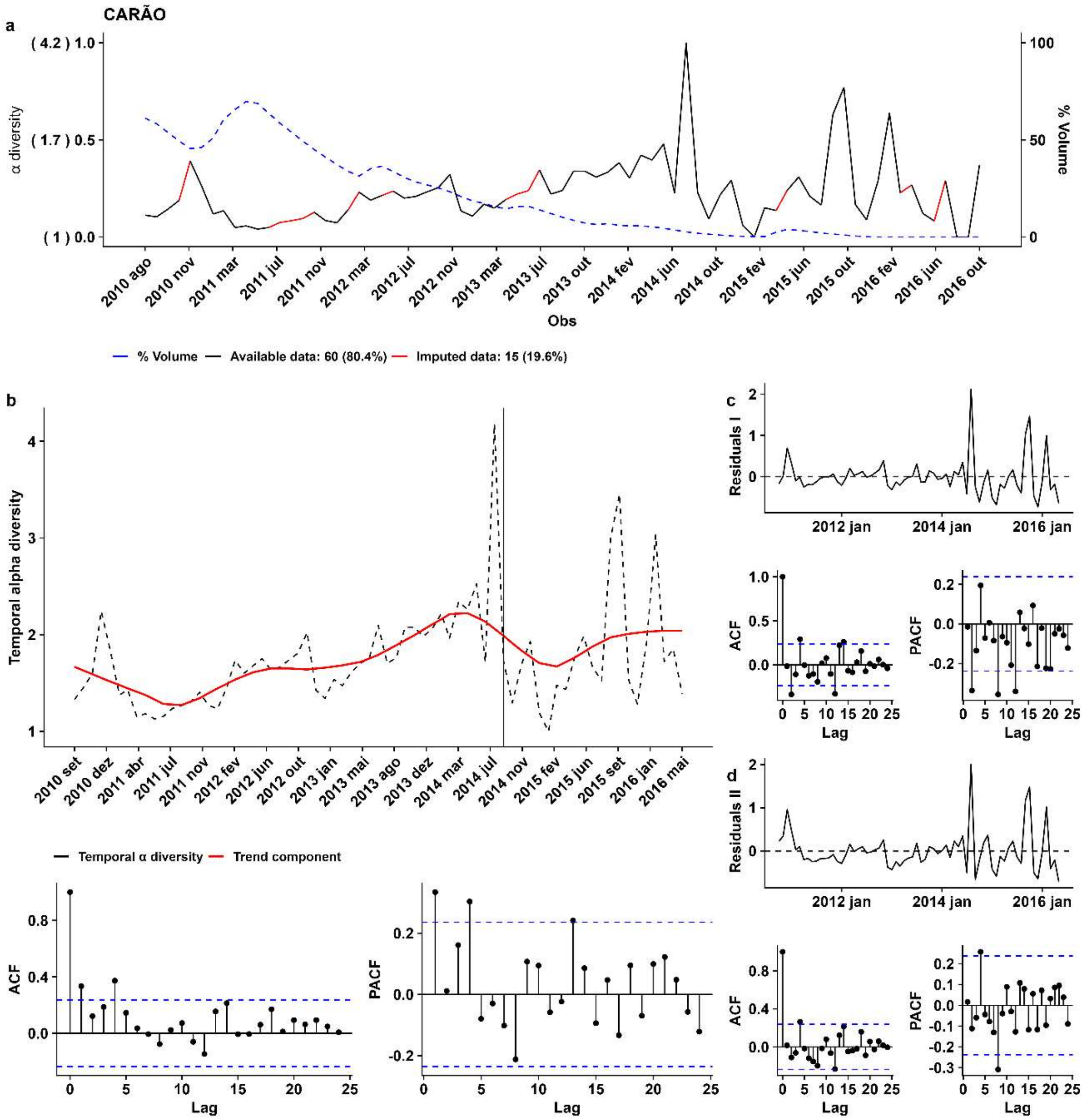


Figure 41 – Temporal α diversity | CASTRO Residuals

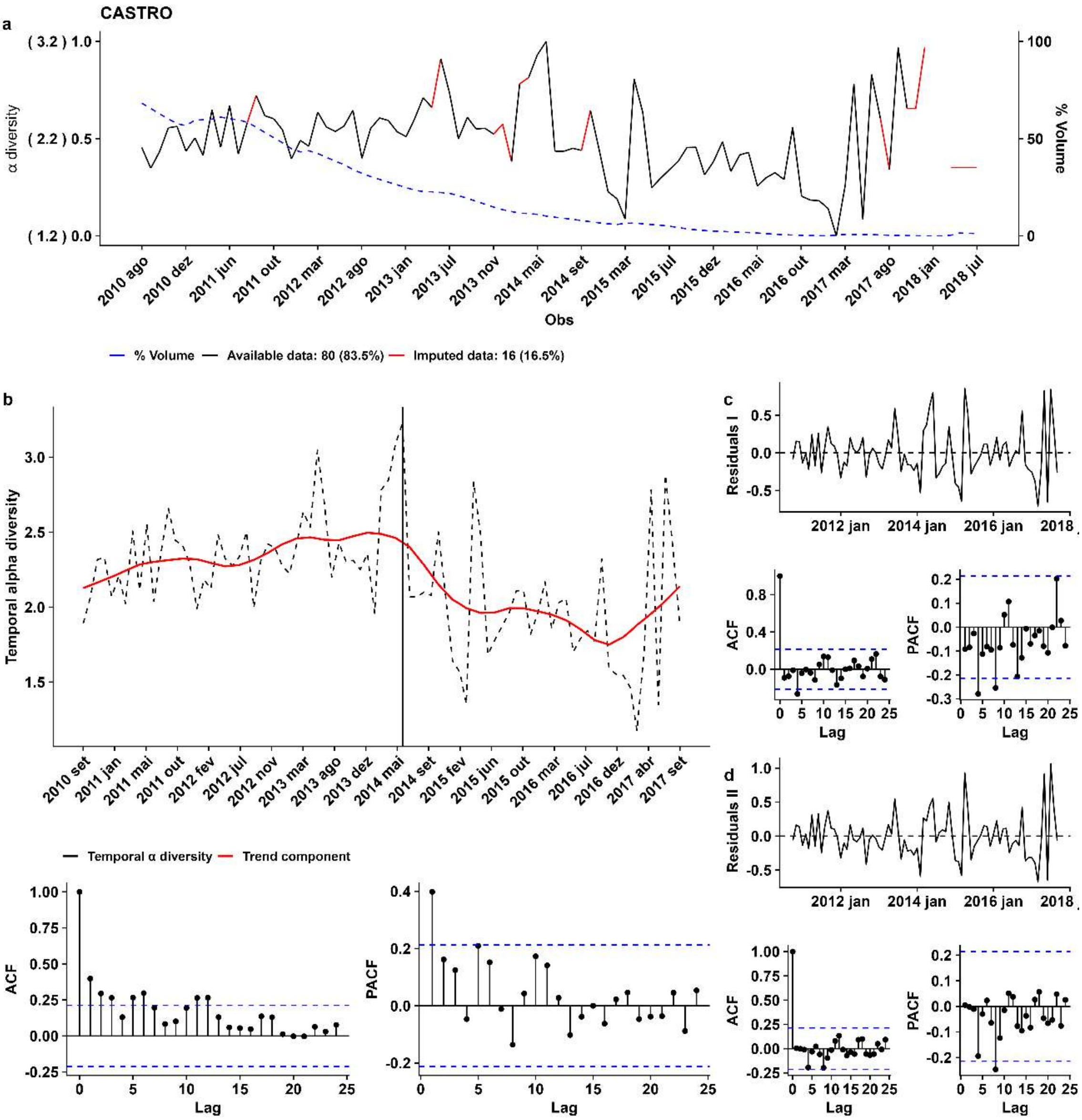


Figure 42 – Temporal α diversity | CAXITORÉ Residuals

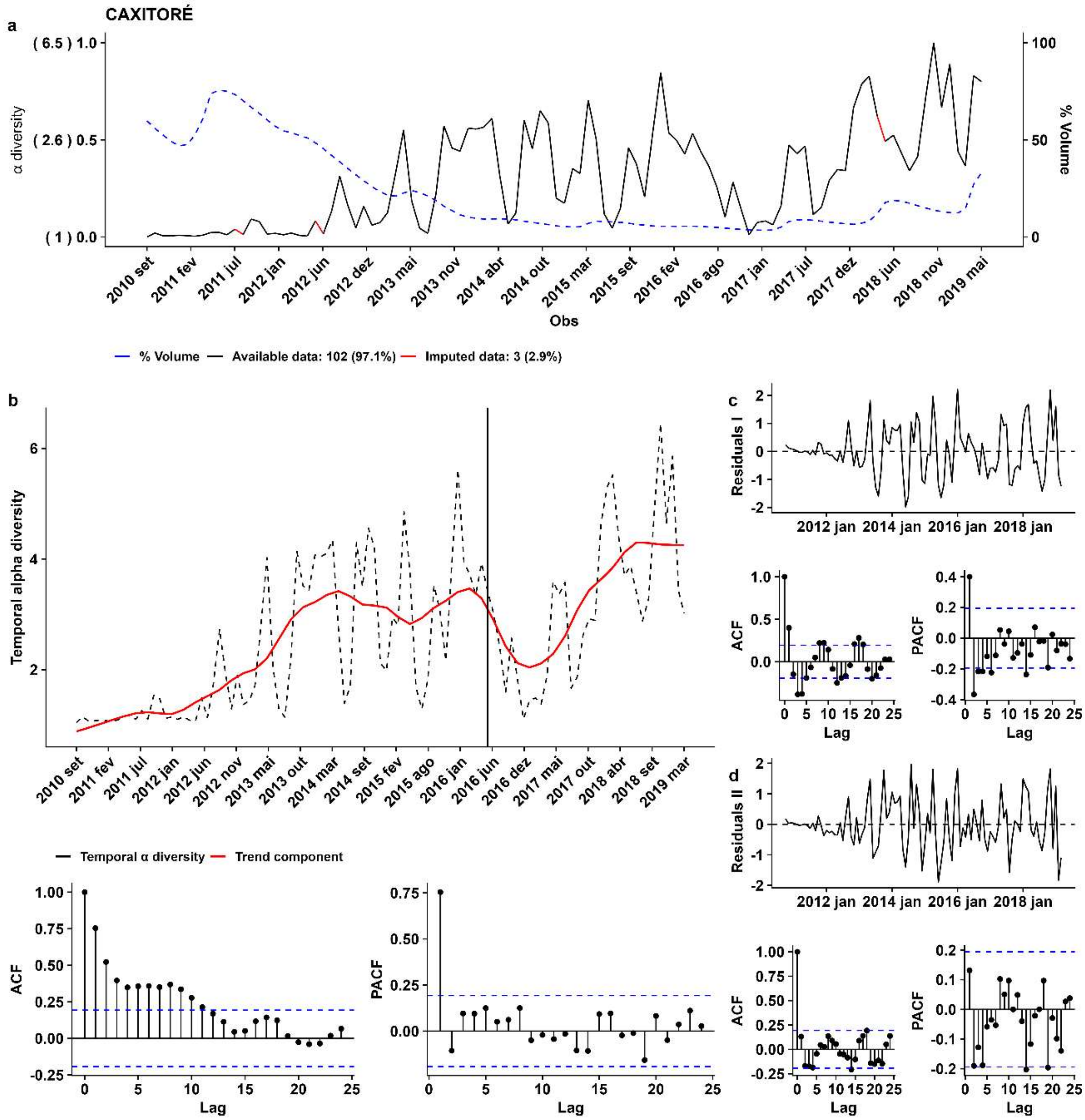


Figure 43 – Temporal α diversity | DO CORONEL Residuals

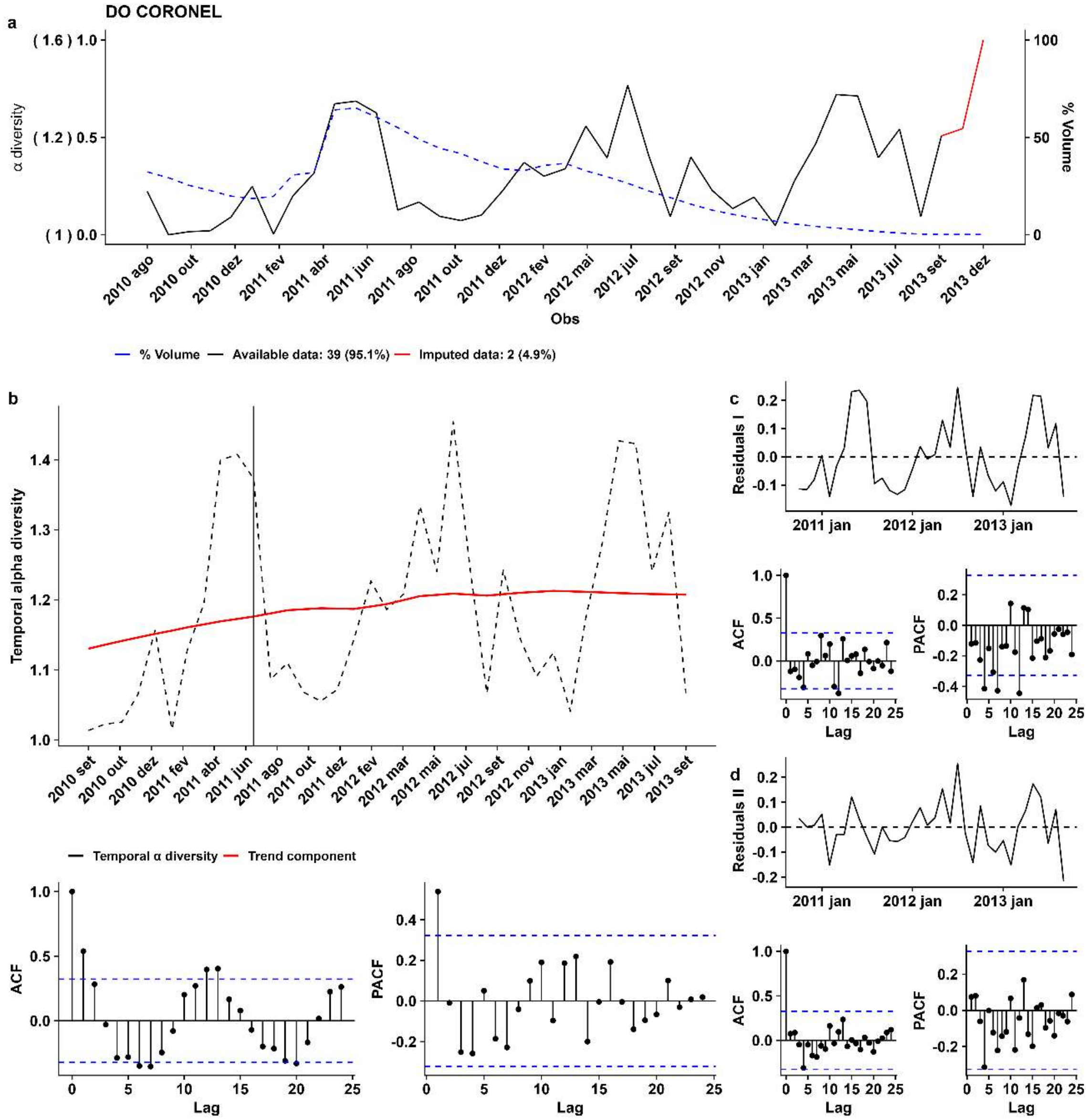


Figure 44 – Temporal α diversity | EDSON QUEIROZ Residuals

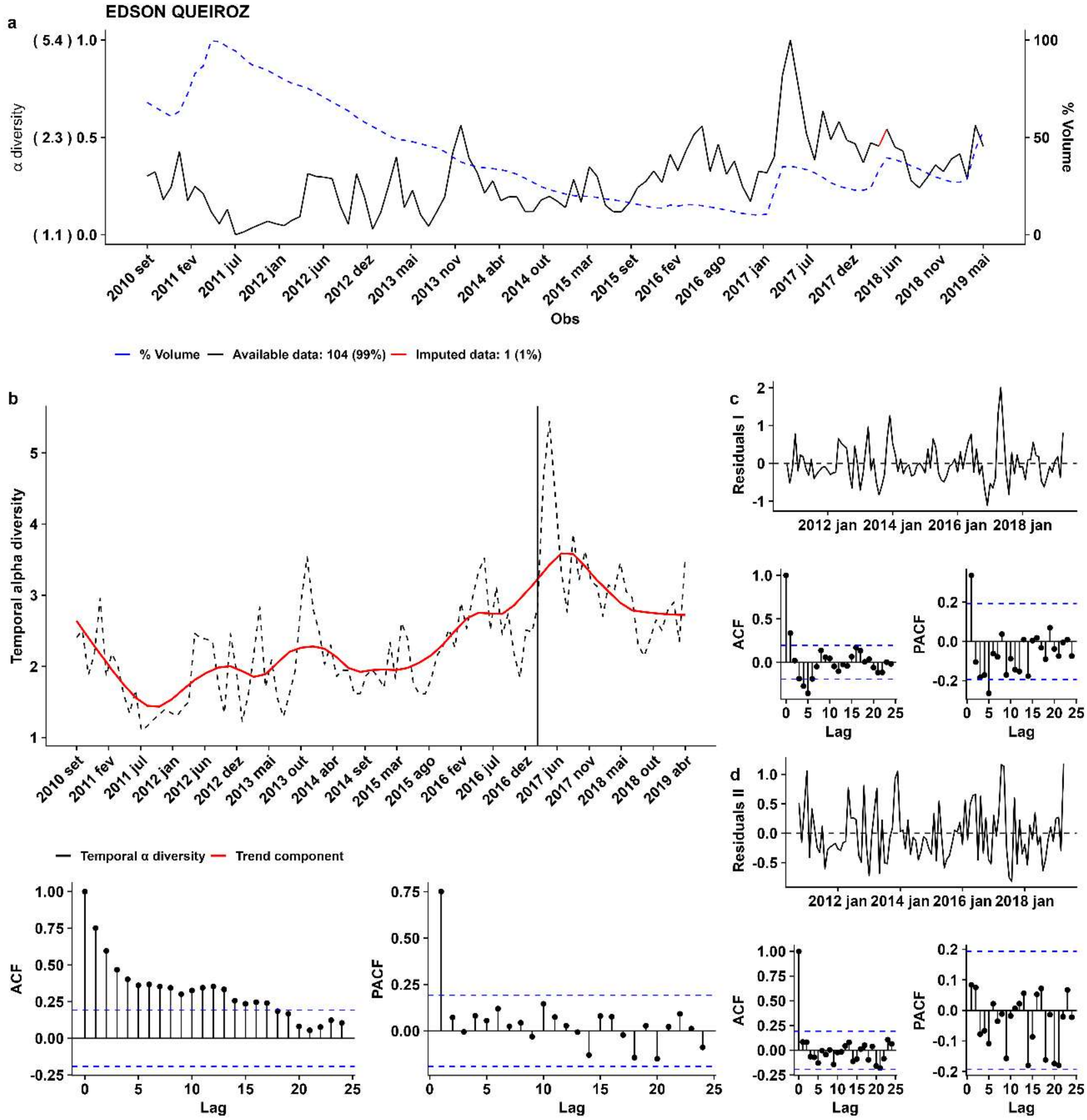


Figure 45 – Temporal α diversity | EMA Residuals

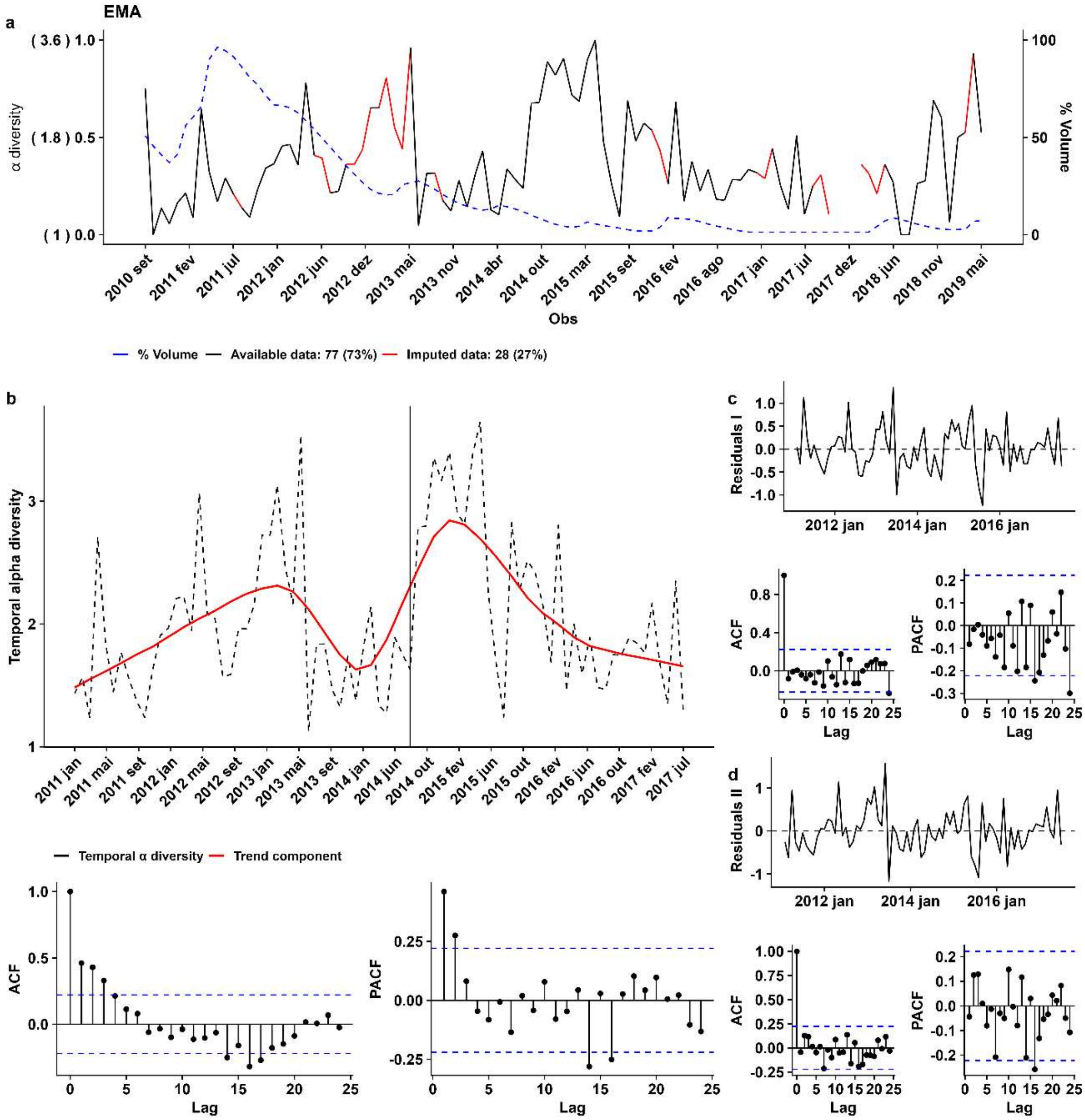


Figure 46 – Temporal α diversity | FIGUEIREDO Residuals

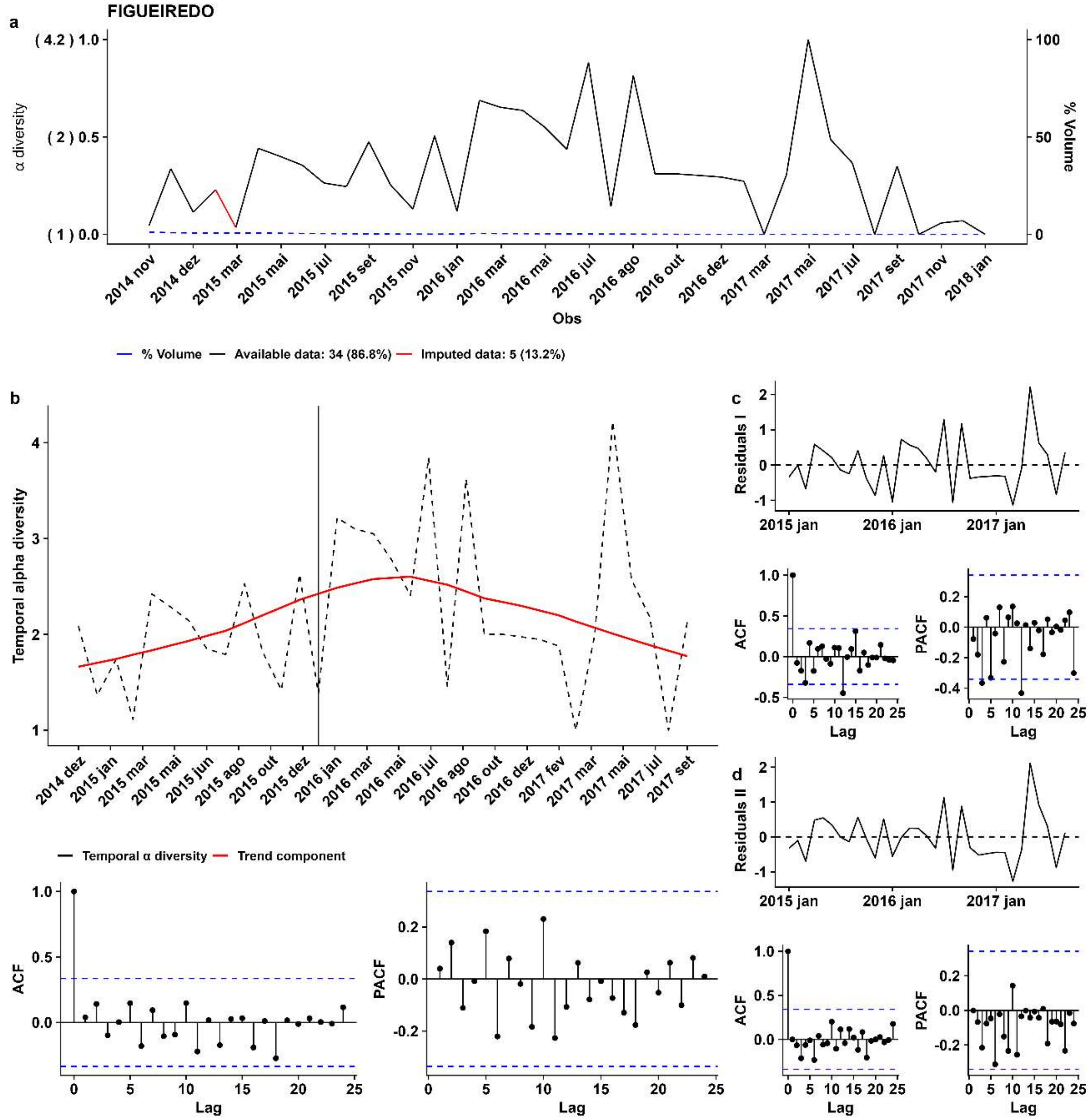


Figure 47 – Temporal α diversity | FLOR DO CAMPO Residuals

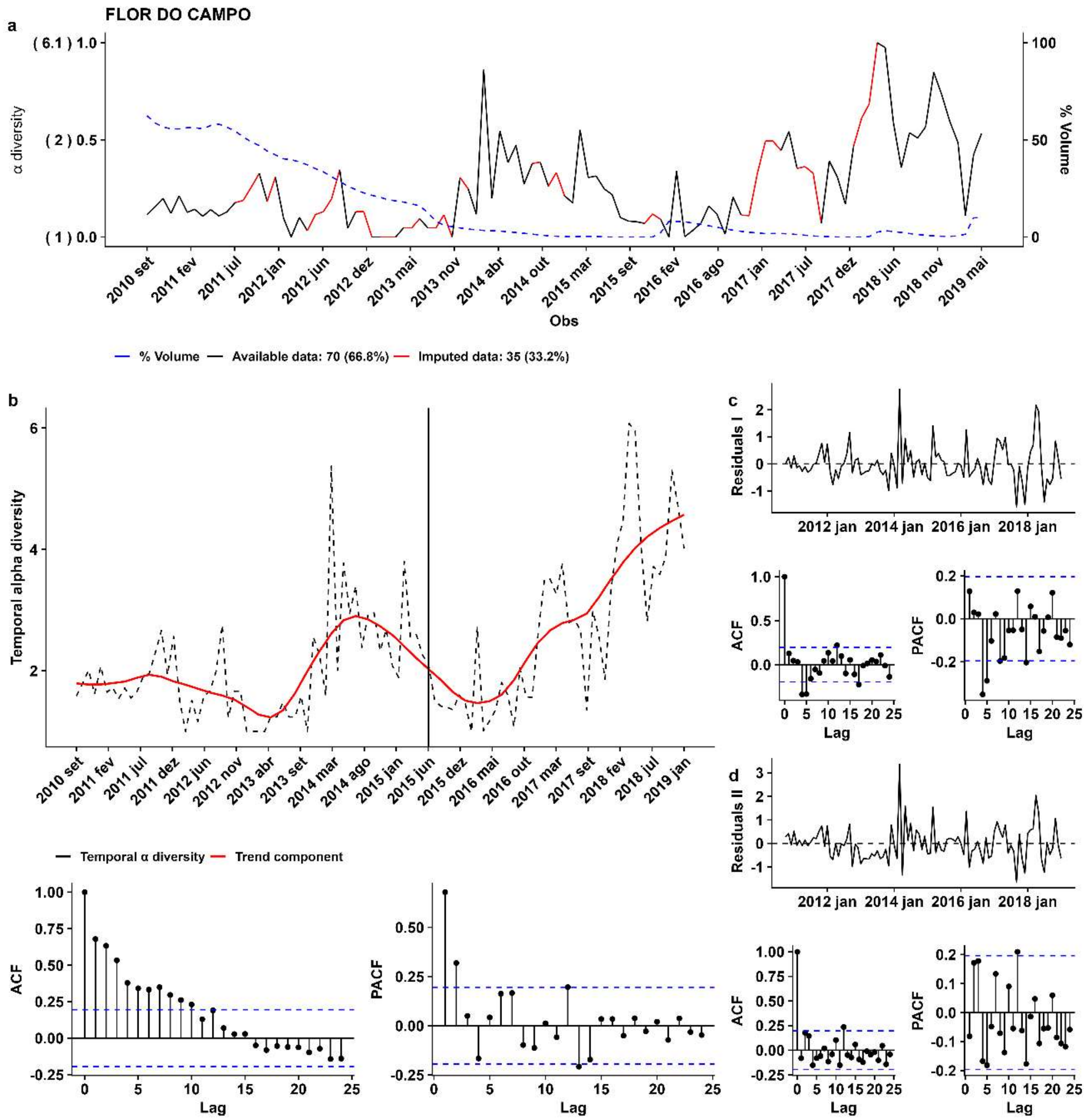


Figure 48 – Temporal α diversity | FORQUILHA Residuals

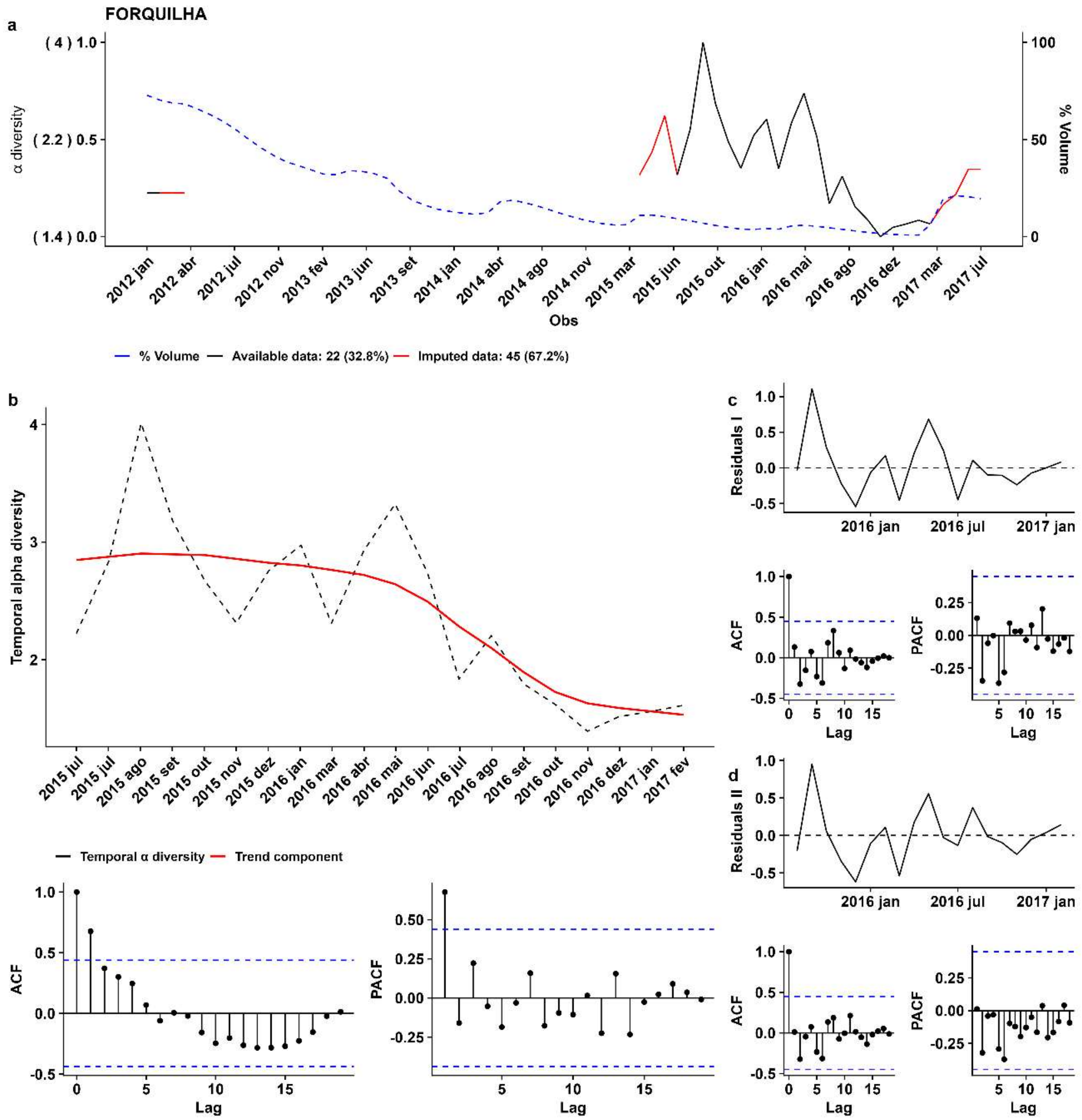


Figure 49 – Temporal α diversity | GAMELEIRA Residuals

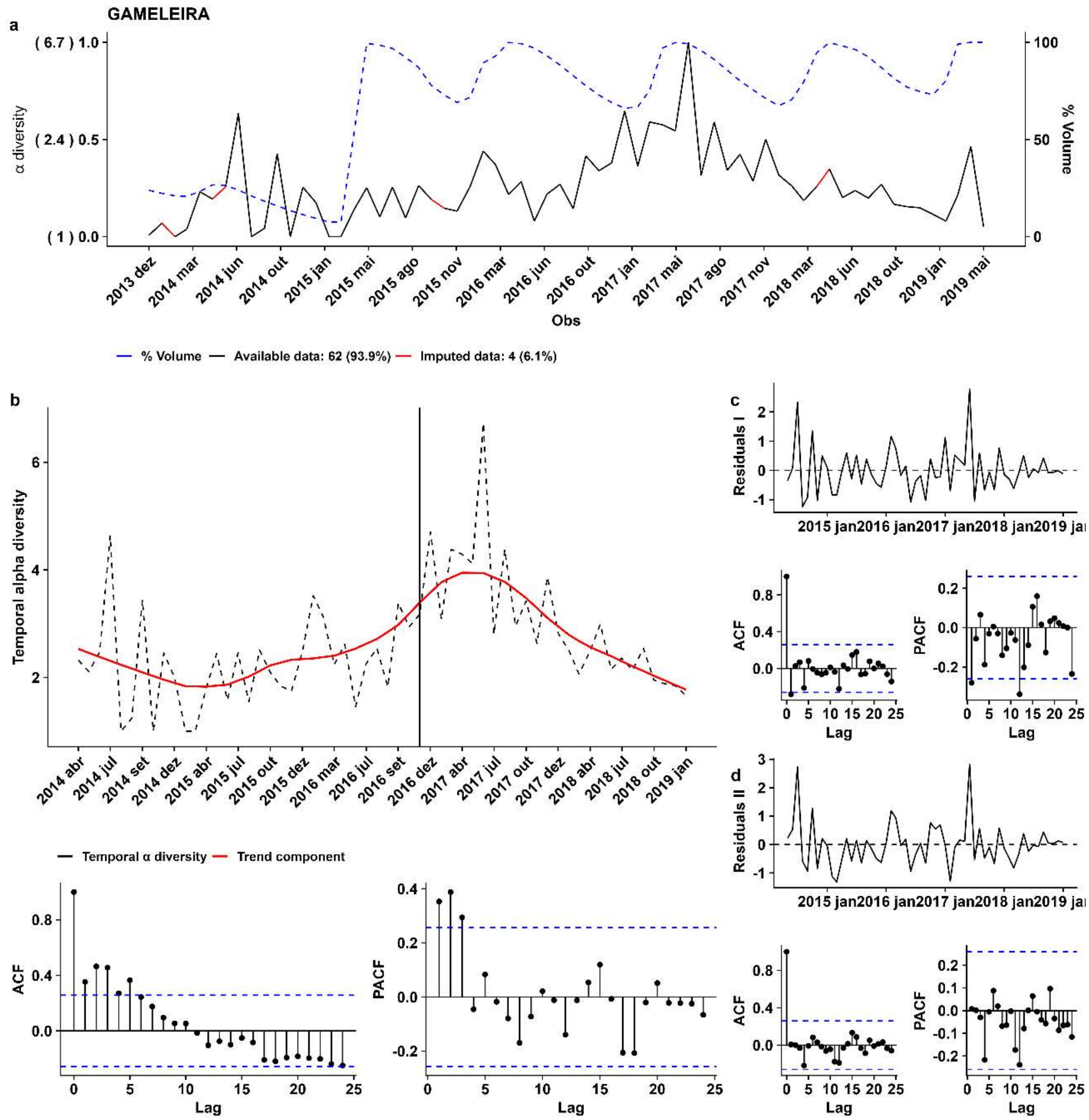


Figure 50 – Temporal α diversity | GAVIÃO Residuals

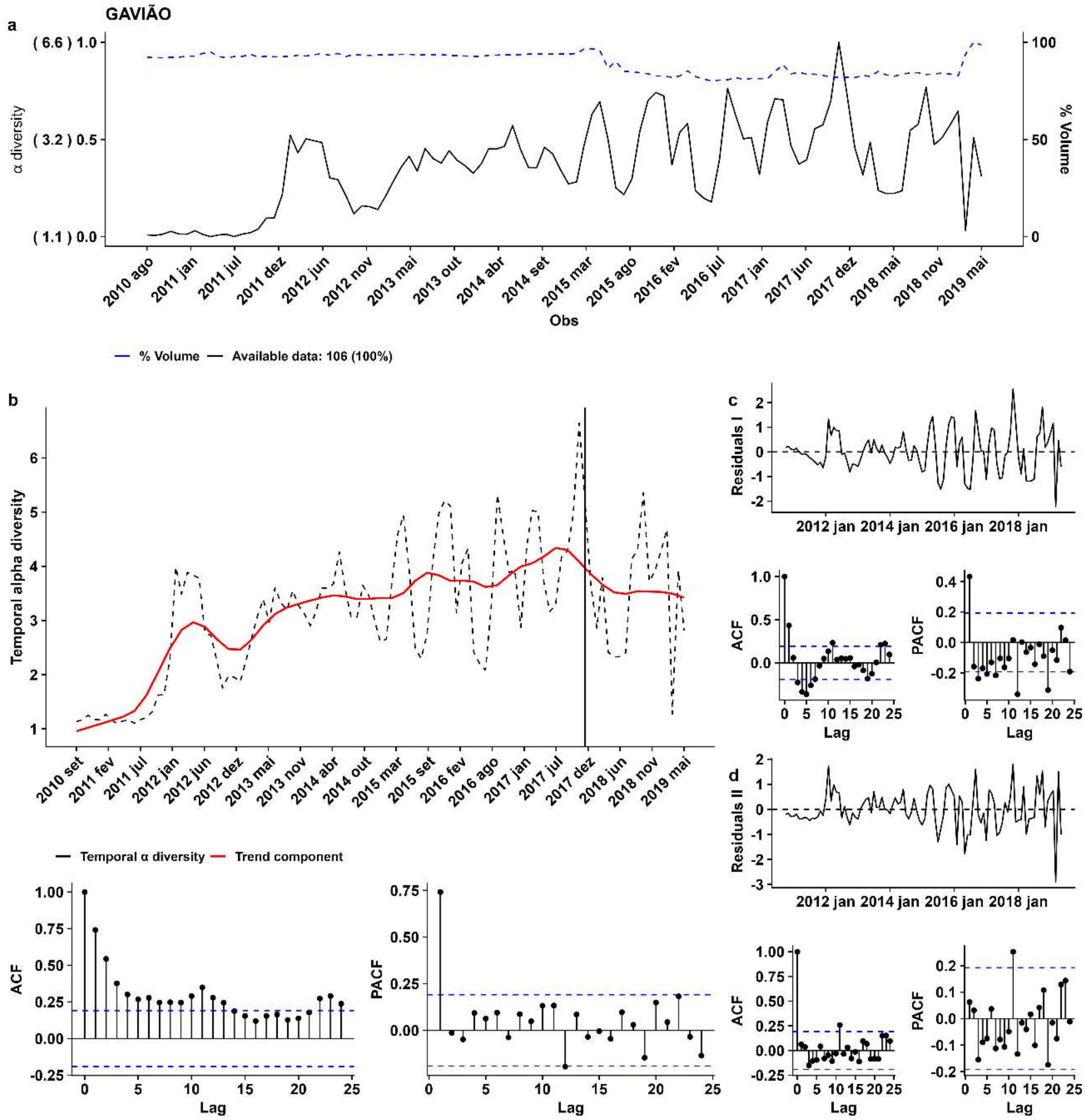


Figure 51 – Temporal α diversity | GENERAL SAMPAIO Residuals

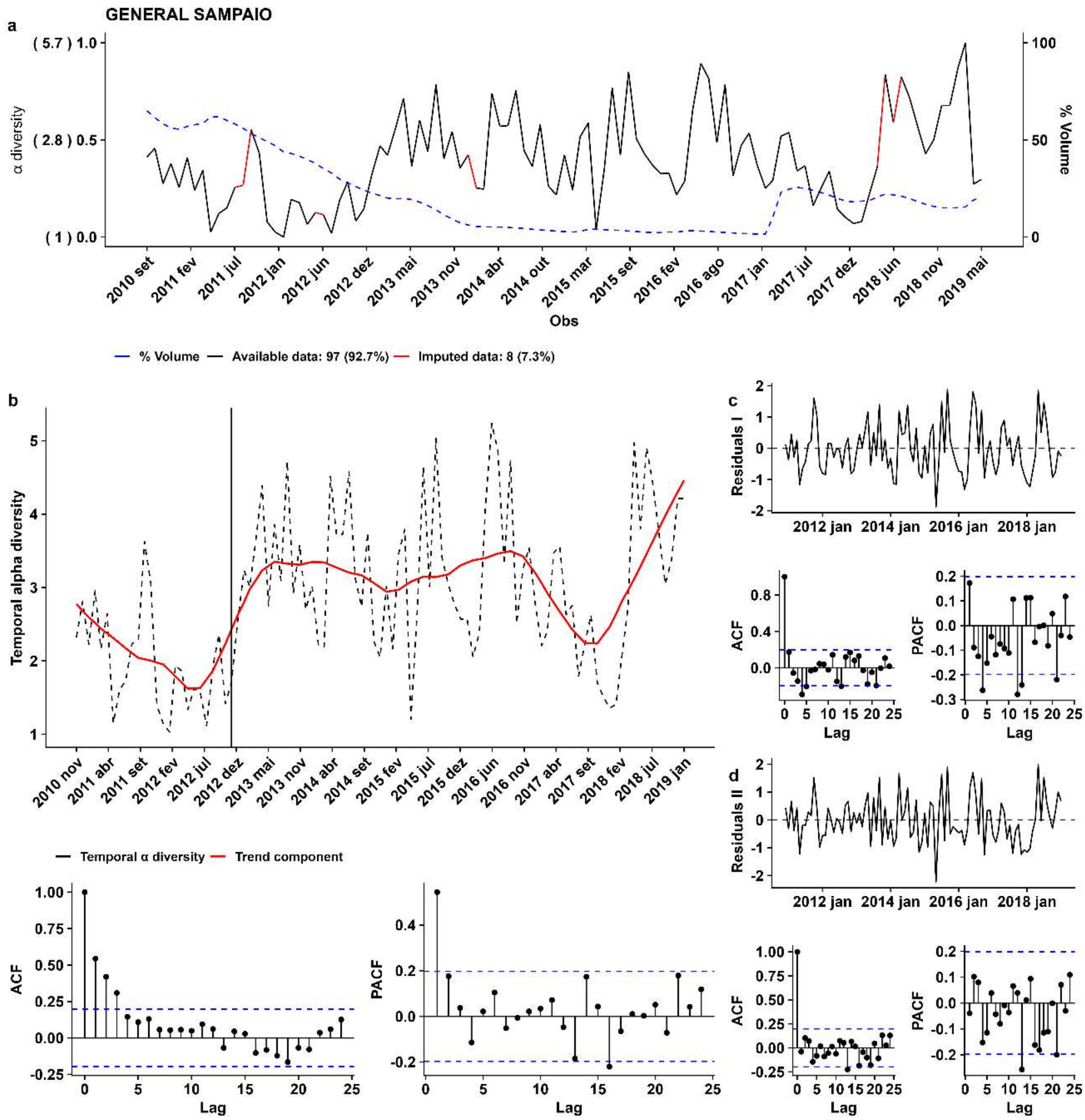


Figure 52 – Temporal α diversity | ITAPEBUSSU Residuals

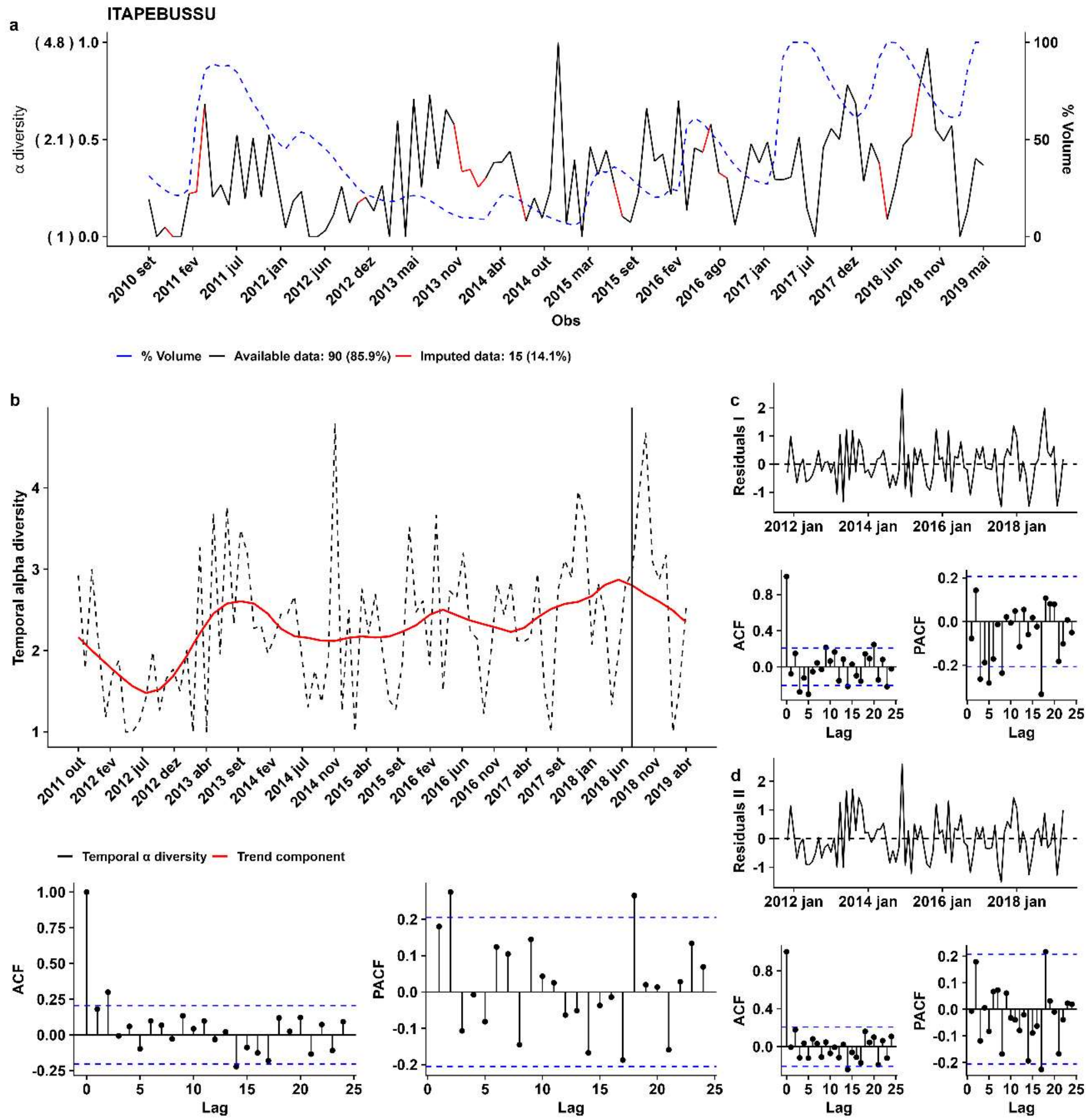


Figure 53 – Temporal α diversity | ITAÚNA Residuals

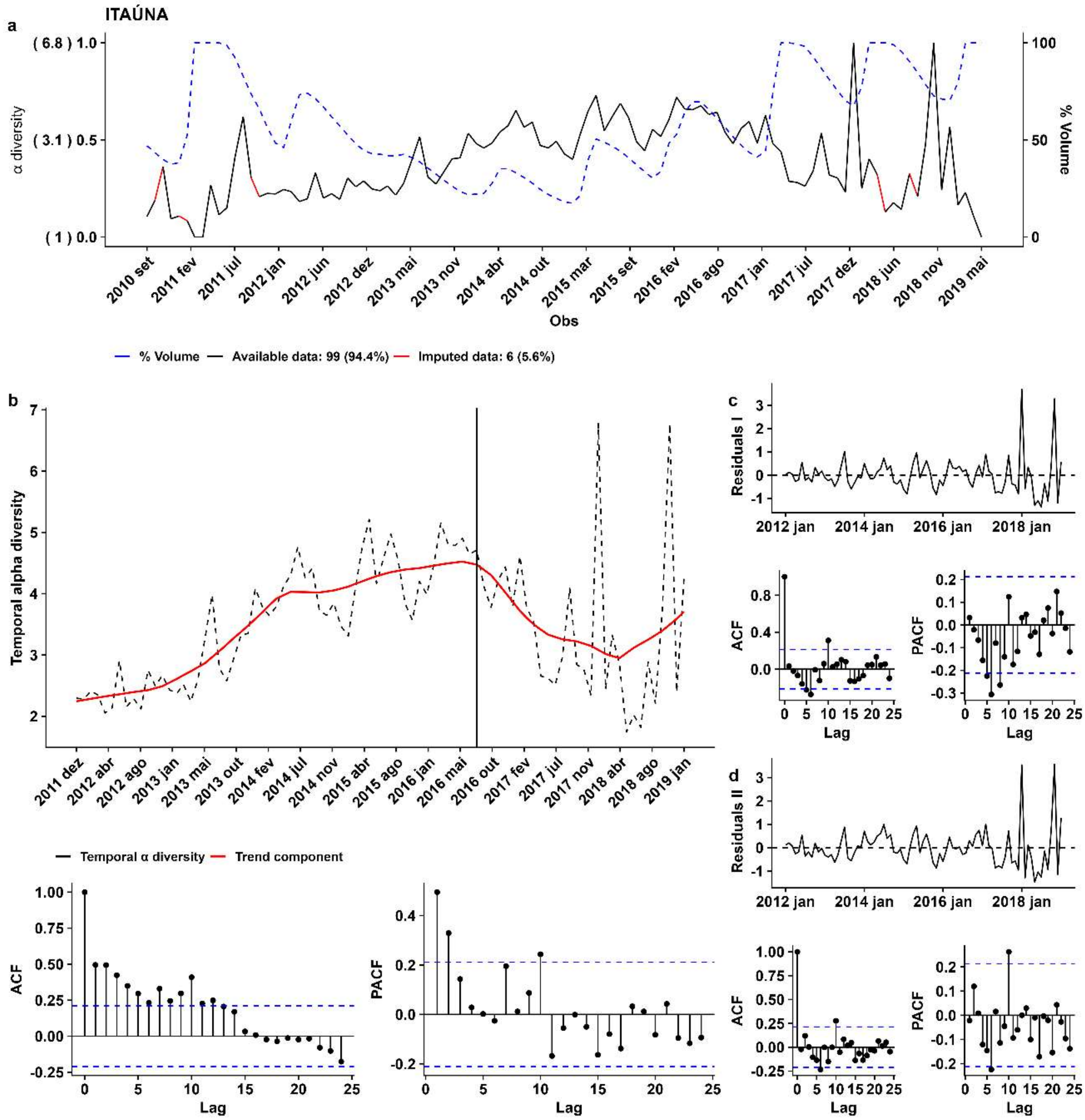


Figure 54 – Temporal α diversity | JABURU I Residuals

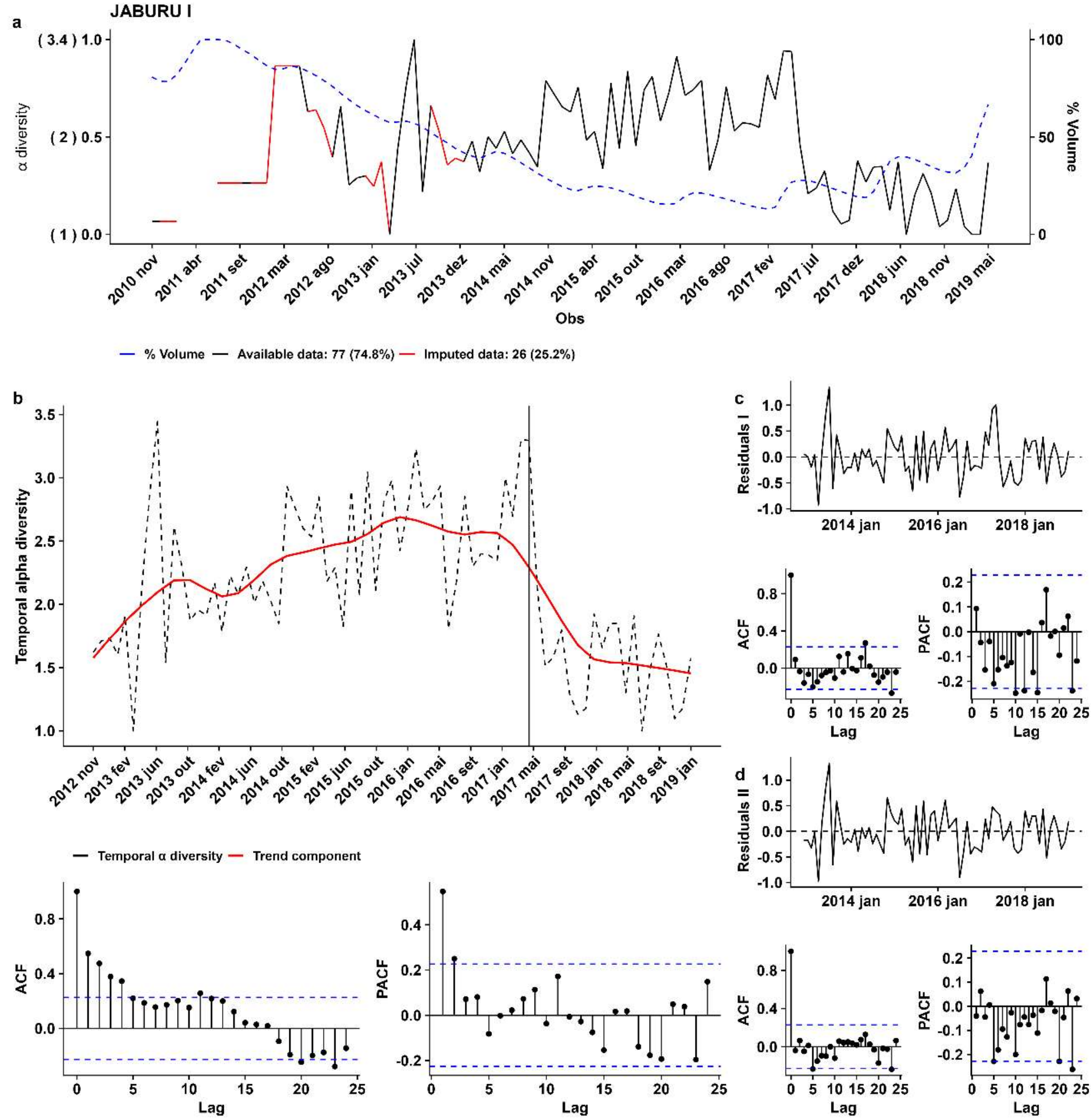


Figure 55 – Temporal α diversity | JENIPAPEIRO Residuals

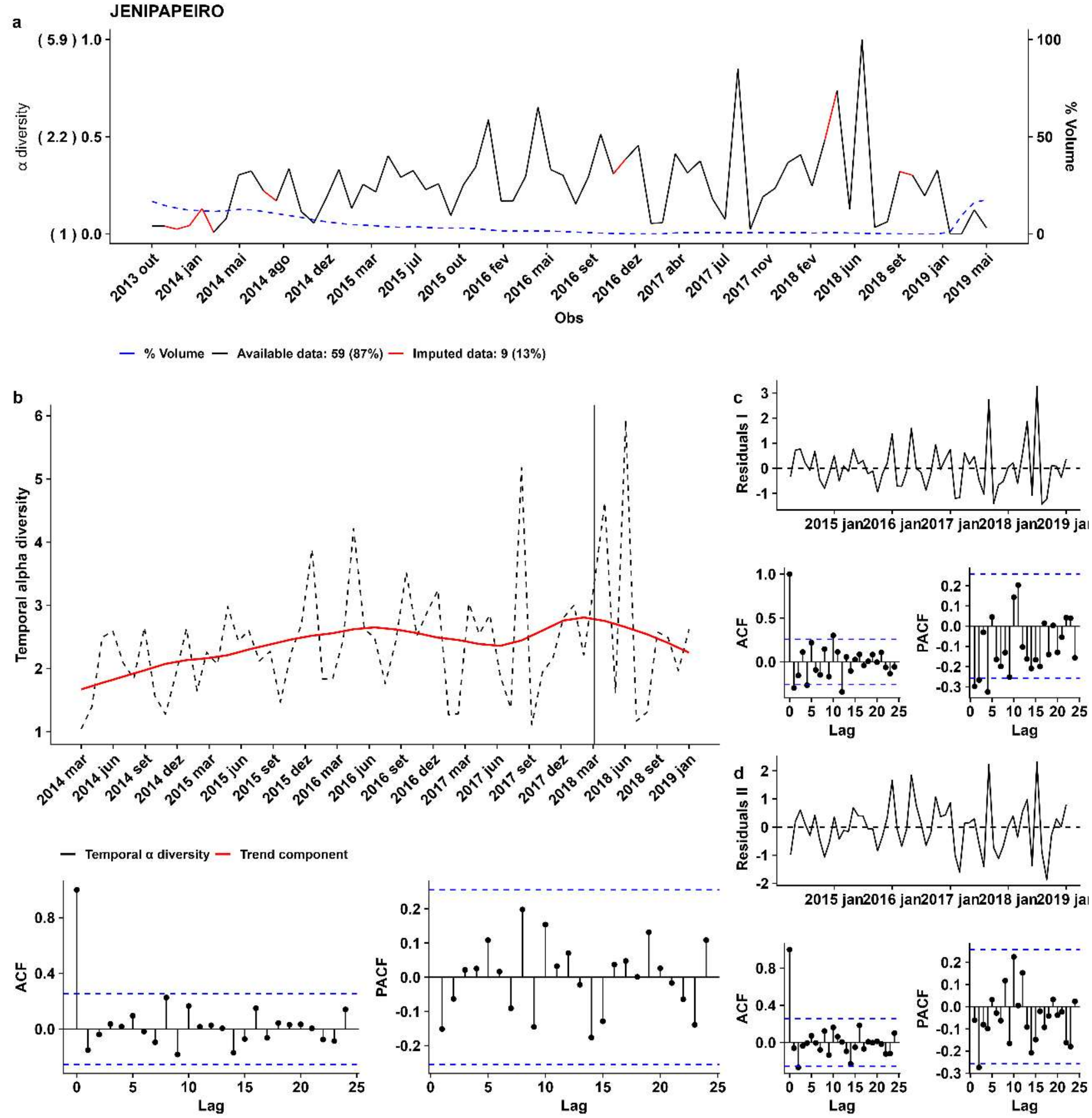


Figure 56 – Temporal α diversity | JOÃO LUÍS Residuals

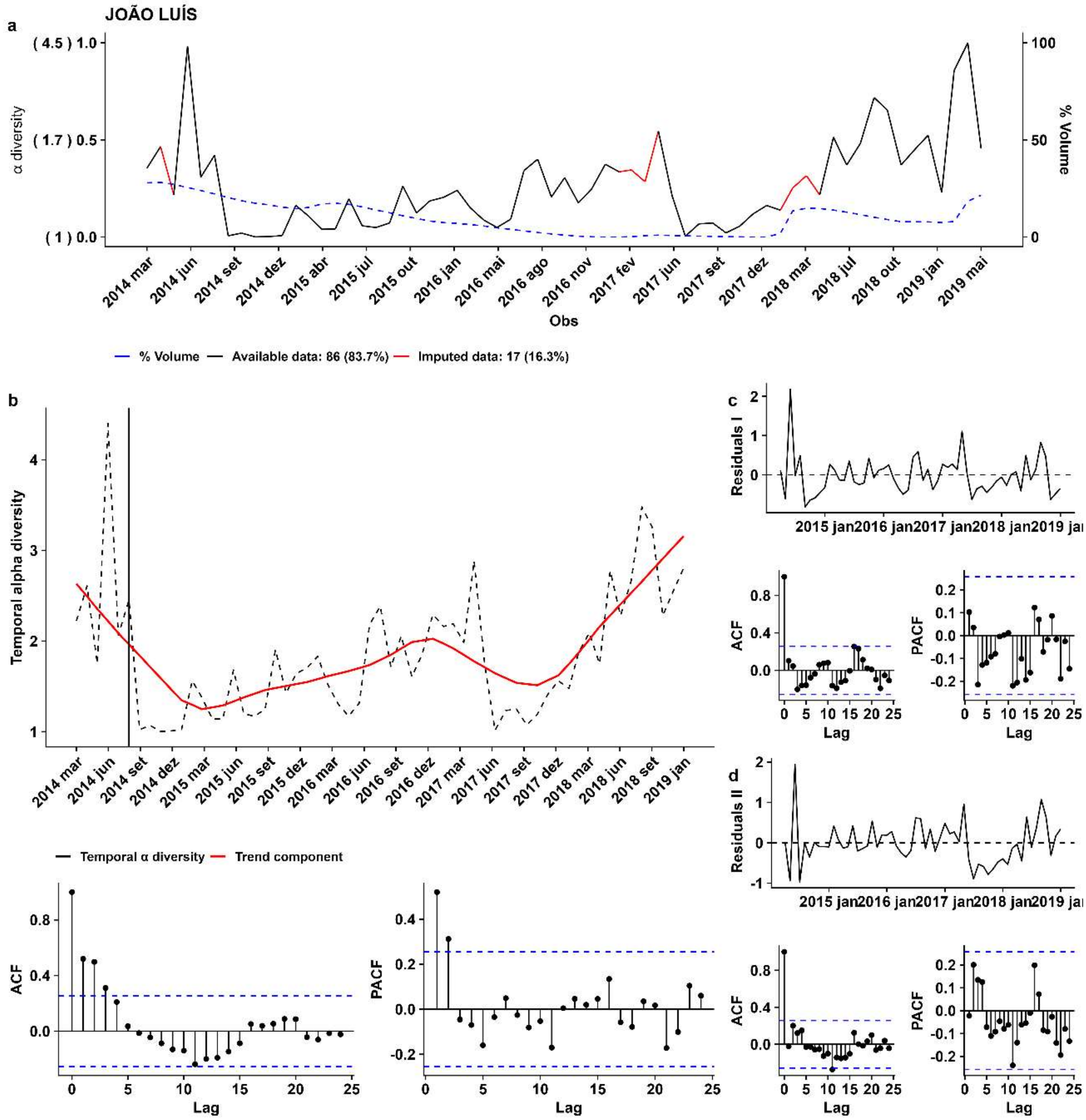


Figure 57 – Temporal α diversity | JUNCO Residuals

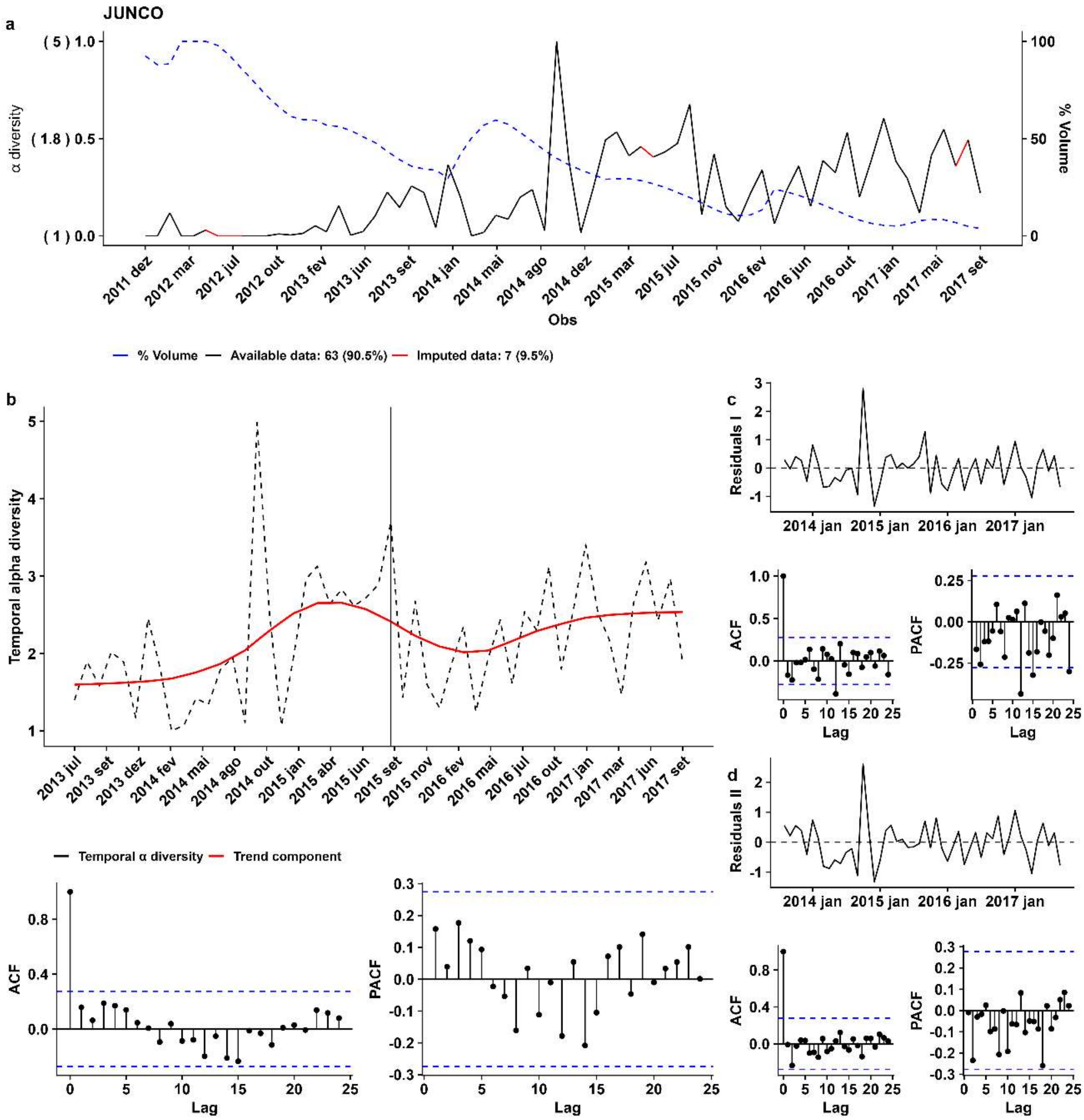


Figure 58 – Temporal α diversity | MADEIRO Residuals

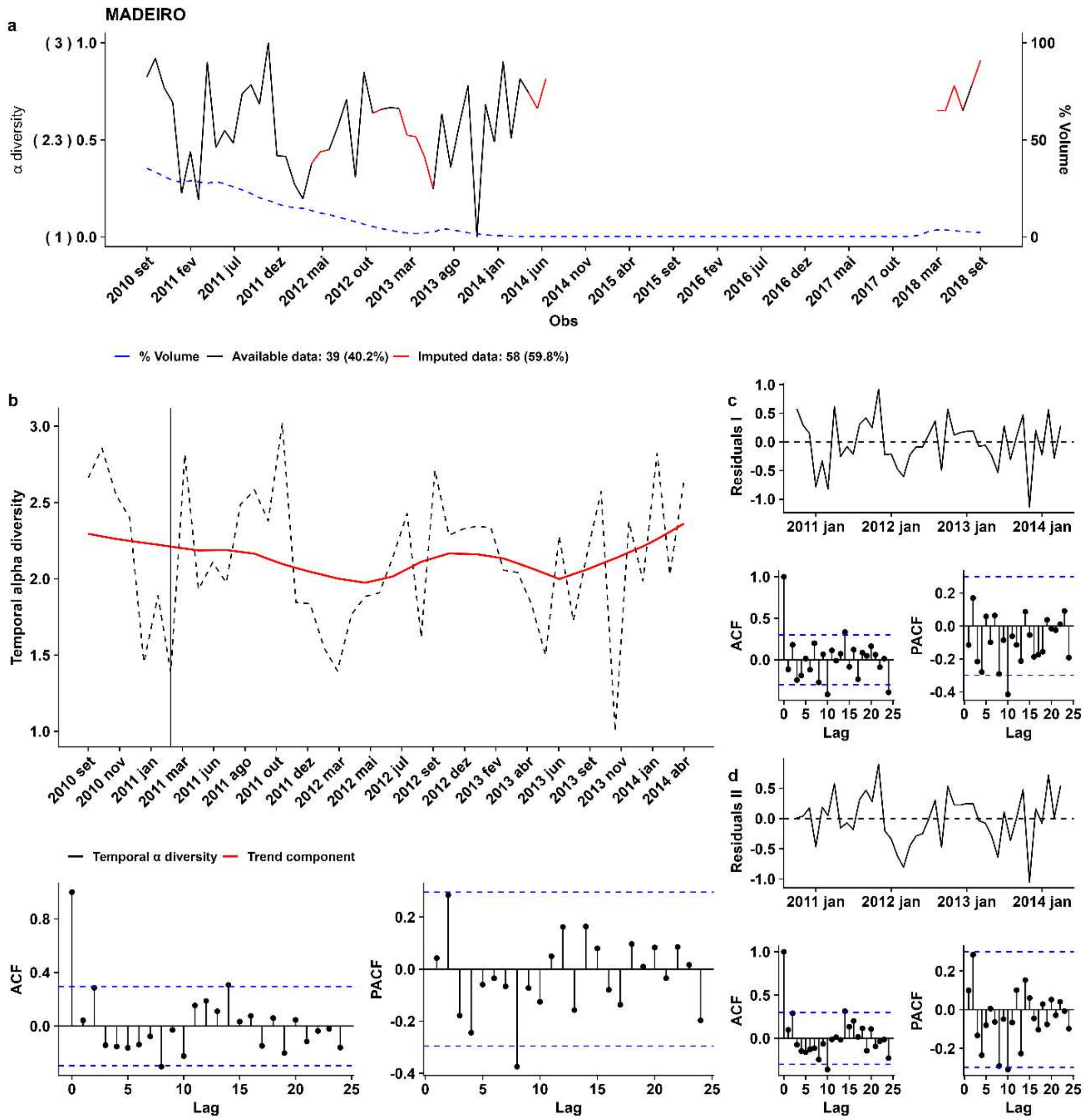


Figure 59 – Temporal α diversity | MALCOZINHADO Residuals

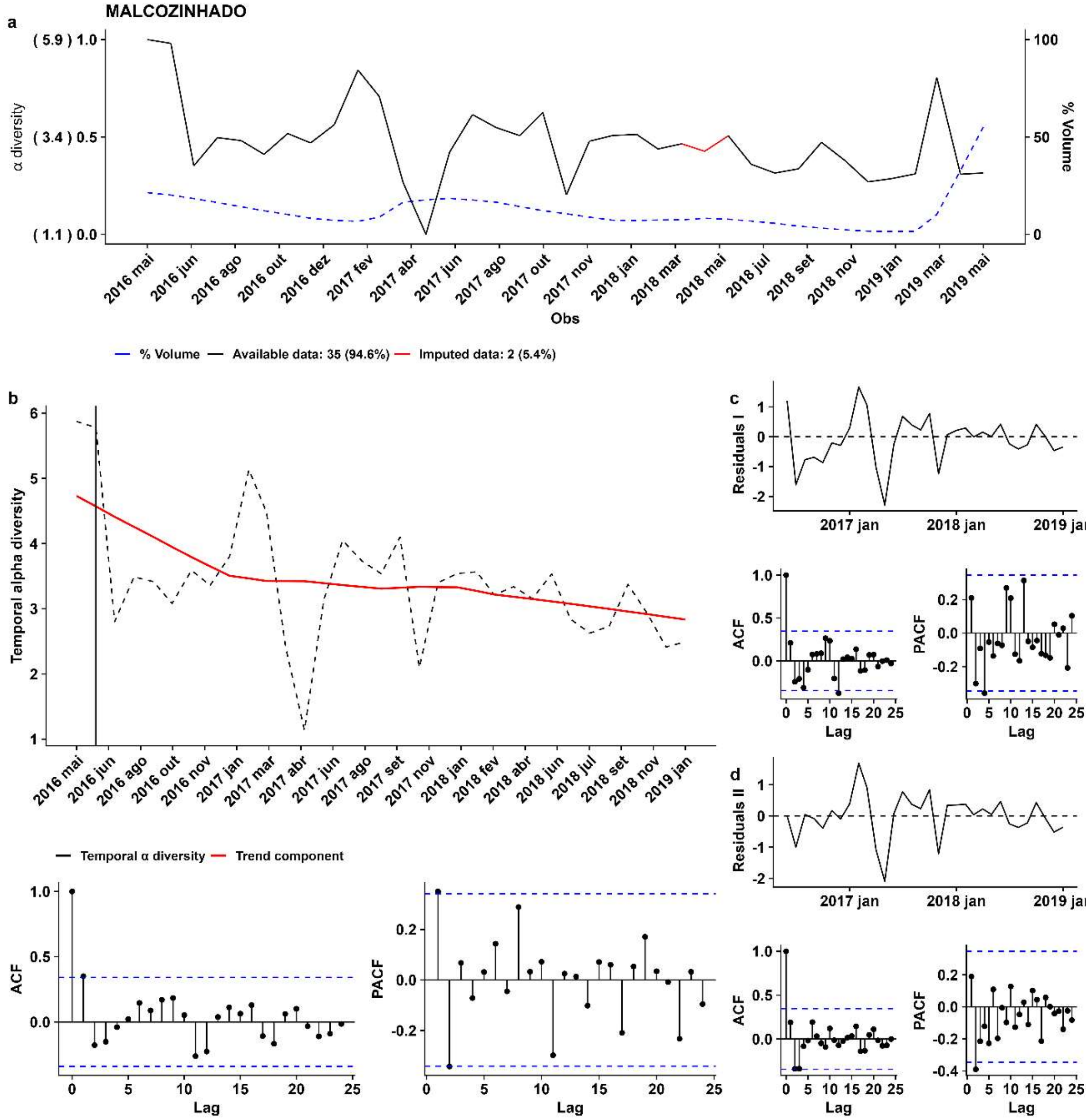


Figure 60 – Temporal α diversity | MAMOEIRO Residuals

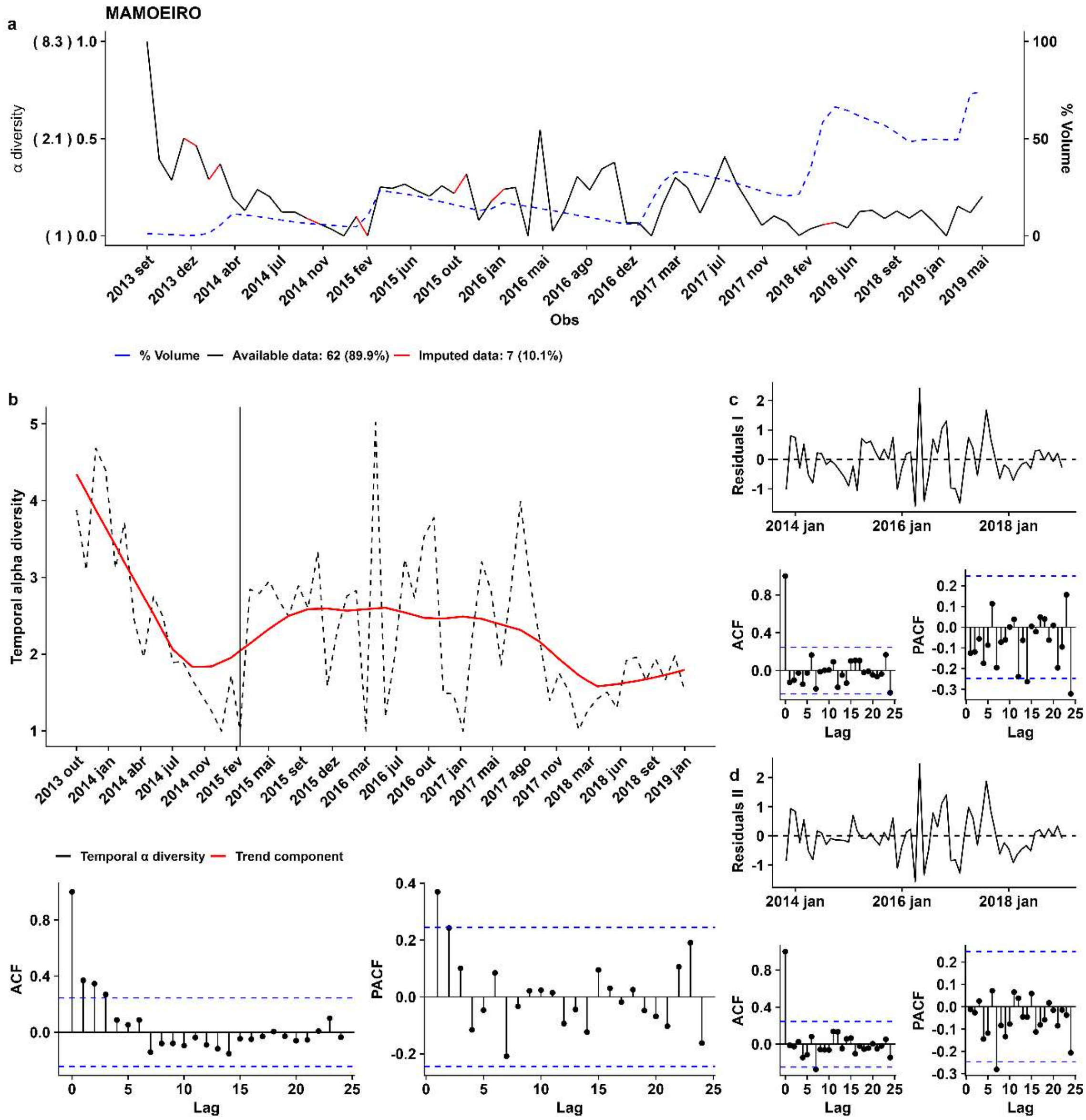


Figure 61 – Temporal α diversity | MARTINÓPOLE Residuals

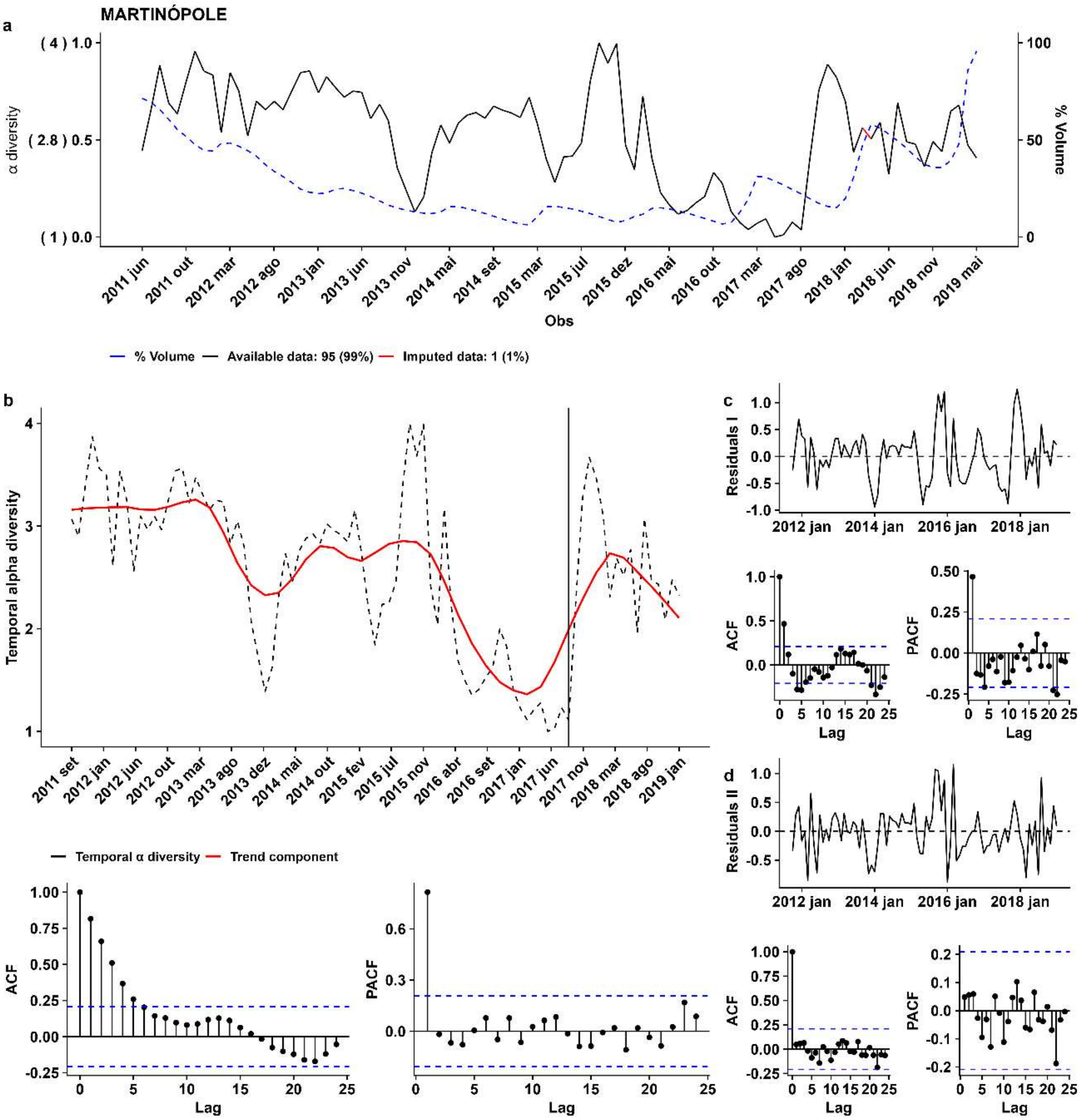


Figure 62 – Temporal α diversity | MISSI Residuals

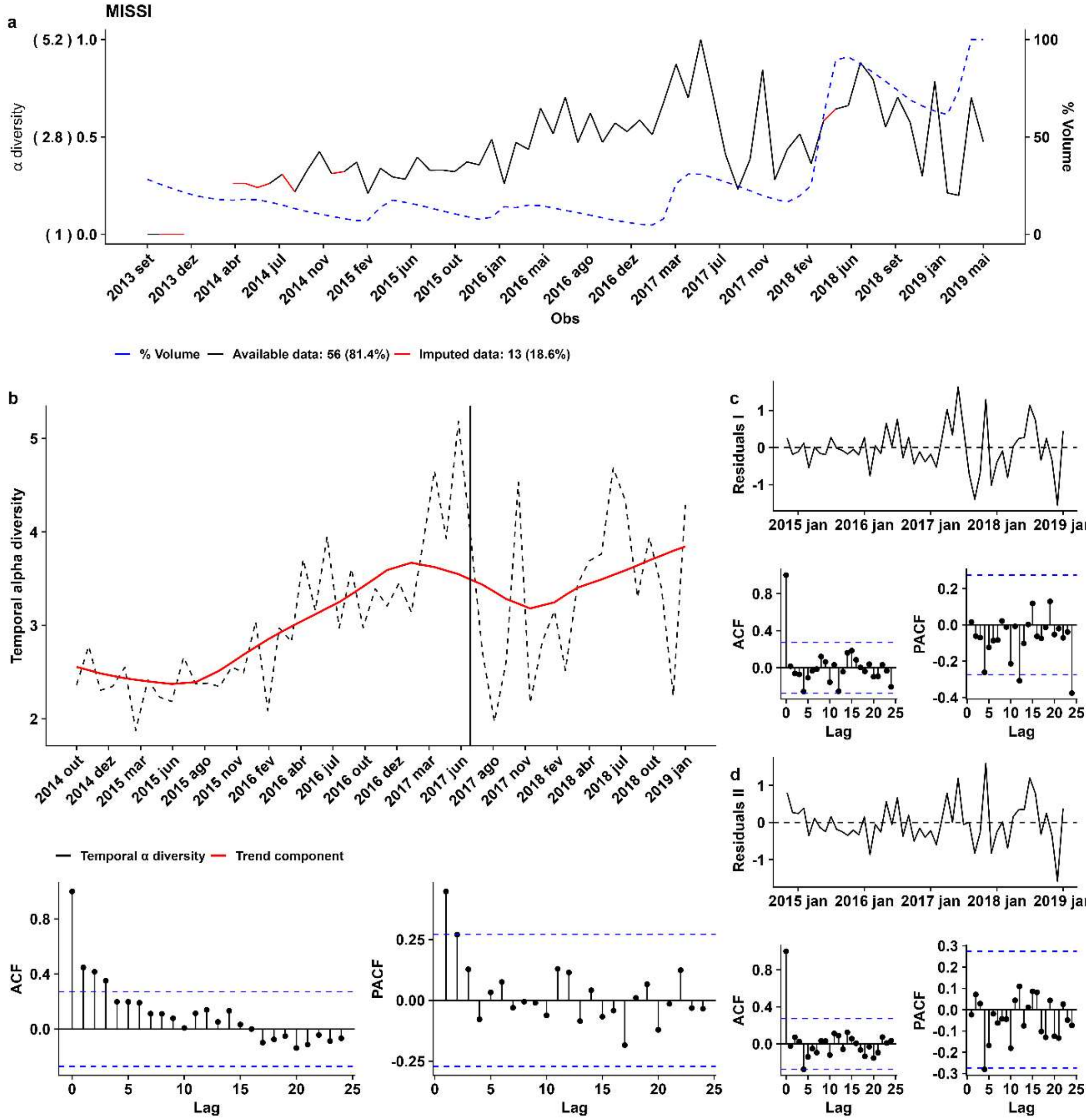


Figure 63 – Temporal α diversity | MONSENHOR TABOSA Residuals

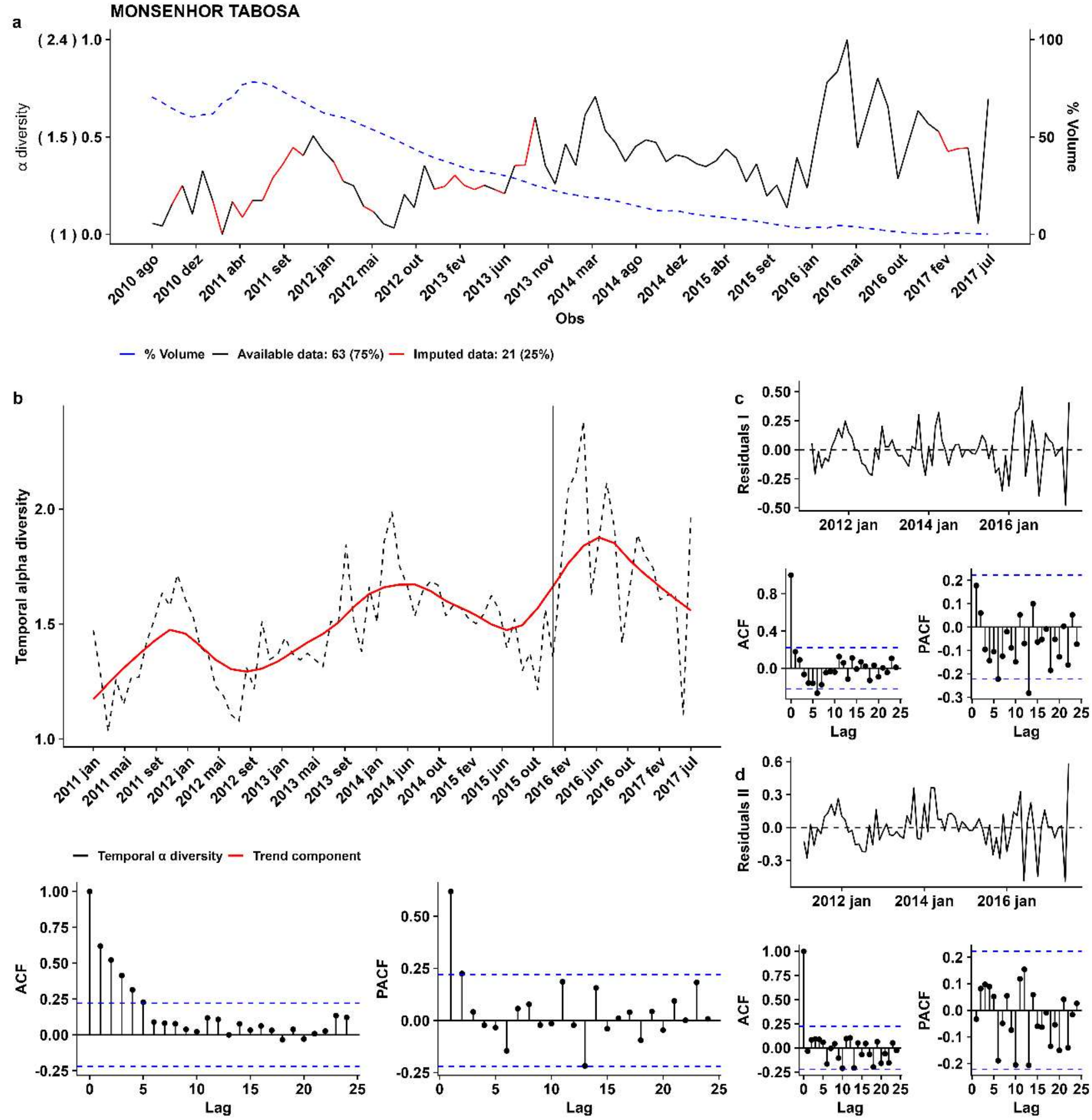


Figure 64 – Temporal α diversity | MUNDAÚ Residuals

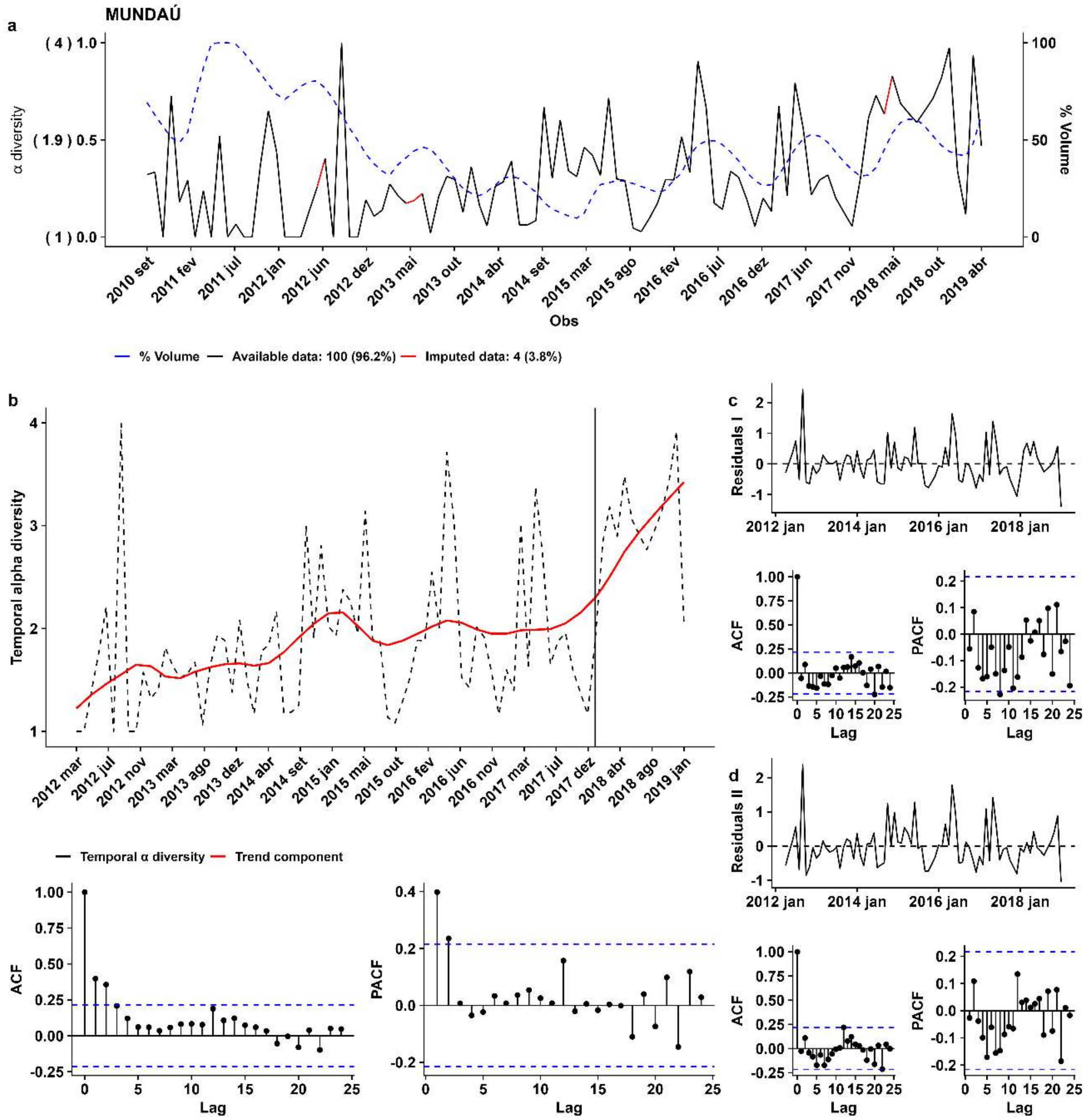


Figure 65 – Temporal α diversity | OLHO D'ÁGUA Residuals

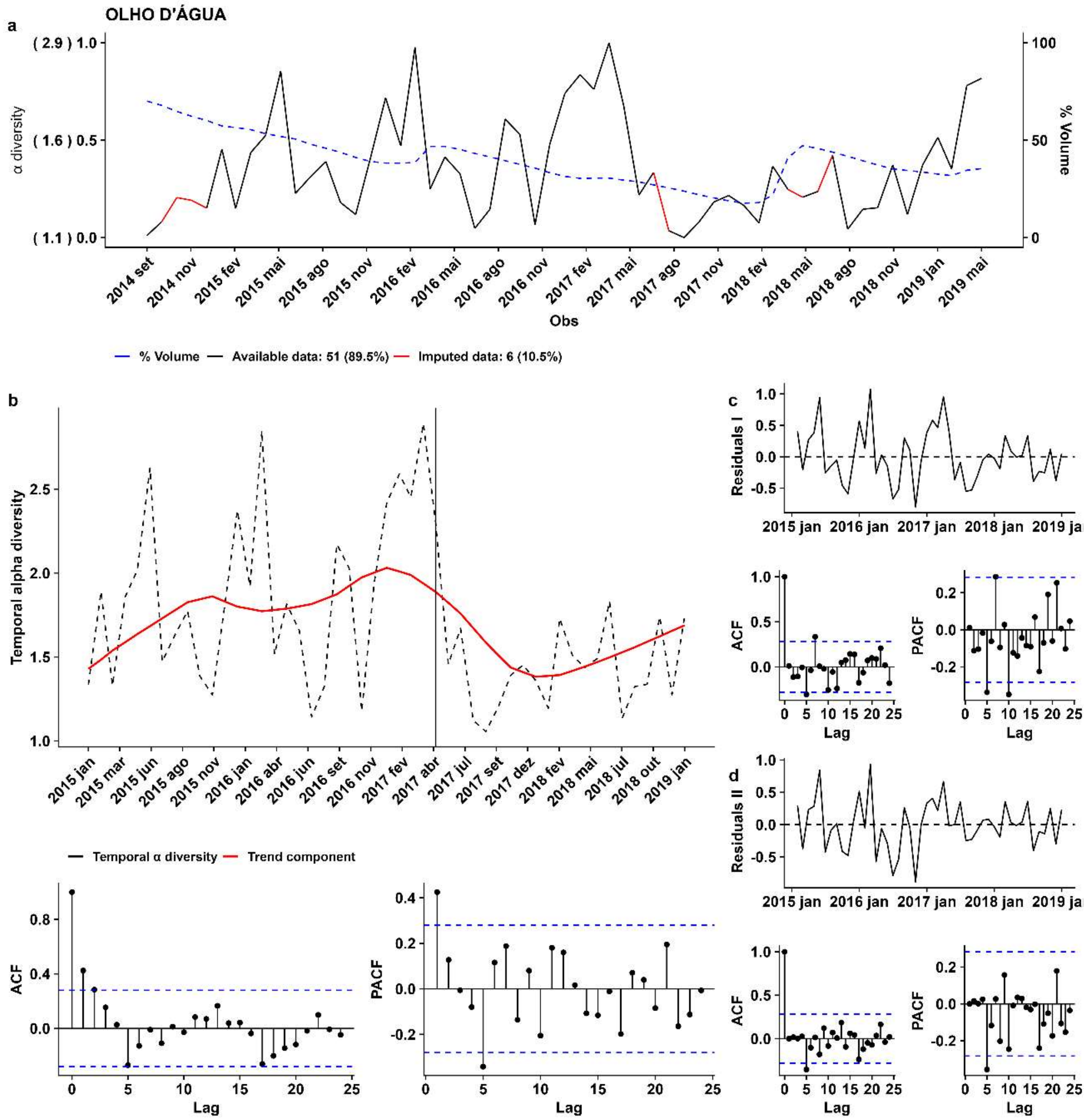


Figure 66 – Temporal α diversity | ORÓS Residuals

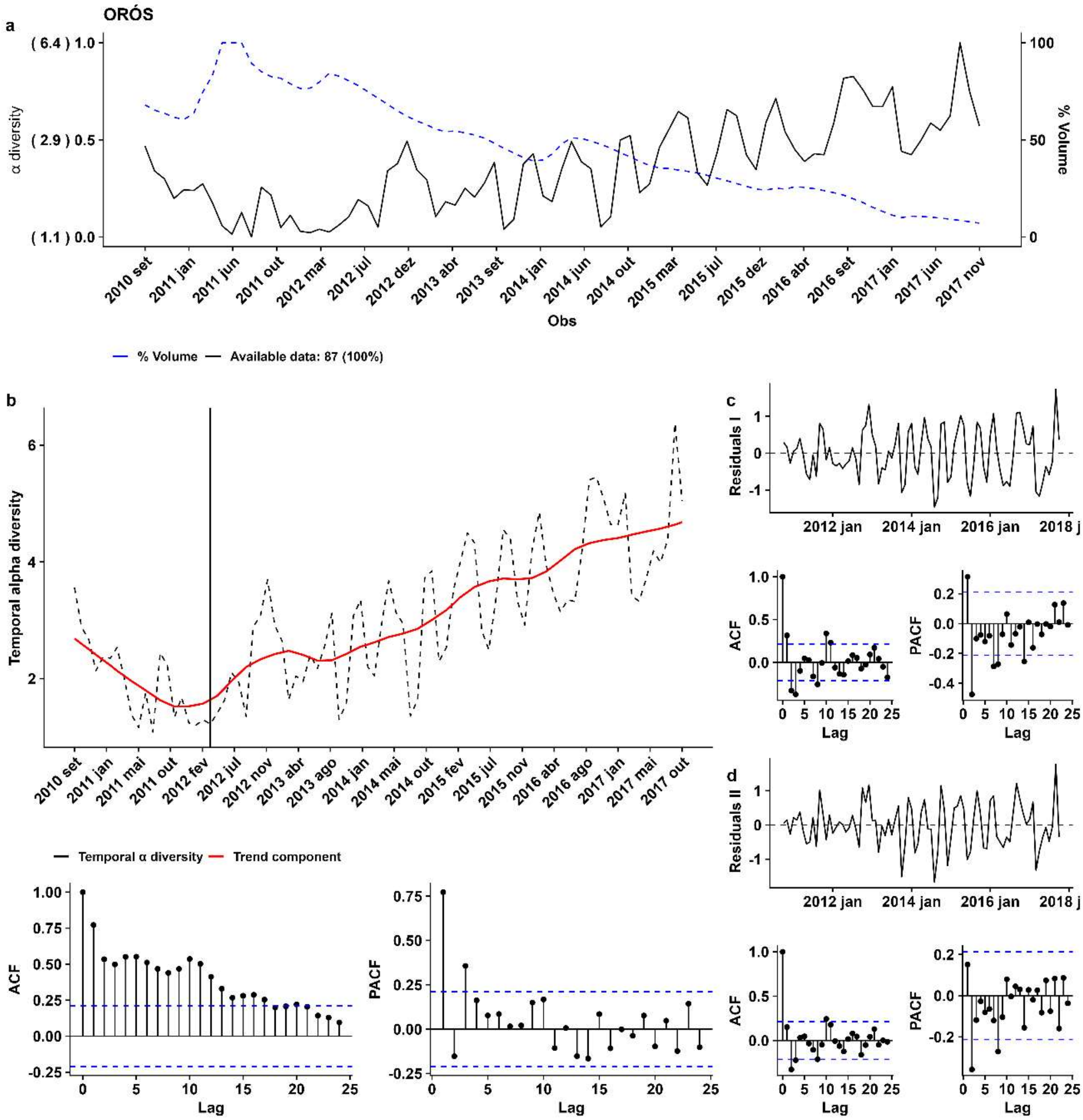


Figure 67 – Temporal α diversity | PACOTI Residuals

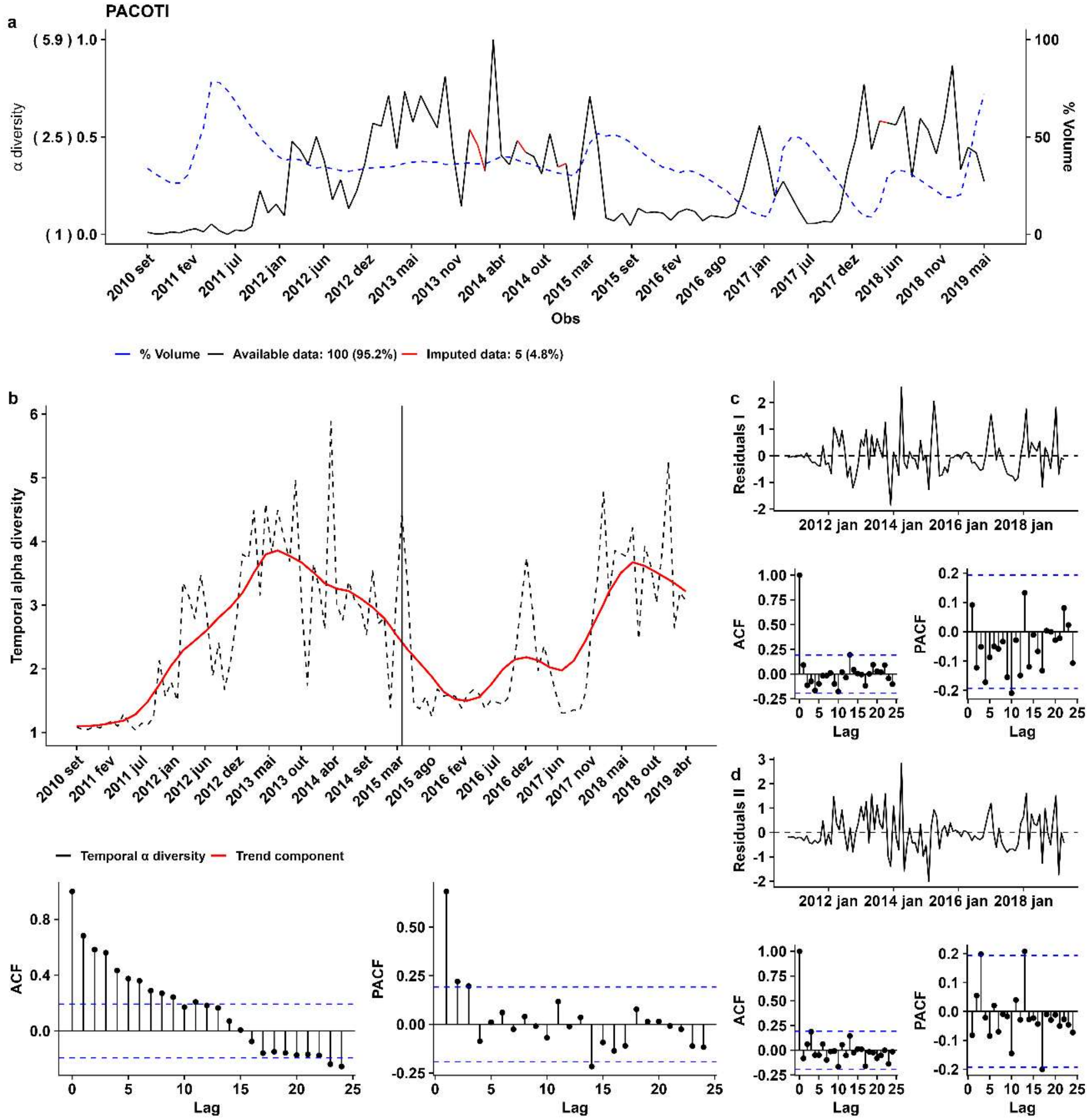


Figure 68 – Temporal α diversity | PATU Residuals

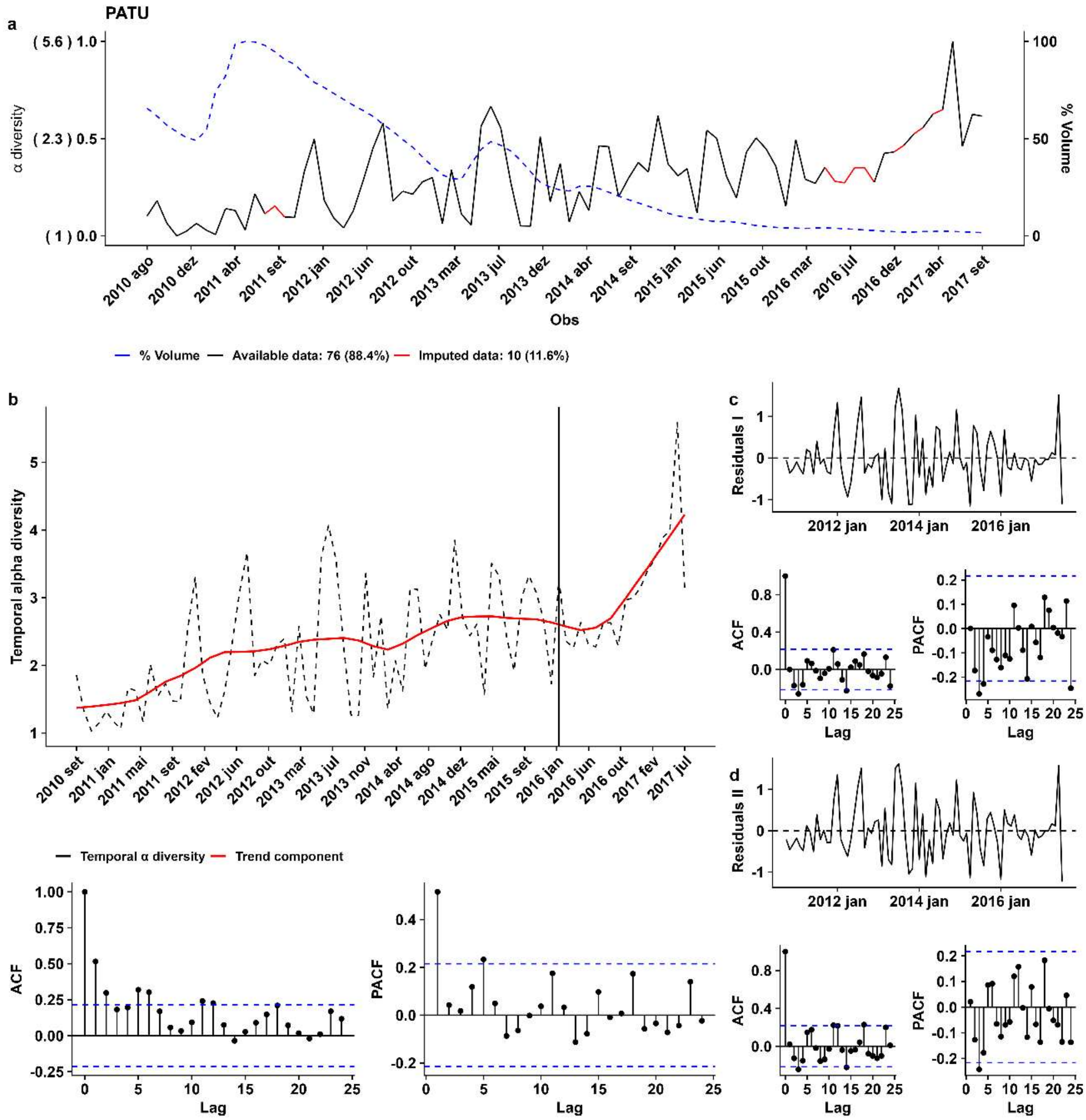


Figure 69 – Temporal α diversity | PEDRAS BRANCAS Residuals

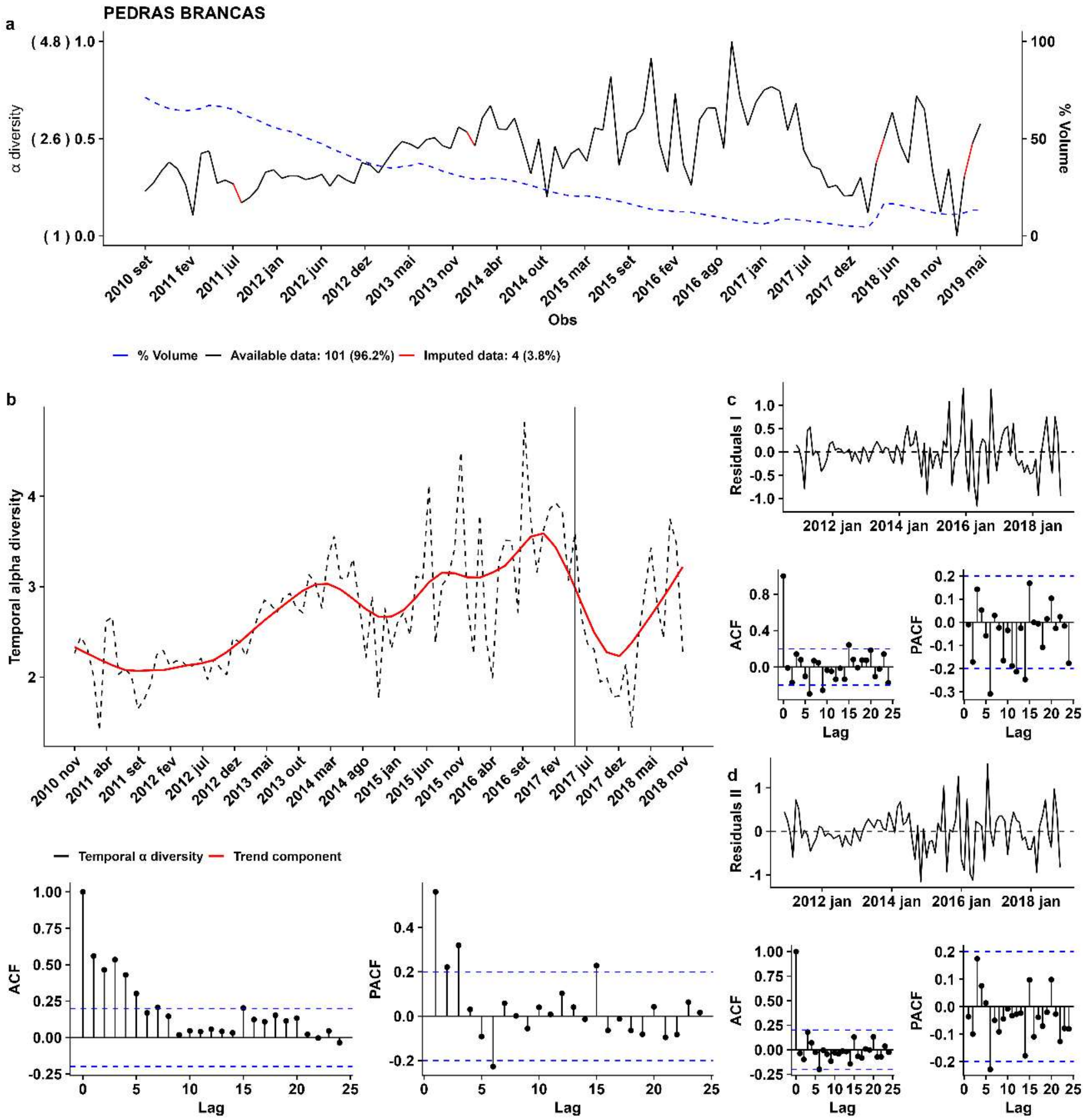


Figure 70 – Temporal α diversity | PENTECOSTE Residuals

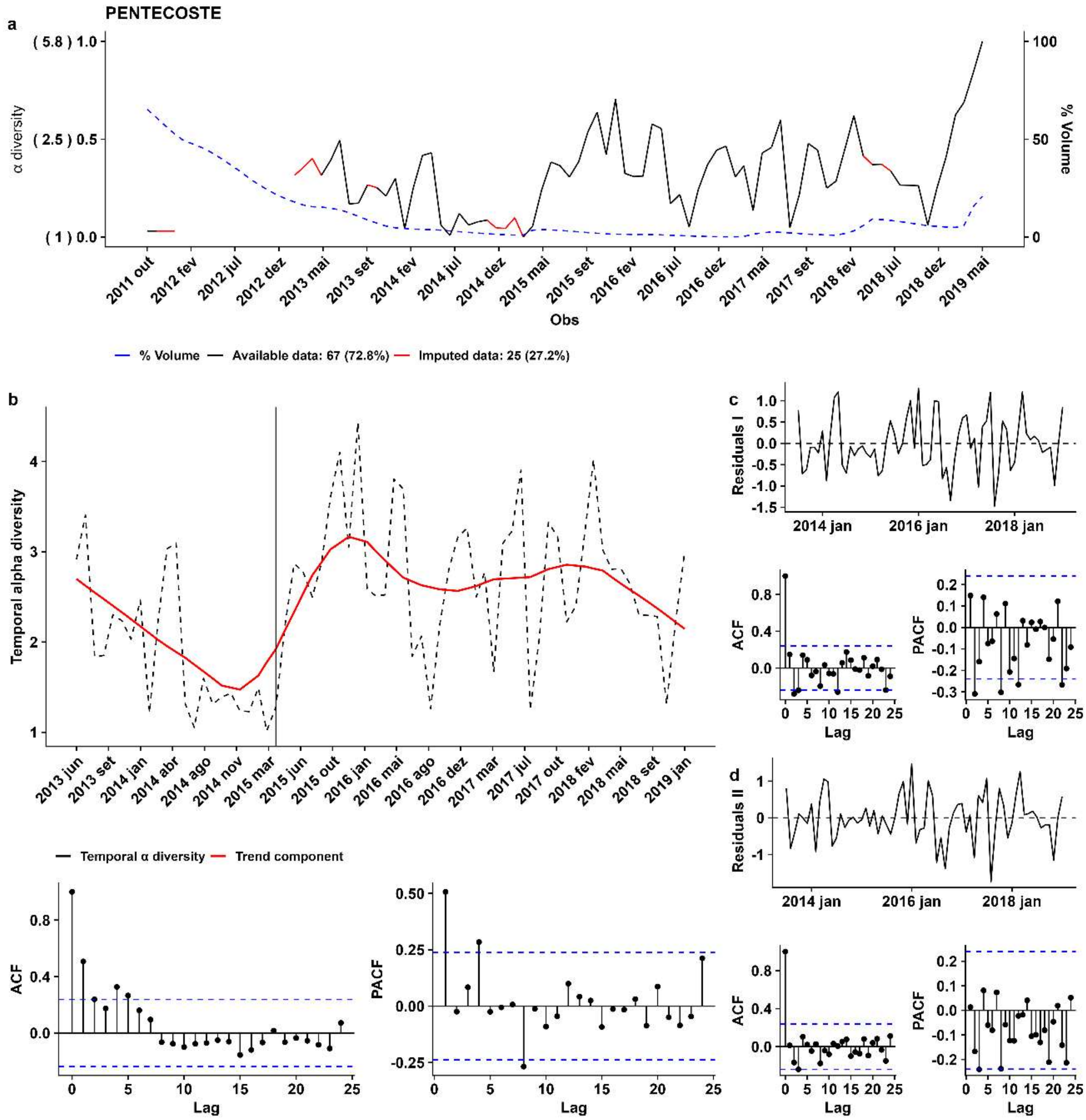


Figure 71 – Temporal α diversity | POMPEU SOBRINHO Residuals

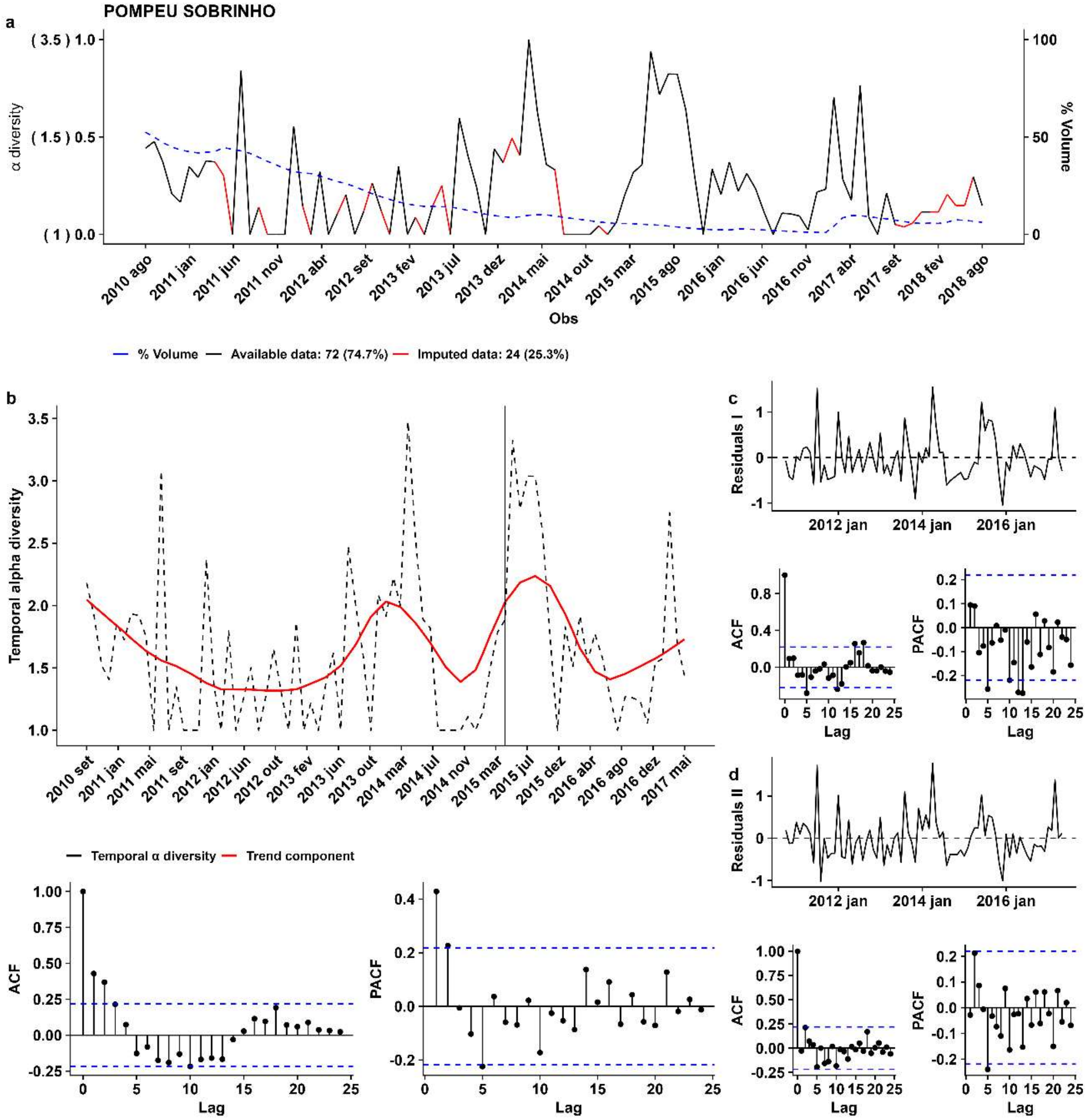


Figure 72 – Temporal α diversity | POTIRETAMA Residuals

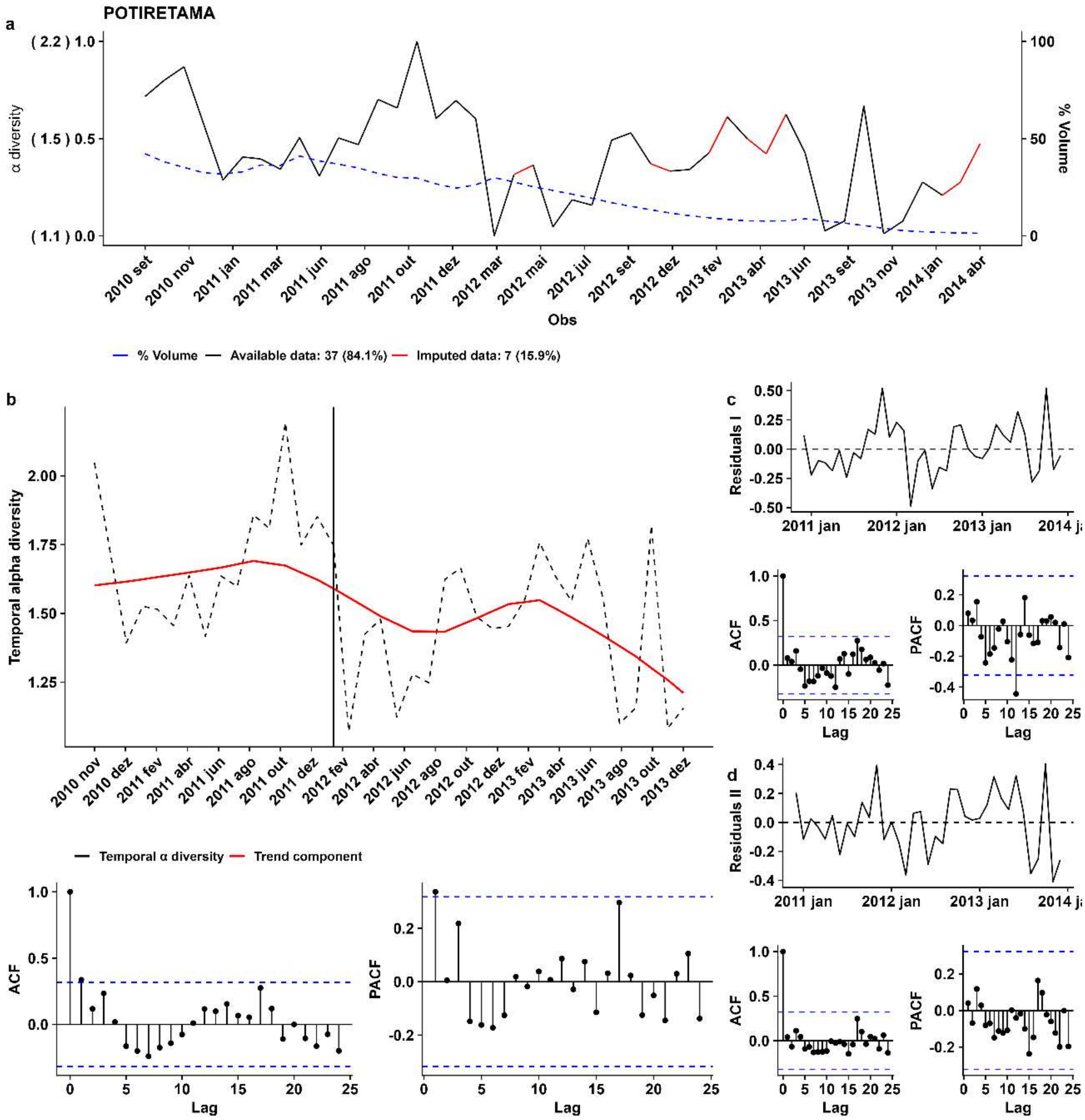


Figure 73 – Temporal α diversity | RIACHÃO Residuals

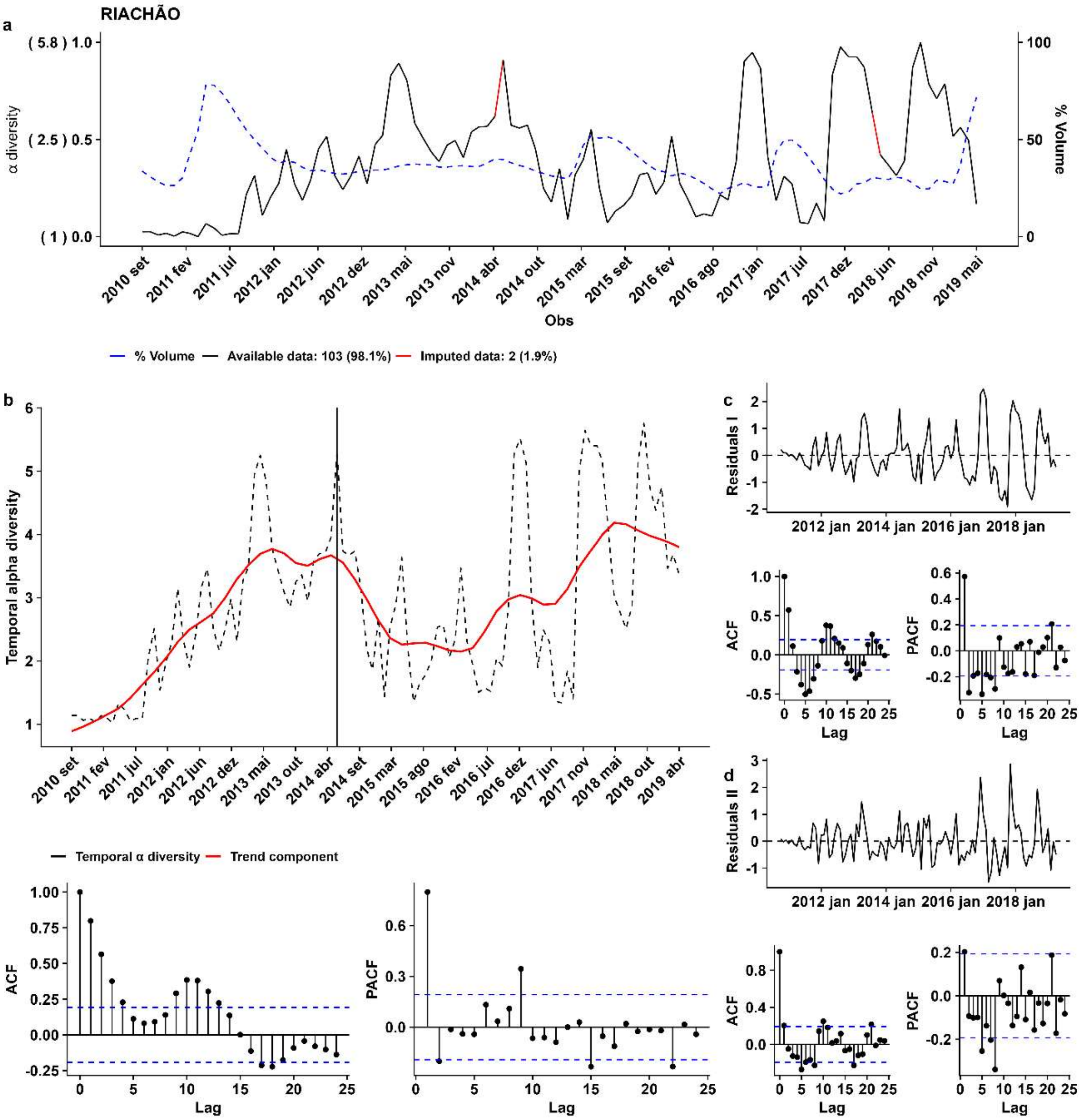


Figure 74 – Temporal α diversity | RIACHO DA SERRA Residuals

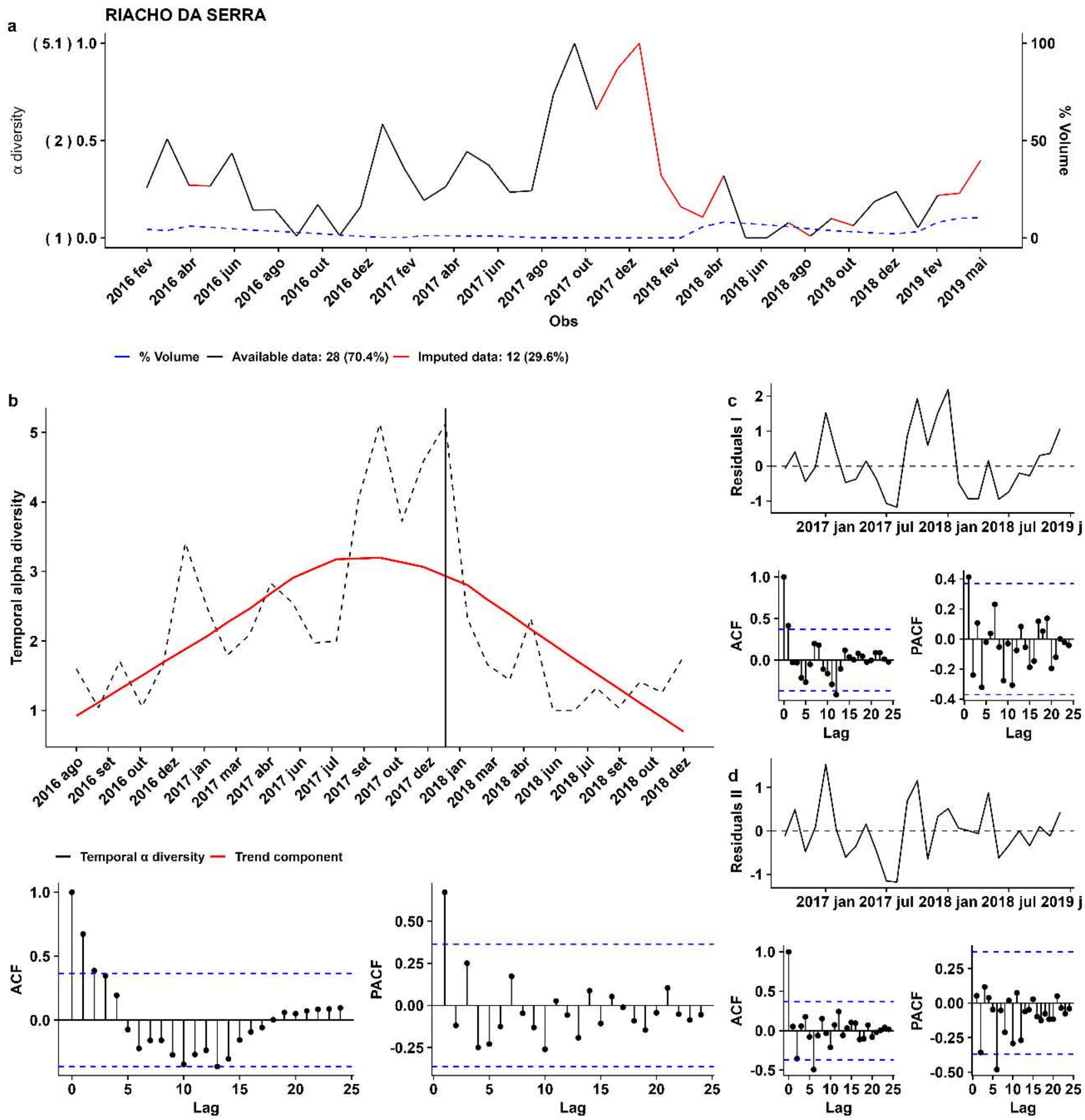

a
RIACHO DA SERRA
α diversity
% Volume
Obs
% Volume
Available data: 28 (70.4%)
Imputed data: 12 (29.6%)
b
Temporal alpha diversity
Temporal α diversity
Trend component
c
Residuals I
d
Residuals II
ACF
PACF
Lag

Figure 75 – Temporal α diversity | ROSÁRIO Residuals

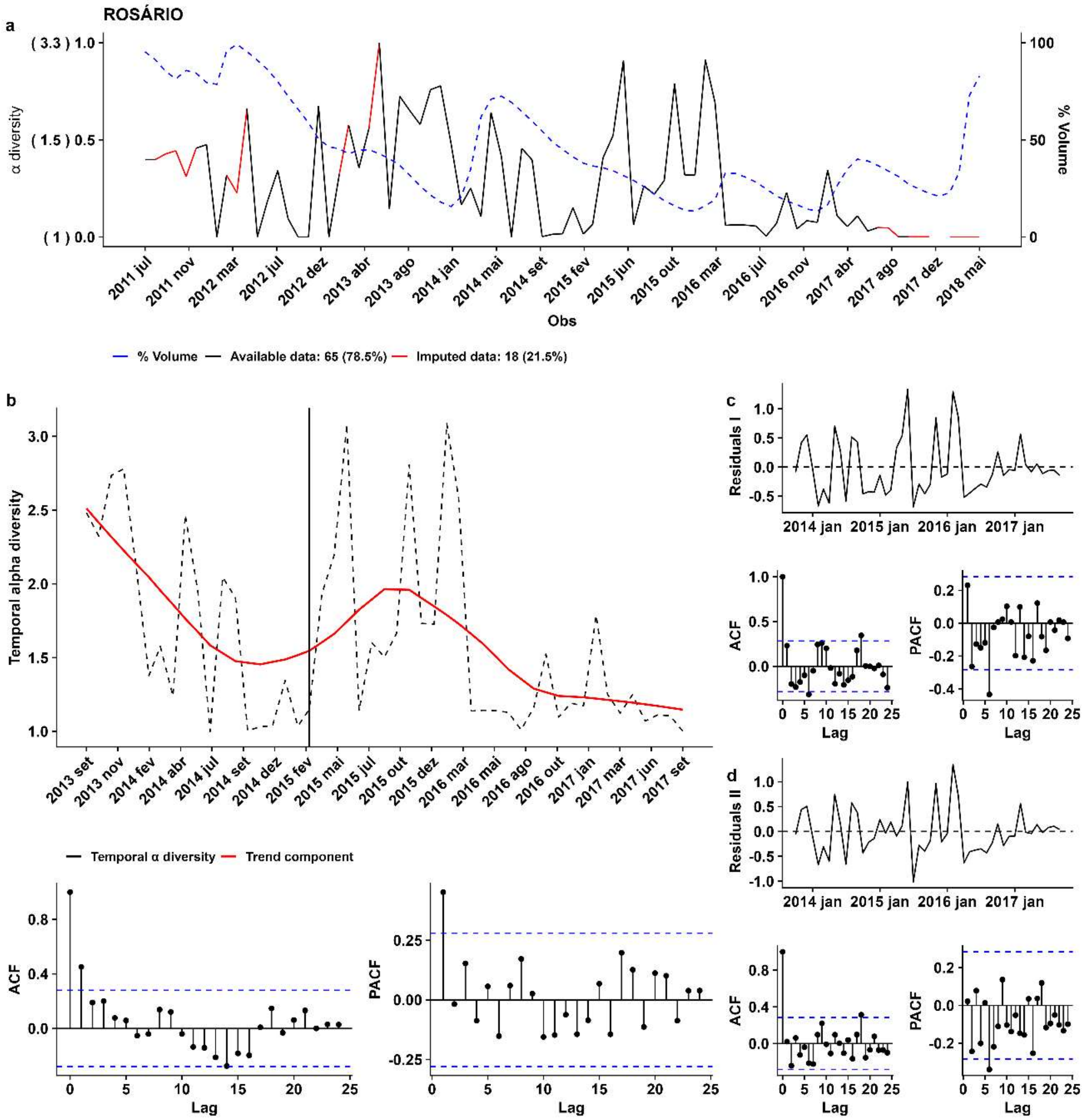


Figure 76 – Temporal α diversity | SÃO DOMINGOS Residuals

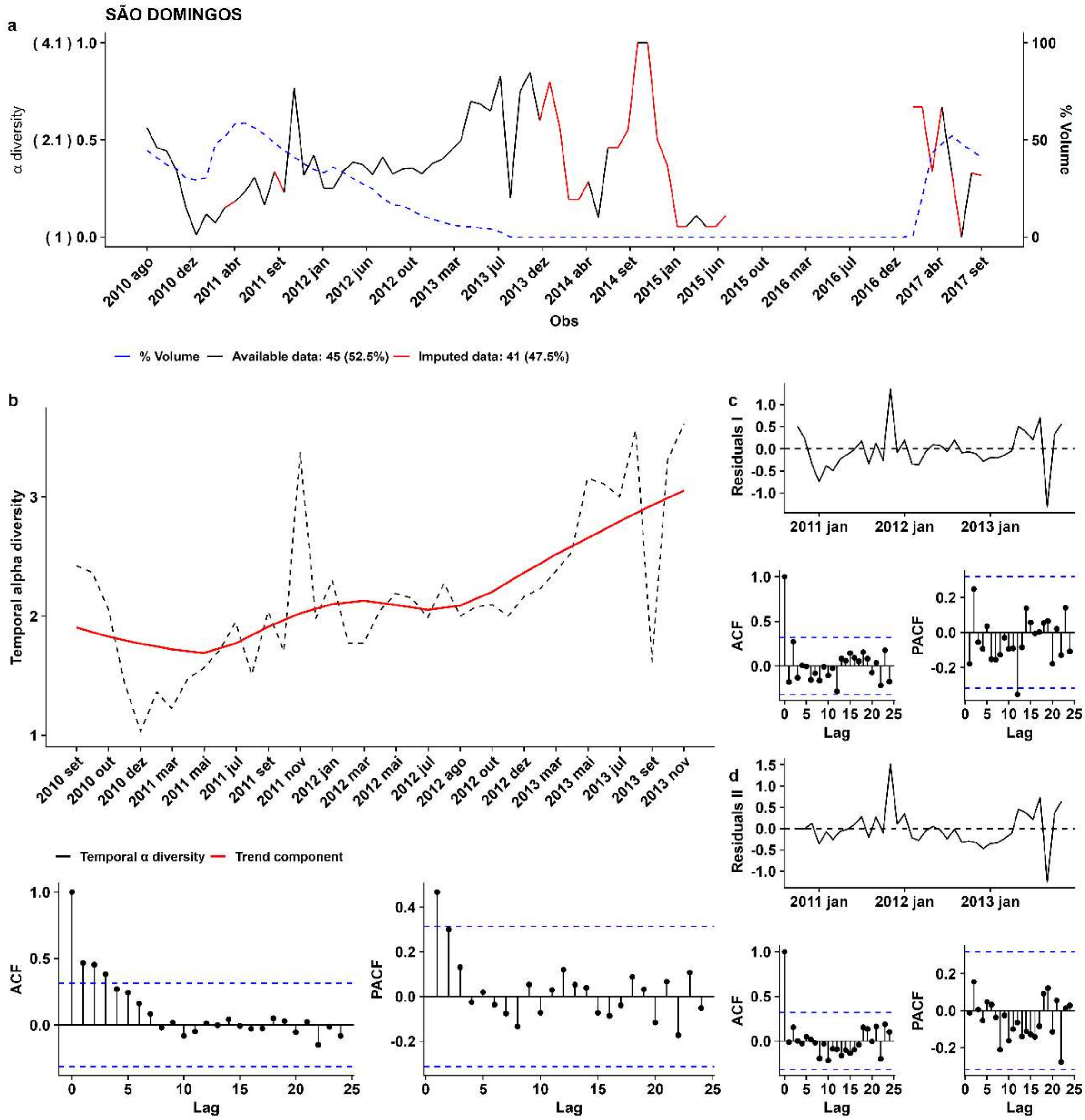


Figure 77 – Temporal α diversity | SÃO JOSÉ I Residuals

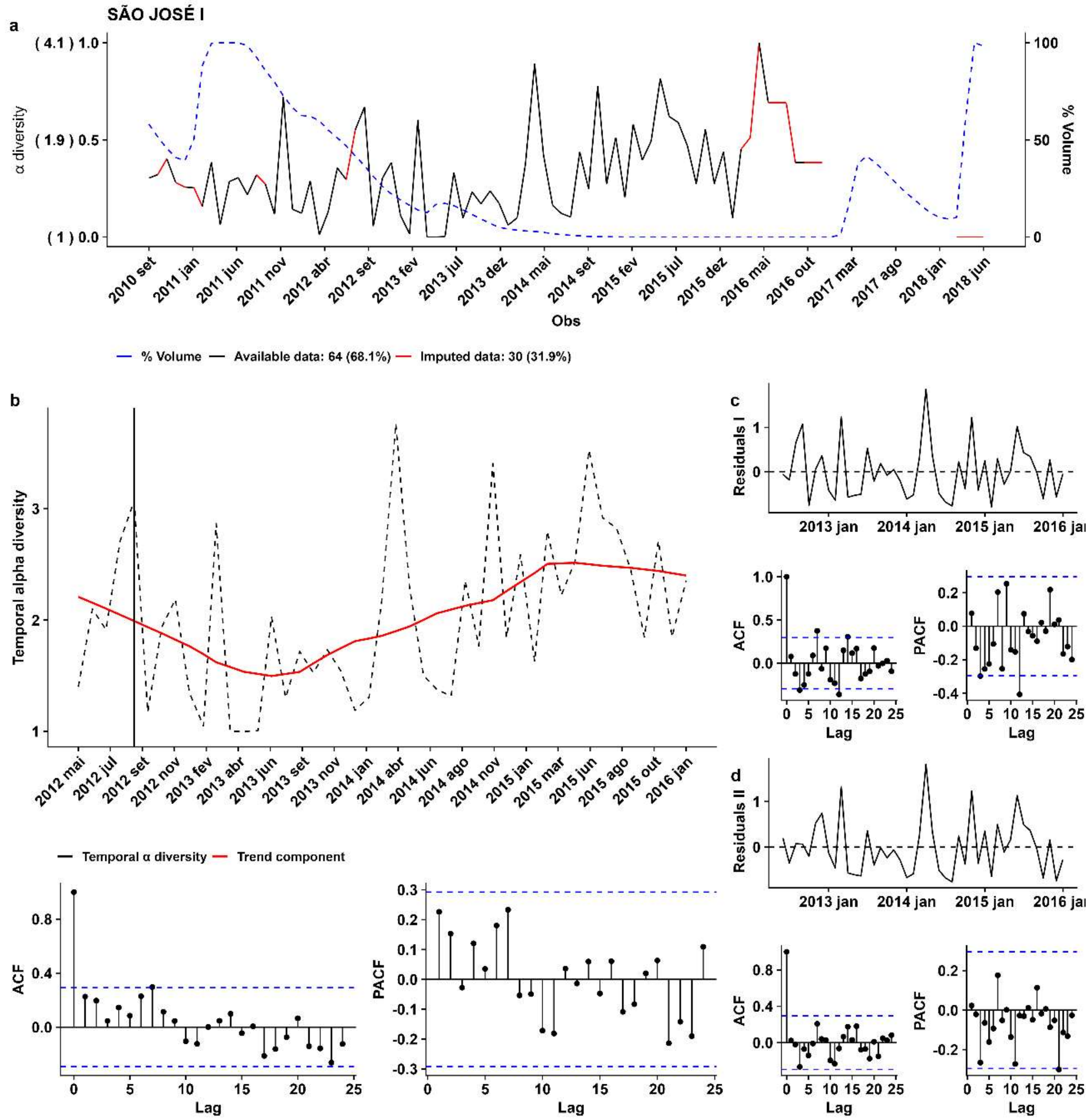


Figure 78 – Temporal α diversity | SÃO PEDRO TIMBAÚBA Residuals

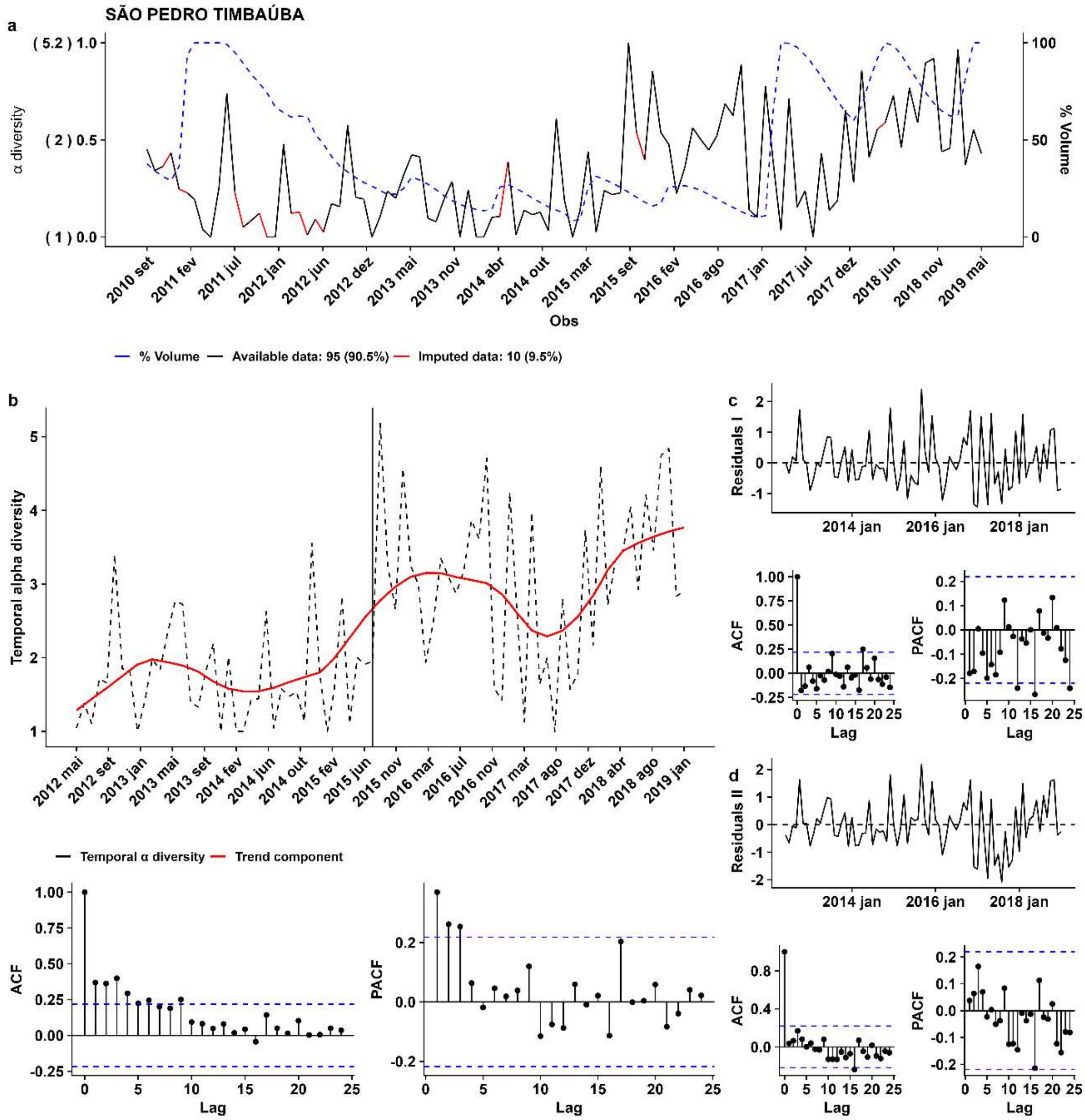


Figure 79 – Temporal α diversity | SITIOS NOVOS Residuals

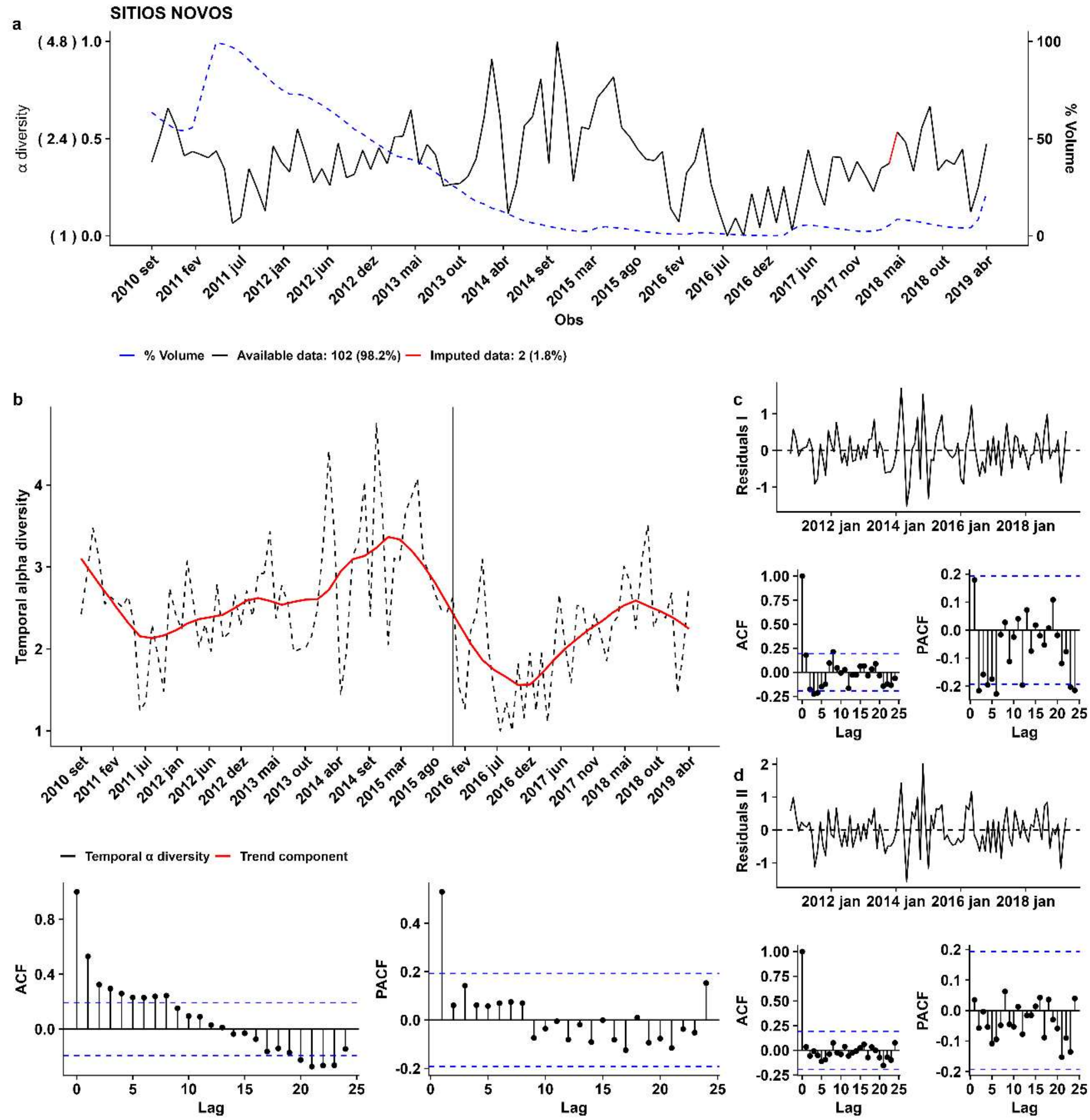


Figure 80 – Temporal α diversity | TRUSSU Residuals

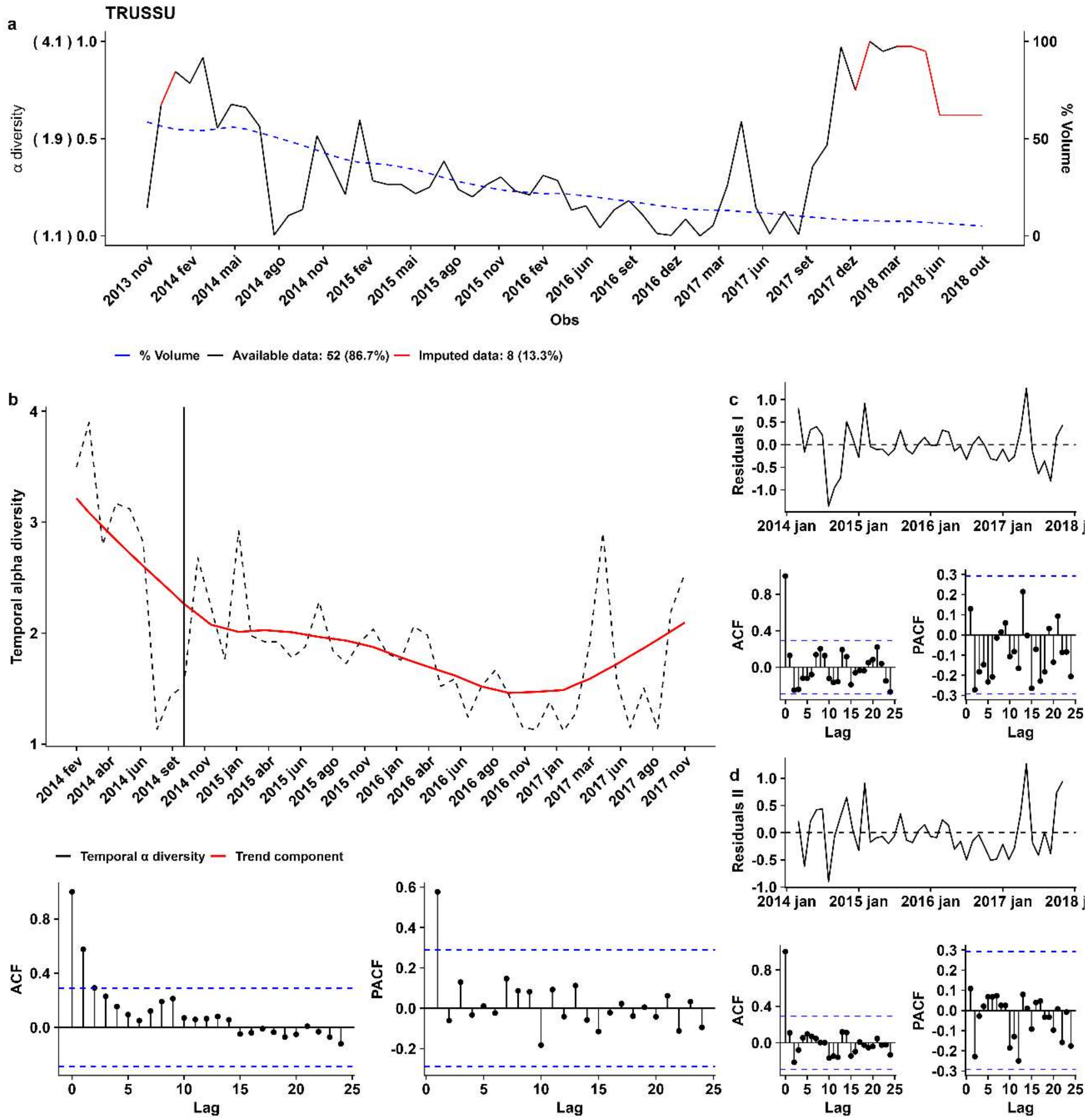


Figure 81 – Temporal α diversity | UBALDINHO Residuals

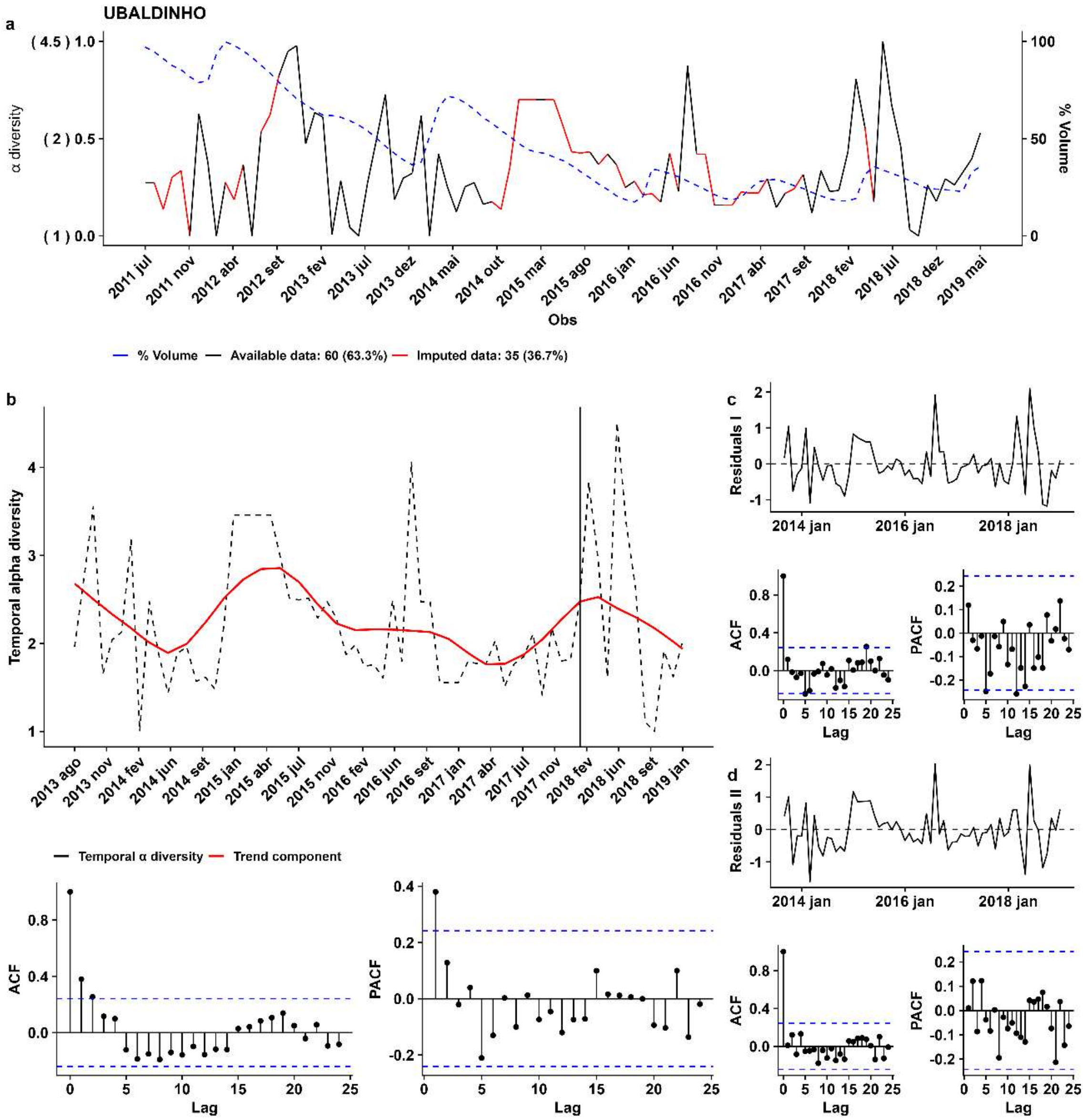


Figure 82 – Temporal α diversity | VALÉRIO Residuals

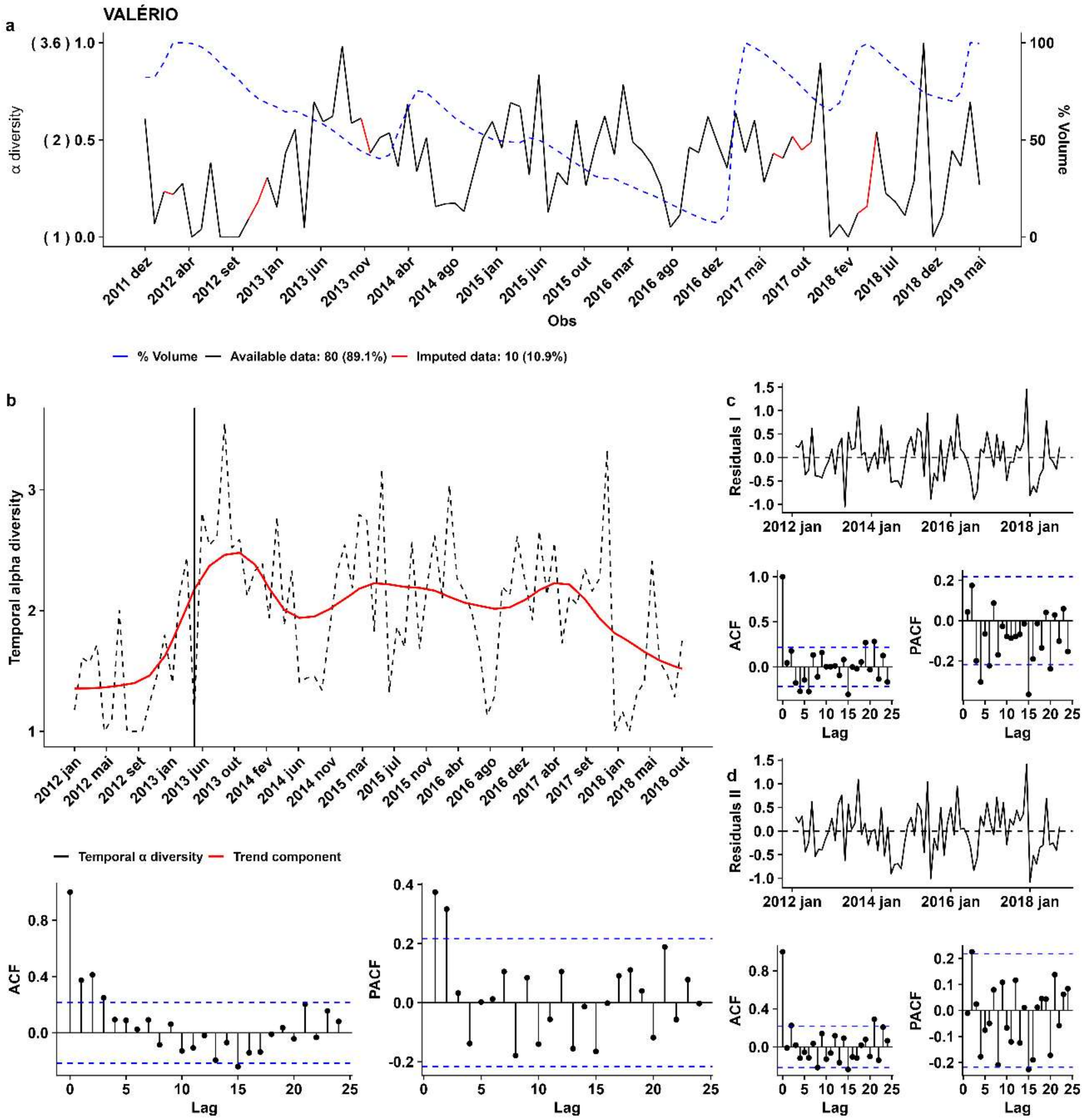


Figure 83 – Temporal α diversity | unmonitored - BOQUEIRÃO Residuals

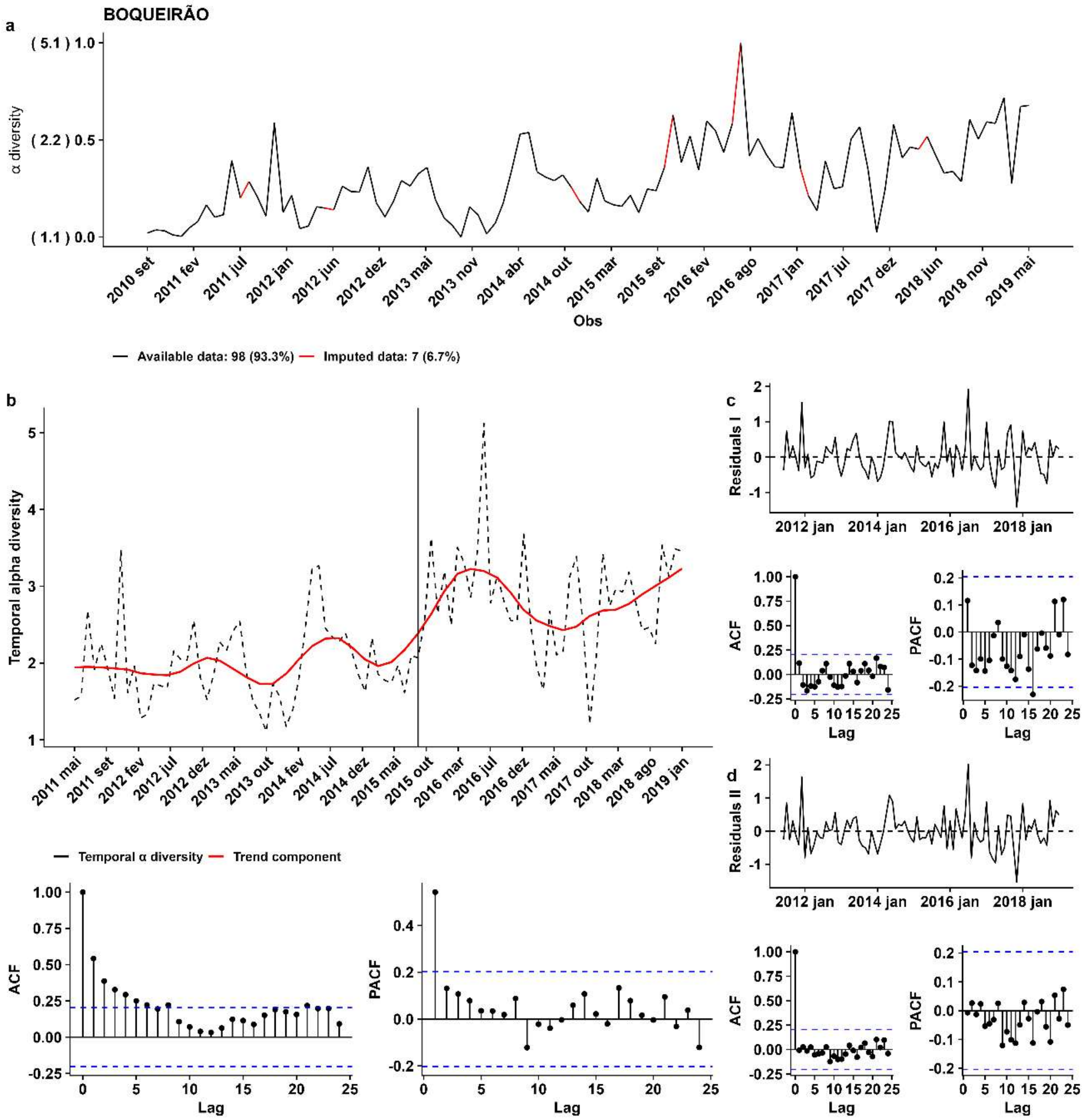


Figure 84 – Temporal α diversity | unmonitored - IPÚ MAZAGÃO Residuals

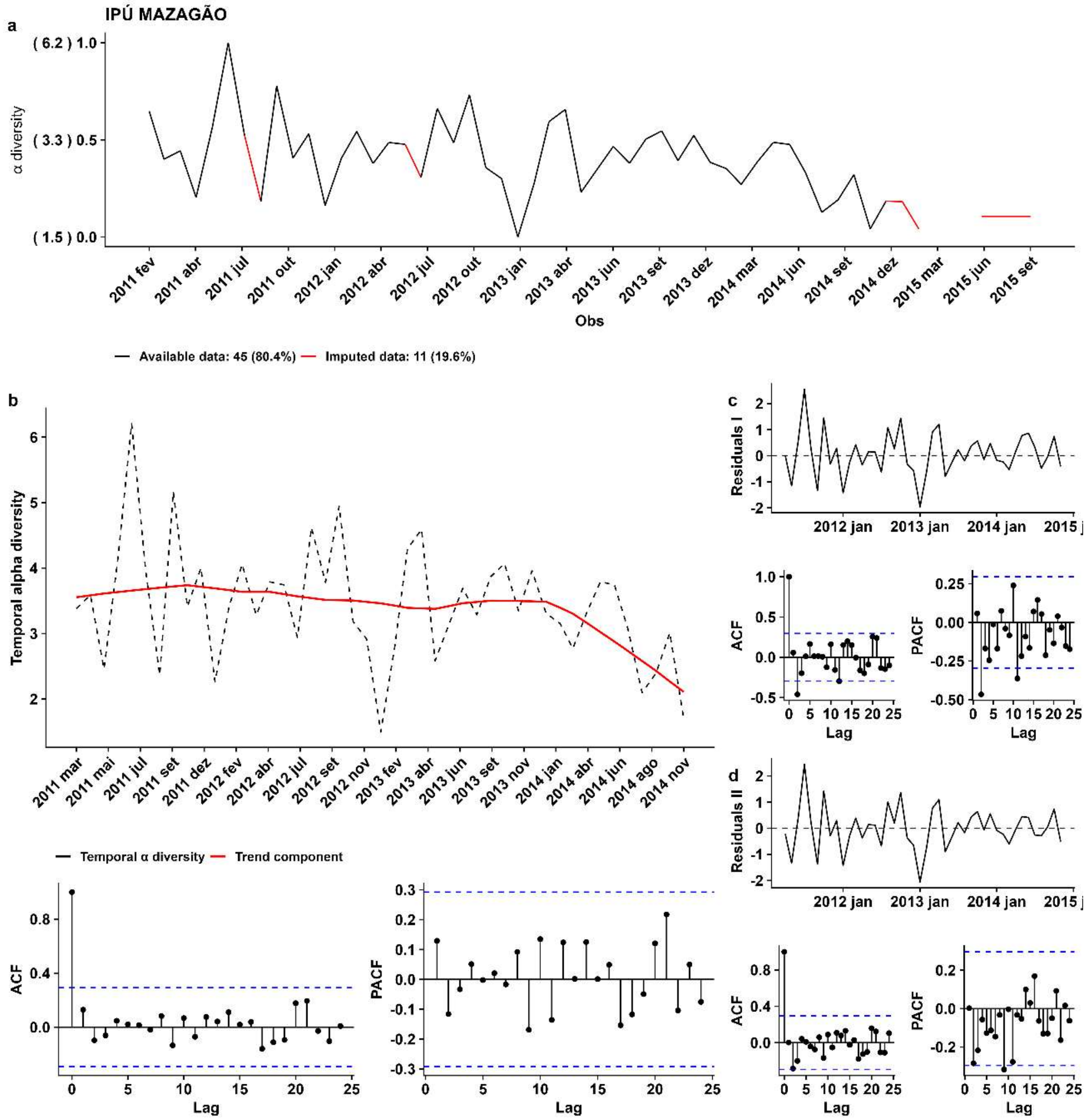


Figure 85 – Temporal α diversity | unmonitored - JOÃO JOSÉ DE CASTRO Residuals

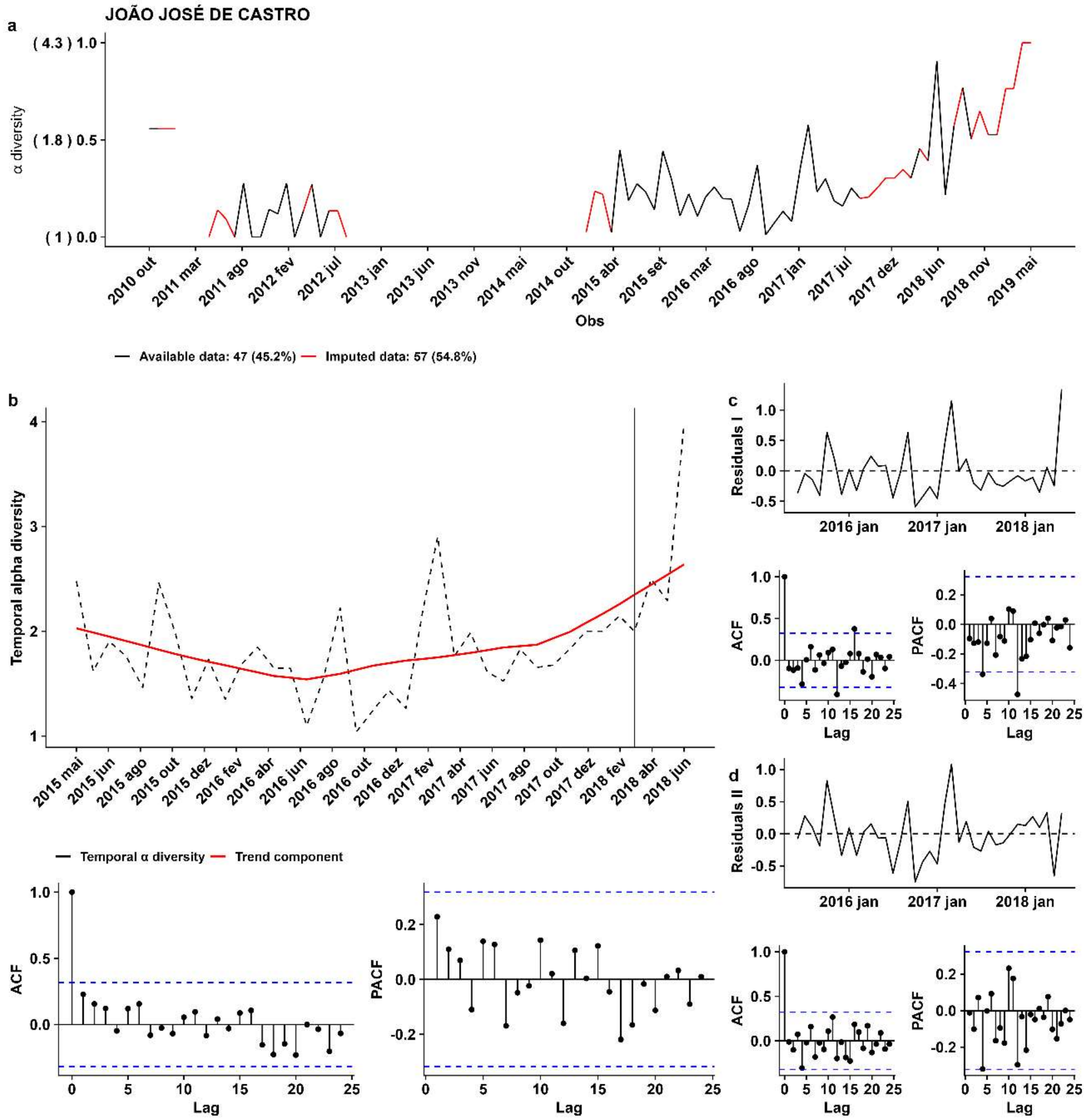


Figure 86 – Temporal α diversity | unmonitored - SÃO JOAQUIM Residuals

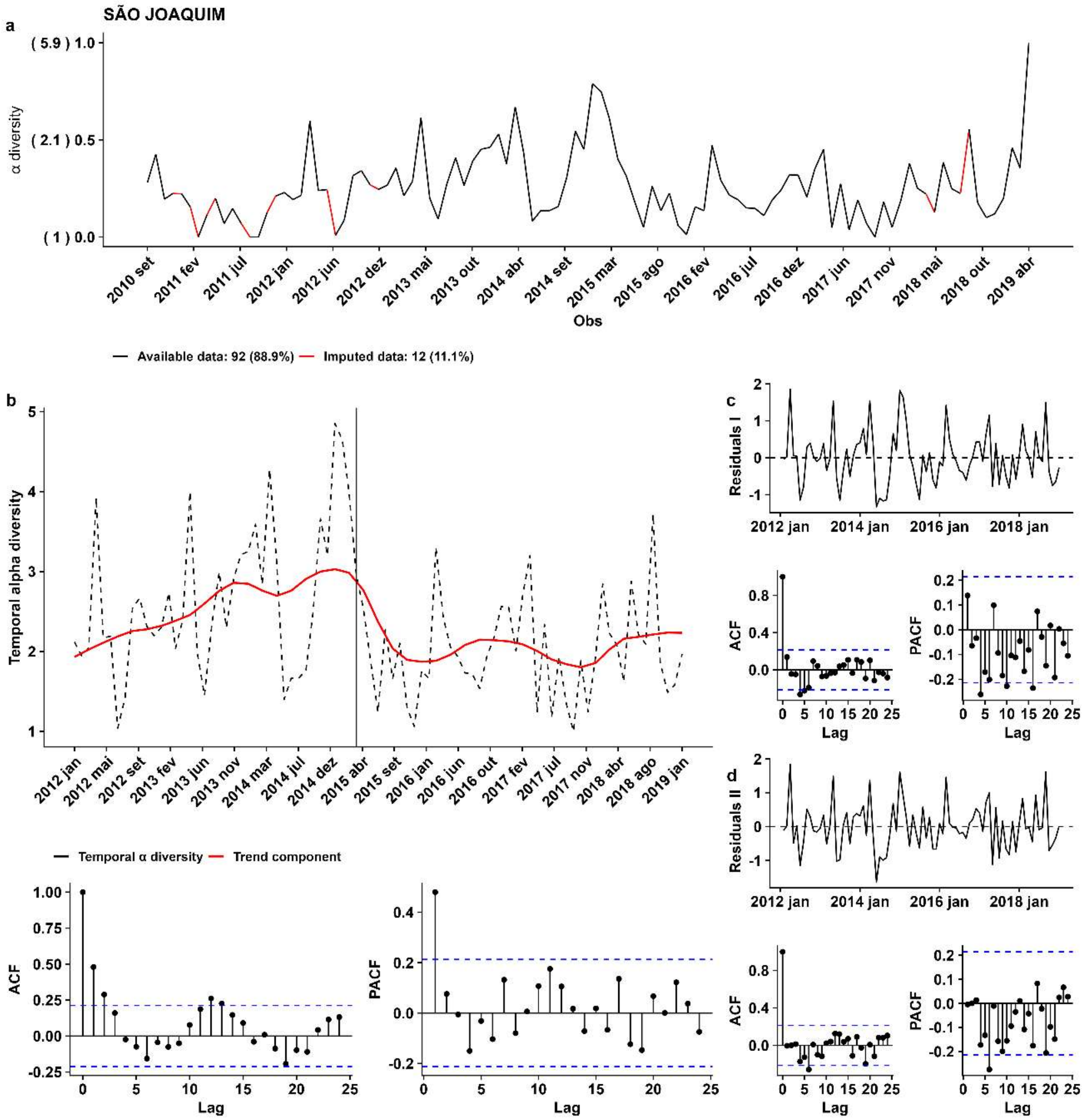
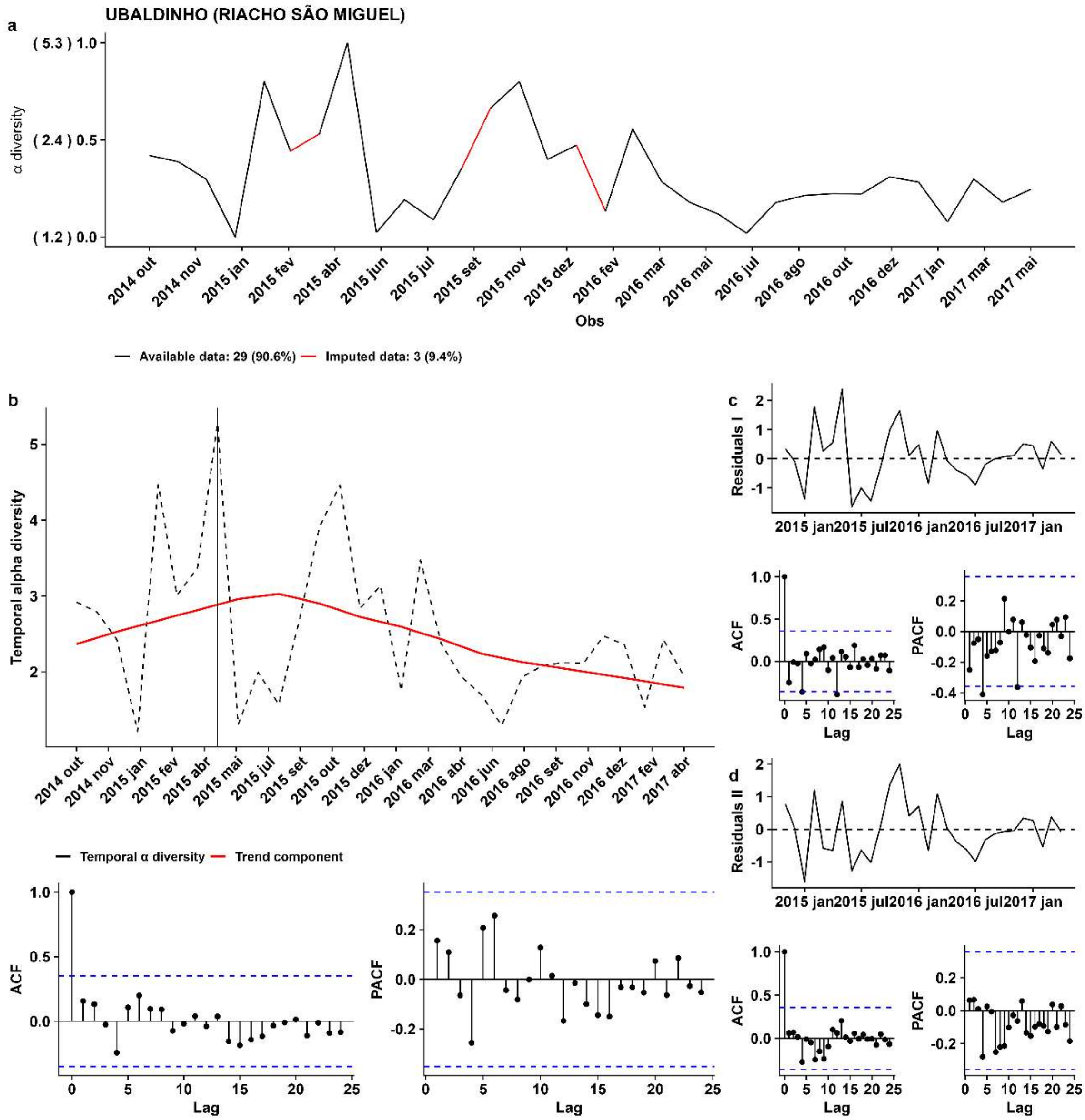


Figure 87 – Temporal α diversity | unmonitored - UBALDINHO (RIACHO SÃO MIGUEL) Residuals



Section 7 TEMPORAL BETA DIVERSITY

- In (a) the time series of beta diversity and %Volume is presented. The values were rescaled to improve the visualization of the series.
- In (b) the time series of beta diversity (dotted black line), the trend component obtained by decomposition using LOESS (red line), the breakpoint (vertical black line) and the ACF and PACF plots of the raw data are shown.
- In (c) the residuals obtained by LOESS decomposition are presented and in (d) the residuals after the ZA test adjustment, both with their respective ACF and PACF plots.

Figure 88 - Temporal β diversity | Acarape do meio Residuals

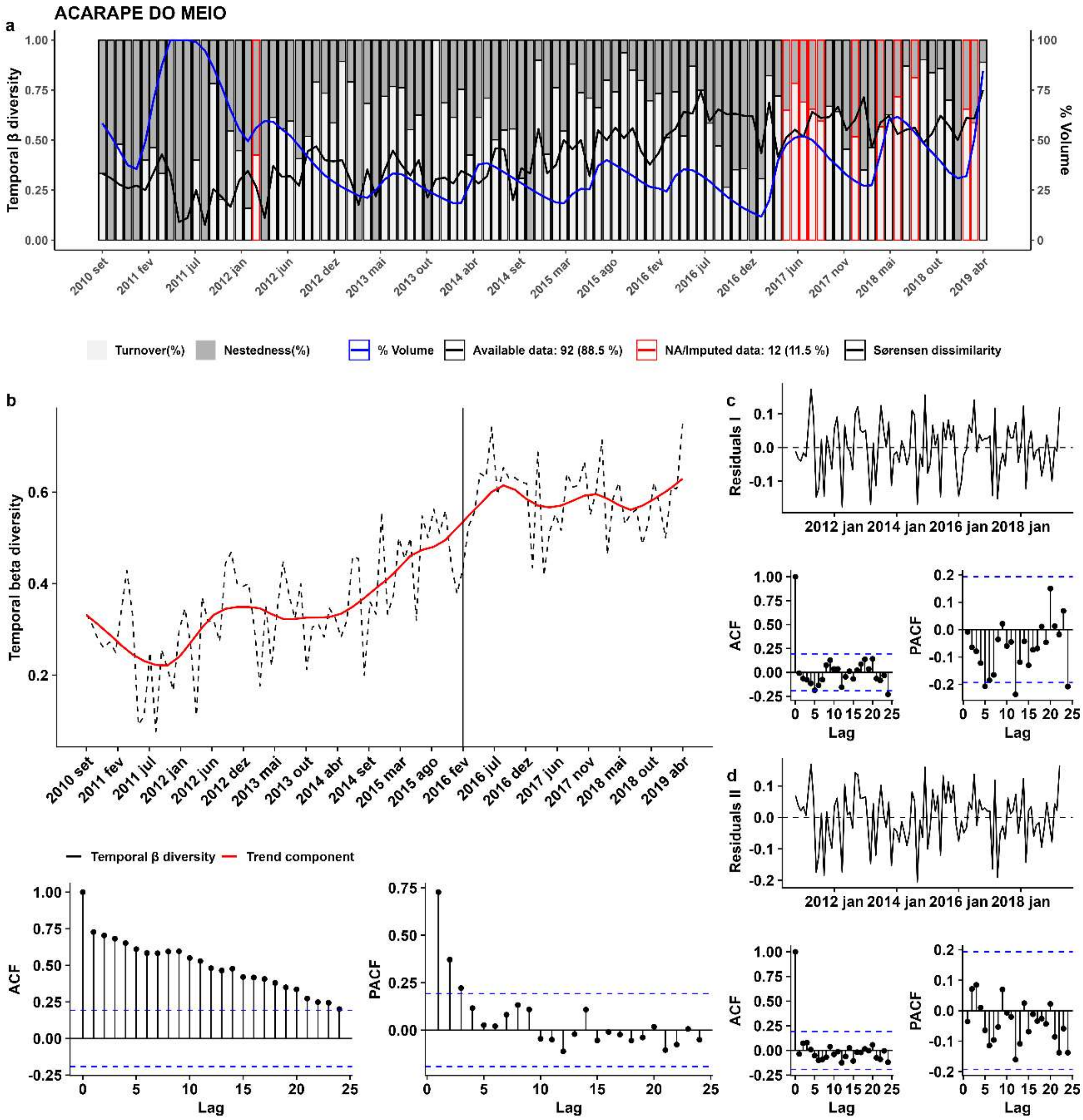


Figure 89 – Temporal β diversity | AMANARY Residuals

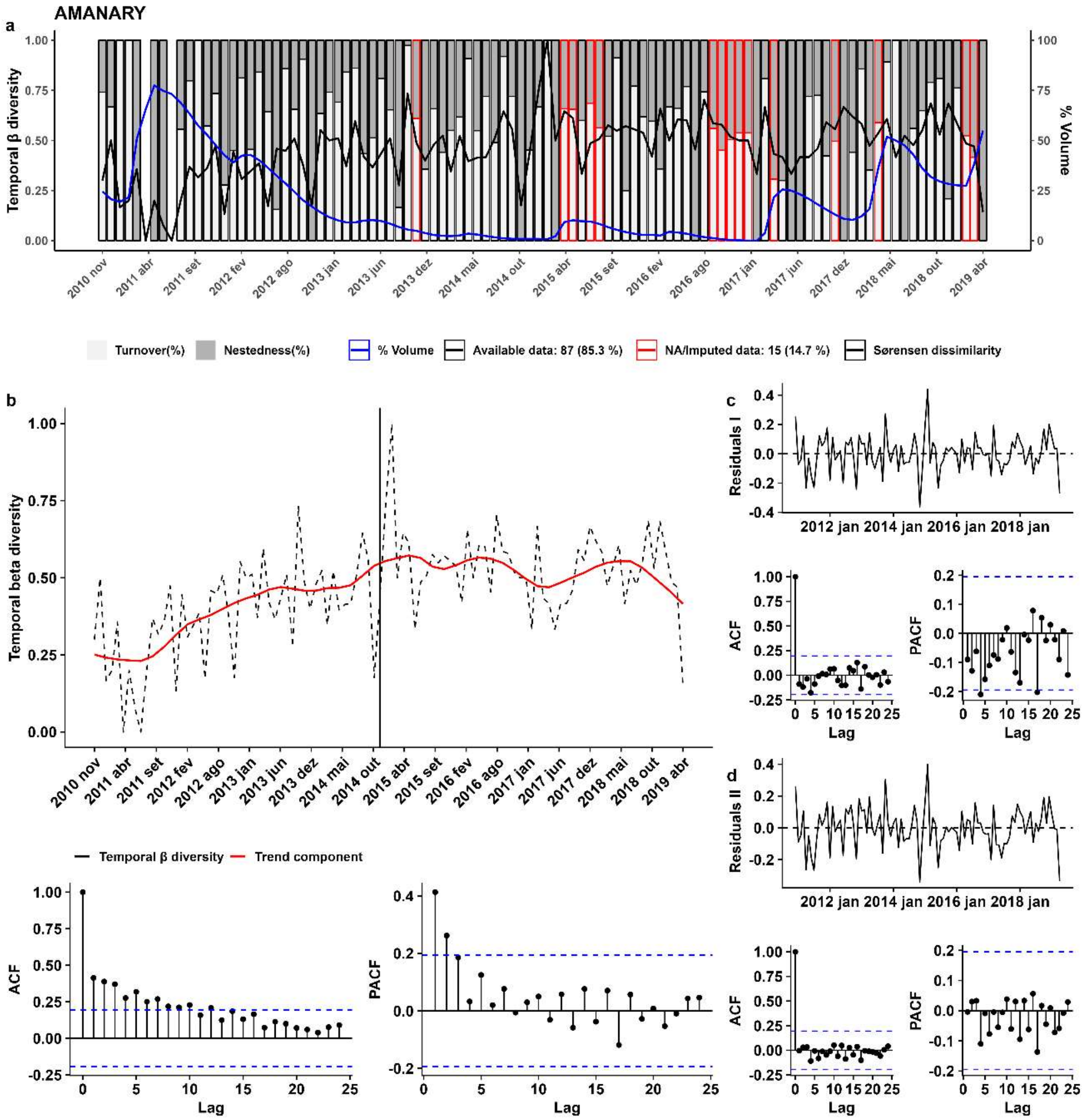


Figure 90 – Temporal β diversity | ANGICOS Residuals

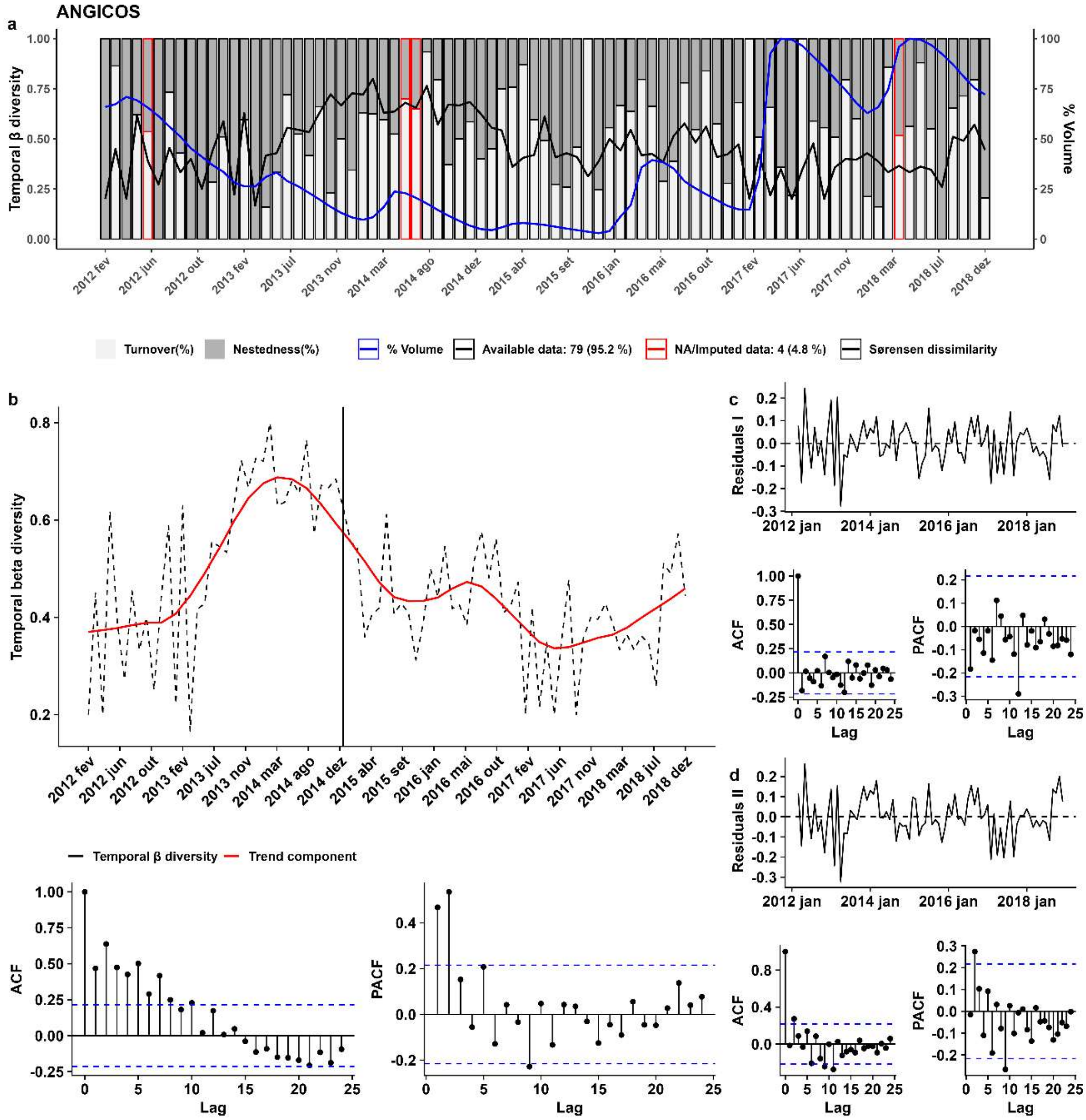


Figure 91 – Temporal β diversity | ARACOIABA Residuals

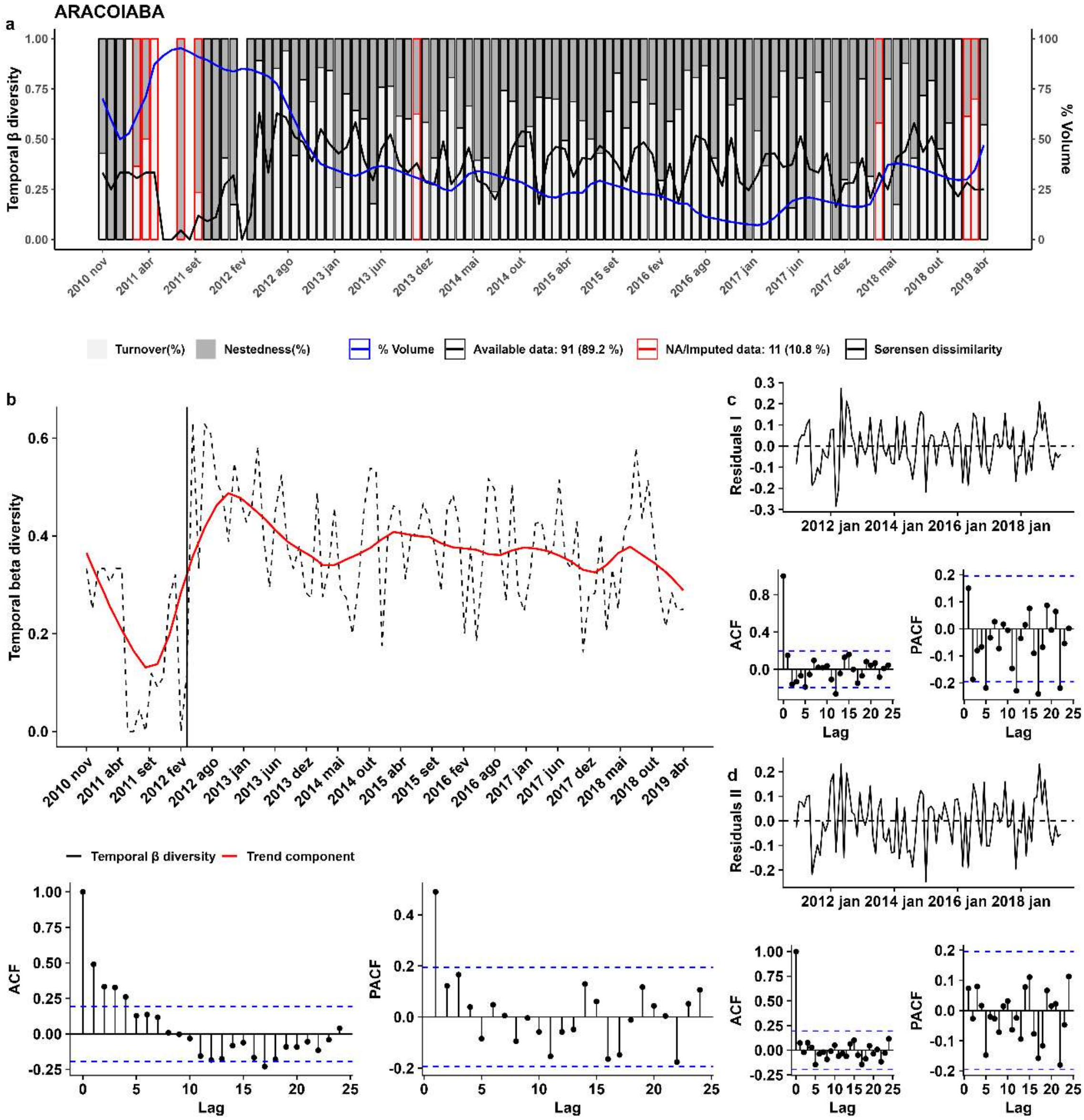


Figure 92 – Temporal β diversity | ARARAS Residuals

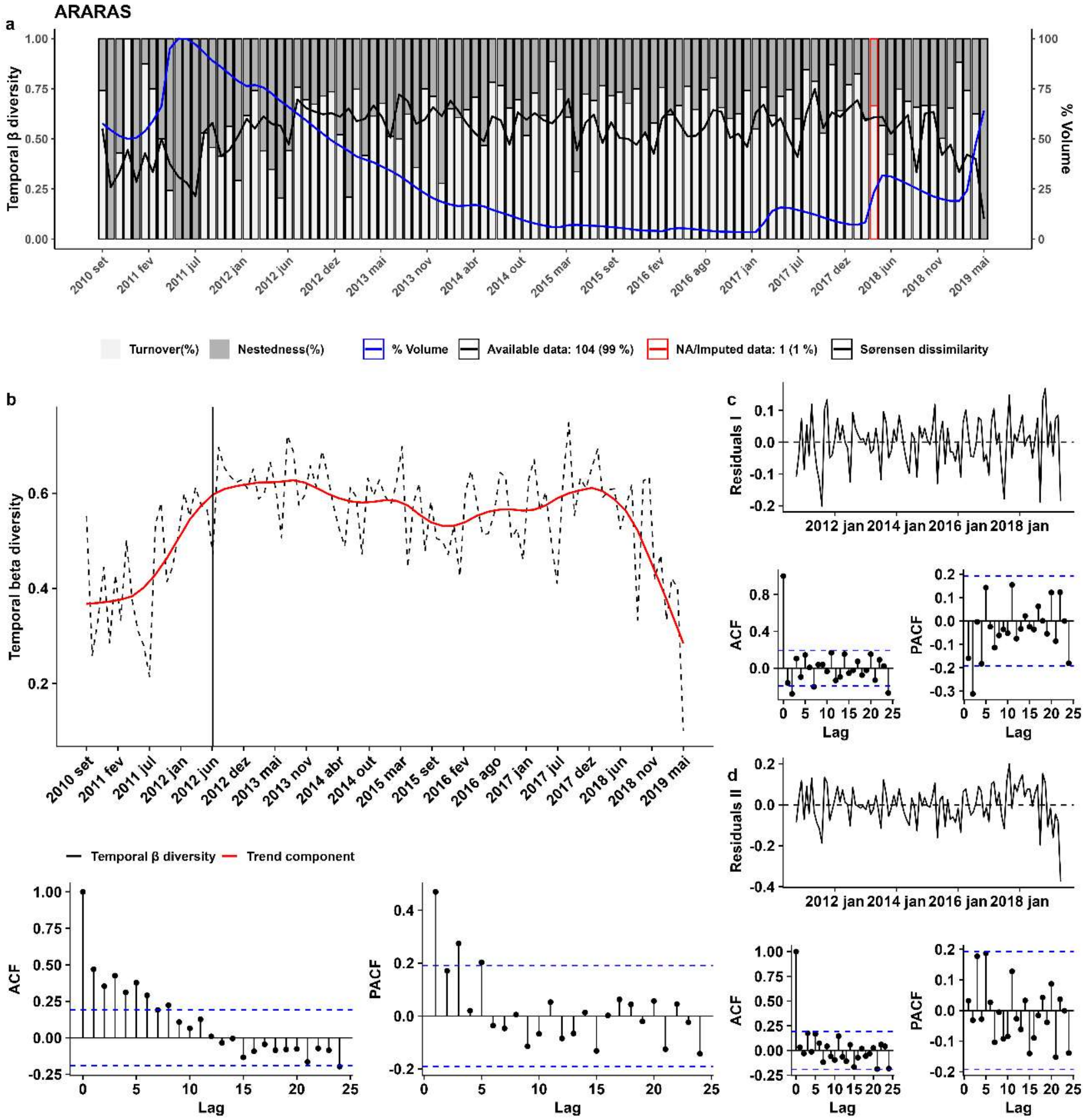


Figure 93 – Temporal β diversity | ARNEIROZ II Residuals

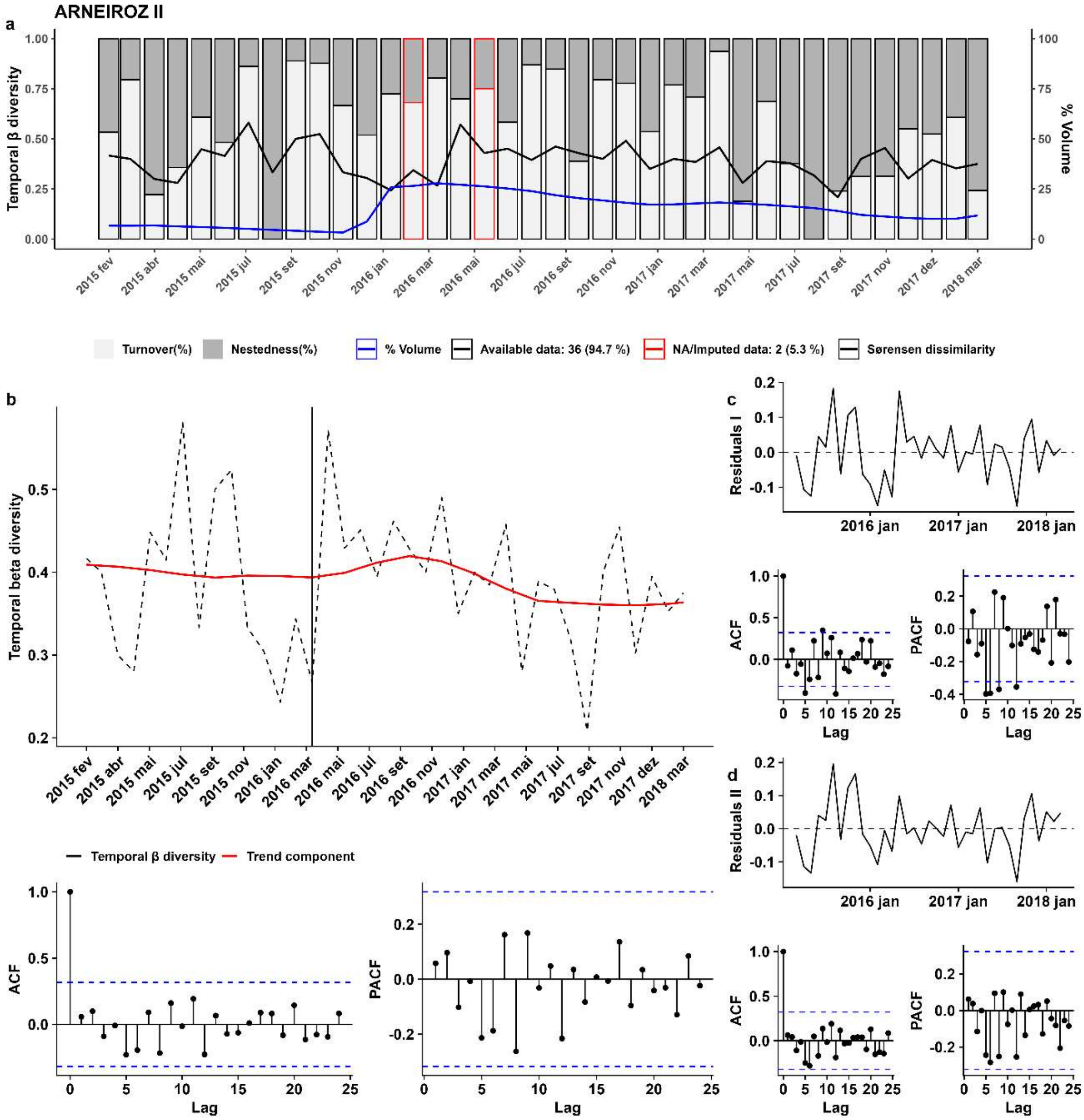


Figure 94 – Temporal β diversity | AYRES DE SOUZA Residuals

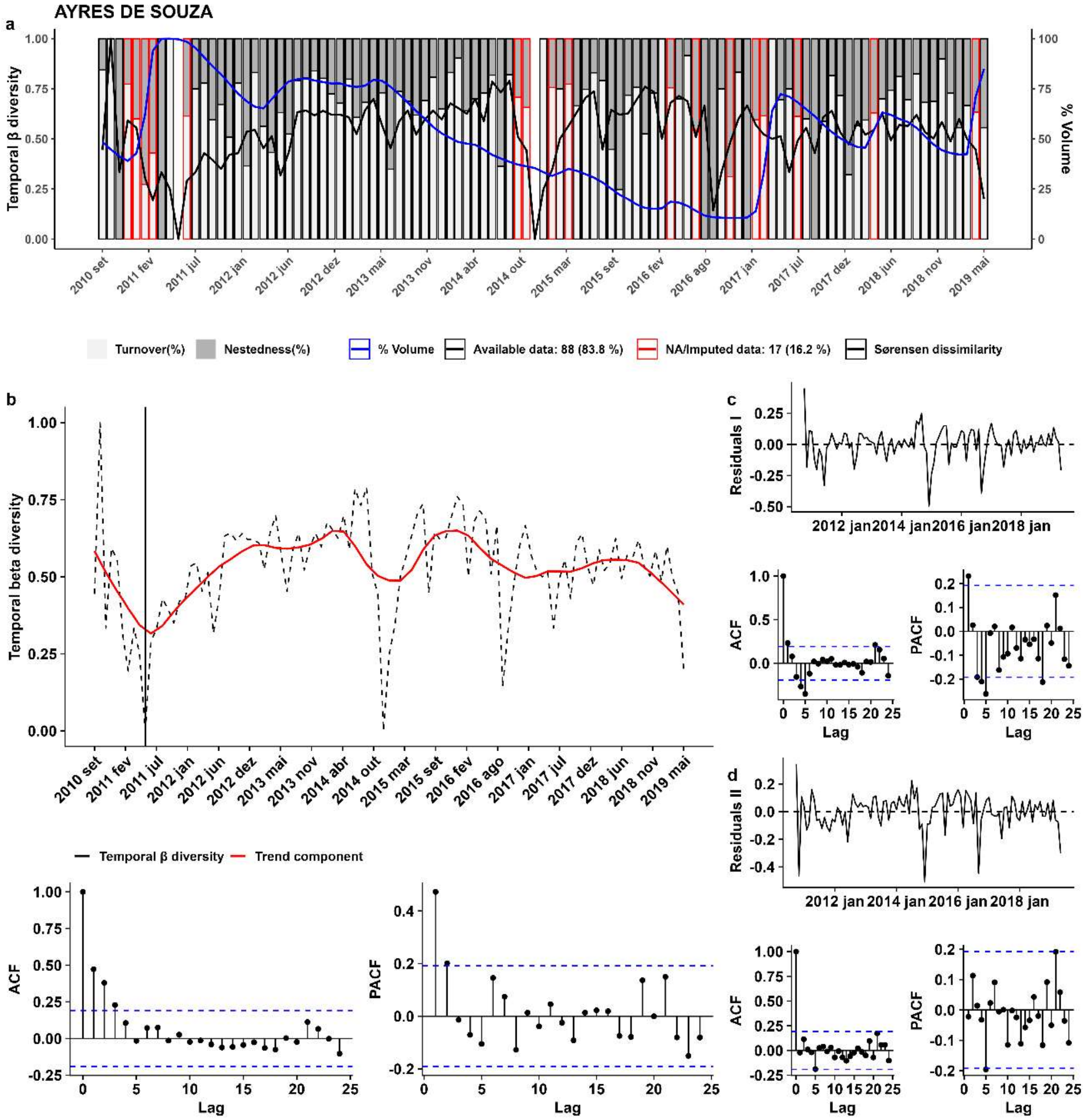


Figure 95 – Temporal β diversity | CANOAS Residuals