

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE ENERGIA E MEIO AMBIENTE  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA  
AMBIENTAL

MARIA GABRIELA LOUZADA MALFATTI

PREVISIBILIDADE DA PRECIPITAÇÃO EM BACIAS  
HIDROGRÁFICAS BRASILEIRAS NA ESCALA SUBSAZONAL E  
POTENCIAL APLICAÇÃO NA GESTÃO DOS RECURSOS  
HÍDRICOS

SÃO PAULO  
2022

MARIA GABRIELA LOUZADA Malfatti

PREVISIBILIDADE DA PRECIPITAÇÃO EM BACIAS  
HIDROGRÁFICAS BRASILEIRAS NA ESCALA SUBSAZONAL E  
POTENCIAL APLICAÇÃO NA GESTÃO DOS RECURSOS  
HÍDRICOS

Tese apresentada ao Programa de Pós-Graduação em  
Ciência Ambiental do Instituto de Energia e Ambiente  
da Universidade de São Paulo para a obtenção do  
título de Doutor em Ciências.

Orientador: Prof. Dr. Humberto Ribeiro da Rocha  
Coorientador: Prof. Dr. Pedro Leite da Silva Dias

Versão Original

SÃO PAULO  
2022

**AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.**

### FICHA CATALOGRÁFICA

Malfatti, Maria Gabriela Louzada.

Previsibilidade da precipitação em bacias hidrográficas brasileiras na escala subsazonal e potencial aplicação na gestão dos recursos hídricos. / Maria Gabriela Louzada Malfatti; orientador: Humberto Ribeiro da Rocha; Co-orientador: Pedro Leite da Silva Dias. -- São Paulo, 2022. 132 f.: il., 30 cm.

Tese (Doutorado em Ciência Ambiental) – Programa de Pós-Graduação em Ciência Ambiental – Instituto de Energia e Ambiente da Universidade de São Paulo.

1. Meteorologia. 2. Recursos Hídricos- aspectos socioeconômicos. 3. Casualidade. 4. Bacia hidrográfica. I. Título

Elaborado por Maria Penha da Silva Oliveira CRB-8/6961

# Agradecimentos

*Primeiramente, agradeço a Deus por me guiar e estar sempre comigo em todos os momentos de minha vida.*

*À minha mãe Silvia Helena, pelo apoio e amor incondicional. Ao meu pai David que mesmo não estando presente, tenho certeza que lá de cima torce muito pelas minhas conquistas. A todos os meus familiares que me incentivam, principalmente meu irmão Gabriel e a minha tia e madrinha Maria Cecília.*

*Também em especial agradeço profundamente aos meus orientadores Prof. Dr. Pedro Leite da Silva Dias e Prof. Dr. Humberto Ribeiro da Rocha, por cada tempo dedicado, por cada correção minuciosa e pela oportunidade de crescer profissionalmente.*

*Aos Prof. Dr. Ricardo de Camargo e Prof.<sup>a</sup> Dr.<sup>a</sup> Maria Assunção Faus da Silva Dias pela oportunidade de participar do projeto da RHAMA Consultoria Ambiental Ltda e todos aos meus colegas que participaram comigo, em especial ao Dr. Lucas Massaroppe e Dr.<sup>a</sup> Yania Souto.*

*Agradeço pelo apoio financeiro do CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e a parceria MASTER-RHAMA por parte deste projeto de doutorado.*

*Agradeço a todos os professores e técnicos do programa de Pós-graduação em Ciência Ambiental (PROCAM), ao Instituto de Energia e Ambiente - IEE/USP e Instituto de Astronomia, Geofísica e Ciências Atmosféricas - IAG/USP, em especial ao Laboratório de Meteorologia Aplicada a Sistemas de Tempo Regionais (MASTER-IAG/USP).*

*Aos membros da banca, por aceitarem o convite e participação da revisão deste trabalho.*

*E por fim, agradeço aos meus amigos que sempre estiveram me apoiando em minhas decisões acadêmicas e da vida.*



# Resumo

MALFATTI, M. G. L. **Modelagem para previsão de precipitação a partir de índices climáticos para planejamento e gestão dos recursos hídricos**. 2022. Tese (Doutorado em Ciência Ambiental) - Instituto de Energia e Ambiente, Universidade de São Paulo, São Paulo, 2022.

Diversas atividades socioeconômicas dependem da disponibilidade hídrica, sendo fundamental o uso de previsões meteorológicas confiáveis para o melhor planejamento de uso desse recurso. Devido à grande extensão territorial do Brasil diversos mecanismos podem atuar inibindo ou favorecendo a precipitação. Esses mecanismos podem ser de origem tropical e subtropical de diferentes escalas e com variações sazonais e subsazonais que produzem grandes reflexos no ciclo hidrológico em bacias brasileiras. Neste contexto, foram desenvolvidos modelos de Regressão Linear Múltipla (RLM) e aprendizado de máquina *Long short-term memory* (LSTM) de precipitação na escala subsazonal a sazonal (2 semanas a 2 meses), visando explorar a previsibilidade nas principais bacias hidrográficas brasileiras. Foram utilizados conjuntos de variáveis preditoras de índices climáticos e/ou precipitação prevista dos modelos ECMWF e NCEP do programa S2S (*Prediction Project Subseasonal-to-Seasonal*). Observou-se que a precipitação nas bacias brasileiras apresenta grande influência dos padrões Oscilação de Madden-Julian (MJO), representado pelos índices RMM1/RMM2, Oscilação Antártica (AAO) e Oscilação do Pacífico América do Sul (PSA), índice de Larga Escala da Monção da América do Sul (LISAM) e o índice Zona de Convergência do Atlântico Sul (ZCAS). Além disso, destaca a importância dos padrões associados ao Oceano Pacífico, El Niño-Oscilação Sul (ENSO) e Oscilação Decadal do Pacífico (PDO), na maioria das bacias brasileiras, exceto na porção noroeste do Brasil, na região Amazônica. As bacias do Norte, Nordeste e Centro-Oeste apresentam maior influência do Atlântico, com os padrões Oscilação Multidecadal do Atlântico (AMO) e Oscilações na Temperatura da Superfície do Mar (TSM) do Atlântico Tropical Norte (TNA) e Sul (TSA). Na avaliação dos modelos baseados em dados, a LSTM, apresentou melhor acurácia após a seleção dos preditores adequada pela busca exaustiva e ajuste dos hiperparâmetros, em quase todas as bacias, exceto na bacia do Rio Madeira e do Rio Iguaçu. Os principais índices selecionados nos períodos mais curtos de previsão (15 a 30 dias) foram o RMM1/RMM2 e LISAM/ZCAS, e os índices relacionados ao ENSO para os períodos mais longos (30 a 60 dias), resultado que evidencia a influência da MJO e do ENSO na precipitação no Brasil. Em adição, os padrões AAO e PSA são selecionados em defasagens longas e curtas em diversas bacias. Nas bacias do Norte e Nordeste são sugeridos índices do Atlântico (AMO, TNA e TSA) em todos os períodos de predição. As séries temporais de precipitação prevista dos modelos NCEP e ECMWF, ajudam a corrigir a previsão e reforçar o padrão da precipitação na bacia, visto que resultados dos MSE apontam melhor desempenho. Conclui-se que os resultados obtidos reforçam a importância da utilização dos índices climáticos e

precipitação prevista de modelos S2S no desenvolvimento de modelagem para previsão de precipitação na escala subsazonal-sazonal auxiliando no planejamento dos recursos hídricos, visto que a precipitação possui um tempo de resposta na vazão de rios variável entre 5 a 40 dias, dependendo da localização da bacia, antecipando seus impactos entre 60 à 100 dias.

**Palavras-chave:** causalidade, modelagem empírica, padrões climáticos e bacias hidrográficas brasileiras.

# Abstract

MALFATTI, M. G. L. **Precipitation forecasting modeling from climate indices for water resources planning and management**. 2022. Tese (Doutorado em Ciência Ambiental) - Instituto de Energia e Ambiente, Universidade de São Paulo, São Paulo, 2022.

Several socioeconomic activities depend on water availability and the use of reliable weather forecasts is essential for better planning for the use of this resource. Due to the large territorial extension of Brazil, several mechanisms can act to inhibit or favor precipitation. These mechanisms can be of tropical and subtropical origin of different scales and with seasonal and sub-seasonal variations that produce great reflexes in the hydrological cycle in Brazilian watersheds. In this context, Multiple Linear Regression (MLR) and machine learning *Long short-term memory* (LSTM) models of precipitation on a sub-seasonal to seasonal scale (2 weeks to 2 months) were developed, aiming to explore the predictability in the main Brazilian watersheds. Sets of predictor variables of climatic indices and/or predicted precipitation from the ECMWF and NCEP models of the S2S program (Prediction Project Subseasonal-to-Seasonal) were used. It was observed that precipitation in Brazilian basins is strongly influenced by the patterns of the Madden-Julian Oscillation (MJO), represented by the indices RMM1/RMM2, Antarctic Oscillation (AAO) and South America Pacific Oscillation (PSA), South America's Large Scale Monsoon Index (LISAM) and the South Atlantic Convergence Zone (ZCAS) index. In addition, it highlights the importance of patterns associated with the Pacific Ocean, El Niño-Southern Oscillation (ENSO) and Pacific Decadal Oscillation (PDO), in most Brazilian basins, except in the northwest portion of Brazil, in the Amazon region. The North, Northeast and Central-West basins are more influenced by the Atlantic, with the Atlantic Multidecadal Oscillation (AMO) and North Tropical Atlantic Ocean (TNA) and South Tropical Atlantic Ocean Surface Temperature Oscillations (SST) patterns. In the evaluation of the data-based models, the LSTM showed better accuracy after the adequate selection of predictors by the exhaustive search and adjustment of hyperparameters, in almost all basins, except for the Madeira River and Iguaçu River watersheds. The main indices selected in the shorter forecast periods (15 to 30 days) were the RMM1/RMM2 and LISAM/ZCAS, and the ENSO-related indices for the longer periods (30 to 60 days), a result that evidences the influence of the MJO and ENSO on precipitation in Brazil. In addition, AAO and PSA patterns are selected in long and short lags in various basins. In the North and Northeast basins, Atlantic indices (AMO, TNA and TSA) are suggested for all prediction periods. The predicted rainfall time series from the NCEP and ECMWF models help to correct the forecast and reinforce the rainfall pattern in the basin, as MSE results indicate better performance. It is concluded that the results obtained reinforce the importance of using climatic indices and predicted precipitation of S2S models in the development of modeling for precipitation



prediction on the sub-seasonal-seasonal scale, helping in the planning of water resources, since precipitation has a response time in river flow varying between 5 to 40 days, depending on the location of the watersheds, anticipating its impacts between 60 to 100 days.

**Keywords:** causality, empirical modeling, weather patterns and Brazilian watersheds

# Sumário

<b>Lista de Abreviaturas</b>	<b>xi</b>
<b>Lista de Figuras</b>	<b>xiii</b>
<b>Lista de Tabelas</b>	<b>xvii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Objetivos . . . . .	1
1.2 Justificativa da Originalidade do trabalho . . . . .	2
1.3 Hipóteses . . . . .	2
1.4 Organização do Trabalho . . . . .	3
<b>2 Revisão Bibliográfica</b>	<b>5</b>
2.1 Modelagem e Previsão de Precipitação . . . . .	5
2.1.1 Modelos Dinâmicos . . . . .	6
2.1.2 Modelos Empíricos . . . . .	6
2.2 Padrões de Variabilidade Climática . . . . .	10
2.2.1 Zona de Convergência do Atlântico Sul (ZCAS) . . . . .	11
2.2.2 Oscilação de Madden-Julian (MJO) . . . . .	12
2.2.3 El Niño Oscilação Sul (ENSO) . . . . .	14
2.2.4 Oscilação Decadal do Pacífico (PDO) . . . . .	16
2.2.5 Anomalias de TSM no Atlântico Tropical Norte (TNA) e Sul (TSA) . . . . .	17
2.2.6 Oscilação Multidecadal do Atlântico (AMO) . . . . .	18
2.2.7 Modo Meridional do Atlântico (AMM) . . . . .	19
2.2.8 Oscilação do Atlântico Norte (NAO) . . . . .	20
2.2.9 Oscilação Pacífico – América do Norte (PNA) e do Sul (PSA) . . . . .	20
2.2.10 Oscilação Ártica (AO) e Antártica (AAO) . . . . .	21
2.2.11 Oscilação Quase-Bienal (QBO) . . . . .	22
2.2.12 Fluxo Solar (SF) . . . . .	22
<b>3 Padrões de Variabilidade Climática</b>	<b>25</b>
3.1 Dados . . . . .	25
3.2 Metodologia . . . . .	26
3.2.1 Desenvolvimento dos Índices Climáticos . . . . .	26
3.2.2 Variabilidade Climática dos Padrões . . . . .	28

3.3	Resultados e Discussão . . . . .	29
3.4	Conclusões . . . . .	36
<b>4</b>	<b>Identificação dos Padrões de Variabilidade Climática nas Principais Bacias Brasileiras</b>	<b>39</b>
4.1	Área de Estudo . . . . .	39
4.2	Dados e Metodologia . . . . .	41
4.3	Resultados e Discussão . . . . .	42
4.4	Conclusão . . . . .	52
<b>5</b>	<b>Desenvolvimento de Modelos Preditivos de Precipitação nas Principais Bacias Brasileiras</b>	<b>55</b>
5.1	Dados e Metodologia . . . . .	55
5.2	Resultados e Discussão . . . . .	57
5.2.1	Seleção dos Preditores Climáticos - Método <i>Stepwise</i> . . . . .	57
5.2.2	Seleção dos Preditores Climáticos - Busca Exaustiva . . . . .	58
5.2.3	Modelagem dos Modelos de RLM e LSTM . . . . .	60
5.3	Conclusões . . . . .	62
<b>6</b>	<b>Modelos com Predição S2S</b>	<b>65</b>
6.1	Dados e Metodologia . . . . .	65
6.2	Resultados e Discussão . . . . .	67
6.3	Conclusões . . . . .	69
<b>7</b>	<b>Conclusões</b>	<b>71</b>
7.1	Considerações Finais . . . . .	73
7.2	Sugestões para Pesquisas Futuras . . . . .	73
<b>A</b>	<b>Padronização de séries temporais</b>	<b>75</b>
<b>B</b>	<b>Transformada de Ondeletas</b>	<b>77</b>
<b>C</b>	<b>Correlação de Pearson</b>	<b>79</b>
<b>D</b>	<b>Análise de Componentes Principais</b>	<b>81</b>
<b>E</b>	<b>Causalidade (PDC e knPDC)</b>	<b>83</b>
<b>F</b>	<b>Método <i>Stepwise</i></b>	<b>85</b>
<b>G</b>	<b>Modelo de Regressão Linear Múltipla</b>	<b>87</b>
<b>H</b>	<b><i>Long short-term memory</i> (LSTM)</b>	<b>89</b>
<b>I</b>	<b>Ilustrações Adicionais - Capítulo 3</b>	<b>91</b>

<b>J Paper: Predictability of Sub-seasonal Precipitation in Brazilian Watersheds based on Teleconnection Patterns and South American Monsoon Modes</b>	<b>93</b>
<b>Referências Bibliográficas</b>	<b>119</b>



# Lista de Abreviaturas

AAO	Oscilação Antártica ( <i>Antarctic Oscillation</i> )
AMM	Modo Meridional do Atlântico ( <i>Atlantic Meridional Mode</i> )
AMO	Oscilação Multidecadal do Atlântico ( <i>Atlantic Multidecadal Oscillation</i> )
AO	Oscilação Ártica ( <i>Arctic Oscillation</i> )
AOGCM	Modelos de Circulação Geral da Atmosfera-Oceano ( <i>Atmosphere-Ocean General Circulation Model</i> )
ARMA	Modelo Autorregressivo com média móvel
ARIMA	Modelo Autorregressivo Integrado com média móvel
ASAS	Anticiclone Subtropical do Atlântico Sul
BOM	<i>Bureau of Meteorology- Australian Government</i>
CC	Coefficiente de Correlação de Pearson
CHIRPS	<i>Climate Hazards Group InfraRed Precipitation with Station Data</i>
CNES	<i>Centre National d'Études Spatiales</i>
CPTEC	Centro de Previsão de Tempo e Estudos Climáticos
DL	<i>Depp learning</i>
ENSO	El Niño Oscilação Sul ( <i>El Niño Southern Oscillation</i> )
ECA&D	<i>European Climate Assessment &amp; Dataset project</i>
ECMWF	<i>European Centre for Medium-Range Weather Forecasts</i>
EOF	<i>Empirical Orthogonal Function</i>
EOFc	<i>Combined Empirical Orthogonal Function</i>
EUMETSAT	Organização Europeia para a Exploração de Satélites Meteorológico
GCP	<i>Global Precipitation Climatology Project</i>
GFS	<i>Global Forecast System Model</i>
GPM	<i>The Global Precipitation Measurement Mission</i>
GTA	Gradiente Meridional de TSM no Atlântico Tropical
INPE	Instituto Nacional de Pesquisas Espaciais
ISRO	Organização de Pesquisa Espacial Indiana
JAMSTEC	<i>Japan Agency for Marine-Earth Science and Technology</i>
JAXA	<i>Japan Aerospace Exploration Agency</i>
JBN	Jato de Baixos Níveis
knPDC	Coerência Parcial Direcionada não-linear kernelizada ( <i>kernel-nonlinear Partial directed Coherence</i> )
LISAM	Índice de Larga Escala da Monção da América do Sul ( <i>Large-scale index for South America Monsoon</i> )

LSTM	<i>Long short-term memory</i>
MEI	Índice Multivariado El Niño Oscilação Sul
MJO	Oscilação de Madden-Julian ( <i>Oscillation Madden-Julian</i> )
MLP	<i>Multilayer Perceptron</i>
MME	Departamento de Informações e Estudos Energéticos de Minas e Energia
NAM	Modo Anular Norte
NAO	Oscilação do Atlântico Norte ( <i>North Atlantic Oscillation</i> )
NASA	<i>National Aeronautics and Space Administration</i>
NCAR	<i>National Center for Atmospheric Research</i>
NCEP	<i>National Centers for Environmental Prediction</i>
NOAA	<i>National Oceanic and Atmospheric Administration</i>
ONS	Operador Nacional do Sistema Elétrico
ORL	Radiação de Onda Longa ( <i>Outgoing longwave Radiation</i> )
PDO	Oscilação Decadal do Pacífico ( <i>Pacific Decadal Oscillation</i> )
PC	Componente Principal ( <i>Principal Component</i> )
PDC	Coerência Parcial Direcionada ( <i>Partial Directed Coherence</i> )
PNA	Oscilação do Pacífico América do Norte ( <i>Pacific North American Pattern</i> )
PSA	Oscilação do Pacífico América do Sul ( <i>Pacific South American Pattern</i> )
QBO	Oscilação Quase-Bienal ( <i>Oscillation Quasi-Biennial</i> )
RCM	Modelos Climáticos Regionais ( <i>Regional Climate Model</i> )
ReLU	Função Retificadora Linear ( <i>Rectified Linear Function</i> )
RMM1/RMM2	<i>Real-time Multivariate MJO</i>
REQM	Raiz do Erro Quadrático Médio
RLM	Regressão Linear Múltipla
RNAs	Redes Neurais Artificiais
RNNs	Redes Neurais Recorrente
SAM	Modo Anular Sul
S2S	Programa de Previsão Subsazonal ( <i>Sub-Seasonal to Seasonal</i> )
SDG	Descida do Gradiente Estocástico
SIN	Sistema Interligado Nacional
SMAS	Sistema de Monção da América do Sul
SOI	Índice Oscilação Sul ( <i>Southern Oscillation Index</i> )
SSN	Número de Manchas Solares ( <i>Sunspot Number</i> )
TNA	Oscilações de TSM no Atlântico Tropical Norte
TNI	Índice Trans-Niño
TRMM	<i>Tropical Rain-fall Measuring Mission</i>
TSA	Oscilações de TSM no Atlântico Tropical Sul
TSM	Temperatura da Superfície do Mar
UHE	Usinas Hidrelétricas
VAR	<i>Vector Autoregressive</i>
ZCAS	Zona de Convergência do Atlântico Sul
ZCIT	Zona de Convergência Intertropical

# Lista de Figuras

2.1	Comparação entre a arquitetura de uma Rede Neural Recorrente (em inglês, Recurrent Neural Network) e uma Rede Neural Feed-forward (em inglês, Feed-Forward Neural Network). . . . .	7
2.2	O vetor de entrada com $n$ variáveis $(x_1, x_2, \dots, x_n)$ que são multiplicados por um vetor de peso do mesmo tamanho e adicionado a um bias $b$ . O somatório deste resultado é passado para a função de ativação $f(\cdot)$ que por sua vez gera uma saída $y$ . . . . .	8
2.3	Arquitetura de um Multilayer Perceptron com duas camadas ocultas e uma camada de saída $(y_1, y_2$ e $y_3)$ . O fluxo de sinal através da rede progride da esquerda para a direita e numa base de camada por camada. . . . .	8
2.4	Ilustração das direções do fluxo de sinal em Multilayer Perceptron e erro backpropagation. . . . .	9
2.5	Funções para treinamento de um perceptron. a) função de ativação sigmoide, $f(\mathbf{x})$ e b) função retificadora linear (em inglês, rectified linear function, ReLU), $\mathbf{y}$ . . . . .	10
2.6	Correlação entre a precipitação a) índice LISAM (em inglês, Large-scale index for South America Monsoon) e b) índice ZCAS (em inglês, South Atlantic Convergence Zone, SACZ) (Fonte: <a href="#">Silva e Carvalho [2007]</a> ). . . . .	13
2.7	Esquema das variações em espaço e tempo da perturbação relacionada com a MJO. Os números de 1 a 5 referem - se aos tempos em que a Ilha de Canton apresenta baixa e alta pressão, respectivamente, e os demais números a tempos intermediários (Fonte: <a href="#">Madden e Julian [1972]</a> ). . . . .	14
2.8	Fases do El Niño Oscilação Sul (ENSO), El Niño e La Niña, com indicação da anomalia de TSM, vermelho e azul escuro indicam anomalias de +0,8 e -0,6 de temperatura, respectivamente, e os vetores indicam a direção e magnitude dos ventos alísios (Fonte: NOAA Climate.gov). . . . .	15
2.9	Regiões do Pacífico Equatorial/Tropical de influência do El Niño Oscilação Sul (ENSO), Niño 1+2, Niño 3, Niño 3.4 e Niño 4 (Fonte: NOAA Climate.gov). . . . .	16
2.10	Fases da Oscilação Decadal do Pacífico (PDO), fase quente (+) e fase fria (-), com indicação da anomalia de TSM, vermelho e azul escuro indicam anomalias de +0,8 e -0,6 de temperatura, respectivamente, e os vetores indicam a direção e magnitude dos ventos alísios (Fonte: NOAA Climate.gov). . . . .	17
2.11	Localização dos Índices TNA (Atlântico Tropical Norte) e TSA (Atlântico Tropical Sul). . . . .	18



2.12 Fases da Oscilação Multidecadal do Atlântico (AMO), fase quente (+) e fase fria (-), com indicação da anomalia de TSM, laranja e azul escuro indicam anomalias de +1,0 e -1,0 de temperatura, respectivamente (Fonte: NOAA Climate.gov). . . . . 19

3.1 Índice América do Sul-Pacífico (PSA) anomalias de altura geopotencial em 500hPa em reanálise do ERA-Interim, segundo modo PSA1 (EOF2 - 13% da variância explicada) e terceiro modo PSA2 (EOF3 - 5% da variância explicada) no período de janeiro de 1981 a julho de 2018 em pêntadas. Os valores em vermelho representam o índice PSA positivo e em azul negativo. . . . . 27

3.2 Média dos valores do segundo e terceiro modo no período de 1981 a 2018, em azul indicam anomalias negativas de altura geopotencial enquanto os valores positivos indicam anomalias positivas (figuras obtidas com os autovetores da EOF2 e EOF3). 28

3.3 Espectro de Potência e Global da análise de wavelet da série temporal dos índices (Morlet) a) Niño 3.4; b) PDO; c) AMO; d) AMM; e) AAO; f) AO. O espectro de potência é delimitado pelo cone de influência, o contorno preto indica o setor do espectro de potência onde o nível de confiança de 95% é satisfeito. O espectro de potência global indica o nível de potência satisfeito para toda a série temporal, ou seja, acima da linha vermelha tracejada. . . . . 30

3.4 Espectro de Potência e Global da análise de wavelet da série temporal dos índices g) NAO; h) PNA; i) QBO; j) GTA; k) SPI; l) PSA1. O espectro de potência é delimitado pelo cone de influência, o contorno preto indica o setor do espectro de potência onde o nível de confiança de 95% é satisfeito. O espectro de potência global indica o nível de potência satisfeito para toda a série temporal, ou seja, acima da linha vermelha tracejada. . . . . 31

3.5 Espectro de Potência e Global da análise de wavelet da série temporal dos índices m) PSA2; n) LISAM; o) ZCAS; p) RMM1 e q) RMM2. O espectro de potência é delimitado pelo cone de influência, o contorno preto indica o setor do espectro de potência onde o nível de confiança de 95% é satisfeito. O espectro de potência global indica o nível de potência satisfeito para toda a série temporal, ou seja, acima da linha vermelha tracejada. . . . . 32

3.6 O gráfico preto representa a (pseudo-) densidade espectral da série, em série em dB; o gráfico vermelho representa os valores PDC / knPDC estatisticamente significativos; a linha preta tracejada representa a aproximação do limiar de Patnaik *Baccalá et al. [2013]*; a linha verde os valores PDC / knPDC estatisticamente não significativos. Portanto, usando o nível de significância de 1%, os números representam, respectivamente: (a) PDC para o conjunto de séries temporais (AMO, Niño 3.4, PDO e GTA) usando um modelo autorregressivo de ordem  $p = 48$ , (b) knPDC para o conjunto de séries temporais (AMO, Niño 3.4, PDO e GTA), usando o kernel polinomial  $[k(x, y) = (x.y)^2]$  e o modelo kernel-autoregressivo de ordem  $p = 24$ . . . . . 36

3.7 O gráfico preto representa a (pseudo-) densidade espectral da série, em série em dB; o gráfico vermelho representa os valores PDC / knPDC estatisticamente significativos; a linha preta tracejada representa a aproximação do limiar de Patnaik (Baccalá et al., 2013); a linha verde os valores PDC / knPDC estatisticamente não significativos. Portanto, usando o nível de significância de 1%, os números representam, respectivamente: (a) knPDC para o conjunto de séries temporais (AAO, PSA1, PSA2 e PDO) e usando um modelo autoregressivo de ordem  $p = 48$ , (b) knPDC para o conjunto de séries temporais (AAO, PSA1, PSA2 e PDO) e usando o kernel polinomial  $[k(x, y) = (x.y)^2]$ do um modelo kernel-autoregressivo de ordem  $p = 24$ . . . . . 37

4.1 Localização das principais bacias brasileiras e os pontos de vazão nas UHE. . . . . 40

4.2 Variância explicada dos seis primeiros componentes principais (CPs) e o erro amostral calculado pelo método de North et al. [1982]. . . . . 43

4.3 Séries temporais das seis primeiras componentes principais da precipitação. . . . . 44

4.4 Espectro de Potência e Global da análise de ondeleta da série temporal dos seis primeiros componentes de precipitação a) PC1; b) PC2; c) PC3; d) PC4; e) PC5; f) PC6. O espectro de potência é delimitado pelo cone de influência, o contorno preto indica o setor do espectro de potência onde o nível de confiança de 90% é satisfeito. O espectro de potência global indica o nível de potência satisfeito para toda a série temporal, ou seja, acima da linha vermelha tracejada. . . . . 45

4.5 Correlação defasada em pêntadas (lag -12 a lag -1, dos tons mais claros aos mais escuros) entre os modos de precipitação ((a) PC1, a (f) PC6) e os índices climáticos, as linhas pontilhadas representa o limite de significância a partir do Teste t de Student (-0,07 e +0,07). . . . . 47

4.6 O gráfico preto representa a (pseudo-) densidade espectral da série, em série em dB (PC1 a PC6); o gráfico vermelho representa os valores PDC estatisticamente significativos; a linha preta tracejada representa a aproximação do limiar de Patnaik Baccalá et al. [2013]; a linha verde os valores PDC estatisticamente não significativos e; os valores do eixo  $x$  em pêntadas. Portanto, usando o nível de significância de 1%, os números representam, a PDC para o conjunto de séries temporais dos PCs x índices usando um modelo autorregressivo de ordem  $p = 24$ . . . . . 48

4.7 Componente principal (PC) com maior correlação em cada bacia brasileira e os respectivos índices climáticos que causam a precipitação. . . . . 49

4.8 Correlação defasada (lag -12 a lag 0, dos tons mais claros aos mais escuros) entre os modos de precipitação ((a) PC1 a (f) PC6) e as vazões nas principais bacias brasileiras, as linhas pontilhadas representa o limite de significância a partir do Teste t de Student (-0,07 e +0,07). . . . . 51

4.9 Diagrama de Causalidades entre as componentes principais da precipitação e os índices climáticos. Os conectores em vermelho indicam causalidade em alta frequência. 53

H.1 Arquitetura de Long short-term memory (LSTM) . . . . . 89

I.1 Espectro de Potência e Global da análise de ondeleta da série temporal dos índices (Morlet) a) MEI; b) SOI; c) TSA e d) TNA. O espectro de potência é delimitado pelo cone de influência, o contorno preto indica o setor do espectro de potência onde o nível de confiança de 95% é satisfeito. O espectro de potencial global indica o nível de potência satisfeito para toda a série temporal, ou seja, acima da linha vermelha tracejada. . . . . 91

J.1 Location of the main Brazilian watersheds and the streamflow points in the Hydro-electric Power Plant . . . . . 99

J.2 Standardized time series of the first six principal components of precipitation. . . . 103

J.3 Power and Global Spectrum of the wavelet analysis of the time series of precipitation a) PC1; b) PC2; c) PC3; d) PC4; e) PC5 e f) PC6. The power spectrum is delimited by the influence cone and the black outline indicates the sector of the power spectrum where the 90% confidence level is satisfied. The global power spectrum indicates the power level satisfied for the complete time series, that is, above the dashed red line. The period is the unit of years ( $0.03125 \approx 2$  pentads,  $0.0625 \approx 4.5$  pentads,  $0.125 \approx 9$  pentads,  $0.25 \approx 18$  pentads, and  $0.5 \approx 36$  pentads). . . . . 104

J.4 Pentad lagged correlation (lag -12 to lag 1, from lightest to darkest tones) between precipitation modes ((**a**) PC1 - (**f**) PC6) and climate variability indices. The dashed lines represent the limit of significance from the student's t-test (-0.07 and +0.07). 105

J.5 The black line is the (pseudo-) spectral density of the series, in dB units (PC1 to PC6); the red graph represents the statistically significant PDC values; the dashed black line represents the approximation of the Patnaik threshold (Baccalá and Sameshima, 2021); the green line shows statistically non-significant PDC values and the values of the x axis in pentads. Therefore, using the significance level of 1%, the numbers represent the PDC for the set of time series of PCs x indices using an autoregressive model of order  $p = 24$ . . . . . 106

J.6 Main components (PC) with the highest correlation in each Brazilian watersheds and the respective climatic indices that cause precipitation at the intraseasonal scale. . . 107

J.7 Lagged correlation (lag -12 to lag 0 in pentads, from lightest to darkest tones) between precipitation modes ((**a**) PC1 to (**f**) PC6) and streamflow in the main watersheds dashed lines represent the limit of significance from the student's t-test (-0.07 and +0.07). . . . . 109

J.8 Causality diagram between the main components of precipitation and climate indices. Red connectors indicate high frequency causality (intraseasonal scale). . . . . 111

# Lista de Tabelas

3.1	Índices climáticos mensais e/ou diários disponibilizados por agências governamentais e seus respectivos períodos e referências. . . . .	26
3.2	Índices climáticos mensais e/ou diários desenvolvidos e seus respectivos períodos e referências. . . . .	26
3.3	Pesos dos índices climáticos selecionados para a análise de causalidade e variância explicada (VE%) para cada componente principal. . . . .	35
4.1	Principais Usinas Hidrelétricas (UHE) em cada bacia brasileira de estudo. Os valores em negrito correspondem as maiores UHE em Potência Instalada (Capacidade Instalada) no Brasil. . . . .	40
4.2	Porcentagens da variância explicada e da variância acumulada dos seis primeiros componentes principais (PCs) da precipitação nas principais bacias brasileiras. . . .	42
4.3	Porcentagens da variância explicada e da variância acumulada dos seis primeiros componentes principais (CPs) da precipitação nas principais bacias brasileiras. . . .	43
4.4	Correlações entre as seis componentes principais e as precipitações nas principais bacias brasileiras. Os valores em negrito corresponde ao nível de significância de 95%, a partir do teste t de Student. . . . .	44
4.5	Correlações entre as seis componentes principais (PCs) e os índices climáticos. Os valores em negrito corresponde ao nível de significância de 95%. . . . .	46
4.6	Índices Selecionados para análise de causalidade (PDC) entre os modos de precipitação (PCs 1 a 6) e os índices climáticos selecionados com maiores correlações. . .	46
5.1	Seleção dos índices climáticos pelo método stepwise para previsão de 15 a 60 dias em cada bacia. . . . .	59
5.2	Índices Climáticos selecionados através da busca exaustiva com modelos de LSTM. . . . .	60
5.3	Erro Quadrático Médio (em inglês, mean squared error - MSE) em mm para modelos desenvolvidos por Regressão Linear Múltipla (RML) e Long Short-Term Memory (LSTM) para cada bacia brasileira selecionada no experimento preliminar. . . . .	61
5.4	Erro Quadrático Médio (em inglês, mean squared error - MSE) em mm para modelos desenvolvidos por Regressão Linear Múltipla (RML) e Long Short-Term Memory (LSTM) para cada bacia brasileira selecionada nas etapas I e II. . . . .	61
5.5	Redução do Erro Quadrático Médio (em inglês, mean squared error - MSE) entre LSTM nas duas etapas (Etapa I e II) e entre o LSTM (Etapa II) com o MRL (Etapa I) para cada bacia brasileira selecionada. . . . .	62

6.1	Informações dos modelos dinâmicos disponíveis pelo programa S2S utilizados no estudo.	65
6.2	Erro Quadrático Médio (em inglês, mean squared error - MSE) em mm para modelos desenvolvidos por Regressão Linear Múltipla (RML) e <i>Long Short-Term Memory</i> (LSTM) para cada bacia brasileira selecionada no experimento preliminar e comparativo com os modelos desenvolvidos com os índices climáticos e os índices climáticos + previsão dos modelos dinâmicos do programa S2S (ECMWF e NCEP).	68
6.3	Índices Climáticos selecionados através da busca exaustiva com modelos de LSTM e precipitação prevista do modelo NCEP.	69
6.4	Erro Quadrático Médio (em inglês, mean squared error - MSE) para modelos desenvolvidos em LSTM para cada bacia brasileira selecionada.	69
I.1	Coefficiente de correlação (CC) entre os índices climáticos para o período instrumental (1980 a 2018). Os valores em negritos são considerados significativos (nível 5%) a partir do teste T de Student (Devore [2014]).	92
J.1	Monthly and/or daily climate indices made available by government agencies and their respective references	97
J.2	Climate indices developed and their respective references.	98
J.3	Main Hydroelectric Power Plants (HPP) in each Brazilian watershed. The values in bold correspond to the highest HPP in installed power (installed capacity) in Brazil.	101
J.4	Percentage of variance and accumulated variance of the first six principal components (PCs) of precipitation in the main Brazilian watersheds.	101
J.5	Correlations between the six main components and rainfall in the main Brazilian watersheds. Values in bold correspond to a significance level of 95%, based on Student's t test.	102
J.6	Correlations between the six principal components (PCs) and climate indices. Values in bold correspond to the significance level of 95%	103
J.7	Selected indices for causality analysis (PDC) between precipitation modes (PCs 1 to 6) and selected climate indices with higher correlations.	106

# Capítulo 1

## Introdução

A utilização das previsões climáticas de precipitação na escala climática, ou seja, da escala temporal sazonal a interanual, é um grande desafio na gestão de recursos hídricos visando a produção de energia, abastecimento urbano, uso agrícola e industrial. O sucesso da previsão implica na associação entre a capacidade de prever a precipitação, como transformar a água que cai na superfície em vazão, em uma bacia hidrográfica, e como gerir os recursos disponíveis de forma a otimizar, por exemplo, a geração de energia e ao mesmo tempo salvaguardar o uso da água nos centros urbanos. A previsão de precipitação, por exemplo, é feita através de: (a) modelos dinâmicos, que usam os princípios de conservação de massa, energia, momentum expressos em equações diferenciais de complexa e custosa solução, do ponto de vista computacional, e (b) modelagem baseada em dados (modelos empíricos), ou seja, baseada no comportamento observado no passado, usando métodos estatísticos ou inteligência artificial, contemplando modelos lineares ou não. A opção (b), baseada em dados, é frequentemente utilizada tendo em vista o custo computacional menor e o fato de os resultados serem bastante razoáveis em algumas aplicações. Porém, frequentemente surgem problemas oriundos da baixa qualidade/quantidade de dados disponíveis para alimentar os modelos estatísticos/estocásticos ou de aprendizado de máquina (em inglês, *machine learning*), e problemas decorrentes da não estacionariedade dos processos climáticos que violam algumas hipóteses fundamentais dos modelos de previsão baseados em dados.

### 1.1 Objetivos

Objetivo Geral: O objetivo geral desse trabalho consiste em desenvolver modelos empíricos de previsão climática de precipitação, utilizando principalmente padrões de variabilidade climática de diferentes escalas (intrasazonal, interanual, interdecadais e multidecadais), visando otimizar o planejamento e a gestão de recursos hídricos.

Objetivos Específicos:

- Caracterizar a variabilidade de índices climáticos de alta (variabilidade intrasazonal) e de mais baixa (variabilidade interanual, interdecadais e multidecadais) frequência, observando os padrões e suas relações;

- Identificar padrões climáticos que influenciam a precipitação em bacias hidrográficas brasileiras;
- Desenvolver modelos empíricos de previsões climáticas de precipitação em bacias hidrográficas brasileiras com antecedência de semanas até 2 meses;
- Auxiliar o agente do setor na tomada de decisão proporcionando indicativos de cheias ou escassez hídrica.

## 1.2 Justificativa da Originalidade do trabalho

Ao longo dos últimos anos, a previsão de precipitação na escala de até 2 semanas de antecedência vem sendo empregada operacionalmente (Kumar *et al.* [2011]; Lin *et al.* [2016]). Após o desenvolvimento da técnica de previsões por conjunto, foi possível explorar as incertezas inerentes ao sistema dinâmico atmosférico, trazendo melhoria de desempenho e inovações metodológicas (Tracton e Kalnay [1993]; Murphy [1988]). A previsão de longo prazo, ou seja, para as próximas estações do ano, também tem sido realizada nas últimas décadas, com o uso de modelos baseados nos princípios físicos de conservação e por conjuntos, quando estes ultrapassaram o desempenho dos modelos puramente estatísticos (Lúcio *et al.* [2010]). Desta forma, as previsões dos fenômenos El Niño e La Niña, no Oceano Pacífico, passaram a ser previstos com cerca de 6 meses de antecedência, permitindo melhor destreza na avaliação do impacto remoto destes fenômenos e de outros que ocorrem no Oceano Atlântico e Índico, principalmente na região equatorial (Newman *et al.* [2011]).

Em resumo, atualmente os modelos de previsão de precipitação produzem boas estimativas até 15 dias e na escala sazonal (entre 3 a 6 meses). Porém, a escala entre 2 semanas e 2 meses de antecedência, a chamada previsão subsazonal (do inglês, sub-seasonal to seasonal ou S2S), ainda representa um grande desafio (White [2017]; Vitart *et al.* [2017]; Min *et al.* [2020]). Este é o principal objetivo desta pesquisa de doutorado, agregar na área de previsão estendida, a partir dos avanços adquiridos com o programa S2S da Organização Meteorológica Mundial (<http://s2sprediction.net/>).

## 1.3 Hipóteses

A pergunta principal que essa pesquisa pretende abordar é: Os padrões de variabilidade climática são bons preditores de modelos de previsão climática de precipitação na escalada de 2 semanas até 2 meses?

Como perguntas secundárias: Os resultados dessas previsões podem produzir indicativos úteis para alerta de semanas secas ou chuvosas em uma bacia hidrográfica, assim contribuindo para o planejamento antecipado dos recursos hídricos? Será que o uso de padrões climáticos, em conjunto com técnicas não lineares de *machine learning*, pode produzir resultados substancialmente melhores que as técnicas estatísticas lineares tradicionais? Como integrar as informações em diferentes escalas de tempo (intrasazonal, interanual, interdecadais e multidecadais)? É possível melhorar a previsão climática da precipitação produzida pelos modelos dinâmicos utilizando os índices climáticos, através de modelos estatísticos com técnicas não lineares?

Além da resposta não-linear aos processos de escoamento superficial e subsuperficial da precipitação em bacias hidrográficas (Chow *et al.* [1994]), os padrões de variabilidade climática não produzem respostas lineares, muitas vezes o comportamento e a interação entre fenômenos são de natureza não-linear, conforme estudos de Palmer [1999], Ramirez *et al.* [2018], Builes-Jaramillo *et al.* [2018] e Raphaldini *et al.* [2021].

Modelos de *machine learning* conseguem reproduzir as relações não lineares das variáveis em análise e, portanto, poderá trazer um ganho em relação as técnicas tradicionalmente utilizadas, baseadas em modelos estatísticos.

## 1.4 Organização do Trabalho

Neste Capítulo Introdutório 1, foram apresentados os objetivos gerais e específicos, a justificativa da originalidade do trabalho e as hipóteses adotadas inicialmente.

No Capítulo 2, apresentamos os conceitos obtidos através da revisão bibliográfica, são abordados os principais modelos de previsão de precipitação utilizado por órgãos de gestão dos recursos hídricos, destacando os principais modelos empíricos lineares e não-lineares. Além disso, são apontados os principais padrões de variabilidade climática em diferentes escalas (intrasazonal, interanual, decadal e multidecadal) na região de estudo.

No Capítulo 3, são realizadas análises e discussões sobre as características temporais dos padrões de variabilidade climática no período instrumental disponível pelos órgãos: *National Oceanic and Atmospheric Administration* (NOAA), *Japan Agency for Marine-Earth Science and Technology* (JAMSTEC), *European Climate Assessment & Dataset project* (ECA&D) e *Bureau of Meteorology-Australian Government* (BOM).

No Capítulo 4, são investigados os potenciais preditores (índices climáticos) de precipitação em cada bacia brasileira delimitada (Madeira, Xingu, Tocantins, Serra da Mesa, São Francisco, Paranaíba, Paraná, Rio Grande, Rio Tietê, Itaipu, Paranapanema, Iguaçu, Jacuí e Uruguai).

No Capítulo 5, foram explorados diversos modelos com técnicas lineares e não lineares para prever a precipitação nas bacias brasileiras delimitadas no capítulo anterior.

No Capítulo 6, são incorporados aos modelos desenvolvidos novas variáveis, compostas por séries temporais de modelos desenvolvidos no programa S2S da Organização Meteorológica Mundial (<http://s2sprediction.net/>).

E, por fim, no Capítulo 7 discutimos algumas conclusões obtidas neste trabalho. Analisando as vantagens e desvantagens dos métodos propostos e sugestões para trabalhos futuros.

A metodologia detalhada das técnicas utilizadas, bem como alguns testes complementares estão disponíveis no Apêndice.





# Capítulo 2

## Revisão Bibliográfica

Neste capítulo 2 são abordados os principais modelos de previsão de precipitação utilizado por órgãos de gestão dos recursos hídricos, destacando os principais modelos empíricos lineares e não-lineares. Além disso, são apontados os principais padrões de variabilidade climática em diferentes escalas (intrasazonal, interanual, decadal e multidecadal) na região de estudo.

### 2.1 Modelagem e Previsão de Precipitação

O Brasil possui uma alta disponibilidade hídrica, que corresponde a aproximadamente 12% do total de água doce do planeta. Além da utilização para o consumo humano e dessedentação de animais, a disponibilidade dos recursos hídricos é fundamental para o sucesso de atividades econômicas agrícolas e de geração de energia elétrica (hidroeletricidade). Neste contexto, a disponibilidade de previsões acuradas de precipitação é uma ferramenta importante para o planejamento operacional do uso do recurso hídrico, auxiliando na tomada de decisão, que visa minimizar os efeitos de cheias excepcionais e de escassez hídrica (Collischonn *et al.* [2005]).

Para a gestão dos recursos hídricos, fatores como fisiologia da bacia hidrográfica e as condições climáticas que se inter-relacionam são essenciais no desenvolvimento de estimativas acuradas para o planejamento da disponibilidade hídrica (Von Sperling [2007]).

Os modelos de previsão possuem várias limitações devido a quantidade e qualidades dos dados, o que dificulta a formulação matemática de alguns processos e simplificação deles. Na hidrologia, diversos modelos utilizam informações climáticas como variável de entrada, por exemplo, o modelo MGB-IPH (Modelo Hidrológico de Grandes Bacias) utiliza informações de precipitação observada em postos pluviométricos e precipitação prevista pelo modelo regional ETA do Centro de Previsão de Tempo e Estudos Climáticos/Instituto Nacional de Pesquisas Espaciais (CPTEC/INPE), para estimar a vazão em bacias brasileiras (Collischonn *et al.* [2007]).

Estudos comprovam a eficiência de diferentes modelos hidrológicos (conceituais e empíricos) de estimativas de vazão que incorporam a precipitação em diversas bacias brasileiras, por exemplo, na bacia do Jacuí (Brun e Tucci [2001]), na bacia do Uruguai (Tucci [2003]), na bacia do Paranaíba (Collischonn *et al.* [2007]), nas bacias do Paraná e Iguazu (Guilhon *et al.* [2003]).

Neste sentido, este trabalho de doutorado visa aperfeiçoar a previsão de precipitação, com a finalidade de melhorar os dados de entrada dos modelos hidrológicos, favorecendo no planejamento

antecipado dos recursos hídricos disponíveis nas grandes bacias brasileiras.

Nos subcapítulos a seguir, serão discutidos duas categorias de modelos de previsão de precipitação: (i) modelos dinâmicos, que usam os princípios de conservação de massa, energia e momentum (Seção 2.1.1), e (ii) modelos empíricos baseado no comportamento estatístico do passado, com ênfase nos modelos de *machine learning* (Seção 2.1.2).

### 2.1.1 Modelos Dinâmicos

Os modelos dinâmicos produzem os movimentos atmosféricos e sua evolução temporal, através dos princípios de conservação de massa, energia e momentum, isto é, baseiam-se nas leis da Mecânica dos Fluidos e da Termodinâmica Clássica. Com o aumento da capacidade computacional ao longo das últimas décadas, proporcionou-se o aperfeiçoamento dos modelos numéricos de circulação geral da oceano-atmosfera (em inglês, *atmosphere-ocean general circulation model*, AOGCM) com resolução espacial igual ou acima de 250 quilômetros e dos modelos climáticos regionais (em inglês, *regional climate model*, RCM) com resolução da ordem de dezenas de quilômetros. No Brasil, foi desenvolvido o modelo atmosférico global do CPTEC/INPE, CPTEC-COLA (Cavalcanti *et al.* [2002]; Marengo *et al.* [2003]) e os modelos atmosféricos regionais Eta e BRAMS (em inglês, *Brazilian Regional Atmospheric Modeling System*, Freitas *et al.* [2009]).

Apesar dos avanços da modelagem dinâmica a solução ainda é bastante custosa, do ponto de vista computacional, sendo necessários supercomputadores para produzir a modelagem numérica. Além disso, para produzir modelos numéricos dinâmicos acurados são necessárias atualizações em tempo real das variáveis atmosféricas e de superfície, porém esses dados são de difícil obtenção, dificultando o processo de modelagem.

No âmbito do programa S2S, escala entre 2 semanas e 2 meses de antecedência (isto é, escala subsazonal a sazonal), a comunidade científica global vem desenvolvendo estudos e modelos para melhorar a previsão e compreensão da escala de tempo subsazonal a sazonal, com ênfase especial em eventos climáticos de alto impacto (isto é, condições climáticas extremas, incluindo ciclones tropicais, secas, inundações, ondas de calor e aumento/diminuição da precipitação de monções).

Diversos centros meteorológico disponibilizam dados dos modelos dinâmicos para estudos na escala subsazonal a sazonal (por exemplo, os modelos: BoM, CMA, CNRM, ECC, ECMWF, HMCR, IAP-CAS, ISAS-CNR, JMA, KMA, NCEP e UKMO) que produzem estimativas de precipitação entre 0 a 65 dias, disponível em <https://confluence.ecmwf.int/display/S2S/Models>. Os dados são fornecidos em uma grade atmosférica regular de  $1.5^\circ \times 1.5^\circ$  e oceânica de  $1.0^\circ \times 1.0^\circ$  com uma frequência de atualização que varia entre diária a semanal, conforme a disponibilidade do centro atmosférico.

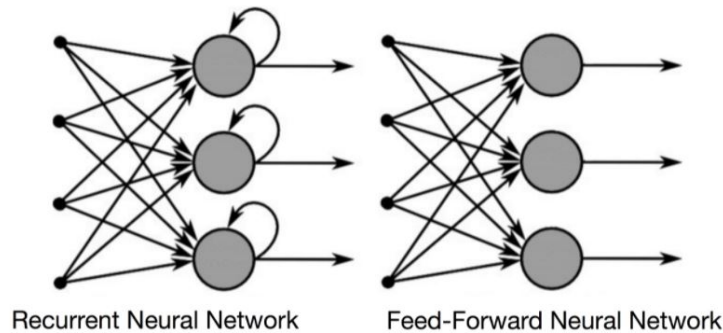
### 2.1.2 Modelos Empíricos

Uma alternativa a modelagem dinâmica, os modelos empíricos são bastante indicados quando se deseja capturar relações entre as variáveis, sem a necessidade de desenvolver modelos físicos complexos. Estes modelos introduzem as variáveis de entrada e utilizam técnicas matemáticas para ajustar parâmetros empíricos que buscam representar a precipitação de saída, sem necessariamente representar o comportamento físico dos processos envolvidos. As técnicas frequentemente utilizadas

são a regressão linear múltipla (RLM), os modelos autorregressivos com média móvel (ARMA, ARIMA) e os modelos não-lineares, como o de redes neurais. Estes modelos geralmente utilizam o conhecimento das estatísticas de séries do passado para estimar as precipitações no futuro, portanto dependem da estacionariedade do comportamento da bacia (Tucci [2005]).

Com a modelagem empírica é possível desenvolver modelos específicos que estabeleçam preditores de precipitação intrínsecos da área de estudo, ou seja, é possível desenvolver modelos para cada região (e.g. bacias, sub-bacias, locais), períodos (épocas do ano) e grupos de preditores. Neste contexto, a análise estatística das séries temporais das variáveis envolvidas fornece um conhecimento empírico que pode conduzir a previsões confiáveis, mesmo que não haja entendimento explícito do processo dinâmico (Guetter e Prates [2002]).

Outra técnica não-linear tradicional, são os modelos de Redes Neurais Artificiais (RNAs) que são compostos de um certo número de elementos não lineares (neurônio artificial) operando em paralelo e dispostos em camadas similarmente às redes neurais biológicas (Haykin [2009]). Através de algoritmos específicos, as RNAs podem aprender certo conhecimento e identificar padrões. As RNAs podem ser grupadas em duas categorias: redes sem realimentação (feedforward networks) e com realimentação (recurrent networks, RNNs) (Graves *et al.* [2013]). As redes sem realimentação são estáticas, isto é, uma dada entrada pode produzir apenas um conjunto de saída; logo, não produzem nenhum tipo de memória. Por outro lado, as redes recorrentes permitem que as características temporais da informação sejam memorizadas (Haykin [2009]). Na Figura 2.1 observa-se a diferença entre as duas arquiteturas de rede.

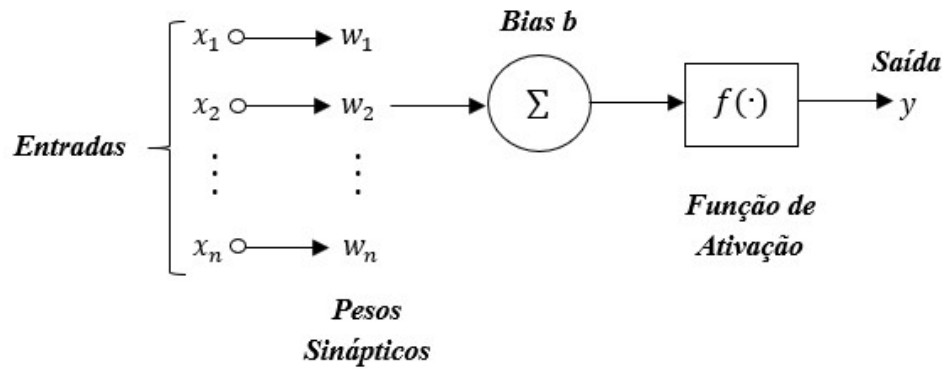


**Figura 2.1:** Comparação entre a arquitetura de uma Rede Neural Recorrente (em inglês, *Recurrent Neural Network*) e uma Rede Neural Feed-forward (em inglês, *Feed-Forward Neural Network*).

Um exemplo mais canônico de RNAs feedforward é o Multilayer Perceptron (MLP), Rosenblatt [1958] introduziu o *perceptron* (Figura 2.2) uma arquitetura simples de rede neural. O modelo otimiza os parâmetros  $w$  (pesos) e  $b$  (bias) para aproximar uma função  $y$ , a partir de um conjunto de variáveis explicativas  $(x_1, x_2, \dots, x_n)$ .

Matematicamente, a saída do perceptron pode ser descrita pela equação 2.1. A função de ativação  $f(\cdot)$  representada pela equação 2.2 (função sigmoide) é geralmente utilizada com o propósito de limitar a saída do neurônio e introduzir não-linearidade ao modelo.

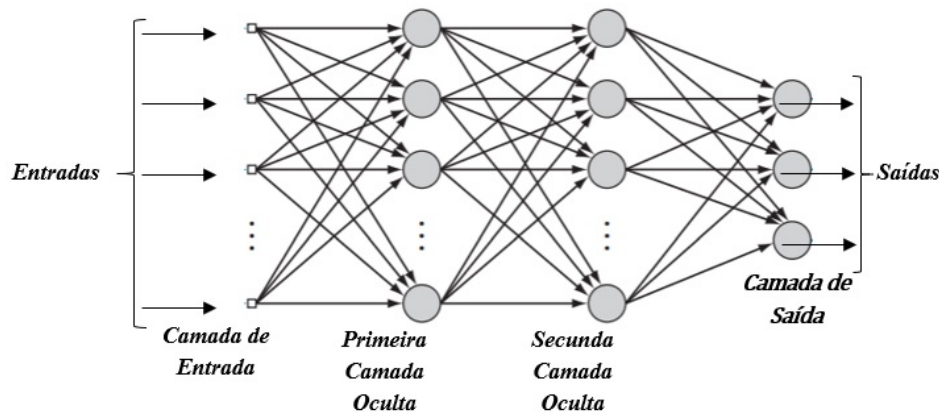
$$y = f\left(\sum_{i=1}^N w_i x_i + b\right) \quad (2.1)$$



**Figura 2.2:** O vetor de entrada com  $n$  variáveis ( $x_1, x_2, \dots, x_n$ ) que são multiplicados por um vetor de peso do mesmo tamanho e adicionado a um bias  $b$ . O somatório deste resultado é passado para a função de ativação  $f(\cdot)$  que por sua vez gera uma saída  $y$ .

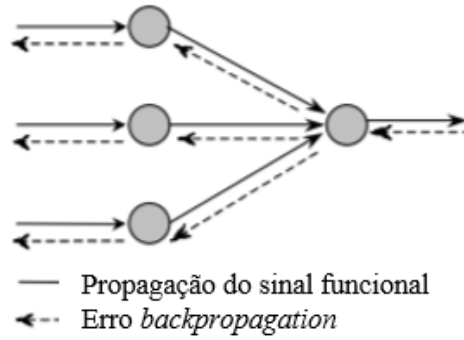
$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.2)$$

A construção de uma rede neural artificial de extensão não linear consiste em várias camadas interconectadas com um número de *perceptron* (ou neurônios) em cada camada. A camada de entrada e a camada de saída são visíveis ao usuário, enquanto as camadas entre elas são chamadas de ocultas. A Figura 2.3 mostra a arquitetura de um MLP com duas camadas ocultas e uma camada de saídas ( $y_1, y_2$  e  $y_3$ ) com conexão total entre os neurônios, ou seja, qualquer camada da rede está conectada a todos os neurônios da camada anterior (Haykin [2009]).



**Figura 2.3:** Arquitetura de um Multilayer Perceptron com duas camadas ocultas e uma camada de saída ( $y_1, y_2$  e  $y_3$ ). O fluxo de sinal através da rede progride da esquerda para a direita e numa base de camada por camada.

O treinamento da rede MLP mais utilizado é o algoritmo de backpropagation, que significa retropropagação do erro. Este algoritmo consiste basicamente de dois passos: *i*) propagação positiva do sinal funcional: todo os pesos da rede são mantidos fixos e *ii*) retropropagação do erro: os pesos da rede são ajustados baseados no erro. O sinal de erro é propagado em sentido oposto ao de propagação do sinal funcional, por isso o nome de retropropagação (*backpropagation*), conforme ilustrado na Figura 2.4.



**Figura 2.4:** Ilustração das direções do fluxo de sinal em Multilayer Perceptron e erro *backpropagation*.

Normalmente, um perceptron de camada única é treinado pelo algoritmo conhecido como Descida do Gradiente Estocástico (SDG) (Robbins e Monro [1951]) combinado com o método *backpropagation* (Rumelhart *et al.* [1986]). A SDG determina um vetor de pesos e minimiza o erro, começando com um vetor inicial de pesos arbitrários e modificando-o repetidamente em pequenos passos. A cada passo, o vetor de pesos é alterado na direção que produz a maior queda ao longo da superfície de erro, este processo continua até atingir um erro mínimo (Robbins e Monro [1951]). O algoritmo *backpropagation* emprega a SDG para minimizar o erro quadrático entre a saída da rede e os valores alvos para estas saídas (Rumelhart *et al.* [1986]).

A técnica não-lineares apresentada anteriormente já está bem consolidada e muitos estudos de previsões de precipitação fazem o uso dessas técnicas. Neste trabalho foram exploradas técnicas de Aprendizado Profundo (do inglês, *Deep Learning*, DL), que envolvem várias camadas ocultas, obtendo resultados melhores que as redes neurais com aprendizado mais raso. O aumento do número de camadas ocultas permite que o MLP obtenha uma melhor incorporação dos dados de entrada que otimiza as previsões (LeCun *et al.* [1986]).

Além disso, o número de funções é diferente para o aprendizado raso e profundo, as redes neurais mais rasas buscam uma única função que possa, a partir de um conjunto de parâmetros/pesos gerar o resultado desejado. Por outro lado, para as DL são obtidas diferentes funções por meio de composições de funções (conforme a equação 2.3), ou seja, cada função faz uso de parâmetros para realizar a transformação dos dados de entrada (conforme a equação 2.4) Ponti e Paranhos da Costa [2017].

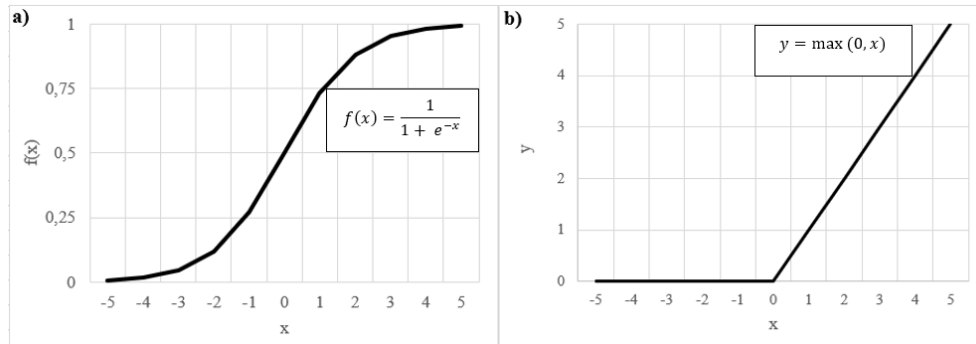
$$f(x) = f_l(\dots f_2(f_1(x_1))\dots) \quad (2.3)$$

onde cada função  $f_l(\cdot)$  (o índice  $l$  se refere a uma camada) assume como entrada um vetor de dados  $x_l$  (preparado para a camada  $l$ ), gerando como saída o próximo vetor  $x_{l+1}$ .

$$f_l(\dots f_2(f_1(x_1, W_1); W_2)\dots), W_l \quad (2.4)$$

Para o treinamento de um perceptron utiliza-se normalmente como função de ativação as funções sigmóides (como a equação 2.2). Porém em DL, não é o ideal, pois as funções sigmóides saturam a partir de um determinado ponto, ocasionando falhas na otimização nas próximas camadas (Ponti e Paranhos da Costa [2017]). Como solução a função retificadora linear (em inglês,

*rectified linear function*, ReLU) tem sido utilizada facilitando o processo de treinamento. A ReLU é simplesmente a função de identidade para os valores positivos (Glorot *et al.* [2011]), conforme a equação 2.4. A Figura 2.5 compara a função sigmoide com a ReLU.



**Figura 2.5:** Funções para treinamento de um perceptron. a) função de ativação sigmoide,  $f(x)$  e b) função retificadora linear (em inglês, *rectified linear function*, ReLU),  $y$ .

Na Figura 2.5.a observa-se que a função de ativação sigmoide comprime a saída para um intervalo curto, enquanto ReLU cancela todos os valores negativos (restrição de não-negatividade), sendo linear para os positivos (Ponti *et al.* [2016]). Na prática, quando se utiliza a ReLU como função de ativação alguns neurônios apresentam saída nula para cada amostra de entrada (Figura 2.5.b), reduzindo a redundância de informações (Glorot *et al.* [2011]). Outra função derivada desta metodologia é a ReLU paramétrica (PReLU) que permite valores negativos com menor ponderação, parametrizado por uma variável  $0 \leq a \leq 1$ , possibilitando um aprendizado de  $a$  durante a fase de treinamento (He *et al.* [2015]).

Alguns autores utilizaram modelos baseados em DL para previsão de precipitação, no Brasil, Perez [2018] desenvolveu modelo de DL para previsão da precipitação para a cidade de São Paulo e observou que o desempenho das redes neurais é significativamente superior à previsão de precipitação acumulada em 24 horas do modelo físico global do NOAA (*Global Forecast System Model*, GFS).

## 2.2 Padrões de Variabilidade Climática

Devido à grande extensão territorial do Brasil diversos mecanismos podem atuar inibindo ou favorecendo a precipitação, esses mecanismos podem ser de origem tropical e subtropical de diferentes escalas (micro, meso, sinótica, intrasazonal, interanual, interdecadal e multidecadal) e com variações sazonais e subsazonais, que produzem grandes reflexos no ciclo hidrológico em bacias brasileiras (Satyamurty *et al.* [1998]; Grimm [2003]; Reboita *et al.* [2010]).

As bacias do extremo Sul do Brasil, apresentam menor variação ao longo dos meses do ano, com picos na primavera e outono, causados principalmente por sistemas frontais (frente fria) e convectivos de mesoescala (Reboita *et al.* [2010]; Malfatti *et al.* [2018a]). Nas outras regiões do Brasil observa-se um ciclo anual bem definido, caracterizado por uma estação seca no inverno e uma estação chuvosa no verão, típico do Sistema de Monção da América do Sul (SMAS), com grande atuação da Zona de Convergência do Atlântico Sul (ZCAS) sobre o Sudeste, Centro-Oeste e faixa norte da região Sul (Grimm *et al.* [2005]), e da Zona de Convergência Intertropical (ZCIT)

sobre faixa norte do Nordeste e Norte do Brasil (Reboita *et al.* [2010]). Além disso, no litoral do Nordeste observam-se alísios de sudeste que sofrem a influência das frentes que se deslocam ao longo da costa até o Nordeste, contribuindo para o aumento da umidade na região.

As variações intrasazonais também podem interferir no padrão de variabilidade da precipitação que conseqüentemente atinge a vazão de rios. A Oscilação de Madden-Julian (MJO), por exemplo, pode favorecer ou desfavorecer períodos mais chuvosos sobre a área de influência do SMAS (Jones e Carvalho [2002]; Carvalho e Cavalcanti [2016]).

Além disso, fenômenos de grande escala podem influenciar remotamente a precipitação sobre a bacia, afetando a variabilidade da precipitação da América do Sul, conseqüentemente a variabilidade do ciclo hidrológico. Alguns estudos sugerem a influência da temperatura da superfície do mar (TSM) sobre diversas regiões do Brasil, no caso do Pacífico ocorre em escala interanual e no Atlântico escala decadal (Robertson e Mechoso [1998]).

Em diversas bacias brasileiras, os fenômenos de variabilidade de baixa frequência podem influenciar a precipitação e conseqüentemente no ciclo hidrológico (Cardoso e Cataldi [2012]; Malfatti *et al.* [2018b]). São destacados os fenômenos: El Niño Oscilação Sul (ENSO), Oscilação Decadal do Pacífico (PDO) e a Oscilação Multidecadal do Atlântico (AMO). Além disso, foi evidenciada a atuação de padrões específicos em meses preferenciais, possivelmente devido à sazonalidade de cada local e principais mecanismos atuantes em cada região do Brasil, tais como: Oscilação Ártica (AO) e Antártica (AAO); Modo Meridional do Atlântico (AMM); Oscilação do Atlântico Norte (NAO); Oscilação do Pacífico América do Norte (PNA) e Sul (PSA); Oscilações de TSM no Atlântico Tropical Norte (TNA) e Sul (TSA); e Oscilações de TSM na região das Malvinas.

Nos subcapítulos a seguir, são abordados com maiores detalhes os padrões de variabilidade intrasazonal, interanual, decadal e multidecadal que podem influenciar a precipitação no Brasil.

### 2.2.1 Zona de Convergência do Atlântico Sul (ZCAS)

A Zona de Convergência do Atlântico Sul é um fenômeno meteorológico caracterizado pela persistência de uma banda de nebulosidade orientada no sentido noroeste-sudeste (NW-SE), associada a uma acentuada região de divergência, em altos níveis e convergência, em baixos níveis (Carvalho *et al.* [2004]). Sua extensão abrange o centro sul da Amazônia, regiões Centro-Oeste e Sudeste, centro sul da Bahia, norte do Estado do Paraná e alongando-se em direção ao Oceano Atlântico Sudoeste (Ferreira *et al.* [2004]). Responsável por intensos e persistentes períodos de precipitação em determinadas regiões (regiões Sudeste, Centro-Oeste e costa sul do Brasil) durante o verão austral (estação chuvosa), assim como redução significativa das chuvas nas adjacências (norte da Argentina, Paraguai e Uruguai) (Carvalho *et al.* [2004]; Garreaud *et al.* [2009]; Reboita *et al.* [2010]; Grimm [2011]).

Robertson e Mechoso [2000] analisaram a influência da ZCAS na bacia dos rios Paraná, Paraguai, Uruguai e Negro, e concluíram que, em episódios de ZCAS intensificados, a precipitação nas bacias dos rios Paraná e Paraguai aumenta, e a dos rios Uruguai e Negro diminui. Seabra *et al.* (2004) avaliou a influência em dois casos de ZCAS em bacias do sudeste e centro-oeste do Brasil, e verificaram que para as bacias na região sudeste (bacia do rio Grande e Paraíba do Sul) a influência da ZCAS é mais proeminente.

Outra grande importância da ZCAS é a relação dela com o SMAS, principal sistema de preci-



itação da América do Sul (Jones e Carvalho [2002]; Carvalho e Cavalcanti [2016]). Climatologicamente o período de atuação da ZCAS está diretamente relacionado com a configuração do SMAS, que tem papel importante para o transporte de umidade para a região Central da América do Sul (Reboita *et al.* [2010]; Grimm [2011]).

A ZCAS pode ser influenciada pela confluência entre os ventos de nordeste do Anticiclone Subtropical do Atlântico Sul (ASAS), que transportam calor e umidade do oceano Atlântico para o interior do Brasil, e o Jato de Baixos Níveis (JBN) a leste dos Andes que transporta calor e umidade da região amazônica para o sudeste da América do Sul (Reboita *et al.* [2010]).

Alguns estudos avaliam a interação da ZCAS com outros fenômenos atmosféricos. Na escala de tempo sinótica, Seluchi e Marengo [2000] observaram que as frentes frias podem se juntar com a ZCAS, intensificando as chuvas, principalmente na região Sudeste. Na escala de tempo interanual, Carvalho *et al.* [2004] sugeriram que em eventos El Niño a configuração da ZCAS é modificada, deslocando a ZCAS para o oceano, diminuindo a precipitação sobre o continente, enquanto em eventos La Niña favorecem a ZCAS sobre o continente, aumentando a precipitação no continente.

Chaves e Nobre [2004] analisaram a relação entre as anomalias de TSM no Oceano Atlântico e eventos ZCAS, e observaram que anomalias positivas de TSM tendem a intensificar a ZCAS, enquanto que anomalias negativas ocorrem o inverso.

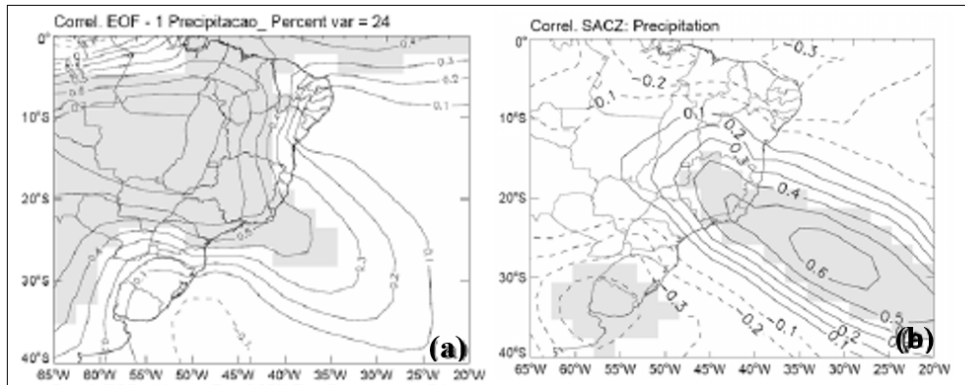
Outras oscilações no Atlântico de diferentes escalas também podem alterar as configurações da ZCAS. Durante a fase negativa da AMO observa-se um aumento na atividade da ZCAS (Chiessi *et al.* [2009]). E durante a fase neutra do ENSO o AMM tem um papel importante na modulação de ciclones (ciclogênese) e nas características da ZCAS. Anomalias negativas de AMM induz um aumento de ciclogênese próximo ao sudeste do Brasil próximo a região de influência da ZCAS, alterando sua organização (Bombardi *et al.* [2014]).

Para identificar os padrões relacionados a SMAS e a ZCAS, Silva e Carvalho [2007] produziram o Índice de Larga Escala da Monção da América do Sul (em inglês, *Large-scale index for South America Monsoon*, LISAM), baseado no primeiro modo de EOF (em inglês, *Empirical Orthogonal Function*) combinado, utilizando dados da precipitação estimada por satélite do GPCP (*Global Precipitation Climatology Project*) e dados em 850 hPa de temperatura do ar, umidade específica e componentes zonal e meridional do vento da reanálise do NCEP/NCAR (*National Centers for Environmental Prediction/ National Center for Atmospheric Research*).

O LISAM representa em um único índice os mecanismos dinâmicos mais importantes do SMAS. Segundo os autores, o primeiro modo da EOF combinada (EOFc-1) representa a maior componente de variabilidade da série temporal e está associado com o SMAS (24% da variância explicada). Além disso, este método fornece uma série contínua (dada pelos coeficientes de expansão da EOFc) com informações sobre o SMAS, podendo ser utilizada para definir as datas de início e fim da monção, assim como a duração da estação chuvosa. A Figura 2.6 mostra a correlação do índice LISAM com a precipitação, indicando correlações positivas sobre grandes áreas da América do Sul tropical organizada em uma banda orientada no sentido noroeste-sudeste que se assemelha à ZCAS.

### 2.2.2 Oscilação de Madden-Julian (MJO)

A Oscilação de Madden-Julian (MJO) com período entre 30 a 60 dias (variabilidade intrasazonal) é um fenômeno que se caracteriza por um deslocamento para leste de uma célula zonal de



**Figura 2.6:** Correlação entre a precipitação a) índice LISAM (em inglês, Large-scale index for South America Monsoon) e b) índice ZCAS (em inglês, South Atlantic Convergence Zone, SACZ) (Fonte: Silva e Carvalho [2007]).

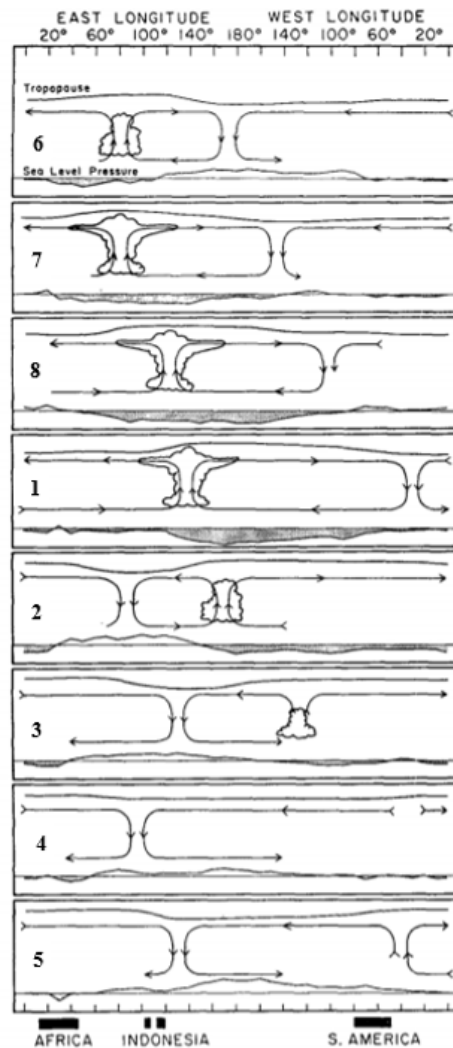
grande escala termicamente direta, que causa variações na convecção tropical (Madden e Julian [1972]; Madden e Julian [1994]). Podendo atuar também na variabilidade subtropical e extratropical (Casarin e Kousky [1986]). Na Figura 2.7 são apresentadas as variações em tempo e espaço (plano zonal) dos distúrbios associados à MJO.

Estudos indicam que quando a convecção é localizada sobre o Oeste do Pacífico (fases 8, 1 e 2, conforme Figura 2.7), há um aumento da convecção sobre a América do Sul, favorecendo sistemas transientes que contribuem para a formação de ZCAS, que possibilita a ocorrência de eventos extremos de precipitação (Kousky e Kayano [1994]; Souza e Ambrizzi [2006]; Grimm e Ambrizzi [2009]).

Shimizu e Ambrizzi [2015] investigaram a ação combinada das fases da MJO e ENSO sobre a América do Sul no verão (novembro a março) e inverno (maio a setembro). Os resultados mostraram que a precipitação e a temperatura anomalias observadas durante os eventos ENSO podem ser fortalecidas ou enfraquecidas quando ocorrem simultaneamente com o MJO, especialmente durante o verão austral. Além disso, o efeito sobre os padrões de anomalias nestes eventos depende da fase da MJO.

Nas regiões Centro-Oeste e Sudeste do Brasil e no Paraguai observa-se uma maior diminuição da precipitação em anos de La Niña durante as fases 3 e 4 da MJO. Em anos neutros há um aumento da precipitação nas fases 1, 2 e 8, onde a convecção MJO é mais fraca sobre a Austrália e Indonésia, e mais forte sobre o Oceano Índico. Enquanto a região Sul do Brasil, norte da Argentina e Uruguai, em anos de El Niño e fases 2 a 6 da MJO, ocorre altas anomalias positivas de precipitação (maiores que 100%), indicando que a convecção MJO sobre o Índico e oeste do Pacífico favorece o aumento da precipitação nesta região. Em anos de La Niña combinada com as fases 2 a 5 da MJO, observa-se um aumento da precipitação, situação semelhante ocorre em anos neutros combinados com as fases 2, 4 e 5 da MJO, indicando que a convecção MJO sobre o oeste do Pacífico (fases 4 e 5) favorece o aumento da precipitação através da propagação de trens de ondas (Shimizu e Ambrizzi [2015]).

Madeira *et al.* [2018] observaram que o subsistema hidro-termo-eólico Sudeste/Centro-Oeste do Sistema Interligado Nacional (SIN) tem mais eventos de chuva acima do normal quando a MJO se encontra nas fases 8, 1 e 2, sendo observado um aumento do nível nos reservatórios dessa região. Na região Sul, o favorecimento da chuva ocorre nas fases 1 e 2 da MJO, enquanto na região Norte esta



**Figura 2.7:** Esquema das variações em espaço e tempo da perturbação relacionada com a MJO. Os números de 1 a 5 referem - se aos tempos em que a Ilha de Canton apresenta baixa e alta pressão, respectivamente, e os demais números a tempos intermediários (Fonte: *Madden e Julian [1972]*).

influência é mais evidente na fase 1, com escassez hídrica na fase 5. E, por fim, na região Nordeste, encontram-se os maiores volumes nos reservatórios quando a MJO se encontra nas fases 8 e 1.

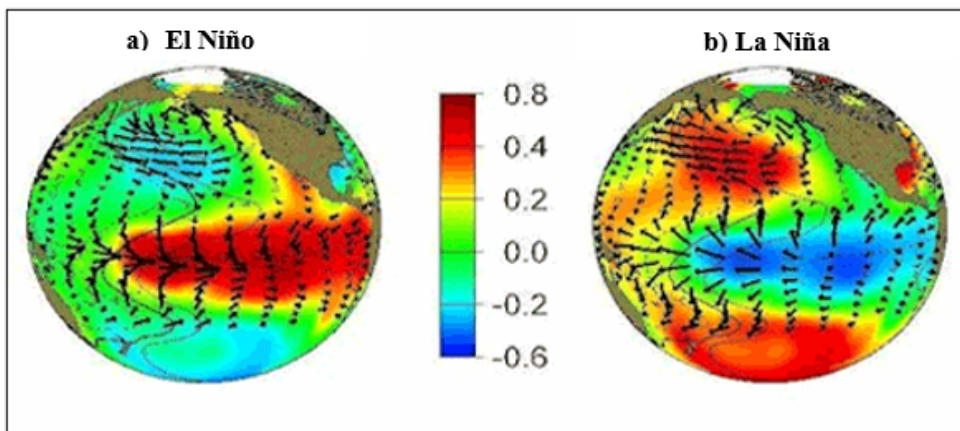
Para monitoramento da MJO são produzidos índices, o principal deles é baseado em uma análise da EOFc de vento zonal em 850 e 200 hPa e Radiação de Onda Longa (em inglês, *Outgoing longwave Radiation*, ORL) entre 15°S - 15°N. As primeiras duas componentes descrevem a propagação da MJO, os índices RMM1 e RMM2 (em inglês, *Real-time Multivariate MJO*) (*Wheeler e Hendon [2004]*).

### 2.2.3 El Niño Oscilação Sul (ENSO)

O El Niño Oscilação Sul (ENSO) corresponde a um fenômeno de grande escala com interação atmosfera-oceano, associado a alterações dos padrões normais da TSM e dos ventos alísios na região do Pacífico equatorial, entre a Costa Peruana e no Pacífico oeste próximo à Austrália. A variação interanual (3-7 anos) do sistema ocorre em torno das condições normais nas componentes oceânicas (El Niño) e atmosféricas (Oscilação Sul) desta região, apresentando duas fases opostas do fenômeno,

El Niño e La Niña (Trenberth [2017]).

Na Figura 2.8a observa-se a mudança dos ventos alísios (de leste) no Pacífico equatorial e o aquecimento da TSM por todo oceano, com destaque para a região leste (anomalias de +0,8), caracterizando o fenômeno El Niño (fase quente). Neste caso, ocorre um deslocamento da região com maior formação de nuvens para o Pacífico equatorial central e a célula de Walker fica bipartida (Oliveira [2000]). Enquanto na Figura 2.8b, os ventos alísios ficam mais intensos, deslocando maior quantidade de água superficial de leste para oeste. Desta forma, há maior quantidade de água represada no Pacífico equatorial oeste e o desnível em relação faixa leste aumenta, aumentando também a ressurgência de águas frias junto à costa oeste da América do Sul (anomalia de -0,6), a célula de Walker fica mais alongada que o normal, caracterizando o fenômeno La Niña (fase fria) (Oliveira [2000]).



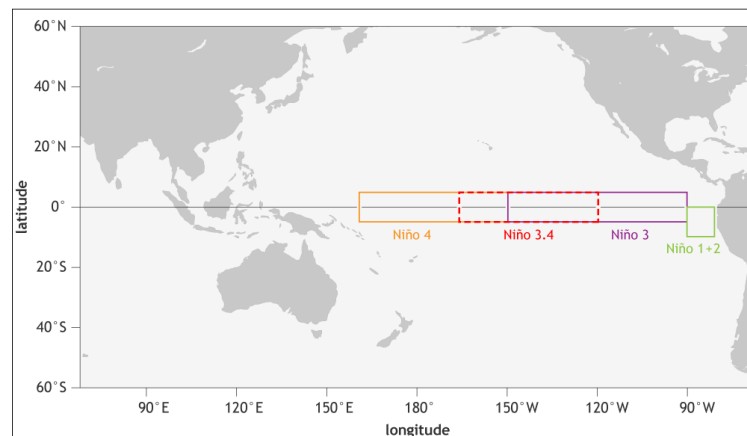
**Figura 2.8:** Fases do El Niño Oscilação Sul (ENSO), El Niño e La Niña, com indicação da anomalia de TSM, vermelho e azul escuro indicam anomalias de +0,8 e -0,6 de temperatura, respectivamente, e os vetores indicam a direção e magnitude dos ventos alísios (Fonte: NOAA Climate.gov).

Em suma, o Brasil é afetado pelo fenômeno ENSO de maneira distinta em duas regiões: o Nordeste e o Sul (Coelho *et al.* [2002]; Grimm [2003]; Tedeschi *et al.* [2016]). Em anos de El Niño a região Sul tem sua estação chuvosa acentuada (primavera), ocasionando grandes enchentes ou até mesmo inundações. Em contrapartida, no mesmo ano, mas em época distinta (outono), o Nordeste sofre com secas severas, devido à inibição da estação chuvosa (Rodrigues *et al.* [2011]; Pinheiro *et al.* [2021]). Nos anos de La Niña, a região Sul sofre com extensivas estiagens. As regiões de transição, Centro-Oeste e Sudeste, não é possível observar um padrão definido entre os impactos opostos do ENSO, entretanto há indícios que a região Sudeste experimenta invernos mais amenos em termos de temperatura em episódios de El Niño (Grimm *et al.* [1998]; Correia Filho *et al.* [2021]; Prado *et al.* [2021]). E, por fim, alguns estudos apontam que em média a região Norte (Amazônia) apresenta comportamento semelhante do Nordeste, porém no período entre janeiro a março (Foley *et al.* [2002]).

Aplicado na gestão de recursos hídricos, a vazão dos rios apresentam correlações opostas entre as regiões Nordeste e Sul do Brasil e o ENSO, no qual a vazão sobre o Nordeste tende a diminuir (aumentar) em anos de El Niño (La Niña), e o contrário sobre a região Sul (Dettinger *et al.* [2000]; Grimm e Tedeschi [2009]). Além disso, observa-se que o segundo modo da componente principal rotacionada da vazão é semelhante ao primeiro componente principal de precipitação, cuja

variabilidade está associada à atuação do ENSO no regime de precipitação da América do Sul (Dettinger *et al.* [2000]).

A intensidade do ENSO é avaliada através de alguns índices climáticos, como: anomalia de TSM no Pacífico Tropical nas regiões Niño 1+2 ( $0^{\circ}$  -  $10^{\circ}$ S,  $90^{\circ}$ W -  $80^{\circ}$ W), Niño 3 ( $5^{\circ}$ N -  $5^{\circ}$ S,  $150^{\circ}$ W -  $90^{\circ}$ W), Niño 3.4 ( $5^{\circ}$ N -  $5^{\circ}$ S,  $170^{\circ}$ W -  $120^{\circ}$ W) e Niño 4 ( $5^{\circ}$ N -  $5^{\circ}$ S,  $160^{\circ}$ E -  $150^{\circ}$ W), conforme a Figura 2.9 (Trenberth e Stepaniak [2001]), Índice Trans-Niño (TNI; Trenberth e Stepaniak [2001]), Índice Multivariado ENSO (MEI; Wolter e Timlin [1998]) e Índice Oscilação Sul (SOI; Ropelewski e Jones [1987]). Neste trabalho serão explorados os índices Niño 3.4, MEI e SOI.



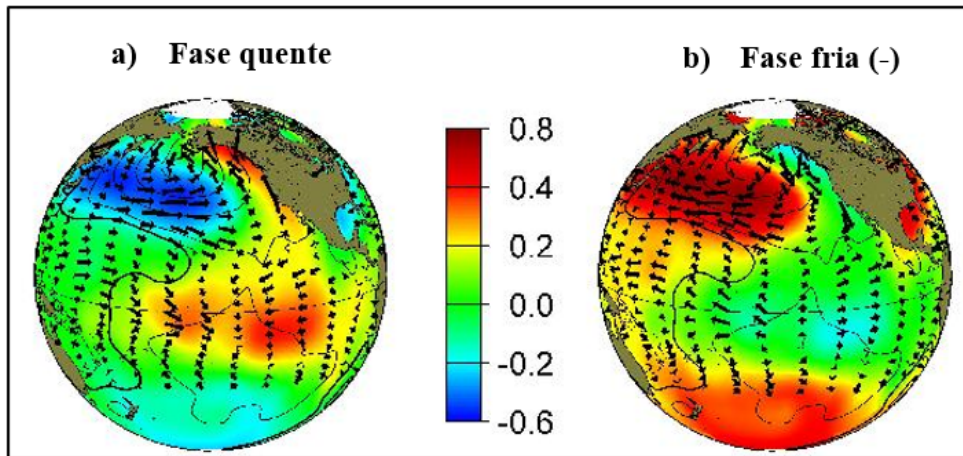
**Figura 2.9:** Regiões do Pacífico Equatorial/Tropical de influência do El Niño Oscilação Sul (ENSO), Niño 1+2, Niño 3, Niño 3.4 e Niño 4 (Fonte: NOAA Climate.gov).

## 2.2.4 Oscilação Decadal do Pacífico (PDO)

A Oscilação Decadal do Pacífico (PDO) é uma variabilidade interdecadal (20 a 30 anos) na configuração das TSM no Oceano Pacífico e apresenta duas fases distintas, conforme a Figura 2.10, a fase positiva ou quente, caracterizada por anomalias de TSM positivas na costa oeste dos Estados Unidos e Pacífico Central, negativas no restante do Pacífico Norte, desde aproximadamente  $20^{\circ}$ N na direção do pólo Figura 2.10a, e a fase negativa ou fria, quando o aquecimento/resfriamento nas duas áreas se inverte (Figura 2.10b). O índice da PDO é obtido a partir da primeira componente principal das anomalias da TSM no Pacífico Norte, entre as latitudes  $20^{\circ}$ N e  $90^{\circ}$ N (Mantua *et al.* [1997]).

De uma maneira geral, observa-se uma forte correlação entre os totais pluviométricos e a PDO, em algumas regiões específicas do Brasil (Sudeste da Amazônia, Centro-Oeste e Sul) (Molion [2008]). Em Goiás (região Centro-Oeste) ocorre uma redução (aumento) de precipitação anual e frequência de dias com precipitação na fase fria (quente) da PDO (Sousa *et al.* [2010]). No Rio Grande do Sul (região Sul) observa-se que na fase positiva da PDO ocorreram chuvas em abundância e na fase negativa houve predomínio de estiagem (Rebello [2006]).

Além disso, verifica-se a relação da PDO com as precipitações da América do Sul ligadas a eventos de ENSO, ou o efeito combinado entre os dois eventos. Durante o verão observa-se evento de El Niño mais pronunciado quando está na fase quente da PDO. Encontra-se também um número menor de eventos La Niña durante a fase quente da PDO, enquanto que na fase fria observa-se maior



**Figura 2.10:** Fases da Oscilação Decadal do Pacífico (PDO), fase quente (+) e fase fria (-), com indicação da anomalia de TSM, vermelho e azul escuro indicam anomalias de +0,8 e -0,6 de temperatura, respectivamente, e os vetores indicam a direção e magnitude dos ventos alísios (Fonte: NOAA Climate.gov).

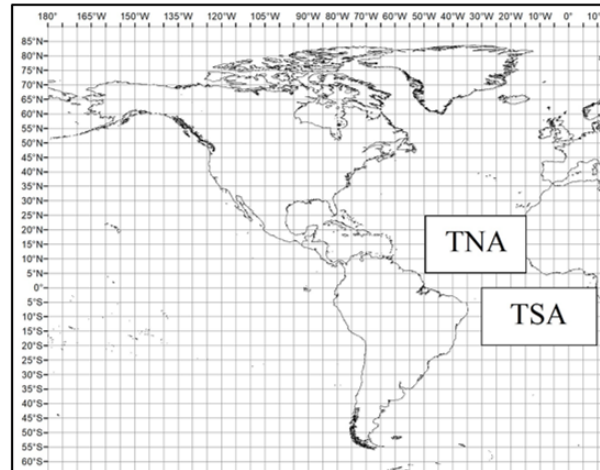
ocorrência de episódios de La Niña intensos, indicando que a resposta do clima da América do Sul é dependente das fases da PDO (Andreoli e Kayano [2005]). Kayano e Andreoli [2009] sugeriram que as diferenças nos padrões de precipitação associados ao ENSO podem também estar relacionadas à fase da PDO, que influencia construtivamente o ENSO, quando estiver na mesma fase. Portanto, pode-se dizer que a PDO pode modular o ENSO (Garcia e Kayano [2008]; Kayano e Andreoli [2009]).

### 2.2.5 Anomalias de TSM no Atlântico Tropical Norte (TNA) e Sul (TSA)

O estudo das influências das anomalias de TSM no Atlântico tropical sobre o clima é facilitado através do uso de índices, tais como: Índice do Atlântico Tropical Norte (TNA), definido pela anomalia de TSM na região do Atlântico Tropical Norte, faixa 5,5°N - 23,5°N e 15°W - 57,5°W e Índice do Atlântico Tropical Sul (TSA), caracterizado pela anomalia de TSM na região do Atlântico Tropical Sul, faixa 0° - 20°S e 10° E 30°W (Figura 2.11). Em alguns casos, durante a fase positiva as águas na faixa do TSA se aquecem, enquanto que na faixa TNA se resfriam (configuração de dipolo), e durante a fase negativa a situação se inverte. Porém essa situação ocorre em apenas 15% do tempo, as situações de não-dipolo são mais frequentes, mas também implicam em um gradiente meridional de TSM no Atlântico Tropical (GTA) (Enfield *et al.* [1999]).

Enfield e Mayer [1997] demonstraram que as anomalias de TSM no Atlântico estão correlacionadas com o ENSO no Pacífico, durante o outono na região TNA e durante quase todo o ano, com maior intensidade no inverno na região TSA. Estas anomalias estariam defasadas em relação às do Pacífico em 4-5 meses. Porém essa relação pode ser desfavorável, enfraquecendo o ENSO durante a fase positiva quando a TSM negativa (Kayano *et al.* [2009]).

Andreoli e Kayano [2005] sugeriram que o ENSO no Pacífico tem maior influência nas anomalias de TSM no TNA e que a variabilidade de TSM no TSA está relacionada principalmente às variações atmosféricas locais ou à própria dinâmica oceânica na escala sazonal.



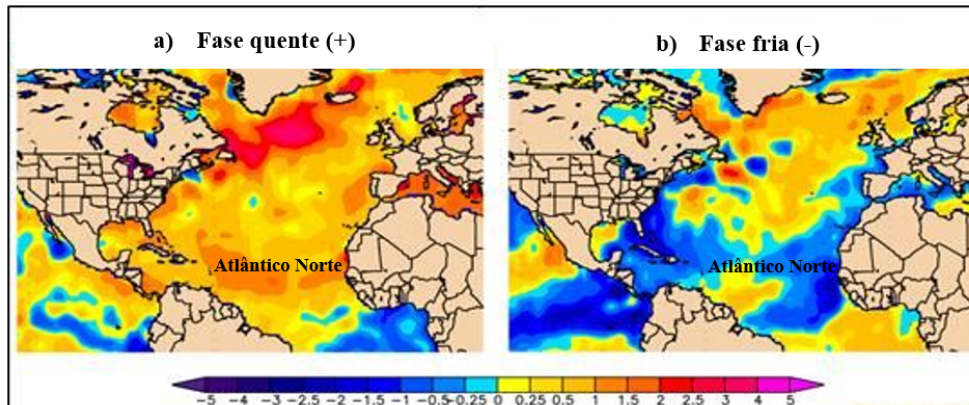
**Figura 2.11:** Localização dos Índices TNA (Atlântico Tropical Norte) e TSA (Atlântico Tropical Sul).

As anomalias do Atlântico Sul podem associar-se a ocorrência de extremos de precipitação sobre o Sul e Sudeste brasileiro (Cardoso e Silva Dias [2004]); Chaves [2011] sugere que as anomalias positivas de TSM no TSA podem influenciar no comportamento da ZCAS, aumentando as chuvas na região Sudeste, Centro-Oeste e parte norte da região Sul. Santos *et al.* [2016b] observaram que quando há um resfriamento (aquecimento) das águas do Oceano Atlântico Norte (TNA) tem-se um aumento (diminuição) dos eventos chuvosos (secos) sobre o Sudeste do Brasil. Para as anomalias negativas (resfriamento) na região TSA tem-se um aumento de tempestades na região Sul e Sudeste do Brasil (Pinto Jr [2015]).

## 2.2.6 Oscilação Multidecadal do Atlântico (AMO)

A Oscilação Multidecadal do Atlântico (AMO) é um fenômeno de variabilidade de baixa frequência na TSM do Atlântico Norte, em um período estimado de 60 a 80 anos, entre a região do Golfo do México e a Baixa da Islândia. A AMO apresenta duas fases, conforme ilustradas pela Figura 2.12, sendo que na fase positiva (fase negativa) ocorrem águas mais quentes (frias) do que o normal no Oceano Atlântico Norte. O índice AMO corresponde ao primeiro modo da EOF rotacionado, considerando as anomalias de TSM da região norte do Atlântico e a região do Golfo do México (Enfield *et al.* [2001]).

Para avaliar as características da variabilidade da AMO, são necessários longos períodos de dados, alguns estudos verificam a ocorrência por registro paleoclimáticos. Moquet *et al.* [2013] verificaram a variabilidade multidecadal do sistema de monções durante os últimos 2000 anos na região Centro-Oeste do Brasil, e observaram variações multidecadais relacionadas à AMO, com frequência aproximada de 63 anos. Nos últimos 1550 anos foi observado um aumento (diminuição) de chuvas durante as fases de AMO negativas (positivas). Santos *et al.* [2016a] analisaram os coeficientes de correlação entre o índice da AMO e a precipitação sobre a América do Sul e identificaram o mesmo padrão em ambos os modos da AMO (negativo e positivo), mostrando que a precipitação sobre a América do Sul aumenta (diminui) quando a AMO está na fase negativa (positiva). Durante o verão na América do Sul, observa-se relação entre a AMO e a ZCAS, durante a fase negativa (fase posi-



**Figura 2.12:** Fases da Oscilação Multidecadal do Atlântico (AMO), fase quente (+) e fase fria (-), com indicação da anomalia de TSM, laranja e azul escuro indicam anomalias de +1,0 e -1,0 de temperatura, respectivamente (Fonte: NOAA Climate.gov).

tiva) AMO aumenta (reduz) a atividade da ZCAS, aumentando (diminuindo) a precipitação sobre a bacia do Rio Prata (Chiessi *et al.* [2009]). Além disso, o ENSO pode se intensificar dependente da fase da AMO, o El Niño é mais pronunciado durante a fase negativa da AMO, enquanto que a La Niña é mais pronunciada durante a fase positiva da AMO (Kayano e Capistrano [2014]). Ao Sul do Brasil, eventos intensos de precipitação estão associados a El Niño, a AMO negativa e PDO positivo (Prado *et al.* [2021]).

### 2.2.7 Modo Meridional do Atlântico (AMM)

O Modo Meridional do Atlântico (AMM; ou Dipolo Atlântico) é um fenômeno acoplado oceano-atmosfera, com variabilidade interanual no Oceano Atlântico tropical (“meridional”). Este modo é caracterizado por variações em TSM e pressão ao nível do mar ao sul e ao norte da ZCIT. O AMM apresenta duas fases, uma fase positiva, AMM (+), apresenta águas mais quentes do que o normal no Atlântico Tropical Norte e águas mais frias no Atlântico Tropical Sul, ocorrendo o oposto na fase negativa, AMM (-). O índice AMM, é obtido combinação das variáveis de TSM e ventos no Atlântico tropical entre 21°S - 32°N e 74°W - 15°E, calcula-se os gradientes meridional de TSM e dos ventos em superfície, com o objetivo de observar as oscilações da ZCIT, que migra para o hemisfério mais quente (Chiang e Vimont [2004]).

As fases do AMM estão associadas com variações na estação chuvosa das regiões Centro-Oeste e Sudeste do Brasil. Quando o AMM está na fase positiva ocorre o atraso na estação chuvosa e término antecipado, diminuindo as chuvas nestas regiões, enquanto que na fase negativa do AMM a estação chuvosa é antecipada e prolongada, aumentando as chuvas nestas regiões (Bombardi e Carvalho [2011]; Cardoso e Wainer [2021]).

Além disso, observa-se que o AMM possui um importante papel na modulação de ciclones (ciclogênese) e nas características da ZCAS em anos de fase neutra do ENSO. Anomalias negativas estão relacionadas ao aumento de ciclogênese próximo ao sudeste do Brasil próximo a região de influência da ZCAS, alterando sua organização, que aumenta a precipitação sobre todo o leste da América do Sul (Wainer e Venegas [2002] Bombardi *et al.* [2014]).



### 2.2.8 Oscilação do Atlântico Norte (NAO)

A Oscilação do Atlântico Norte (NAO) é um fenômeno de variabilidade sazonal e interanual caracterizada por uma alternância da intensidade dos centros de baixas e altas pressões na região da baixa polar na Islândia e a Alta Subtropical dos Açores. Apresenta duas fases, na fase positiva da NAO, observam-se pressões mais baixas próximas a Islândia e pressões mais altas sobre Açores. Na fase negativa a situação se inverte, pressões mais altas próximas a Islândia e pressões mais altas sobre Açores. O índice é composto pelo primeiro modo da EOF derivado de médias mensais de anomalias de altura geopotencial em 500 hPa nas latitudes de 0°N - 90°N (Hurrell *et al.* [2003]).

O índice NAO positivo pode levar ao resfriamento anômalo (pelo mecanismo vento – evaporação e fluxo de calor latente - TSM) das águas superficiais oceânica em boa parte da região (Atlântico Norte), próximo ao TNA e o índice negativo pode levar um aquecimento dessa área (Teixeira [2002]). Além da influência no Atlântico tropical a NAO pode se associar ao ENSO (Siqueira e Molion [2015]). A NAO pode estar associada as chuvas na região centro-leste e sul do Brasil (grande parte da bacia do Paraná), principalmente nas estações de transição (outono e primavera) e com menor intensidade no verão e inverno (Severo [2008]; Siqueira e Molion [2015]).

### 2.2.9 Oscilação Pacífico – América do Norte (PNA) e do Sul (PSA)

O padrão atmosférico PSA (América do Sul-Pacífico ou Pacific-South American Pattern) é um padrão de multi-escala temporal de variabilidade de baixa frequência, caracterizado por trens de onda emanados da costa leste da Austrália em direção ao Pacífico Sul e sobre a América do Sul (Mo [2000]).

Segundo Mo e Higgins [1998], o modo PSA apresenta dois padrões: PSA1 e PSA2 e apresentam períodos entre 16 a 40 dias. O PSA1 está associado com a intensificação da convecção no Pacífico Central entre 140°E e 170°W e a supressão de convecção sobre o oceano Índico, padrão que se assemelha com o do fenômeno ENSO. Já o PSA2 está associado com a intensificação da convecção no Pacífico Central estendendo-se de 160°E a 150°W, próximo ao sul do equador, e com a supressão no Pacífico Oeste. O índice PSA é caracterizado pelo segundo (PSA1) e terceiro (PSA2) modo da EOF utilizando dados de anomalia média mensal da altura geopotencial em 500 hPa ou da função de corrente em 200 hPa (Mo [2000]).

Hill *et al.* [2011] observou o comportamento da precipitação da América do Sul em respostas as anomalias no Oceano Pacífico. Nos subtropicais, as anomalias da precipitação são opostas em sinal entre o Pacífico ocidental e oriental (PSA). Quando o aquecimento da TSM ocorre no oeste (leste) equatorial do Oceano Pacífico, a precipitação é significativamente reduzida (amentada) em torno de 30°S e 50°W (região Sul do Brasil). Bodnariuk *et al.* [2021] avaliou a variabilidade interanual no Rio da Prata, e destacou a influência entre os padrões PSA1 e PSA2 no Sul do Brasil.

Grimm [2003] destacou que o deslocamento da convecção ao longo do Oceano Pacífico modula o padrão gangorra entre as regiões Sul e Sudeste do Brasil, esta combinação é realizada através dos trens de ondas associados com o PSA.

O padrão atmosférico PNA (América do Norte-Pacífico ou Pacific-North American) é um dos principais modos de variabilidade de baixa frequência do Hemisfério Norte, sendo observado em quase todos os meses, exceto em junho e julho. O padrão PNA é caracterizado por um padrão

quadripolo de anomalias na altura de 500 hPa, que ligam as regiões de latitudes baixas às áreas das latitudes médias e altas do Hemisfério Norte. O índice PNA é caracterizado pelo segundo modo da EOF rotacionado utilizando dados de anomalia média mensal da altura geopotencial em 500 hPa nas latitudes de 0 - 90°N (Wallace e Gutzler [1981]). De uma maneira geral, a variabilidade no ENSO, na PDO e na NAO está relacionado com a variabilidade do padrão de teleconexão PNA (Barry e Chorley [2013]).

### 2.2.10 Oscilação Ártica (AO) e Antártica (AAO)

A Oscilação Antártica (AAO; ou Modo Anular Sul, SAM) é definida como a oscilação entre cinturões de pressão de latitudes médias e altas no Hemisfério Sul, apresentando duas fases distintas. Quando a AAO está em sua fase positiva, há uma maior diferença de pressão entre latitudes médias e altas do Hemisfério Sul, uma baixa pressão prevalece em latitudes altas e alta pressão nas latitudes médias. Na fase negativa essa condição se altera, uma alta pressão prevalece em latitudes altas e baixa pressão nas latitudes médias (Gong e Wang [1999]). O índice da AAO é obtido a partir da primeira componente principal das anomalias de altura geopotencial, entre as latitudes 20°S e 90°S (Mo [2000]).

Observa-se que a AAO pode influenciar no posicionamento das trajetórias dos sistemas ciclônicos, afetando o regime de precipitação da América do Sul. Durante a fase negativa, há ocorrência de anomalias positivas de precipitação sobre o Sul do Brasil, principalmente no verão e no outono (Carvalho *et al.* [2005]). No Sudeste do Brasil, nota-se verões muito chuvosos (muito secos) que estão associados à intensificação de trens de ondas do tipo PSA pela fase positiva (negativa) da AAO (Vasconcellos e Cavalcanti [2010]; Machado *et al.* [2021]).

Mendes e Cavalcanti [2014] estudaram a relação entre a AAO e eventos de bloqueios ao longo do Pacífico e Atlântico, verificando que a frequência de dias bloqueados é diferente para cada fase da AAO, apresentando na fase positiva maior duração e maior número de eventos no Pacífico sudoeste e Atlântico sul, havendo uma persistência maior de bloqueios no Pacífico sudeste na fase negativa da AAO. Além disso, observa-se a ação combinada entre o ENSO e AAO, nas fases negativas (positivas) da AAO predominam eventos El Niño (La Niña), havendo uma maior frequência de secas no Rio Grande do Sul quando há a combinação da La Niña com a fase positiva da AAO (Schossler *et al.* [2018]).

A Oscilação Ártica (AO; ou Modo Anular Norte, NAM) é um modo hemisférico de variabilidade que representa as variações dominantes do campo da pressão a Norte do paralelo 20°N sendo caracterizada por anomalias de pressão no Ártico, simultâneas com anomalias da intensidade do jacto polar e correspondentes anomalias simétricas numa faixa de latitudes médias no intervalo 37°N - 45°N (Thompson e Wallace [1998]). A AO é equivalente da AAO no Hemisfério Norte, responsável por induzir anomalias nas características climáticas da América do Norte (Deser [2000]).

Segundo Thompson e Wallace [2000] as fases positivas da AO, em todas as outras estações do ano (exceto verão) estão associadas com anomalias positivas de temperatura em superfície na maior parte da América do Norte e anomalias negativas sobre o extremo leste do Canadá. De forma análoga ao AAO, o índice da AO é obtido a partir da primeira componente principal das anomalias de altura geopotencial, entre as latitudes 20°N e 90°N (Thompson e Wallace [1998]).

Poucos trabalhos sugerem a influência da AO no regime de precipitação na América do Sul.

Ferreira *et al.* [2012] avaliaram a sazonalidade da precipitação na Amazônia e observaram que durante episódios de ENSO, a TSM do Atlântico pode estar associadas com a Oscilação do Ártico, através do centro de alta pressão do Atlântico Norte.

### 2.2.11 Oscilação Quase-Bienal (QBO)

A Oscilação Quase-Bienal (QBO) é uma oscilação periódica do vento zonal equatorial entre os ventos de oeste (anomalia positiva) e de leste (anomalia negativa) na estratosfera tropical com um período médio de 28 a 29 meses (Randel e Wu [2007]). O índice QBO é calculado a partir da média do vento zonal 30mb no Equador (Baldwin *et al.* [2001]). A QBO pode interferir na estratosfera tropical, especificamente na atividade dos ciclones tropicais do Atlântico (Camargo e Sobel [2010]).

Estudos mostram que a interferência construtiva do ENSO na QBO pode conduzir anomalias estratosféricas do vento zonal mais fortes nas latitudes altas de Hemisfério Sul em novembro e dezembro, ocorrendo a desagregação precoce do vórtice polar estratosférico e o início precoce de nuvens mesosféricas polares de verão, durante eventos quentes de ENSO e QBO leste, e o oposto durante a fase fria do ENSO e QBO oeste (Li *et al.* [2016]).

Medina e Nunes [2011] avaliaram a influência da QBO em fase neutra de ENSO (ciclo de dezembro de 1979 a maio de 1982) na América do Sul, e verificaram um aumento da precipitação nas regiões sul e nordeste do Brasil em períodos de ventos de oeste, enquanto que na região sudeste apresentou aumento na fase de vento de leste.

Outros padrões de diferentes variabilidades são influenciados pela QBO. Raphaldini *et al.* [2021] sugerem mecanismos físicos que podem estar associado a interferência da QBO na variabilidade intranual da MJO.

### 2.2.12 Fluxo Solar (SF)

O Sol é a principal fonte de energia para o planeta Terra, sua variabilidade tem impacto na atmosfera e no clima em diferentes escalas e padrões (Gray *et al.* [2010]). Observa-se que as variações no ciclo solar podem causar variações ou mudanças na temperatura, ventos e ozônio estratosférico. Além disso, pode influenciar na circulação troposférica, que afeta os movimentos verticais e consequentemente a convecção, por exemplo, associados a circulação nas células de Hadley e Walker (Van Loon *et al.* [2004]).

A principal variação na atividade solar, ocorre em um período de aproximadamente 11 anos, essas variações são provenientes do número de manchas solares (em inglês, sunspot number, SSN) ou densidade de fluxo solar a 10,7 cm de comprimento (F10.7) e podem afetar a variabilidade de alguns padrões climáticos de alta e baixa frequência (Gray *et al.* [2010]).

Van Loon e Meehl [2014] verificaram respostas distintas nas anomalias de pressão no Oceano Pacífico equatorial e norte relacionadas às flutuações no ciclo de SSN de 11 anos, de acordo com a fase da PDO, sendo também notada a intensificação da NAO positiva (negativa) quando está em fase (fora de fase) com os picos de SSN. A atuação combinada do ciclo solar e da QBO causa efeitos distintos na estratosfera inferior e na troposfera, quando comparado com seus efeitos individuais, sendo significativo em altas latitudes. Os modos polares são fortalecidos quando há uma combinação da baixa atividade solar com a QBO oeste e da alta atividade solar com a QBO leste, ocorrendo o

enfraquecimento no caso de combinações oposta. O sinal da variabilidade solar, modulada pela fase do QBO é detectável na pressão ao nível do mar em latitudes altas (Roy e Haigh [2011]).

Matthes *et al.* [2013] mostrou através de experimentos de sensibilidade com modelo que a resposta solar média anual na estratosfera superior tropical é independente da presença do QBO, mas a resposta na estratosfera inferior difere dependendo da presença do QBO e do ciclo solar. Ambos os forçamentos (solar e QBO) são importantes para gerar a resposta sazonal observada. Maior aquecimento estratosférico ocorre durante o máximo solar e QBO oeste. Foi sugerida uma modulação da fase QBO no sinal solar na estratosfera superior subtropical, sendo realçada uma pequena antecipação do início do sinal solar invernal na QBO leste e reduzido durante a QBO oeste. Consequentemente, há influência na transferência do sinal solar-QBO durante o inverno, resultando em diferenças observadas durante o final do inverno.

Estudos de relações entre as variações na atividade solar e na TSM apresentam correlações significativas, principalmente em três áreas do oceano Pacífico, com maiores correlações positivas entre julho a dezembro, que correspondem aos padrões da PDO, do ENSO clássico e Modoki, sendo destacado também relações com a pressão atmosférica e a circulação (Yamakawa *et al.* [2016]; Lihua e Zhiqiang [2017]; Zhai [2017]). Além disso, estudos sugerem El Niño mais fraco e com menor frequência em períodos de maior atividade solar (Zhai [2017]; Lin *et al.* [2021]).



# Capítulo 3

## Padrões de Variabilidade Climática

Neste capítulo são realizadas análises e discussões sobre as características dos padrões de variabilidade climática no período instrumental disponível por agências governamentais de diversos países (Estados Unidos, União Europeia, Japão e Austrália), entretanto, alguns índices relevantes para caracterizar o padrão de variabilidade climática no Brasil não se encontram disponíveis na *web*. Neste caso, foram desenvolvidos índices climáticos através de variáveis atmosféricas conforme a literatura.

As análises a seguir auxiliaram no entendimento dos padrões quanto a sua variabilidade temporal e frequência dos acontecimentos.

### 3.1 Dados

Os dados de indicadores climáticos foram obtidos dos sites da National Oceanic and Atmospheric Administration (NOAA, <https://www.esrl.noaa.gov/psd/data/climateindices/list/>), Japan Agency for Marine-Earth Science and Technology (JAMSTEC, <http://www.jamstec.go.jp/aplinfo/sintexf/e/index.html>), European Climate Assessment & Dataset project (ECA&D, <https://www.ecad.eu/> ou <https://climexp.knmi.nl>) e Bureau of Meteorology- Australian Government (BOM, <http://www.bom.gov.au/>). Nestes sites são disponibilizados dados mensais e/ou diários de índices climáticos que representam diversos padrões de variabilidade climática de diversas escalas (intrasazonal, interanual, decadal ou multidecadal). A Tabela 3.1 indica os índices obtidos pelos órgãos citados, o período utilizado e a referência do cálculo do índice.

Alguns índices relevantes para caracterizar o padrão de variabilidade climática no Brasil não se encontram disponíveis na *web*. Neste caso, foram desenvolvidos índices climáticos conforme a literatura destacada na Tabela 3.2.

Para o PSA foram gerados os índices conforme a referência de Mo [2000] que caracteriza o índice para a Oscilação Pacífico América do Sul a partir da segunda (PSA1) e terceira (PSA2) componente da EOF utilizando dados de anomalia média mensal da altura geopotencial em 500 hPa ou da função de corrente em 200 hPa, no domínio do hemisfério sul (90°S - Equador).

Da mesma forma, os índices LISAM e ZCAS foram desenvolvidos segundo a metodologia de Silva e Carvalho [2007], baseado no primeiro e segundo modo da EOFc (Zang e Moore [2015]) dos dados de precipitação estimada por satélite do GPCP (*Global Precipitation Climatology Project*) e

**Tabela 3.1:** Índices climáticos mensais e/ou diários disponibilizados por agências governamentais e seus respectivos períodos e referências.

Sigla	Índice	Período	Referência
AAO	Oscilação Antártida	1851 a 2018	Gong e Wang [1999]
AMM	Modo Meridional do Atlântico	1948 a 2018	Chiang e Vimont [2004]
AMO	Oscilação Multidecadal do Atlântico	1871 a 2018	Enfield <i>et al.</i> [2001]
AO	Oscilação Ártica	1950 a 2018	Thompson e Wallace [1998]
FS	Fluxo Solar (10.7cm)	1948 a 2018	Gray <i>et al.</i> [2010]
GTA	Gradiente Interhemisférico de TSM do Atlântico Tropical	1948 a 2018	Servain [1991]
MEI	Índice Multivariado do ENSO	1950 a 2018	Wolter e Timlin [1998]
MJO	Oscilação de Madden-Julian	1979 a 2018	Wheeler e Hendon [2004]
NAO	Oscilação do Atlântico Norte	1821 a 2018	Hurrell <i>et al.</i> [2003]
NIÑO 3.4	TSM na região 3.4 do ENSO	1870 a 2018	Rayner <i>et al.</i> [2003]
PDO	Oscilação Decadal do Pacífico	1900 a 2018	Mantua <i>et al.</i> [1997]
PNA	Oscilação Pacífico América do Norte	1899 a 2018	Trenberth e Hurrell [1994]
QBO	Oscilação Quase-Bienal	1948 a 2018	Baldwin <i>et al.</i> [2001]
SPI	Índice Sunspot	1749 a 2018	Gray <i>et al.</i> [2010]
SOI	Índice Oscilação Sul	1866 a 2018	Ropelewski e Jones [1987]
TNA	Anomalias de TSM no Atlântico Tropical Norte	1948 a 2018	Enfield <i>et al.</i> [1999]
TSA	Anomalias de TSM no Atlântico Tropical Sul	1948 a 2018	Enfield <i>et al.</i> [1999]

**Tabela 3.2:** Índices climáticos mensais e/ou diários desenvolvidos e seus respectivos períodos e referências.

Sigla	Índice	Período	Referência
LISAM	Índice de Larga Escala da Monção da América do Sul	1979 a 2018	Silva e Carvalho [2007]
PSA	Oscilação Pacífico América do Sul	1979 a 2018	Mo [2000]
ZCAS	Zona de Convergência do Atlântico Sul Tropical	1979 a 2018	Silva e Carvalho [2007]

dados em 850 hPa de temperatura do ar, umidade específica e componentes zonal e meridional do vento.

Na próxima seção são detalhadas as etapas de desenvolvimento dos índices PSA1/PSA2 e LISAM/ZCAS.

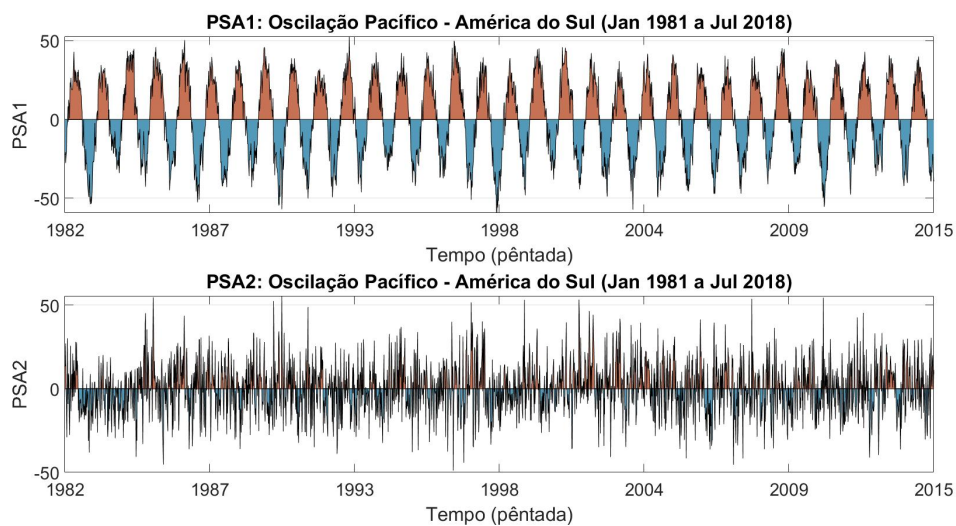
## 3.2 Metodologia

### 3.2.1 Desenvolvimento dos Índices Climáticos

Para o desenvolvimento dos índices climáticos PSA1/PSA2 (Mo [2000]) e LISAM/ZCAS (Silva e Carvalho [2007]) foi utilizada a técnica de EOF. A EOF é uma técnica estatística multivariada que transforma linearmente um conjunto de  $p$  variáveis em um conjunto com um número menor de variáveis ( $k$ ), que explicam uma grande parcela da variabilidade do conjunto inicial. A redução do número de variáveis é realizada pela construção de novas variáveis sintéticas, obtidas pela combinação linear das variáveis iniciais. As combinações são escolhidas para maximizar a variância original dos dados. Isto é, dado múltiplas observações  $x$  ( $K \times 1$ ), a EOF determina um vetor  $u$  ( $M \times 1$ ) cujos elementos são combinações lineares de elementos de  $xs$ , que contém a maior parte das informações do conjunto inicial de dados, esses novos vetores  $u$  são chamados componentes principais (em inglês, *principal component* - PCs). Para que a análise tenha eficiência máxima a compressão dos dados pode ser atingida com  $M \ll K$ , esta situação é atingida quando existem correlações substanciais entre as variáveis de  $x$ , ou seja, caso onde as variáveis contemplem informações redundantes (Zang e Moore [2015]). Para mais informações sobre o cálculo da EOF, disponível no Apêndice D.

Segundo Mo [2000], o modo PSA é caracterizado pelo segundo (PSA1) e terceiro (PSA2) modo da EOF. Neste estudo, foram utilizados dados de reanálise (ERA-Interim) de altura geopotencial média diária em 500 hPa para no período de 1979 a 2018 para o hemisfério sul (90°S - Equador), em uma grade regular de 2,5° x 2,5° (dado original disponível 0,5° x 0,5°, como índice representa um padrão de grande escala, optou-se por diminuir a resolução espacial por eficiência computacional), disponível pelo ECMWF (*European Centre for Medium-Range Weather Forecasts*) em <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets>. Os dados foram transformados em anomalias diárias e acumulados em pântadas e/ou mensais de altura geopotencial em 500 hPa e utilizados para a análise de EOF (Zang e Moore [2015]). Como resultado, são obtidos os índices PSA1 e PSA2, representados pela série temporal (em inglês, *scores*) da segunda e terceira componente principal, respectivamente EOF2 e EOF3.

A Figura 3.1 mostra os índices PSA1 e PSA2 no período de janeiro de 1981 a julho de 2018 em pântadas, os valores em vermelho indicam anomalias positivas, enquanto os valores em azul as anomalias negativas. Na Figura 3.2, o padrão médio do período analisado dos índices PSA1 e PSA2 são representados.



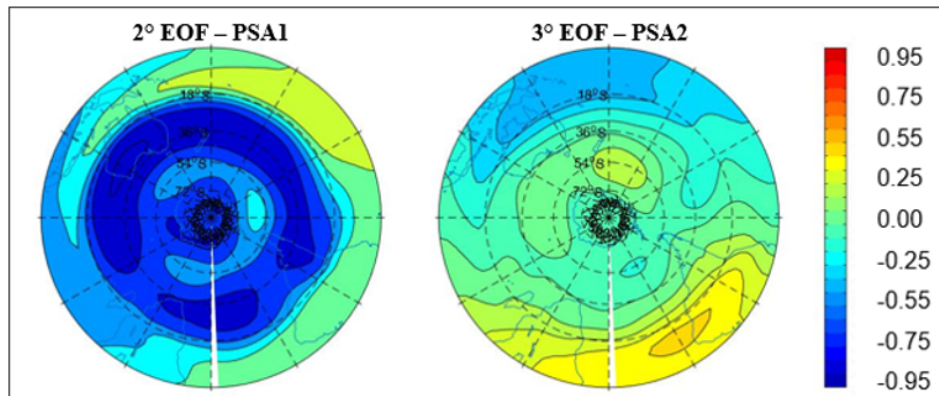
**Figura 3.1:** Índice América do Sul-Pacífico (PSA) anomalias de altura geopotencial em 500hPa em reanálise do ERA-Interim, segundo modo PSA1 (EOF2 - 13% da variância explicada) e terceiro modelo PSA2 (EOF3 - 5% da variância explicada) no período de janeiro de 1981 a julho de 2018 em pântadas. Os valores em vermelho representam o índice PSA positivo e em azul negativo.

O segundo modo PSA1 possui variância explicada de 13% e o terceiro modo PSA2 5%. Esses resultados obtidos neste trabalho corroboram com o padrão PSA, caracterizado pelos índices PSA1 e PSA2 observados por Mo [2000].

Os índices LISAM e ZCAS foram desenvolvidos segundo a metodologia de Silva e Carvalho [2007], baseado no primeiro e segundo modo da EOFc (Zang e Moore [2015]). A principal diferença entre uma análise de EOF x EOFc, é que na primeira utiliza-se a matriz de covariância para medir o grau de relacionamento linear entre duas variáveis com dimensões iguais, enquanto a segunda utiliza a matriz de correlação, ideal para variáveis de dimensões ou escalas distintas, metodologia completa da EOF disponível no Apêndice D.

Foram utilizados dados diários em 850 hPa de temperatura do ar, umidade específica e compo-





**Figura 3.2:** Média dos valores do segundo e terceiro modo no período de 1981 a 2018, em azul indicam anomalias negativas de altura geopotencial enquanto os valores positivos indicam anomalias positivas (figuras obtidas com os autovetores da EOF2 e EOF3).

mentos zonal e meridional do vento da reanálise do ERA-INTERIM para o período de 1979 a 2018. O primeiro modo LISAM possui variância explicada 34% e o segundo modo ZCAS 13%. Os resultados obtidos neste trabalho corroboram com os padrões observados por [Silva e Carvalho \[2007\]](#).

Detalhes sobre a variabilidade climática dos índices desenvolvidos serão apresentados neste capítulo.

### 3.2.2 Variabilidade Climática dos Padrões

Para compreender a variabilidade climática dos padrões de alta e baixa frequência no período disponível dos dados, conforme apresentado nas Tabelas 3.1 e 3.2, foi utilizada a transformada em ondeletas (em inglês, *wavelet*) ([Torrence e Compo \[1998\]](#)) nas séries temporais dos índices climáticos após remoção do ciclo anual, conforme a metodologia de padronização de séries temporais, disponível no Apêndice A.

A transformada em ondeletas é utilizada para transformar uma função no domínio do tempo em uma nova função que converte nas escalas de frequência e tempo, possibilitando identificar o tempo em que ocorre as variabilidades de maior frequência da série, obtendo um diagnóstico mais preciso das variáveis a serem estudadas, para mais informações detalhadas do método Apêndice B.

Para avaliar se os índices climáticos possuem relação entre si, foi realizada a análise de correlação simples, calculando o coeficiente  $r$  de Pearson (Apêndice C) entre as séries dos índices climáticos, testada a significância da correlação utilizando o teste t-Student ao nível de significância de 5%. Porém, a análise de correlação é uma relação simplificada entre os índices, a correlação não implica em causalidade, ou seja, relação de causa e efeito, entre os índices climáticos.

Neste caso, utilizam-se métodos que identificam causalidade, isto é relação de causa e efeito entre as variáveis. Inicialmente [Granger \[1969\]](#) propôs um método que busca causalidade entre duas variáveis, a partir de modelos lineares vetoriais autorregressivos (em inglês, *Vector Autoregressive*, VAR).

De uma forma geral, os modelos avaliam quando uma variável  $X$  causa-Grange uma variável  $Y$  se, na média, o evento  $Y$  é verificado toda vez que o evento  $X$  ocorreu algum período antes. Uma das principais propriedades da causalidade de Granger é que ela é direcional, ou seja, a série  $X$

causa a série  $Y$ , não implica em  $Y$  causar  $X$ .

No domínio da frequência a Causalidade de Granger pode ser inferida a partir da Coerência Parcial Direcionada (em inglês, *Partial Directed Coherence*, PDC) (Baccalá e Sameshima [2001]; Baccalá e Sameshima [2021]) que permite detecção de conexões diretas entre as variáveis lineares. Além disso, as variáveis climáticas em questão possuem comportamento não-linear, então optou-se por também utilizar a Coerência Parcial Direcionada não-linear kernelizada (em inglês, *kernel-nonlinear Partial directed Coherence*, *knPDC*), método que transforma a série não-linear, em “quase” linear (Massaroppe e Baccalá [2019]), a descrição da metodologia se encontra no Apêndice E.

Para atingir significância estatística, o método consegue capturar relações entre um conjunto reduzido de variáveis, como são dezessete índices climáticos foi necessário reduzir o conjunto de variáveis de entrada. Para definir quais variáveis seriam testadas em conjunto foi utilizada a análise de componentes principais ou EOF, disponível no Apêndice D.

Na etapa anterior, os índices foram avaliados separadamente no período de dados disponível (Tabelas 3.1 e 3.2), enquanto na análise de causalidade o período escolhido deve coincidir para todos os índices. Os dados foram padronizados (Apêndice A) no período selecionado de 1979 a 2018 e aplicada a técnica de EOF para selecionar os índices climáticos que apresentam maiores pesos em cada componente (Apêndice D).

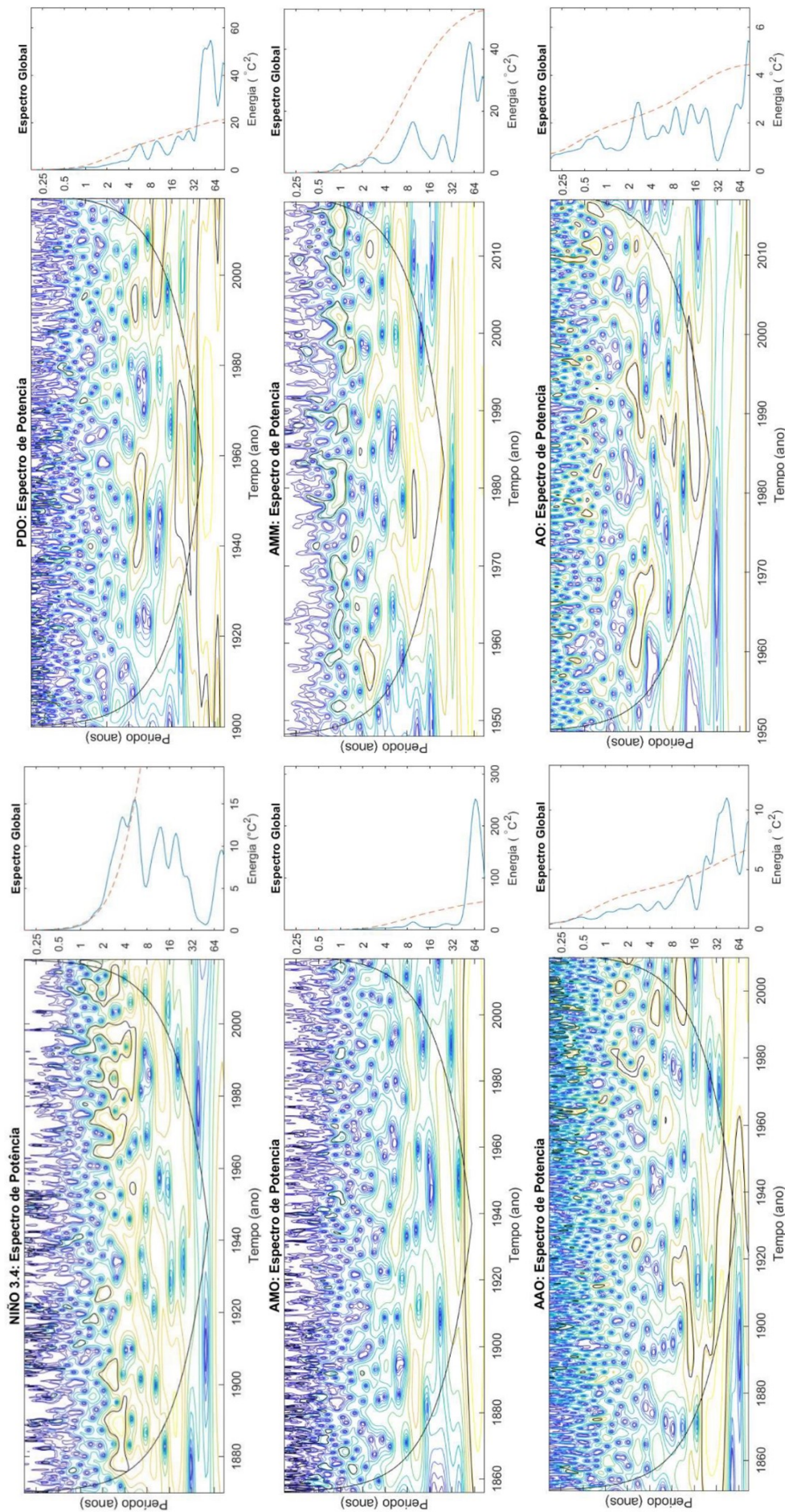
Para o método PDC e *knPDC*, os índices selecionados foram padronizados sem remover o ciclo anual, ou seja, média e desvio padrão da série completa dos dados. No caso do índice que caracteriza o fluxo solar (SPI) aplicou-se filtro de *denoising* pelo método de *wavelet* (Donoho e Johnstone [1994]), pois o sinal do SPI apresenta uma relação sinal-ruído muito alto, que compromete as análises de causalidade.

### 3.3 Resultados e Discussão

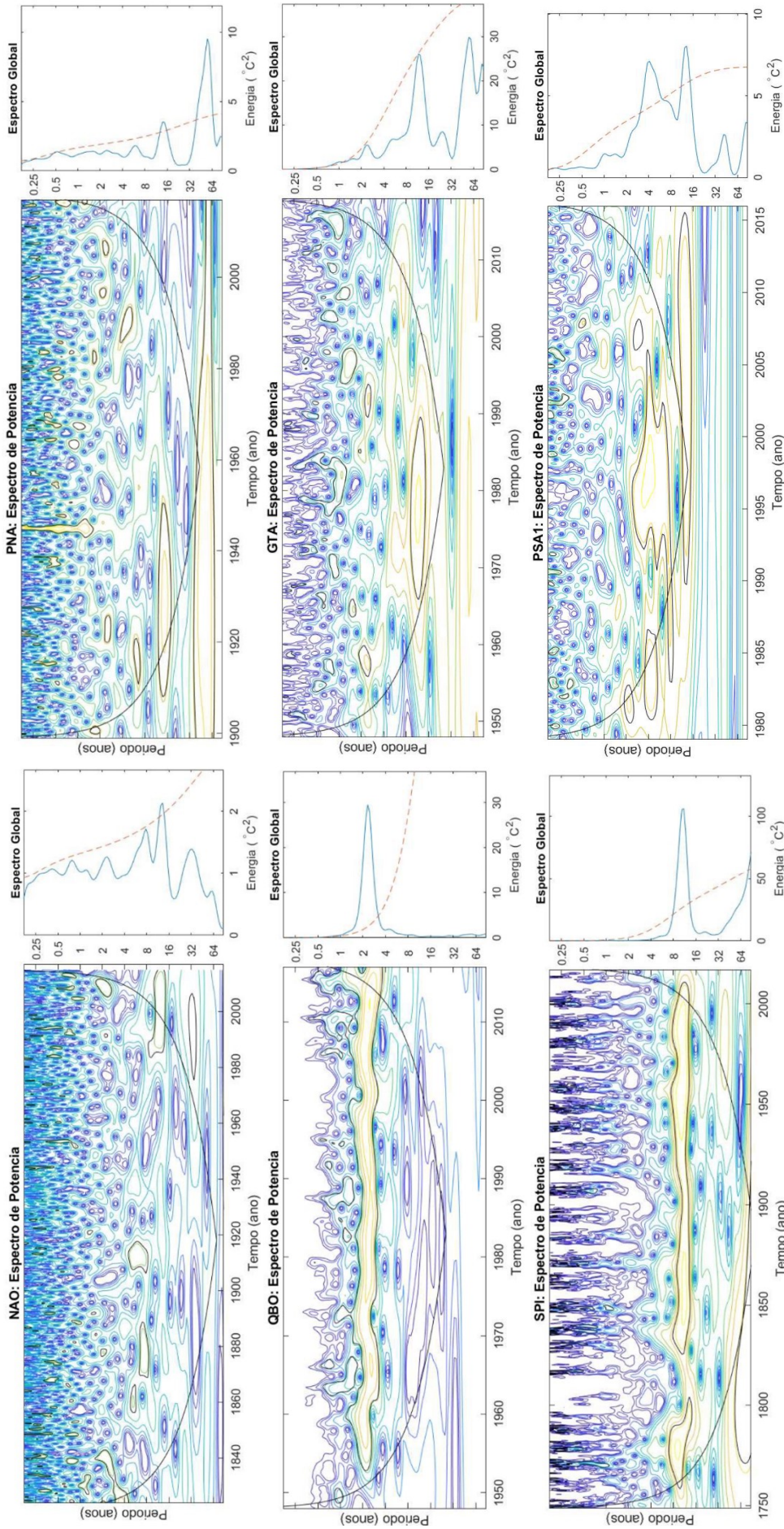
A variabilidade dos padrões climáticos de baixa e alta frequência foi investigada através da transformada de *wavelet* das séries temporais dos índices climáticos. A análise de *wavelet* fornece informações sobre a escala de variabilidade temporal dos padrões de interesse, determinando os principais picos espectrais ao longo do período de análise (Torrence e Compo [1998]).

As Figuras 3.3, 3.4 e 3.5 apresentam as análises de *wavelet* para o período de dados disponível de cada índice, Tabelas 3.1 e 3.2, após a remoção do ciclo anual. Nestas figuras, o espectro de potência da transformada de *wavelet* (Morlet) é delimitado pelo cone de influência, o contorno preto indica o setor do espectro de potência onde o nível de confiança de 95% é satisfeito. O espectro de potência global (*Global Wavelet Spectrum*) indica o nível de potência satisfeito para toda a série temporal, ou seja, acima da linha vermelha tracejada.

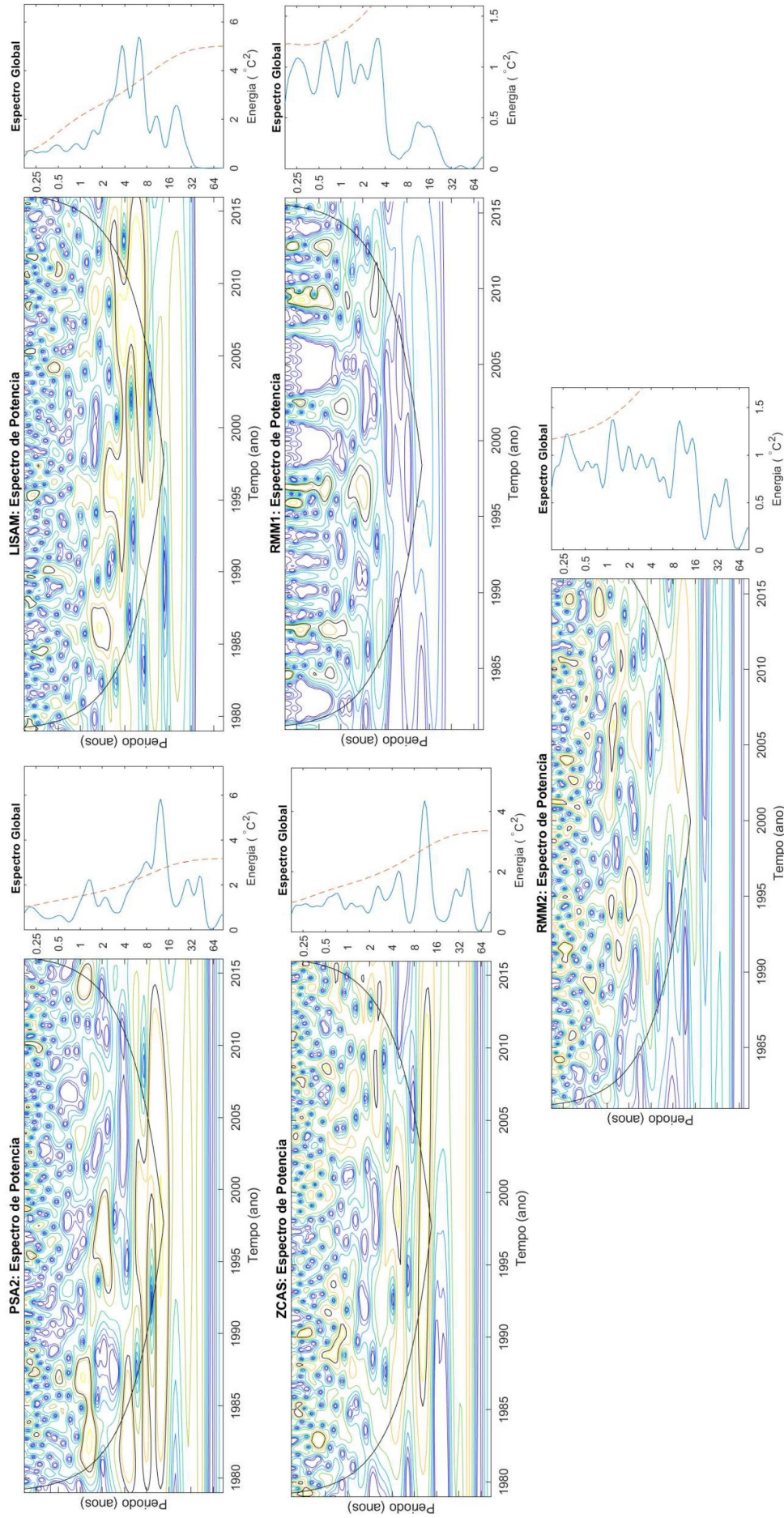
De uma maneira geral, alguns índices apresentaram pico espectral significativo em 11 anos, sugerindo a influência da atividade solar, ou seja, o ciclo solar de 11 anos (índice SPI, Figura 3.4 (k)) em padrões climáticos, como: PDO, AAO, NAO, PNA, PSA1/PSA2, GTA e ZCAS, respectivamente Figura 3.3 (b) e (e), Figura 3.4 (g), (h), (j), (l) e Figura 3.5 (m) e (o), os resultados obtidos coincidem com outros trabalhos (Gray *et al.* [2010]; Van Loon e Meehl [2014] e Lihua e Zhiqiang [2017]).



**Figura 3.3:** Espectro de Potência e Global da análise de wavelet da série temporal dos índices (Morlet) a) Niño 3.4; b) PDO; c) AMO; d) AMM; e) AAO; f) AO. O espectro de potência é delimitado pelo cone de influência, o contorno preto indica o setor de potência onde o nível de confiança de 95% é satisfeito. O espectro de potência global indica o nível de potência satisffeito para toda a série temporal, ou seja, acima da linha vermelha tracejada.



**Figura 3.4:** Espectro de Potência e Global da análise de wavelet da série temporal dos índices g) NAO; h) PNA; i) QBO; j) GTA; k) SPI; l) PSA1. O espectro de potência é delimitado pelo cone de influência, o contorno preto indica o nível de potência onde o nível de confiança de 95% é satisfeito. O espectro de potência global indica o nível de potência satisffeito para toda a série temporal, ou seja, acima da linha vermelha tracejada.



**Figura 3.5:** Espectro de Potência e Global da análise de wavelet da série temporal dos índices m) PSA2; n) LISAM; o) ZCAS; p) RMM1 e q) RMM2. O espectro de potência é delimitado pelo cone de influência, o contorno preto indica o nível de confiança de 95% é satisffeito. O espectro de potência global indica o nível de potência satisffeito para toda a série temporal, ou seja, acima da linha vermelha tracejada.

Para caracterizar o fenômeno ENSO foi aplicada a análise de *wavelet* nos índices Niño 3.4, MEI e SOI. Foram observadas periodicidades significativas entre 1 e 5 anos, com picos máximos em 3 e 5 anos para o índice Niño 3.4 (Figura 3.3 (a)) e MEI (Figura I.1 (a), disponível no Apêndice I), enquanto o índice SOI (Figura I.1 (b), disponível no Apêndice I), os valores significativos apresentam uma faixa mais extensa, entre 2 e 7 anos, com picos em 3,5 e 6 anos. Esses resultados caracterizam o ENSO como fenômeno de variabilidade interanual entre 2 a 7 anos, que corresponde ao obtido na bibliografia (Trenberth [2017]).

Para o índice PDO (Figura 3.3 (b)) foram observados valores significativos acima de 30 anos, indicando variabilidade decadal, porém na literatura sugere variações entre 20 a 30 anos (Mantua *et al.* [1997]), situação que ocorre no período analisado, entre os anos de 1940 a 1980. Outra variabilidade de baixa frequência ocorrendo em períodos mais longos, é a AMO que apresenta variabilidade multidecadal entre 30 a 60 anos (Enfield *et al.* [2001]), como mostrado na Figura 3.3 (c).

Para o índice AAO, quando se utiliza o período instrumental (1979 a 2018) não são verificados valores significativos, indicando a necessidade de uma maior quantidade de dados na realização da análise de *wavelet*. Já para o valor modelado da AAO (20CRV2c) (Figura 3.3 (e)) no período de 1851 a 2011, observa-se valores significativos para valores inferiores a 0,5 ano e acima de 20 anos, com maior energia no período de 20 a 40 anos, sugerindo relação da AAO com a PDO (Pezza *et al.* [2007]). A AO apresenta variabilidade diferente da sua equivalente AAO, com picos em 2,5 anos e 80 anos (Figura 3.3 (f)), indicando relações entre episódios de ENSO e TSM do Atlântico (Ferreira *et al.* [2012]).

Para o índice PNA (Figura 3.4 (h)) observa-se periodicidade significativa em 11 anos (ciclo solar) e entre 30 a 60 anos, indicando forte influência da AMO (Wallace e Gutzler [1981]). O padrão similar ao PNA no Hemisfério Sul, o PSA, apresenta duas componentes PSA1 e PSA2, com picos entre 3 a 5 anos e 1,5 ano respectivamente. A primeira (PSA1), apresenta forte associação aos padrões ENSO, enquanto a segunda (PSA2) com o AAO, conforme sugerido por Yu *et al.* [2015].

Para os índices AMM (Figura 3.3 (d)), TNA e TSA (respectivamente, Figura I.1 (c) e (d), disponível no Apêndice I) e o GTA, observa-se variabilidade interanual que variam entre 0,5 a 2 anos (Enfield *et al.* [1999]; Chiang e Vimont [2004]). Esses padrões têm um papel importante na modulação de fenômenos intrasazonais, principalmente na modulação da ZCAS (Chaves [2011]; Bombardi *et al.* [2014]), conforme o pico espectral do índice ZCAS (Figura 3.5 (o)).

Os fenômenos intrasazonais, LISAM, ZCAS, RMM1 e RMM2, apresentam escalas intrasazonais que variam de 4 a 8 meses, respectivamente Figura 3.5 (n) a (q). Além disso, é possível observar picos ao longo dos anos entre 3 a 7 anos, representa a influência do fenômeno ENOS no SMAS (Grimm [2003]; Grimm [2004]; Garreaud *et al.* [2009]) e MJO (Shimizu e Ambrizzi [2015]). Outro pico interessante encontrado para os índices RMM1/RMM2, com variação entre 1 a 3 anos que podem estar associados a relações não-lineares entre a MJO e a QBO (Raphaldini *et al.* [2021]).

E, por fim, o índice QBO apresenta periodicidade significativa entre 1 a 3 anos, com pico máximo em 2,3 anos (Figura 3.4 (i)), como demonstrado por Randel e Wu [2007]. Na literatura foi ressaltada a influência da atividade solar na QBO (Roy e Haigh [2011]; Matthes *et al.* [2013]), porém na análise de *wavelet* não foi considerada significativa a periodicidade do ciclo solar de 11 anos.

Após determinar a frequência dos eventos dos principais padrões de variabilidade climática, a partir da análise de *wavelet*, foi investigado o coeficiente de correlação de Pearson (CC) entre os

índices. Para isso foi considerado o nível de significância de 5%, a partir do teste t de Student ( $CC \geq 0,10$ ).

Conforme a Tabela I.1, disponível no Apêndice I, foram observadas altas correlações ( $CC > 0,65$ ) entre os índices AMM, AMO, TNA e GTA, todos esses padrões são observados ao norte do Oceano Atlântico e podem influenciar em diversas bacias brasileiras (Chiessi *et al.* [2009]; Bombardi e Carvalho [2011]; Pinto Jr [2015]). Além disso, outras oscilações, como NAO e AO são altamente correlacionadas ( $CC = 0,60$ ) e exerce maior influência no Hemisfério Norte (Deser [2000]; Hurrell *et al.* [2003]).

Para o Oceano Pacífico são observadas correlações moderadas (entre  $CC = 0,35$  a  $0,47$ ) entre o ENSO, expresso por todos os índices analisados (Niño 3.4, MEI e SOI) com a PDO, destacando a correlação entre o índice multivariado (MEI),  $CC = 0,47$  (Garcia e Kayano [2008]; Kayano *et al.* [2009]). Já para o padrão PNA são observadas correlações significativas ( $CC > 0,10$ ) com a AO e a PDO (Barry e Chorley [2013]).

Os índices intrasazonais apresentam correlações significativas com os índices PSA, destaques para o LISAM ( $CC = 0,93$ ) e ZCAS ( $CC = 0,41$ ) com o primeiro padrão (PSA1), RMM1 ( $CC = 0,25$ ) e RMM2 ( $CC = 0,26$ ) com o segundo padrão (PSA2) que corroboram com os estudos de Arias *et al.* [2015] e Mori e Watanabe [2008]. Além disso, o índice RMM1 apresenta correlação significativa com o SOI, indicando influencia do fenômeno ENSO com a MJO, conforme observado por Shimizu e Ambrizzi [2015].

Correlações significativas ( $CC > 0,10$ ) com a atividade solar foram observadas para os índices AMM, AMO, AO, NAO, PDO, TNA e TSA, com destaque para as altas correlações com os padrões dos trópicos (TNA e TSA), que corroboram com diversos estudos que sugerem a influência do fluxo solar em alguns padrões climáticos (Yamakawa *et al.* [2016]; Lihua e Zhiqiang [2017]; Zhai [2017]).

Conforme os resultados obtidos nas análises do coeficiente de correlação, optou-se por reduzir o número de variáveis, primeiramente foram removidos índices que apresentam informações redundantes, como Niño 3.4, MEI e SOI, ou índices que a partir da revisão bibliográfica não foram identificadas fortes influências na precipitação em bacias brasileiras, como NAO, AO e PNA. Além disso, optou-se por transformar os índices TNA e TSA, em um gradiente (GTA) para agregar as informações. E, por fim, optou-se por analisar apenas índices de variabilidade de baixa frequência. Portanto, os índices AAO, AMO, Niño 3.4, PDO, QBO, SPI, GTA e PSA1/PSA2 foram selecionados para as análises de causalidade entre os padrões climáticos de baixa frequência.

A Tabela 3.3 mostra os pesos de cada índice e a variância explicada das componentes principais, utilizada para selecionar os padrões climáticos na análise de causalidade. Geralmente na EOF o ciclo anual dos dados é removido, porém decidiu-se não remover o ciclo anual para preservar as relações entre os índices.

Na Tabela 3.3 observa-se que as duas primeiras componentes apresentam variância explicada similares, juntas elas explicam quase 40% da variabilidade, enquanto as três componentes seguintes possuem variância explicada menor, mas igualmente significativas. As cinco componentes principais somadas resultam em uma variância explicada de 72,92%, portanto as variáveis que apresentam maior peso (acima de 0,30) nessas componentes são adotadas para a análise de causalidade. Para tanto, nesta análise foram introduzidas no algoritmo as seguintes relações: 1) AMO, Niño 3.4, PDO e GTA (maiores pesos das componentes CP1 e CP2); 2) AAO, PSA1, PSA2 e PDO (merge entre

**Tabela 3.3:** Pesos dos índices climáticos selecionados para a análise de causalidade e variância explicada (VE%) para cada componente principal.

Índice	CP1	CP2	CP3	CP4	CP5
VE(%)	18,32	16,83	14,46	12,48	10,83
AAO	0,30	-0,22	<b>0,56</b>	0,06	-0,11
AMO	<b>0,54</b>	<b>0,40</b>	-0,05	0,08	0,21
Niño3.4	<b>-0,33</b>	<b>0,56</b>	0,17	0,13	-0,06
PDO	<b>-0,44</b>	<b>0,39</b>	<b>0,36</b>	-0,01	-0,27
QBO	-0,14	-0,09	0,16	<b>0,50</b>	<b>0,70</b>
SPI	-0,05	-0,22	<b>-0,40</b>	<b>0,49</b>	<b>-0,52</b>
GTA	<b>0,40</b>	<b>0,44</b>	-0,11	<b>0,44</b>	-0,16
PSA1	-0,28	-0,20	0,08	<b>0,52</b>	0,02
PSA2	0,26	-0,18	<b>0,56</b>	0,15	-0,29

as componentes CP3 e CP4).

As Figuras 3.6 (a) e (b) mostram os resultados da PDC/ $knPDC$  para os índices AMO, Niño 3.4, PDO e GTA, enquanto as Figuras 3.7 (a) e (b) os índices AAO, PSA1, PSA2 e PDO. O traçado em preto representa a (pseudo-) densidade espectral das séries em  $dB$ ; o traçado em vermelho representa os valores estatisticamente significantes da PDC/ $knPDC$ ; o traçado preto descontínuo representa o limiar da aproximação de Patnaik (Baccalá *et al.* [2013]), com nível de significância igual a  $\alpha$ ; o traçado em verde os valores estatisticamente não-significantes da PDC/ $knPDC$  e; os valores do eixo  $x$  em anos. A direção da causalidade é observada do eixo  $x$  (abscissa) para o eixo  $y$  (ordenada).

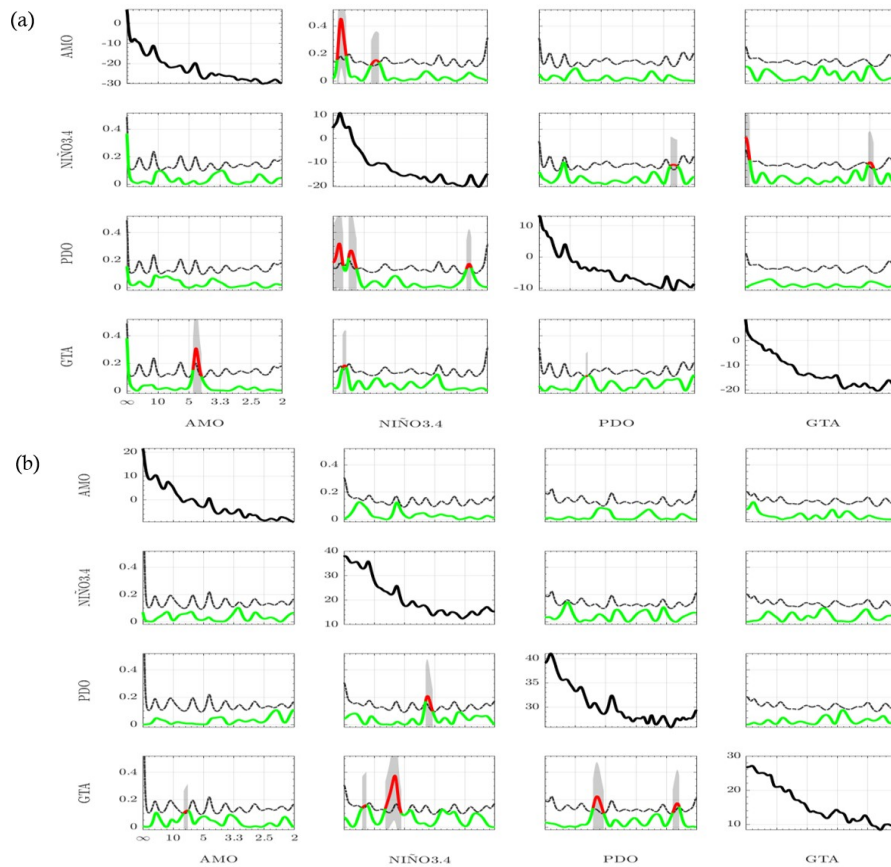
Na Figura 3.6 (a) observa-se relação de causalidade linear (PDC) entre: AMO  $\rightarrow$  GTA; Niño 3.4  $\rightarrow$  AMO; Niño 3.4  $\rightarrow$  PDO; Niño 3.4  $\rightarrow$  AMO; Niño 3.4  $\rightarrow$  GTA; PDO  $\rightarrow$  GTA. A causalidade não-linear (Figura 3.6 (b)) é observada entre Niño 3.4  $\rightarrow$  GTA; Niño 3.4  $\rightarrow$  PDO; PDO  $\rightarrow$  GTA.

Conforme a Figura 3.6, nota-se que o ENSO influencia todos os outros padrões de variabilidade climática analisados (AMO, PDO e GTA), estudos sugerem teleconexões extratropical entre os oceanos Atlântico e Pacífico, onde o ENSO exerce maior influência no Atlântico Norte (Rodríguez-Fonseca *et al.* [2016]; García-Serrano *et al.* [2017]).

Trabalhos similares observaram a relação de causalidade entre o Pacífico (Niño 3) e o Atlântico (TNA) (Builes-Jaramillo *et al.* [2018]). Em adição, a Figura 3.6 mostra também que a PDO possui relação de causalidade não-linear com o GTA, evidenciando que o Pacífico exerce influência no Atlântico. Além disso, os resultados obtidos da PDC e  $knPDC$  sugerem que o ENSO causa a PDO e são relacionados linearmente (Figura 3.6 (a)) e não-linearmente (Figura 3.6 (b)). Estudos teóricos corroboram com os resultados obtidos (Ramirez *et al.* [2018]).

Na Figura 3.7 observa-se relação de causalidade linear e não linear (PDC/ $knPDC$ ) entre: AAO  $\leftrightarrow$  PSA1 e AAO  $\leftrightarrow$  PSA2. E não linear ( $knPDC$ ) entre: PDO  $\rightarrow$  AAO. Os resultados da PDC/ $knPDC$  indicam causalidade entre a PDO e AAO, observa-se que para a  $knPDC$  os valores significativos estão na escala de baixa frequência, indicando que a PDO causa a AAO não linearmente em escala longa. Alguns estudos sugerem a influência da PDO na AAO, no verão e final do inverno (Pezza *et al.* [2007]; Goodwin *et al.* [2016]).





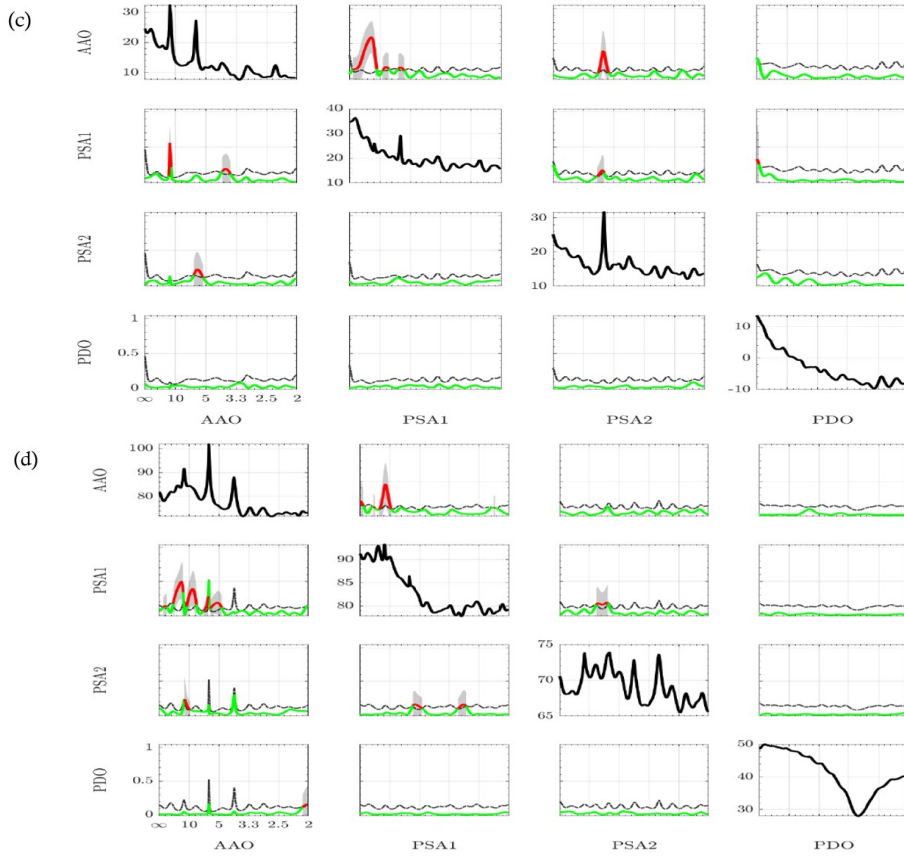
**Figura 3.6:** O gráfico preto representa a (pseudo-) densidade espectral da série, em série em dB; o gráfico vermelho representa os valores PDC / knPDC estatisticamente significativos; a linha preta tracejada representa a aproximação do limiar de Patnaik *Baccalá et al. [2013]*; a linha verde os valores PDC / knPDC estatisticamente não significativos. Portanto, usando o nível de significância de 1%, os números representam, respectivamente: (a) PDC para o conjunto de séries temporais (AMO, Niño 3.4, PDO e GTA) usando um modelo autorregressivo de ordem  $p = 48$ , (b) knPDC para o conjunto de séries temporais (AMO, Niño 3.4, PDO e GTA), usando o kernel polinomial  $[k(x, y) = (x \cdot y)^2]$  e o modelo kernel-autoregressivo de ordem  $p = 24$ .

### 3.4 Conclusões

Neste capítulo foi possível verificar as características dos padrões de variabilidade climática para o período instrumental, as análises auxiliaram no entendimento dos padrões quanto a sua variabilidade temporal e frequência dos acontecimentos. Além disso, foi possível observar como os padrões se relacionam, determinando causalidade entre eles.

A análise de *wavelet* mostrou as diferentes variabilidades dos padrões climáticos, sendo destacada a influência da atividade solar nos padrões PDO, AAO, NAO, PNA, PSA1/PSA2, GTA e ZCAS. Os resultados mostraram a periodicidade do ENSO (2 a 7 anos), PDO (20 a 30 anos), AMO (entre 30 a 60 anos), AAO, AMM, TNA e TSA (0,5 a 2 anos), QBO (2,3 anos), PSA1 (3 a 5 anos) e PSA2 (1,5 ano). Em adição, foram observadas relações entre os índices AAO e PDO (pico de 20 anos na AAO).

Os fenômenos intrasazonais, LISAM, ZCAS, RMM1 e RMM2, apresentam escalas intrasazonais que variam de 4 a 8 meses, respectivamente Figura 3.5 (n) a (q). Além disso, é possível observar picos ao longo dos anos entre 3 a 7 anos, representa a influência do fenômeno ENOS no SMAS.



**Figura 3.7:** O gráfico preto representa a (pseudo-) densidade espectral da série, em série em dB; o gráfico vermelho representa os valores PDC / knPDC estatisticamente significativos; a linha preta tracejada representa a aproximação do limiar de Patnaik (Baccalá et al., 2013); a linha verde os valores PDC / knPDC estatisticamente não significativos. Portanto, usando o nível de significância de 1%, os números representam, respectivamente: (a) knPDC para o conjunto de séries temporais (AAO, PSA1, PSA2 e PDO) e usando um modelo autoregressivo de ordem  $p = 48$ , (b) knPDC para o conjunto de séries temporais (AAO, PSA1, PSA2 e PDO) e usando o kernel polinomial  $[k(x, y) = (x \cdot y)^2]$  do um modelo kernel-autoregressivo de ordem  $p = 24$ .

Na análise de correlação entre os índices verifica-se correlação entre os padrões do Atlântico (AMM, AMO e TNA; NAO e AO) e entre os padrões do Pacífico (ENSO e PDO), porém não foram observadas correlações significativas entre os oceanos. Além disso, foram observadas correlações fracas entre a atividade solar e os padrões AMM, AMO, AO, NAO, PDO, TNA e TSA, com destaque para as correlações moderadas com os padrões dos trópicos (TNA e TSA). Os índices intrasazonais apresentam correlações significativas com os índices PSA, destaques para o LISAM e ZCAS com o primeiro padrão (PSA1), RMM1 e RMM2 com o segundo padrão (PSA2). Além disso, o índice RMM1 apresenta correlação significativa com o SOI, indicando influência do fenômeno ENSO com a MJO.

Nas análises de causalidade, sugere-se que o ENSO causa os padrões de variabilidade do Atlântico, AMO e GTA, e do Pacífico, PDO, destacando a relação não-linear entre o ENSO  $\rightarrow$  PDO e ENSO  $\rightarrow$  GTA. Além disso, observa-se a influência da PDO  $\rightarrow$  GTA e PDO  $\rightarrow$  AAO, evidenciando a influência do Oceano Pacífico no Atlântico.

Em adição, os resultados da PDC e knPDC sugerem que alguns índices possuem interação não-linear, ressaltando o uso de técnicas que capturem essas variações.

O trabalho desenvolvido neste capítulo foi apresentado no AGU 2019 Fall Meeting, em dezembro de 2019 em São Francisco, CA, USA, disponível em [Malfatti \*et al.\* \[2019\]](#).

# Capítulo 4

## Identificação dos Padrões de Variabilidade Climática nas Principais Bacias Brasileiras

Neste capítulo são realizadas análises e discussões sobre as características dos padrões de variabilidade climática e como eles influenciam a precipitação nas principais bacias brasileiras. As análises a seguir servem para dar suporte à definição de um conjunto de variáveis preditoras e para fundamentar as interpretações físicas das soluções encontradas pelos modelos empíricos de previsão de precipitação nos próximos capítulos.

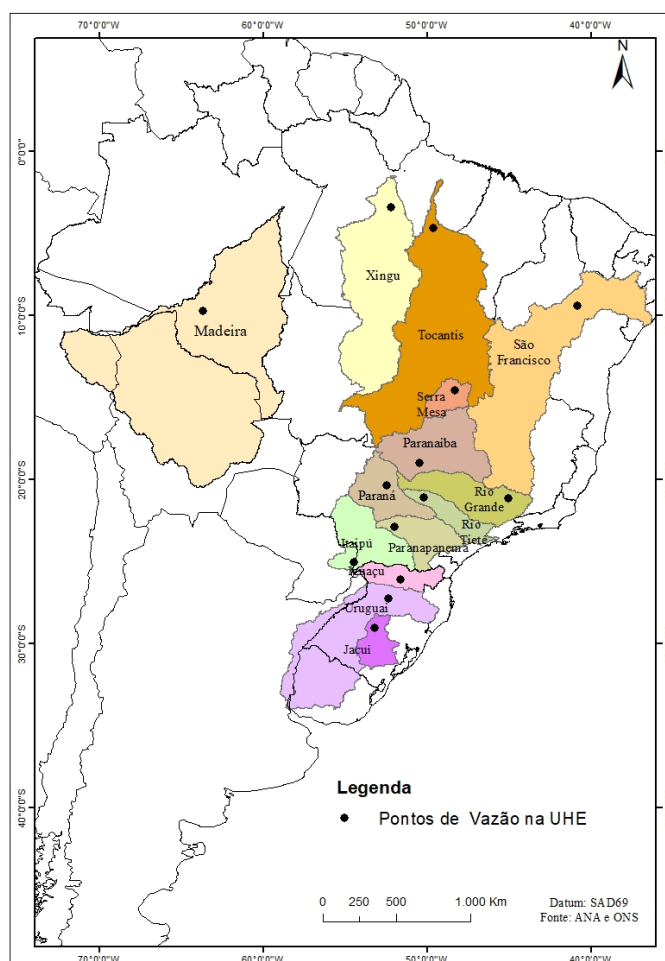
Além disso, séries históricas de vazão de rios nas bacias de interesse foram comparadas com os resultados da precipitação nessas bacias, facilitando as discussões sobre o planejamento dos recursos hídricos.

O trabalho desenvolvido neste capítulo foi submetido para a revista *International Journal of Climatology*, com o título *Predictability of Sub-seasonal Precipitation in Brazilian Watersheds based on Teleconnection Patterns and South American Monsoon Modes* e aguarda aprovação para publicação (versão original submetida no Anexo J).

### 4.1 Área de Estudo

Associado a grande variabilidade espacial e temporal dos sistemas meteorológicos que atuam no Brasil com a não linearidade da resposta fluvial a precipitação, identificar padrões climáticos relevantes para as variações de reservatórios de água, se torna um problema de grande complexidade (Raynaut [2011]; Jacobi *et al.* [2015]), visto que os fenômenos climáticos exercem influências em regiões e meses preferenciais (Capozzoli *et al.* [2017]; Malfatti *et al.* [2018a]). Neste contexto, bacias brasileiras que apresentam intenso uso múltiplo das águas, como: abastecimento público, irrigação (atividade agrícola), lazer, transporte hidroviário e principalmente aproveitamentos hidrelétricos (Usinas Hidrelétricas, UHE) foram selecionadas para estudo, conforme a Figura 4.1.

Essas bacias foram selecionadas por apresentar alta complexidade, devido aos seus usos múltiplos (isto é, maior complexidade no planejamento e gestão de recursos hídricos) e grandes variabilidades climáticas de diferentes escalas (intrasazonal, interanual, interdecadais e multidecadais).



**Figura 4.1:** Localização das principais bacias brasileiras e os pontos de vazão nas UHE.

Segundo dados do governo brasileiro do Departamento de Informações e Estudos Energéticos de Minas e Energia (MME), em 2020 a matriz energética brasileira oriunda de fontes renováveis representa 87,4% - sendo 60,2% hídrica, 8,9% eólica, 8,4% biomassa, 8,2% gás natural e 1,7% solar. Na Tabela 4.1 são informados detalhes sobre os pontos de vazão de rios em UHE selecionadas (Figura 4.1) para análises deste capítulo.

**Tabela 4.1:** Principais Usinas Hidrelétricas (UHE) em cada bacia brasileira de estudo. Os valores em negrito correspondem as maiores UHE em Potência Instalada (Capacidade Instalada) no Brasil.

Macro Região	UF	Bacia	Sub-bacia	UHE	Latitude	Longitude
Norte	RO	Madeira	Madeira	<b>Santo Antônio</b>	-8,80°	-63,95°
	PA	Xingu	Xingu	<b>Belo Monte</b>	-3,26°	-51,56°
	PA	Tocantins	Baixo Tocantins	<b>Tucuruí</b>	-3,83°	-49,65°
Nordeste	BA	São Francisco	Médio São Francisco	Sobradinho	-9,43°	-40,83°
Centro-Oeste	GO	Serra da Mesa	Alto Tocantins	Serra da Mesa	-13,84°	-48,3°
	GO	Paranaíba	Paranaíba	<b>São Simão</b>	-19,02°	-50,5°
Sudeste	RJ	Rio Grande	Grande	Funil Grande	-21,14°	-45,04°
	SP	Rio Tietê	Tietê	Nova Avanhandava	-21,12°	-50,2°
	SP	Paranapanema	Paranapanema	Rosana	-22,6°	-52,87°
	SP	Paraná	Médio Paraná	<b>Ilha Solteira</b>	-20,38°	-51,36°
Sul	PR	Itaipu	Baixo Paraná	<b>Itaipu</b>	-25,43°	-54,59°
	PR	Iguaçu	Iguaçu	Foz da Areia	-26,08°	-51,65°
	RS	Jacuí	Jacuí	Jacuí	-29,07°	-53,21°
	RS	Uruguai	Uruguai	Itá	-27,28°	-52,38°

## 4.2 Dados e Metodologia

Para a análise de padrões espaciais da precipitação pode-se usar o TRMM (*Tropical Rainfall Measuring Mission*), para observações de 1997 a 2014, e GPM (*The Global Precipitation Measurement Mission*), de 2015 a dias atuais, que fornecem informações de chuva no continente e no oceano, por observações por satélite. O GPM, iniciado pela *National Aeronautics and Space Administration* (NASA) e pela *Japan Aerospace Exploration Agency* (JAXA) como um sucessor global do TRMM, compreende um consórcio de agências espaciais internacionais, incluindo o *Centre National d'Études Spatiales* (CNES), a Organização de Pesquisa Espacial Indiana (ISRO), a NOAA, a Organização Europeia para a Exploração de Satélites Meteorológicos (EUMETSAT) e outros.

No caso do CHIRPS (*Climate Hazards Group InfraRed Precipitation with Station Data*), a análise espacial só pode ser feita para sub-regiões continentais, pois seus registros de precipitação são somente sobre o continente. O CHIRPS é um conjunto de dados de precipitação quase global com resolução espacial alta ( $0,05^\circ \times 0,05^\circ$ ) e cobertura temporal de longo prazo (de 1981 a dias atuais), que combina estimativas de precipitação por satélite e informações de estações meteorológicas Funk *et al.* [2015]. Os dados são disponibilizados em diferentes escalas temporais (intervalos de tempo: mensal, 10 dias, 5 dias e diários) pela Universidade da Califórnia em Santa Bárbara (UCSB), <https://www.chc.ucsb.edu/data/chirps>.

Os dados fornecidos foram submetidos a diversas avaliações, no Brasil. O conjunto de dados de precipitação do CHIRPS possui um desempenho relativamente bom tanto em escala regional quanto global, principalmente em termos de viés e coeficiente de correlação de Pearson quando são comparados ao GPM e TRMM (Paredes-Trejo *et al.* [2017]; Nogueira *et al.* [2018]; Costa *et al.* [2019]; Oliveira-Júnior *et al.* [2021]). Por essa razão para as análises a seguir, são utilizados os registros de precipitação proveniente do CHIRPS.

Para cada bacia foram calculadas a precipitação acumulada em pântadas (5 dias) no período de 1981 a 2018. Para observar como os padrões de variabilidade climática podem interferir na precipitação nas principais bacias brasileiras foram utilizados os mesmos dados de índices climáticos do Capítulo 3 (Padrões de Variabilidade Climática), no período entre 1981 a 2018. Alguns índices não possuem dados diários, e foram transformados em pântadas, neste caso, foi realizada a interpolação cúbica (polinômio de grau 3) (Boor [2001]) dos dados mensais.

Como o conjunto de dados apresenta um grande número de variáveis, o qual dificulta a compreensão dos resultados e aumenta os esforços computacionais. Foi necessário diminuir a dimensão das variáveis, porém sem perder informações essenciais para sua compreensão. Uma ferramenta útil para este fim é a ACP ou conhecido também por EOF, que reduz o número de variáveis do conjunto inicial de dados e explica grande parte da variância do conjunto inicial (Zang e Moore [2015]), detalhes da metodologia utilizada Apêndice D.

Na identificação da variabilidade da precipitação em bacias brasileiras foram selecionados, os componentes principais (PCs) que representam a maior parte da variabilidade do conjunto de variáveis de entrada. As séries temporais resultantes dos primeiros PCs (*scores*) foram filtradas (Filtro de Lanczos, passa alta acima de 6 meses) para eliminar as variabilidades de baixa frequência (Thomson e Emery [2014]) e aplicada a transformada de ondeletas (Torrence e Compo [1998], Apêndice B) para identificar os padrões existentes. Além disso, foi calculado o coeficiente  $r$  de Pearson (correlação), correlação cruzada (Stoica e Moses [2005]) e a causalidade (Baccalá e Sameshima

[2021], Apêndice E) entre as precipitações e os seis primeiros PCs.

Para investigar a relação entre a precipitação e a vazão dos principais rios brasileiros (isto é, vazão afluente ( $m^3/s$ ) nas UHEs com maior capacidade instalada de geração de energia) foram correlacionados as séries temporais dos seis primeiros PCs de precipitação e as séries temporais das vazões afluentes naturalizadas, disponível pelo Operador Nacional do Sistema Elétrico (ONS), <http://www.ons.org.br>, em aproveitamentos hidroelétricos no Brasil.

Os dados disponíveis pelo ONS são diários e não contém falhas, os dados são tratados previamente e naturalizados, ou seja, a vazão afluente é restaurada conforme ela ocorreria naturalmente sem os reservatórios da cascata do SIN.

Para cada bacia foram calculadas a vazão afluente média em pântadas (5 dias) no período de 1981 a 2018 e correlacionadas com os modos de precipitação no mesmo período.

### 4.3 Resultados e Discussão

Para investigar os padrões de variabilidade climática na precipitação de bacias brasileiras, o conjunto inicial de dados foi reduzido através da análise de EOF. A Tabela 4.2 mostra em porcentagem a variância explicada e a variância acumulada para os seis primeiros componentes principais (PCs) da precipitação das principais bacias brasileiras, e observa-se que o primeiro PC explica aproximadamente 33,4% da variabilidade da precipitação no Brasil e os seis primeiros PCs juntos explicam quase 90%. Isto é, foi reduzido para seis o conjunto de dados de precipitação das quatorze bacias brasileiras, preservando sua variabilidade em aproximadamente 89%.

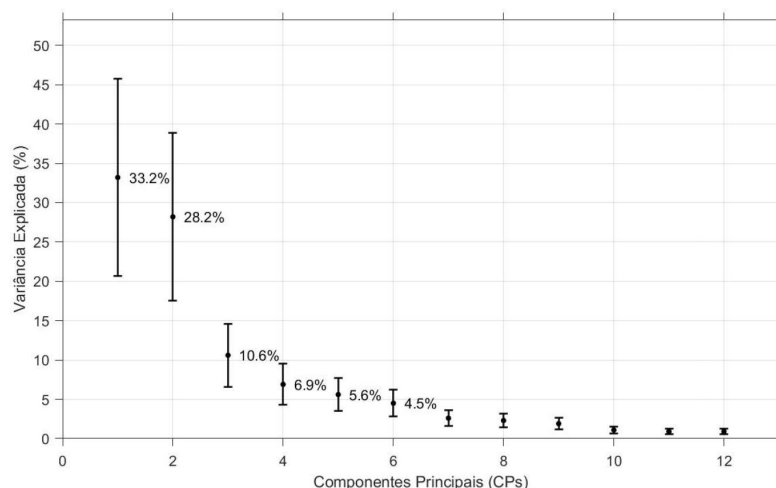
**Tabela 4.2:** Porcentagens da variância explicada e da variância acumulada dos seis primeiros componentes principais (PCs) da precipitação nas principais bacias brasileiras.

PCs	Variância Explicada (%)	Variância Acumulada (%)
1	33,2%	33,2%
2	28,2%	61,4%
3	10,6%	72,0%
4	6,9%	78,9%
5	5,6%	84,5%
6	4,5%	89,0%

Para verificar a dependência entre cada componente foi calculado o erro amostral, a Figura 4.2 mostra a variância explicada e o erro amostral (North *et al.* [1982]) dos seis primeiros PCs. Conforme as barras do erro amostral, observa-se forte dependência entre os PCs (PC1 ↔ PC2, PC3 ↔ PC4, PC3 ↔ PC5 e PC5 ↔ PC6).

De modo, a maximizar a variância dos resultados foi usada a EOF rotacionada pelo método de rotação ortogonal Varimax (Kaiser [1958]). Após a rotação das PCs, a porcentagem de variância total explicada pelas duas componentes permanece a mesma, mas cada uma delas explica diferentes porcentagens das que eram explicadas pelas componentes antes da rotação.

Conforme a Tabela 4.3, observa-se que as componentes PC1 a PC6, apresentam variação total explicada pelos dados sem rotação respectivamente, 33,2%, 28,2%, 10,6%, 6,9%, 5,6% e 4,5%, enquanto com rotação 22,1%, 12,1%, 13,0%, 7,3%, 18,0% e 14,5%. A Variância acumulada das seis primeiras PCs rotacionadas sofreram uma redução de 2% relação a não rotacionada (observação: as



**Figura 4.2:** Variância explicada dos seis primeiros componentes principais (CPs) e o erro amostral calculado pelo método de North et al. [1982].

14 PCs continuam com o total de 100% da variância explicada). Isto é, as novas PCs rotacionadas preservam sua variabilidade em aproximadamente 87%.

**Tabela 4.3:** Porcentagens da variância explicada e da variância acumulada dos seis primeiros componentes principais (CPs) da precipitação nas principais bacias brasileiras.

PCs	Variância Explicada (%)	Variância Explicada (%)	Variância Acumulada (%)
	Sem rotação	Rotacionada	Rotacionada
<b>1</b>	33,2%	22,1%	22,1%
<b>2</b>	28,2%	12,1%	34,2%
<b>3</b>	10,6%	13,0%	72,0%
<b>4</b>	6,9%	7,3%	78,9%
<b>5</b>	5,6%	18,0%	84,5%
<b>6</b>	4,5%	14,5%	87,0%

Nota-se, que após a rotação dos vetores pelo método ortogonal Varimax as componentes apresentam variância explicada mais próximas entre si, apresentando variabilidades com menor dependência do que o método sem rotação, isto é ameniza a dependência entre as componentes.

Após a análise entre os índices e os PCs foram relacionados os resultados com a precipitação em cada bacia brasileira selecionada. A Tabela 4.4 destaca os valores significativos de CC (valores em negrito com 95% de significância estatística), as bacias foram divididas conforme a localização geográfica (de Norte a Sul).

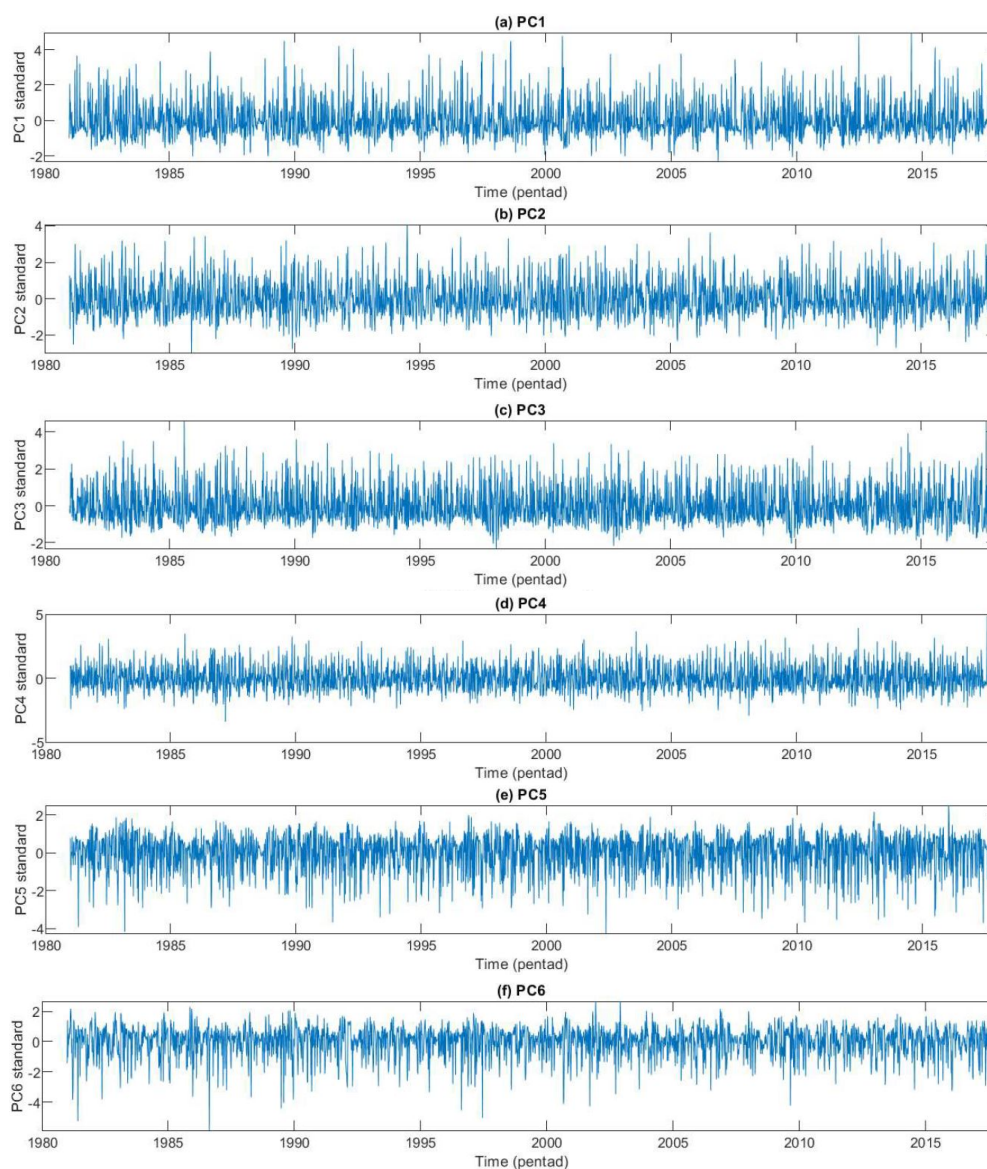
A Tabela 5 mostra que a distribuição espacial dos PCs, a componente PC1 tem forte correlação com as bacias do Sudeste, PC2 com a parte central da região Norte, PC3 tem grande peso para o extremo sul, PC4 com a bacia do Rio Madeira (próximo a Andes), PC5 com o norte da região Sul e PC6 nas bacias hidrográficas das regiões Centro-Oeste e Nordeste.

Em seguida, foi analisada a variabilidade das séries temporais padronizadas e filtradas (Figura 4.3) na escala intrasazonal (até 6 meses) das primeiras seis PCs por meio da transformada de *wavelet*, a Figura 4.4 mostra os resultados obtidos. No geral, observa-se que os picos espectrais significativos no espectro global de potência entre as PCs são distintos entre si, indicando padrões distintos que interferem na escala intrasazonal.



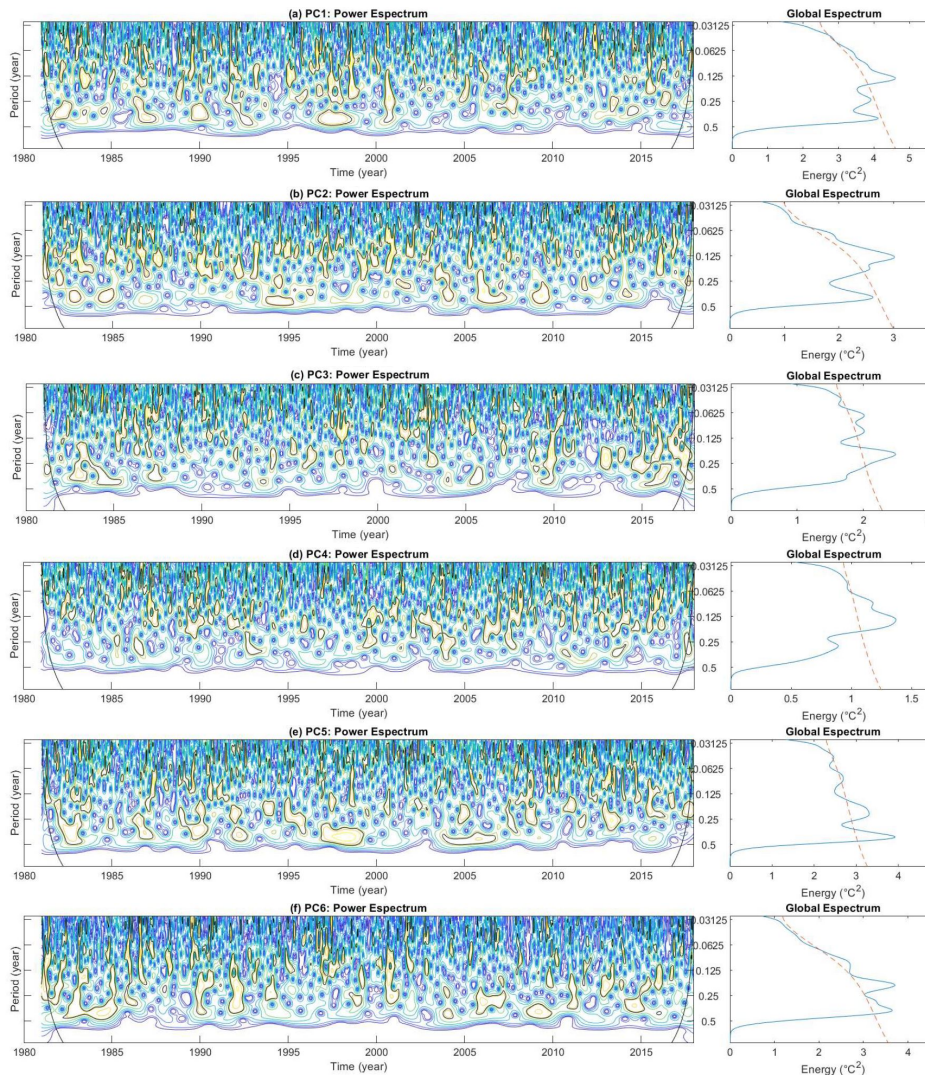
**Tabela 4.4:** Correlações entre as seis componentes principais e as precipitações nas principais bacias brasileiras. Os valores em negrito corresponde ao nível de significância de 95%, a partir do teste *t* de Student.

Macro Região	Bacia	PC1	PC2	PC3	PC4	PC5	PC6
Norte	Madeira	<b>0,21</b>	<b>0,24</b>	0,05	<b>0,99</b>	<b>-0,17</b>	<b>-0,15</b>
	Xingu	<b>0,10</b>	<b>0,96</b>	<b>-0,26</b>	<b>0,21</b>	<b>0,17</b>	<b>-0,42</b>
	Tocantins	<b>0,27</b>	<b>0,91</b>	<b>-0,30</b>	<b>0,23</b>	<b>0,14</b>	<b>-0,77</b>
Nordeste	São Francisco	<b>0,23</b>	<b>0,51</b>	<b>-0,32</b>	<b>0,09</b>	<b>0,17</b>	<b>-0,89</b>
Centro-Oeste	Serra da Mesa	<b>0,36</b>	<b>0,60</b>	<b>-0,24</b>	<b>0,17</b>	<b>0,09</b>	<b>-0,90</b>
	Paranaíba	<b>0,80</b>	<b>0,40</b>	<b>-0,19</b>	<b>0,25</b>	<b>-0,14</b>	<b>0,71</b>
Sudeste	Rio Grande	<b>0,93</b>	<b>0,20</b>	<b>-0,13</b>	<b>0,15</b>	<b>-0,33</b>	<b>-0,43</b>
	Rio Tietê	<b>0,93</b>	<b>0,11</b>	0,00	<b>0,16</b>	<b>-0,57</b>	<b>-0,18</b>
	Paranapanema	<b>0,73</b>	-0,03	<b>0,19</b>	<b>0,17</b>	<b>-0,87</b>	<b>0,03</b>
	Paraná	<b>0,82</b>	0,06	0,05	<b>0,34</b>	<b>-0,66</b>	<b>-0,14</b>
Sul	Itaipu	<b>0,48</b>	<b>-0,14</b>	<b>0,33</b>	<b>0,25</b>	<b>-0,95</b>	<b>0,09</b>
	Iguaçu	<b>0,30</b>	<b>-0,19</b>	<b>0,49</b>	<b>0,13</b>	<b>-0,91</b>	<b>0,18</b>
	Jacuí	-0,06	<b>-0,26</b>	<b>0,96</b>	0,03	<b>-0,34</b>	<b>0,28</b>
	Uruguai	<b>-0,07</b>	<b>-0,31</b>	<b>0,96</b>	0,06	<b>-0,41</b>	<b>0,31</b>



**Figura 4.3:** Séries temporais das seis primeiras componentes principais da precipitação.

Nas primeiras quatro PCs, na Figura 4.4 (a) a (d), observa-se como principal pico significativo em 0,125 anos, que representa um período de aproximadamente 9 pântadas, também observa picos significativos na PC3 em 0,0625 e 0,25 (4,5 e 18 pântadas) e a PC4 em 0,0625 (4,5 pântadas), conforme a Figura 4.4 (c) e (d). Enquanto para as PC5 e PC6 os picos significativos são observados em 0,25 e 0,5, que em pântadas representa respectivamente 18 e 36,5 pântadas, conforme a Figura 4.4 (e) e (f).



**Figura 4.4:** Espectro de Potência e Global da análise de ondeleta da série temporal dos seis primeiros componentes de precipitação a) PC1; b) PC2; c) PC3; d) PC4; e) PC5; f) PC6. O espectro de potência é delimitado pelo cone de influência, o contorno preto indica o setor do espectro de potência onde o nível de confiança de 90% é satisfeito. O espectro de potência global indica o nível de potência satisfeito para toda a série temporal, ou seja, acima da linha vermelha tracejada.

Para evidenciar as relações entre as componentes e os índices climáticos foram calculadas as correlações de Pearson nas séries temporais padronizadas, na Tabela 4.5 são apresentadas os índices que apresentaram correlações significativas de 95%, a partir do teste t de Student.

De uma maneira geral, conforme a Tabela 4.5, os índices que apresentam maior correlação com os PCs da precipitação nas principais bacias brasileiras são os que representam os fenômenos interanuais ENOS (índice MEI), PSA e AAO, interdecadal a PDO, multidecadal a AMO, as variabilidades intrasazonal do SMAS e MJO.

**Tabela 4.5:** Correlações entre as seis componentes principais (PCs) e os índices climáticos. Os valores em negrito corresponde ao nível de significância de 95%.

PCs	AAO	PSA1	PSA2	LISAM	ZCAS	RMM1	RMM2	AMO	MEI	PDO	PNA	TNA
<b>1</b>	-0,01	-0,02	<b>-0,08</b>	<b>-0,30</b>	<b>0,13</b>	<b>-0,10</b>	0,01	-0,02	<b>0,08</b>	0,02	0,01	-0,01
<b>2</b>	<b>-0,09</b>	<b>0,16</b>	<b>-0,19</b>	<b>-0,32</b>	<b>0,34</b>	<b>-0,23</b>	-0,02	-0,04	<b>-0,10</b>	<b>-0,07</b>	<b>-0,07</b>	<b>-0,07</b>
<b>3</b>	0,03	<b>-0,07</b>	0,00	0,01	<b>-0,45</b>	<b>0,08</b>	-0,01	0,03	<b>0,14</b>	<b>0,09</b>	0,04	0,03
<b>4</b>	<b>-0,09</b>	0,02	<b>-0,14</b>	<b>-0,20</b>	0,04	<b>-0,15</b>	<b>0,09</b>	<b>-0,07</b>	-0,05	-0,03	-0,05	<b>-0,08</b>
<b>5</b>	0,01	<b>0,07</b>	0,03	<b>0,15</b>	<b>0,14</b>	0,00	-0,02	-0,01	<b>-0,12</b>	-0,06	-0,03	-0,02
<b>6</b>	0,01	<b>-0,07</b>	<b>0,12</b>	<b>0,43</b>	<b>-0,36</b>	<b>0,15</b>	-0,01	0,05	0,00	0,01	0,02	0,04

Para identificar potenciais preditores climáticos na precipitação das bacias brasileiras, foi aplicada correlação cruzada, entre as componentes principais de precipitação (PC1 a PC6) e os índices climáticos. As séries dos índices foram defasadas de -1 a 12 pênadas, isto é, 2 meses de defasagem máxima entre os modos de precipitação e os índices climáticos, conforme os resultados disponíveis na Figura 4.5.

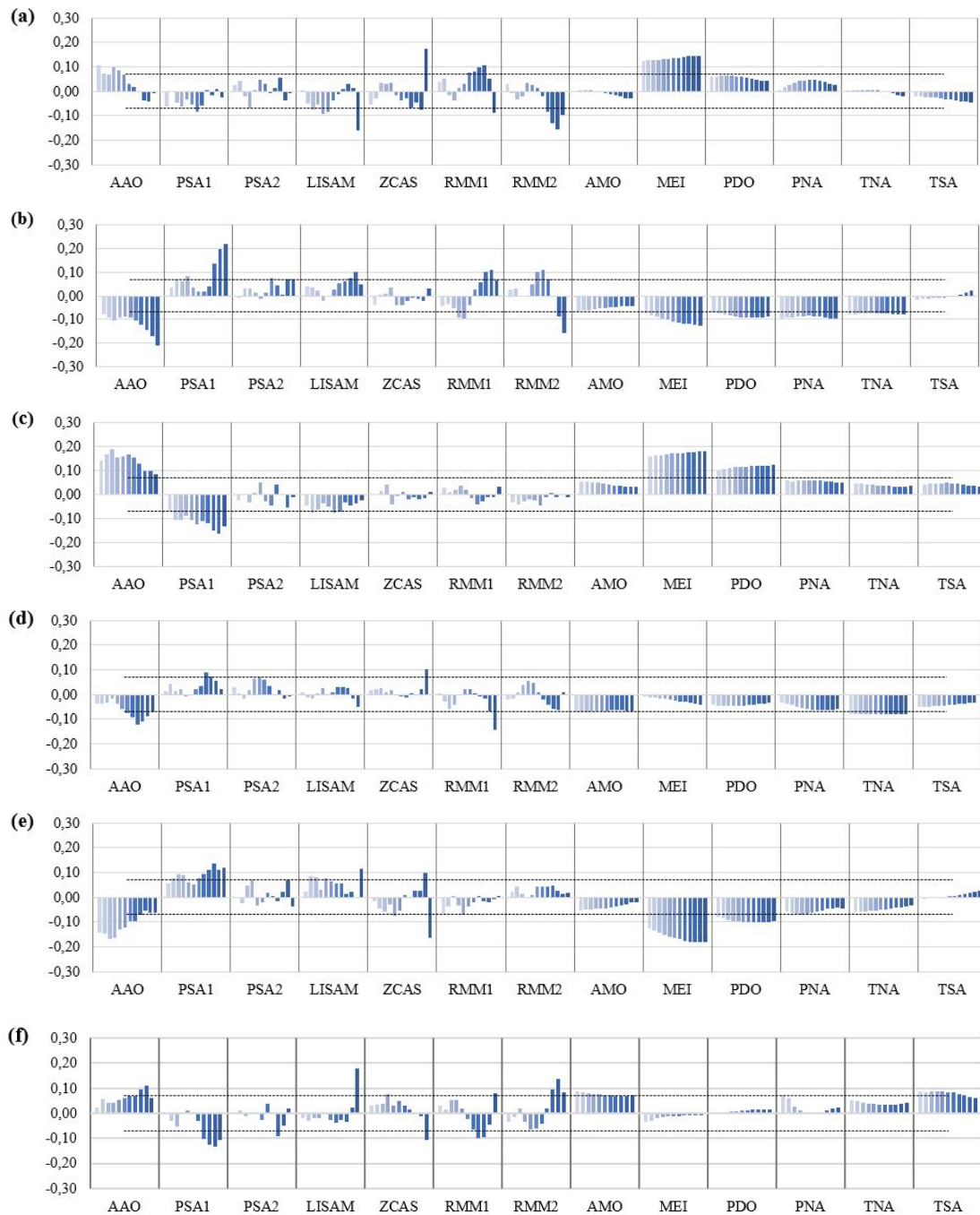
Na Figura 4.5 as defasagens mais longas possuem tons mais claros, enquanto as mais curtas tons mais fortes (barras da esquerda para a direita). Na primeira componente PC1 (Figura 4.5 (a)), observa-se CC significativos para os padrões LISAM e ZCAS, RMM1 e RMM2, indicando relação com o padrão SMAS e MJO, sendo bons preditores para períodos mais curtos, isto é para até 30 dias. Enquanto os padrões MEI e AAO são significativos respectivamente para todas as defasagens analisadas (5 a 60 dias) e nas defasagens mais longas (entre 30 a 60 dias). Na PC2 (Figura 4.5 (b)), nota-se CC significativos para os índices de baixa frequência, AAO, MEI, PDO, PNA e TNA, em todas as defasagens, ou seja, bons preditores no período de 5 a 60 dias, à medida que os índices PSA1, RMM1 e RMM2, nas defasagens curtas a intermediárias (até 45 dias). Na PC3 e PC5 (Figura 4.5 (c) (e)), os CC significativos são observados para os índices AAO, PSA1, MEI e PDO, em todas as defasagens, indicando bons preditores de precipitação até 60 dias. E, por fim, as componentes PC4 (Figura 4.5 (d)) e PC6 (Figura 4.5 (f)) apresentam padrão similar, os índices do Atlântico AMO, TNA ou TSA (PC6) em todas as defasagens e para os índices AAO e RMM1/RMM2 nas defasagens mais curtas, sugerindo bons preditores para o período até 60 dias e até 20 dias, respectivamente.

Conforme os resultados das correlações entre os índices climáticos e os PCs (Tabela 4.5), e entre os PCs e as precipitações médias em cada bacia brasileira (Tabela 4.4), foram selecionados os índices para a análise de causalidade. A Tabela 4.6 apresenta os índices selecionados para cada modo de precipitação (PC1 a PC6).

**Tabela 4.6:** Índices Selecionados para análise de causalidade (PDC) entre os modos de precipitação (PCs 1 a 6) e os índices climáticos selecionados com maiores correlações.

Índices Selecionados para análise de causalidade													
PCs	LISAM	ZCAS	AAO	PSA1	PSA2	RMM1	RMM2	MEI	PDO	AMO	PNA	TNA	TSA
<b>1</b>	x	x	x		x	x	x	x					
<b>2</b>	x	x		x	x	x	x	x	x		x	x	
<b>3</b>	x	x	x	x	x	x		x	x				
<b>4</b>	x		x		x	x	x			x		x	
<b>5</b>	x	x	x	x				x	x		x		
<b>6</b>	x	x			x					x			x

A Figura 4.6 mostram os resultados da PDC para as seis PCs e os índices selecionados conforme



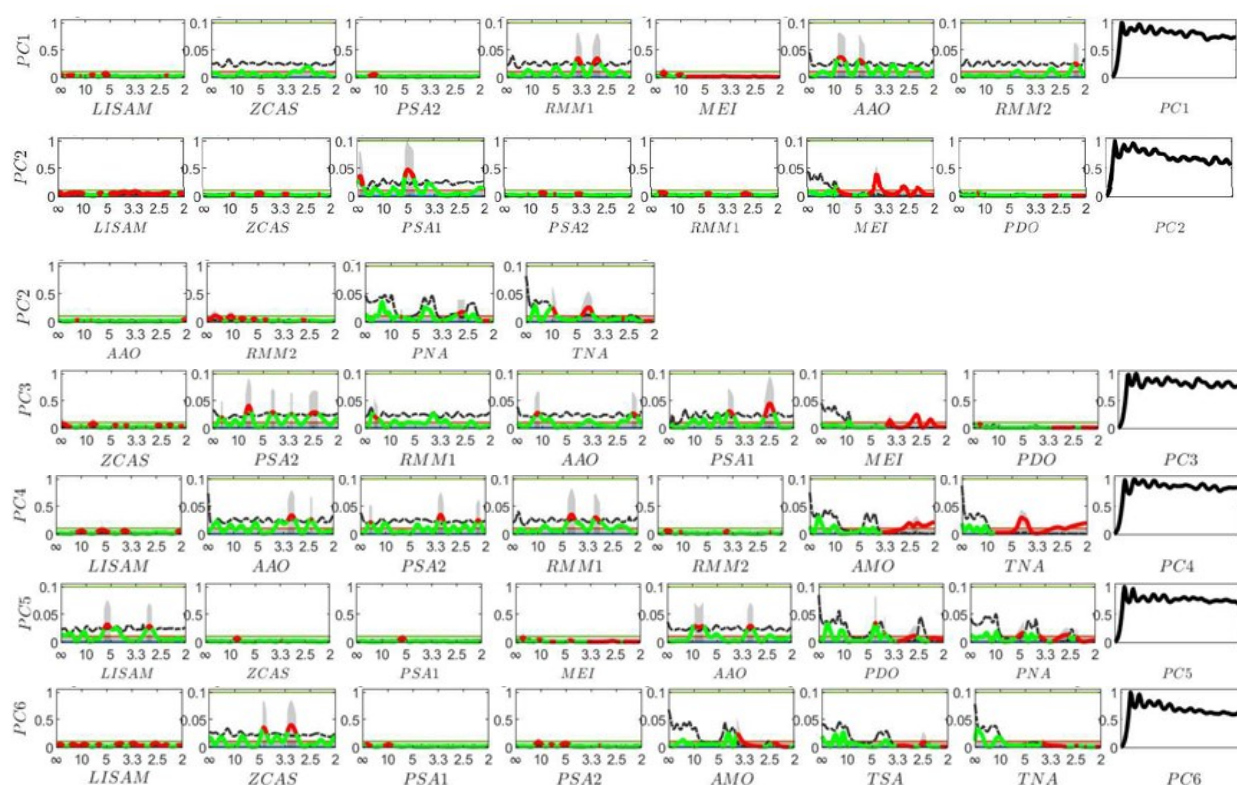
**Figura 4.5:** Correlação defasada em pântadas (lag -12 a lag -1, dos tons mais claros aos mais escuros) entre os modos de precipitação ((a) PC1, a (f) PC6) e os índices climáticos, as linhas pontilhadas representa o limite de significância a partir do Teste t de Student (-0,07 e +0,07).

a Tabela 4.6. O traçado em preto representa a (pseudo-) densidade espectral das séries em  $dB$ ; o traçado em vermelho representa os valores estatisticamente significantes da PDC; o traçado preto descontínuo representa o limiar da aproximação de Patnaik (Baccalá *et al.* [2013]), com nível de significância igual a  $\alpha$ ; o traçado em verde os valores estatisticamente não-significantes da PDC e; os valores do eixo  $x$  em pântadas. A direção da causalidade é observada do eixo  $x$  (abscissa) para o eixo  $y$  (ordenada), neste trabalho optou-se por expor apenas os resultados onde os índices causam a precipitação nas PCs.

Conforme a Figura 4.6, observa-se valores estatisticamente significativos de causalidade na alta frequência (intrasazonal) entre todas as PCs e os índices LISAM, ZCAS, MJO (índices RMM1 e/ou RMM2), AAO e PSA (índices PSA1 e/ou PSA2), indicando forte relação entre esses padrões e a precipitação no Brasil em todas as bacias brasileiras analisadas.

Nas componentes são observadas valores significativos em todas as frequências analisadas (2, 2.5, 3.3, 5, 10 pântadas e acima de 10 pântadas) para os fenômenos do Pacífico ENSO (índice MEI) e PDO. Esses resultados reforçam a influência do Pacífico na precipitação da América do Sul (Figura 4.5).

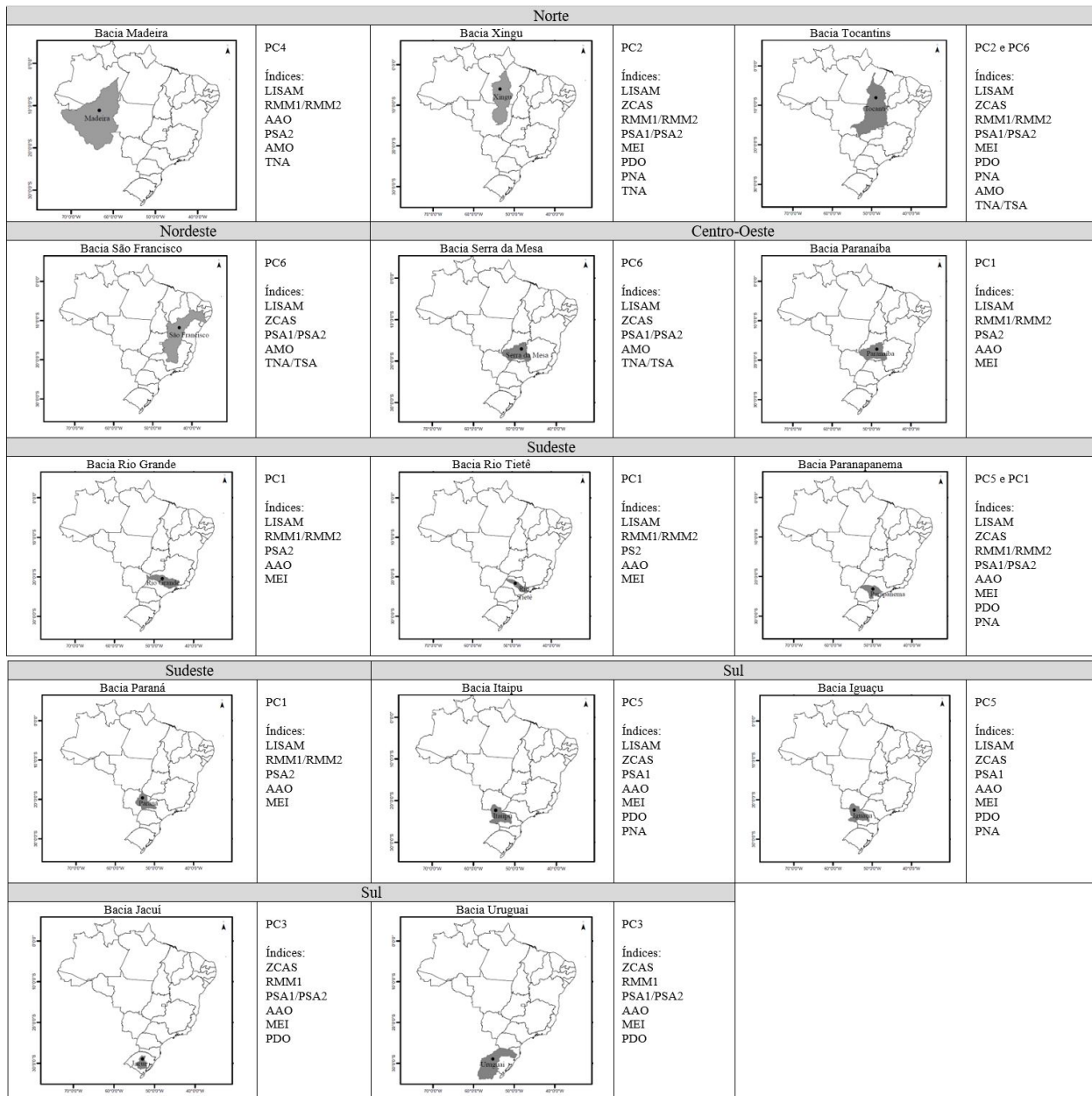
Para as componentes PC2, PC4 e PC6 observa-se na Figura 4.6, os picos espectrais significativos sugerem influência da de padrões do Atlântico, AMO, TNA e TSA, entre 2 a 10 pântadas. Além disso, nas componentes PC2 e PC5 observa-se em períodos mais curtos (2 a 5 pântadas) a influência do padrão do Hemisfério Norte (PNA).



**Figura 4.6:** O gráfico preto representa a (pseudo-) densidade espectral da série, em série em dB (PC1 a PC6); o gráfico vermelho representa os valores PDC estatisticamente significativos; a linha preta tracejada representa a aproximação do limiar de Patnaik *Baccalá et al. [2013]*; a linha verde os valores PDC estatisticamente não significativos e; os valores do eixo x em pântadas. Portanto, usando o nível de significância de 1%, os números representam, a PDC para o conjunto de séries temporais dos PCs x índices usando um modelo autorregressivo de ordem  $p = 24$ .

Na Figura 4.7 é possível observar as diferenças nos padrões de precipitação para cada bacia brasileira, que associada aos resultados das correlações entre os índices climáticos e os PCs (Tabela 4.5), podemos inferir sobre os padrões que influenciam em cada bacia brasileira.

Conforme a Figura 4.7, no geral, a precipitação das bacias possuem grande influência do SMAS (índices LISAM e/ou ZCAS), MJO (índices RMM1 e/ou RMM2), AAO e PSA (índices PSA1 e/ou PSA2).



**Figura 4.7:** Componente principal (PC) com maior correlação em cada bacia brasileira e os respectivos índices climáticos que causam a precipitação.

De acordo com Kousky e Kayano [1994] a MJO exercer influência nas variações na posição e intensidade da ZCAS, pois quando ocorre a passagem da MJO sobre a América do Sul há o favorecimento de sistemas transientes que contribuem para a formação de ZCAS, principalmente nos meses de verão nas regiões centro sul da Amazônia (bacias Serra da Mesa e Tocantins), regiões Centro-Oeste e Sudeste (bacias do Paranaíba, Paranapanema, Paraná, Rio Grande e Rio Tietê), centro sul da Bahia, norte do Estado do Paraná (bacias de Itaipu e Iguazu) e alongando-se em direção ao Oceano Atlântico Sudoeste (Carvalho *et al.* [2004]; Ferreira *et al.* [2004], Marengo *et al.* [2012]). O padrão PSA, como MJO (RMM1/RMM2), também contribui para a convecção das ZCAS (Vasconcellos e Cavalcanti [2010]), enquanto a AAO pode influenciar no posicionamento das trajetórias dos sistemas ciclônicos, podendo contribuir para a intensificação de trens de ondas do tipo PSA pela fase positiva (negativa) da AAO (Vasconcellos e Cavalcanti [2010]; Machado *et al.* [2021]).

A bacia do Madeira no Noroeste do Brasil (Bacia Amazônica) possui forte correlação com o PC4 ( $CC = 0,99$ ) (Figura 4.7), que por sua vez, possui correlação com os padrões SMAS, MJO, AAO, PSA e do Atlântico (AMO e TNA). Os resultados obtidos corroboram com vários estudos da Bacia Amazônica que reforçam a influência da MJO, PSA, AAO (Cavalcanti e Shiumizu [2012]; Mayta *et al.* [2020]) e nos modos de variabilidade do Atlântico (Jones e Carvalho [2018]; Marengo e Nobre [2009]) no SMAS.

Ao Norte do Brasil, as bacias do Xingu e Tocantins se correlacionam fortemente com o PC2 ( $CC = 0,96$  e  $CC = 0,91$ , respectivamente) (Figura 4.7) e moderadamente com as bacias Nordeste e Centro-Oeste (São Francisco, Serra da Mesa e Paranaíba). Os resultados sugerem, além dos citados anteriormente (LISAM/ZCAS, MJO, PSA e AAO), relação com os padrões do Pacífico (ENSO e PDO) e do Atlântico Norte (PNA e TNA). Estudos corroboram a relação entre a precipitação no Norte com os fenômenos de baixa frequência do Pacífico (Foley *et al.* [2002]; Tedeschi *et al.* [2016]; Molion [2008]) e Atlântico (Andreoli e Kayano [2006]). Além disso, observa-se  $CC$  alta ( $CC = -0,77$ ) para a bacia do Tocantins com a PC6 (Figura 4.7) e com as bacias do Nordeste e Centro-Oeste (São Francisco, Serra da Mesa e Paranaíba), indicando correlação com os padrões do Atlântico Sul (AMO e TSA) (Jones e Carvalho [2018]; Andreoli e Kayano [2006]), que por sua vez, podem influenciar no posicionamento da ZCAS fortalecendo ou desfavorecendo a precipitação ao longo de sua extensão (Chaves [2011]).

Para as bacias do Sul (Itaipu, Iguaçu, Jacuí e Uruguai) e Sudeste do Brasil (Rio Grande, Rio Tietê, Paranapanema e Paraná) as maiores correlação são observadas para a PC1 e PC6 (Figura 4.7), indicando forte relação com os fenômenos do Pacífico ENSO e PDO, que conferem com os resultados da literatura (Rebello [2006]; Molion [2008]; Tedeschi *et al.* [2016]). Vale ressaltar que para os fenômenos do Pacífico ENSO e PDO, são observados padrões opostos entre Norte/Nordeste e Sul/Sudeste.

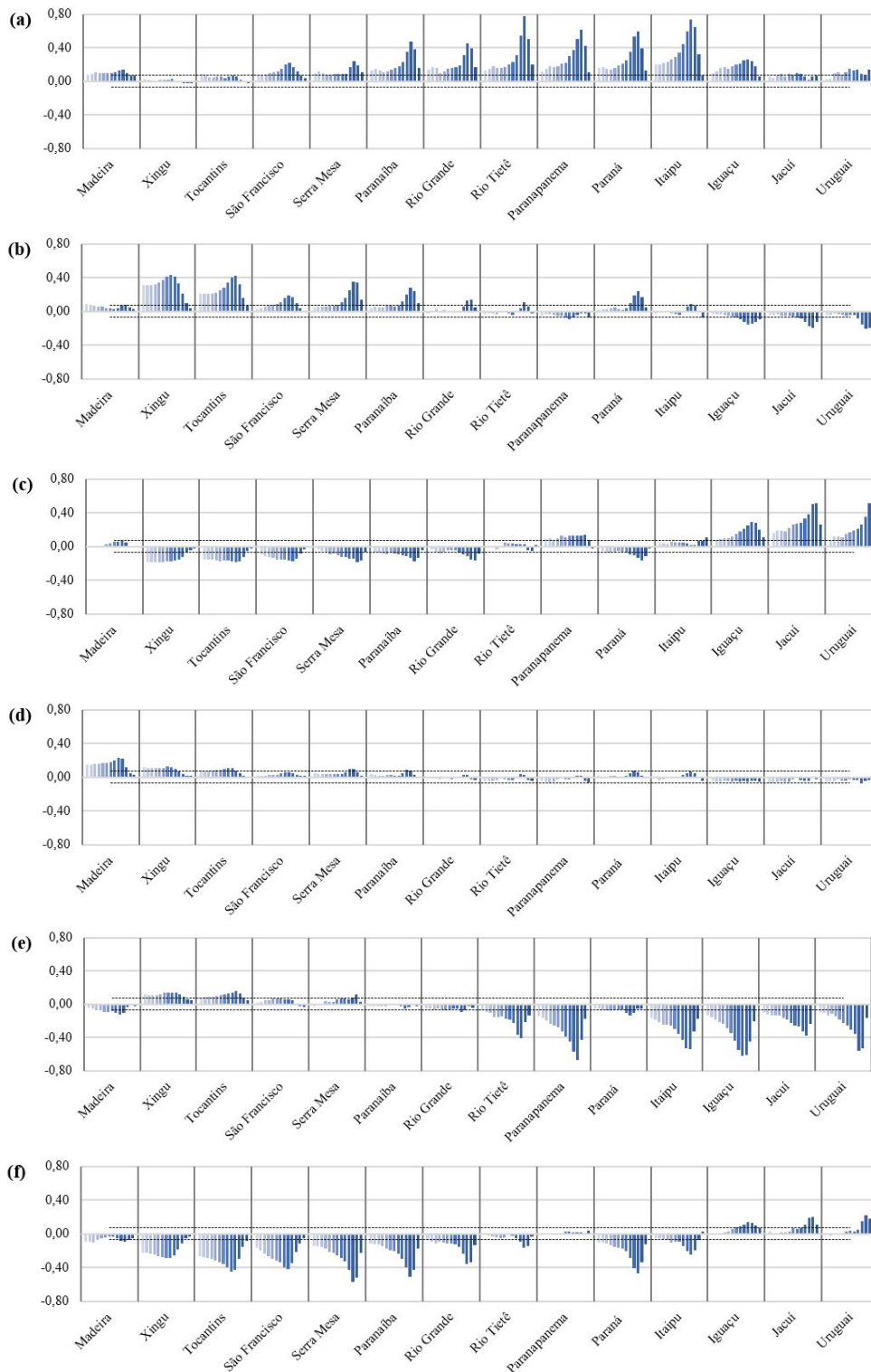
Com o intuito de analisar o tempo de resposta da precipitação nas vazões de rios, para planejamento estratégico dos recursos hídricos, foram calculados os coeficientes de correlação para as vazões afluentes naturalizadas nos principais rios de cada bacia, isto é, nos pontos das principais UHE do país.

A Figura 4.8 mostra as correlações defasadas de lag -12 a lag 0 em pântadas entre os modos de precipitação (PC1 a PC6) e as vazões nas principais bacias brasileiras. As defasagens mais longas possuem tons mais claros, enquanto as mais curtas tons mais fortes (barras da esquerda para a direita).

Nesta Figura 4.8, observa-se que o tempo de resposta da precipitação para a vazão de rios podem alterar conforme a bacia. Na bacia do Rio Madeira (Noroeste do Brasil), o tempo de resposta varia entre 15-30 dias (expresso pelos valores de  $CC$  ao longo das defasagens da PC4). Ao Norte (Xingu e Tocantins) o tempo de resposta é mais lento variando entre 15-40 dias, conforme os valores de  $CC$  ao longo das defasagens da PC2 (influência positiva  $CC > 0$ ) e PC6 (influência negativa  $CC < 0$ ).

Nas bacias do Centro-Oeste (Serra Mesa e Paranaíba) o tempo de resposta é mais rápido do que as regiões anteriores, entre 5-20 dias, segundo os valores observados na PC2 e PC6. Comportamento similar observado nas bacias do Sul (Itaipu, Iguaçu, Jacuí e Uruguai) e Sudeste (Rio Grande, Rio Tietê, Paranapanema e Paraná), conforme os valores de  $CC$  ao longo das defasagens da PC1 e PC5.

Lembrando que a precipitação incidente na bacia é apenas uma das componentes que compõem



**Figura 4.8:** Correlação defasada (lag -12 a lag 0, dos tons mais claros aos mais escuros) entre os modos de precipitação ((a) PC1 a (f) PC6) e as vazões nas principais bacias brasileiras, as linhas pontilhadas representa o limite de significância a partir do Teste t de Student (-0,07 e +0,07).

a vazão de rios. Além dela, fatores relacionados a fisiologia (isto é, as características físicas da bacia hidrográfica, como a área da bacia de contribuição, sua forma, condições de superfície do solo e constituição geológica do subsolo, e a conformação topográfica) e antrópicos (isto é, as obras hidráulicas construídas na bacia e ocupação humana) da bacia hidrográfica também interferem no compor-



tamento hidrológico dos cursos d'água (Von Sperling [2007]). Neste contexto, previsões acuradas de precipitação auxiliaram na tomada de decisão, porém fatores não relacionados a variabilidade climática também devem ser considerados no planejamento estratégico dos recursos hídricos.

## 4.4 Conclusão

Neste capítulo foi possível identificar as relações entre os padrões de variabilidade climática, intrasazonal, interanual, interdecadal e multidecadal, que interferem na escala intrasazonal da precipitação em diferentes bacias brasileiras, através da análise de correlação dos modos de variabilidade da precipitação (PC1 a PC6, que correspondem a 87% da variabilidade total) e índices climáticos. Além disso, foi possível identificar possíveis preditores climáticos para as bacias em estudo.

No geral, os modos de precipitação das bacias possuem grande influência dos padrões MJO (índices RMM1 e/ou RMM2), AAO e PSA (índices PSA1 e/ou PSA2) no SMAS (índices LISAM e/ou ZCAS). Além disso, destaca a importância dos padrões do Pacífico, ENSO e PDO, para quase todas as bacias brasileiras, exceto para a Bacia a Noroeste no Rio Madeira.

As bacias do Norte (Madeira, Xingu e Tocantins), Nordeste (São Francisco) e Centro-Oeste (Serra da Mesa e Paranaíba) apresentam maior influência do Atlântico, com os padrões AMO e TNA/TSA. Enquanto os padrões do Pacífico, ENSO e PDO, possuem maior influência na precipitação nas bacias do Sul (Itaipu, Iguazu, Jacuí e Uruguai) e Sudeste do Brasil (Rio Grande, Rio Tietê, Paranapanema e Paraná).

Para identificar em quais faixas de frequência a precipitação são ocasionadas por esses padrões de variabilidade climática, aplicou-se a análise de causalidade PDC. A Figura 4.9 sintetiza as relações de causalidade entre os índices climáticos e as componentes de precipitação. Os conectores em vermelho indicam causalidade em alta frequência (escala intrasazonal).

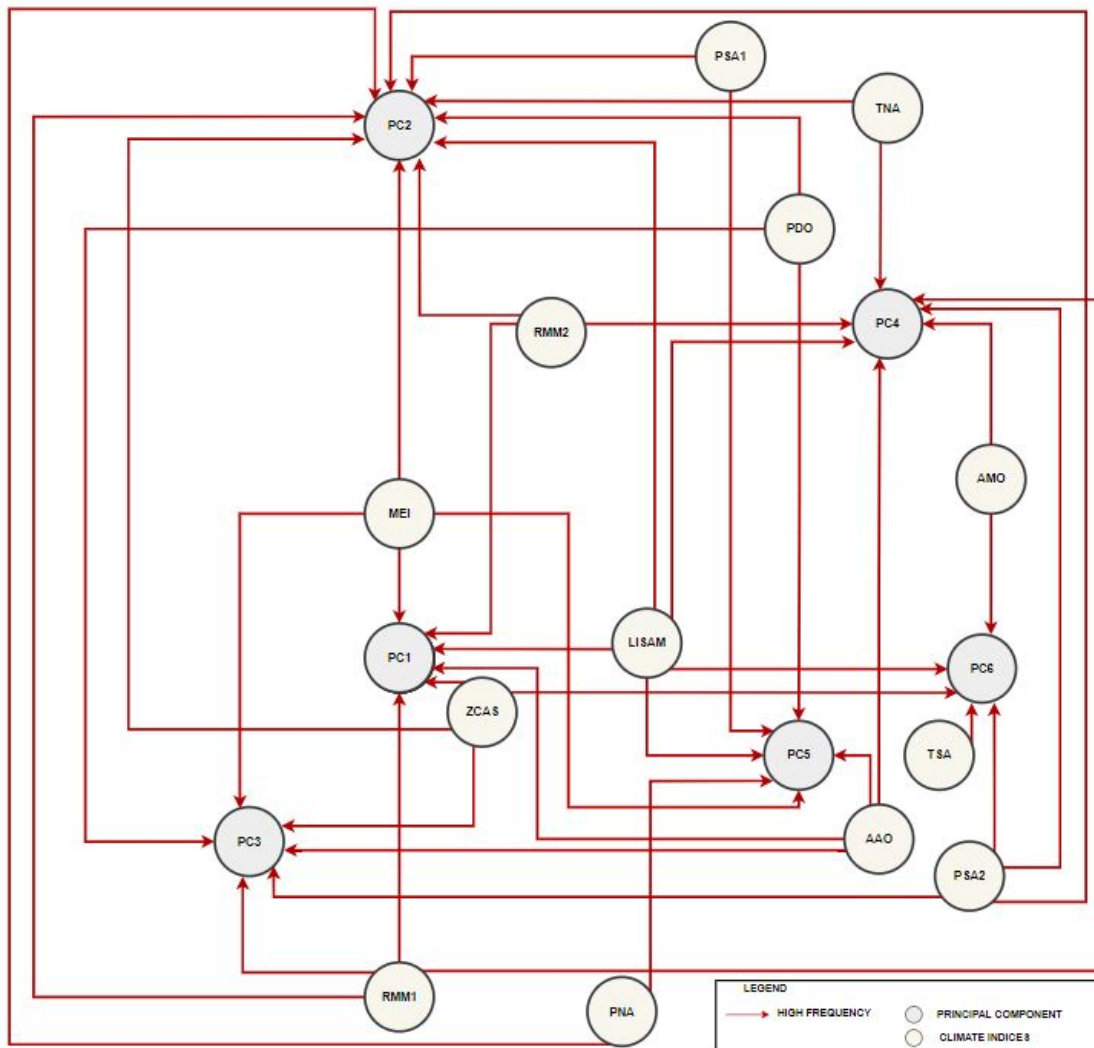
Os resultados da PDC validam a influência de diversos padrões de variabilidade climática distintas (intrasazonal, interanual, decadal e multidecadal) em na escala de tempo intrasazonal.

Além disso, as análises de correlação cruzada indicam que esses padrões climáticos são potenciais preditores climáticos na escala sub-sazonal. Os índices LISAM e ZCAS apresentam potencial papel previsor em períodos mais curtos até 20 dias, enquanto os índices RMM1/RMM2 para períodos intermediários (até 45 dias). Os índices AAO, MEI, PDO, AMO, PNA e TNA/TSA possuem correlações significativas em todas as defasagens, indicando preditores de precipitação em todo o período analisado, isto é, até 60 dias.

Nesta Figura, observa-se que tempo de resposta da precipitação para a vazão de rios podem alterar conforme a bacia. Na bacia do Rio Madeira (Noroeste do Brasil), o tempo de resposta varia entre 15-30 dias (expresso pelos valores de CC ao longo das defasagens da PC4). Ao Norte (Xingu e Tocantins) o tempo de resposta é mais lento variando entre 15-40 dias, conforme os valores de CC ao longo das defasagens da PC2 (influência positiva  $CC > 0$ ) e PC6 (influência negativa  $CC < 0$ ).

Nas bacias do Centro-Oeste (Serra Mesa e Paranaíba) o tempo de resposta é mais rápido do que as regiões anteriores, entre 5-20 dias, segundo os valores observados na PC2 e PC6. Comportamento similar observado nas bacias do Sul (Itaipu, Iguazu, Jacuí e Uruguai) e Sudeste (Rio Grande, Rio Tietê, Paranapanema e Paraná), conforme os valores de CC ao longo das defasagens da PC1 e PC5.

Os resultados indicam que os padrões de baixa frequência, ou seja, os modos de variabilidade



**Figura 4.9:** Diagrama de Causalidades entre as componentes principais da precipitação e os índices climáticos. Os conectores em vermelho indicam causalidade em alta frequência.

climática mais lenta, com energia significativa em períodos acima da escala S2S, também influenciam o comportamento da precipitação na escala S2S. Isto é, a evolução da precipitação na escala S2S é dependente, por exemplo, da influência de fenômenos como a PDO e a AMO que são caracterizados por períodos de vários anos. Deste modo, no próximo capítulo serão utilizados os índices dos padrões climáticos como preditores de precipitação, visto que os índices apresentam correlações significativas em defasagens mais longas, que auxiliam no aprimoramento de modelos preditores de precipitação.

Como o tempo de resposta da precipitação na vazão de rios pode variar de 5 a 40 dias, dependendo da localização da bacia, a previsão da precipitação na escala S2S tem potencial para auxiliar no planejamento e gerenciamento dos recursos hídricos em algumas bacias com tempo de resposta mais lento, antecipando seus impactos entre 60 a 100 dias.



# Capítulo 5

## Desenvolvimento de Modelos Preditivos de Precipitação nas Principais Bacias Brasileiras

Neste capítulo, são apresentados e discutidos os resultados da aplicação do modelo de Regressão Linear Múltipla (RLM) e modelo não-linear de *Long Short-Term Memory* (LSTM), a partir de métodos diferentes de seleção de preditores. O interesse desses modelos é contribuir para escala entre 2 semanas e 2 meses de antecedência, a chamada previsão *sub-seasonal to seasonal* (S2S).

### 5.1 Dados e Metodologia

Para selecionar os índices climáticos que contribuem para a previsão de precipitação nas principais bacias brasileiras foram aplicados dois métodos de análise: *i*) Método *Stepwise* (Draper e Smith [1998]) e *ii*) Busca Exaustiva com LSTM (Hochreiter e Schmidhuber [1997]).

A seleção de preditores foi dividida em duas etapas, na primeira etapa de seleção de preditores, foi utilizado o método *stepwise* (detalhe Apêndice F) para a escolha de índices para cada modelo de RLM, isto é, foram explorados os índices climáticos como preditores de modelo em diferentes períodos (previsões que variam entre 15 a 60 dias, isto é, previsões de 3 a 12 pântadas). Optou-se por utilizar o modelo de RLM para avaliar diversos períodos, pois o modelo apresenta fácil ajuste, robustez e baixo custo/tempo computacional.

Na segunda etapa, foi implementada uma busca exaustiva gerando um espaço de possíveis combinações de índices pré-selecionados no método *stepwise* em diferentes períodos (de 15 a 60 dias) e construindo modelos LSTM para avaliar a qualidade da combinação (Hochreiter e Schmidhuber [1997]), metodologia detalhada do modelo LSTM disponível no Apêndice H.

Os dados de precipitação utilizados foram os mesmos do Capítulo 4 (Identificação dos Padrões de Variabilidade Climática nas Principais Bacias Brasileiras) e os índices climáticos do Capítulo 3 (Padrões de Variabilidade Climática), no período de 1981 a 2018. Isto é, os dados diários de precipitação do CHIRPS e dos índices climáticos foram acumulados em pântadas, enquanto os dados mensais foram interpolados (polinômio de grau 3) (Boor [2001]).

Na primeira etapa, seleção pelo método *stepwise*, os índices climáticos e as precipitações em cada

bacia foram padronizados conforme a metodologia disponível no Apêndice A. Na segunda etapa, seleção pelo método de busca exaustiva com modelos de LSTM, os dados não foram previamente padronizados.

Após a seleção de preditores, foram desenvolvidos modelos multivariados, que utiliza diversas variáveis para prever a precipitação em uma bacia e captura relações entre índices e a precipitação na bacia. Para capturar as relações lineares, foi utilizado o Modelo de Regressão Linear (RLM - detalhado no Apêndice G) (Wilks [2006]) e as relações não lineares o Modelo de *Long Short-Term Memory* (LSTM - detalhado no Apêndice H) (Hochreiter e Schmidhuber [1997]).

Com intuito de comparação preliminar, foram utilizados os mesmo índices climáticos selecionados no método *stepwise* para o desenvolvimento dos modelos RML e LSTM para todas as bacias na etapa de ajuste, ou seja, modelos no instante  $t = 0$ .

Para os modelos de RML foram utilizados os índices climáticos selecionados no método *stepwise* em cada período de previsões (15 a 60 dias) para cada bacia brasileira, conforme as bacias selecionadas no Capítulo 4 (Identificação dos Padrões de Variabilidade Climática nas Principais Bacias Brasileiras, Figura 4.1). O conjunto de dados de entrada foram divididos em dois:

1. Conjunto de treinamento (85% do total de pântadas - período de 1981 a 2013): os dados utilizados nesta etapa são utilizados para a calibração e ajuste da equação de regressão;
2. Conjunto de teste (15% do total de pântadas - período de 2014 a 2018): os dados empregados durante a fase de teste, completamente ignorado durante a fase de treinamento.

Para o modelo de LSM foram utilizados os índices climáticos selecionados pelo método de busca exaustiva, em cada período de previsões (15 a 60 dias) para cada bacia brasileira. O conjunto de dados de entrada foram divididos em três:

1. Conjunto de treinamento (60% do total de pântadas - período de 1981 a 2002): os dados usados efetivamente para treinar o sistema de rede neural;
2. Conjunto de validação (25% do total de pântadas - período de 2003 a 2013): os dados usados durante a fase de treinamento para assegurar-se de que não ocorreu nenhum *overfitting*, ou seja, fenômeno que a rede neural memoriza os exemplos do conjunto de treinamento em vez de generalizar e aprender (Theodoridis e Koutroumbas [2009]);
3. Conjunto de teste (15% do total de pântadas - período de 2014 a 2018): os dados empregados durante a fase de teste, ou seja, dados não utilizados na fase de treinamento e validação.

Além disso, foi implementado um algoritmo de otimização de hiperparâmetros (em inglês, *Model Tuning*) (Yu e Zhu [2020]) que busca a combinação ótima de conjuntos de hiperparâmetros dentro de um espaço de busca, por exemplo, o número de neurônios, tamanho de lote, número de épocas, etc. Como principal objetivo, a técnica reduz a intervenção humana no processo de seleção de hiperparâmetros, e com isso obter um aumento da reprodutibilidade e melhora do desempenho dos modelos durante treinamento dos algoritmos.

Para avaliar a desempenho dos modelos desenvolvidos foi utilizado o erro quadrático médio (em inglês, *mean squared error*, MSE) como parâmetro de avaliação do modelo no conjunto de teste.

$$MSE = \frac{1}{n} \sum (F - O)^2 \quad (5.1)$$

onde, “ $F$ ” é o valor modelado e “ $O$ ” o valor observado.

## 5.2 Resultados e Discussão

### 5.2.1 Seleção dos Preditores Climáticos - Método *Stepwise*

Na primeira etapa, um refinamento da seleção de preditores foi realizado aplicando-se o método *stepwise* no conjunto de dados em pântadas de índices climáticos padronizados para cada bacia selecionada. Os índices explorados foram indicados previamente pelas análises desenvolvidas no Capítulo 4 (Identificação dos Padrões de Variabilidade Climática nas Principais Bacias Brasileiras), isto é, foi aplicada a seleção apenas os índices que tiveram relação de causalidade para as componentes principais de precipitação nas bacias brasileiras, portanto neste capítulo não foram explorados os índices AO, AMM, FS, SPI e NAO.

Os resultados dessa seleção para as defasagens testadas (defasagens 3, 6, 9 e 12 pântadas) estão detalhados na Tabela 5.1, que apresentam os índices climáticos selecionados para defasagens no tempo em cada bacia brasileira selecionada, que corresponde a previsão no período de 15, 30, 45 e 60 dias da precipitação. Além disso, a ordem dos preditores na tabela indicam a ordem de contribuição do índice climático para previsão de precipitação na bacia em cada defasagem.

A partir da Tabela 5.1, nota-se, que a quantidade de preditores diminui ao logo das defasagens, ou seja, os índices que contribuem para previsões mais curtas (15 dias) são maiores que para previsões mais longas (30 a 60 dias). Observa-se também, que em bacias ao sul e sudeste o número de preditores foi bem menor que os encontrados nas bacias ao norte, nordeste e central.

No geral, verifica-se o predomínio dos índices RMM1 e RMM2 para os períodos mais curtos de previsão (15 a 30 dias) e MEI, Niño 3.4 ou SOI para os períodos mais longos (30 a 60 dias). Resultado que sugere a influência da MJO e do ENSO na precipitação nas bacias em estudo, conforme os estudos de Kousky e Kayano [1994] e Tedeschi *et al.* [2016]. Além disso, observa-se que o índice AAO foi selecionado em defasagens longas e curtas em diversas bacias, corroborando com os estudos de Vasconcellos e Cavalcanti [2010] e Machado *et al.* [2021].

Para as bacias ao sul (Itaipu, Iguaçu, Jacuí e Uruguai), evidencia-se a contribuição do fenômeno ENSO na previsão de precipitação nas bacias, nota-se na Tabela 5.1 que o índice MEI foi selecionado para todas as bacias e períodos (de 15 a 60 dias). Enquanto os índices RMM1/RMM2 e LISAM aparecem em defasagens mais curtas (15 a 30 dias). Além disso, observa-se a contribuição do índice PDO na defasagem mais longa analisada para a bacia do Iguaçu, destacando a importância de fenômenos de baixíssima frequência em defasagens mais longas (12 pântadas), estudos de Malfatti *et al.* [2018b] são observadas correlações para a bacia de Itaipu em períodos mais longos (acima de 60 dias).

Observa-se que os fenômenos ENSO, MJO, AAO e PSA influenciam todas as bacias do sudeste, ENSO, AAO e PSA para os períodos mais longos e MJO períodos mais curtos. Além disso, os índices PSA e AAO são selecionados nos períodos mais curtos para as bacias Paranaíba e Rio Grande, enquanto os índices RMM1 e/ou RMM2 são observados para períodos mais longos apenas

para as bacias Paranapanema e Rio Grande.

Para as bacias do Sudeste, observam-se alguns padrões distintos: *i*) a sudoeste (Bacia do Paranapanema), em quase todos os períodos o método *stepwise* não selecionou nenhum índice para previsão de precipitação na bacia, exceto para 45 dias o índice AAO; *ii*) ao sul (bacia Paranapanema), possui característica semelhante da região Sul, com forte evidência do fenômeno ENSO em todas as defasagens e PDO mais longa; e *iii*) ao norte, (bacia do Rio Grande) observam-se características similares a região central, influência dos índices MEI e RMM1/RMM2 nas defasagens mais curtas, nas mais longas AAO e o índice PSA1 em longas e curtas; e *iv*) na bacia do Rio Tietê com características das bacias do Rio Grande e Paranapanema, evidência do ENSO mais curtas e padrões AAO e PSA1 mais longas.

Para as bacias centrais, o método selecionou índices RMM1/RMM2, AAO e PSA1/PSA2 em defasagens mais curtas (15 a 30 dias), destacando a influência da MJO (Kousky e Kayano [1994]), a AAO e PSA (Vasconcellos e Cavalcanti [2010]). Além disso, nota-se importância da AMO em defasagens longas e curtas, evidenciando a influência do Atlântico em bacias mais ao norte do Brasil, corroborando com os estudos de Jones e Carvalho [2018].

No nordeste (bacia do Rio São Francisco), observa-se a contribuição dos fenômenos MEI e AMO nas defasagens mais longas e os padrões PSA, LISAM, MJO e AAO para defasagens mais curtas, sendo bons preditores de precipitação. E, por fim, nas bacias ao norte, sugerem-se os índices do Atlântico (AMO, GTA, TNA e TSA) para todos os períodos, enquanto MJO, LISAM e AAO para os períodos mais curtos. Além disso, observa-se a seleção do ENSO e PNA em todas as bacias do norte para períodos curtos e longos.

Os resultados obtidos nas bacias do norte e nordeste corroboram com os observados por diferentes autores da influência do Pacífico (Foley *et al.* [2002]; Tedeschi *et al.* [2016]; Molion [2008]), Atlântico (Andreoli e Kayano [2006]), a MJO (Mayta *et al.* [2020]) e Padrões AAO e PSA Cavalcanti e Shiumizu [2012].

As vantagens de aplicação do método *stepwise* consiste na facilidade de ajuste e validação para diferentes combinações de preditores do modelo e baixo custo/tempo computacional, como desvantagens as variáveis de entrada precisam de tratamento prévio para capturar variabilidades diferentes do ciclo anual e por se tratar de um método linear não conseguem capturar variações não-lineares.

### 5.2.2 Seleção dos Preditores Climáticos - Busca Exaustiva

A partir da seleção pelo método *stepwise* foram considerados índices climáticos importantes para explicar linearmente o comportamento da precipitação na bacia. Porém, foram identificadas algumas limitações do método, principalmente em função da existência de padrões de comportamento não linear. O método de busca exaustiva busca mitigar o problema, gerando combinações de índices climáticos e experimentando em modelos LSTM que conseguem capturar variabilidades não-lineares.

No espaço de possibilidades, foram selecionados índices climáticos adicionais aos observados no método *stepwise*. A partir da busca exaustiva, notou-se que alguns índices climáticos melhoram o desempenho dos modelos, isto é, diminuição do MSE sem gerar *overfitting* dos modelos. Por esta razão também foram selecionados para essa bacia no desenvolvimento dos modelos de LSTM. A Tabela 5.2 destaca os índices selecionados para cada bacia independente do período de previsão.

**Tabela 5.1:** Seleção dos índices climáticos pelo método *stepwise* para previsão de 15 a 60 dias em cada bacia.

Bacia	15 dias	30 dias	45 dias	60 dias
Madeira	AMO, RMM2, LISAM RMM1, MEI	AAO, PNA, AMO	PNA, MEI, AMO RMM1	MEI, TNA
Xingu	RMM2, MEI, GTA RMM1, PNA, PDO AAO, LISAM	MEI, RMM1, PNA GTA	MEI, GTA, PNA	MEI, GTA, PNA AAO
Tocantins	RMM2, MEI, GTA RMM1, PNA, PDO AAO	AAO, RMM1, MEI	MEI, PNA	MEI, AMO, TSA
São Francisco	PSA2, LISAM	AAO, RMM1	AMO	MEI
Serra da Mesa	RMM2, PSA2, AAO LISAM, AMO	RMM1	AMO	AMO
Paranaíba	RMM2, RMM1, MEI PSA1, AAO	RMM2	PSA1	—
Rio Grande	MEI, RMM1, RMM2 PSA1	RMM2, MEI	AAO	PSA1
Rio Tietê	MEI	MEI	AAO	PSA1
Paraná	—	—	AAO	—
Paranapanema	MEI	MEI	MEI	MEI, PDO
Itaipu	MEI, RMM2	MEI	MEI	MEI, AAO
Iguaçu	MEI, RMM1	MEI, RMM1, LISAM	MEI	MEI, PDO
Jacuí	MEI	MEI, RMM1, LISAM	MEI	MEI
Uruguai	MEI, RMM1	MEI	MEI	MEI

Para as bacias do Norte, geralmente, observa-se a adição do índice ZCAS e padrão PSA (índices PSA1 e PSA2), e ao invés do índice TNA ou TSA unicamente, optou-se pelo Gradiente do Atlântico (GTA), isto é, a variação entre o Atlântico Tropical norte e sul. Além disso, foram excluídos do conjunto selecionado os índices como PNA, PDO e AAO, enquanto na bacia do Nordeste a exclusão do padrão MJO (índice RMM1) e adição do índice ZCAS.

As bacias do Centro-Oeste adicionou-se o índice ZCAS e remoção do padrão do Atlântico de baixíssima frequência para um padrão de maior frequência (índice TSA).

No Sudeste, a seleção anterior (*stepwise*) uma menor quantidade de índice foi selecionada para preditores de precipitação na bacia, enquanto nesta seleção, os padrões MJO e PSA são adicionados em todas as bacias. Observa-se a sobreposição dos índices Niño3.4 e MEI, evidenciando a influência do fenômeno ENSO nas bacias do Sudeste do Brasil. Situação semelhante observada para as bacias do Sul, com a adição do índice SOI que também representa o fenômeno ENSO. Porém o SOI caracteriza a diferença de pressão entre Tahiti e Darwin, enquanto o índice Niño 3.4 descreve



apenas as condições da TSM no Oceano Pacífico Tropical na região do Niño 3.4. Já o índice MEI é uma combinação de diversas variáveis como temperatura, pressão e radiação (para maiores detalhes, ver Capítulo 2 - Revisão Bibliográfica). Além disso, um detalhe observado nas bacias do Sul, é a adição de padrões do Atlântico (AMO e TSA), corroborando com estudos de [Cardoso e Silva Dias \[2004\]](#), [Chaves \[2011\]](#) e [Prado \*et al.\* \[2021\]](#).

**Tabela 5.2:** Índices Climáticos selecionados através da busca exaustiva com modelos de LSTM.

Macro Região	Bacia	Índices Climáticos
Norte	Madeira	LISAM, PSA1, PSA2, RMM1, RMM2, GTA, AMO, AAO
	Xingu	LISAM, PSA1, ZCAS, RMM1, GTA, PSA2
	Tocantins	LISAM, ZCAS, RMM1, PSA1, MEI, TSA
Nordeste	São Francisco	LISAM, ZCAS, MEI, PSA2, PSA1, AAO
Centro-Oeste	Serra da Mesa	LISAM, ZCAS, PSA1, PSA2, AAO, MEI, TSA
	Paranaíba	LISAM, ZCAS, PSA1, MEI, RMM1
Sudeste	Rio Grande	LISAM, ZCAS, PSA1, PSA2, RMM1, RMM2, AAO
	Rio Tietê	LISAM, ZCAS, RMM1, RMM2, MEI, NINO3.4
	Paranapanema	LISAM, ZCAS, RMM1, MEI, NINO3.4, PSA1, PSA2
	Paraná	LISAM, PSA1, PSA2, RMM1, RMM2, AAO
Sul	Itaipu	LISAM, MEI, PSA1, RMM1, PSA2, AMO, NINO3.4, PDO
	Iguaçu	LISAM, ZCAS, MEI, AMO, RMM1, RMM2, NINO3.4, TSA, PNA
	Jacuí	LISAM, ZCAS, MEI, NINO3.4, RMM1, PDO, AMO, SOI
	Uruguai	LISAM, ZCAS, RMM1, MEI, NINO3.4, SOI

As vantagens desta técnica são relacionadas ao tratamento das variáveis e captura de padrões não lineares, isto é para adicionar as variáveis de entrada no modelo, não é necessária padronização (remoção do ciclo anual) dos preditores e nem preditas. Porém, como desvantagens estão no custo/tempo computacional mais elevado, como são 14 bacias com 16 índices climáticos, que geram diferentes combinações (espaço de busca).

### 5.2.3 Modelagem dos Modelos de RLM e LSTM

Para avaliar a contribuição de cada conjunto de preditores selecionado pelos métodos, *stepwise* e busca exaustiva, foram desenvolvidos modelos de regressão linear múltipla e de LSTM para todas as bacias selecionadas.

Em um primeiro momento, foi realizado um teste preliminar de ajuste do modelo no instante  $t = 0$  (isto é, sem defasagem dos preditores, ocorre no mesmo instante que a precipitação na bacia), com os índices selecionados pelo método *stepwise* para comparação inicial entre os resultados dos modelos. Na Tabela 5.3 são apresentados os resultados deste experimento inicial.

Nota-se, na Tabela 5.3, que o modelo RLM apresenta melhor desempenho, expresso pelo MSE baixo em todas as bacias selecionadas em comparação com o LSTM, indicando que o método *stepwise* seleciona os preditores linearmente e captura os padrões lineares dos fenômenos climáticos em estudo. Comportamento esperado, visto que o método *stepwise* utiliza o modelo RLM para definição das variáveis preditoras.

Em busca de melhores resultados para o modelo LSTM, foram utilizados os preditores selecionados pelo método de busca exaustiva e ajuste de hiperparâmetros de rede.

Na Etapa I os modelos foram desenvolvidos sem ajuste dos hiperparâmetros, enquanto na Etapa II a rede de LSTM foi ajustada interativamente para seleção da melhor configuração de rede [Bergstra e Bengio \[2012\]](#). Isto é, são utilizadas várias configurações de rede até atingir o menor

**Tabela 5.3:** Erro Quadrático Médio (em inglês, mean squared error - MSE) em mm para modelos desenvolvidos por Regressão Linear Múltipla (RLM) e Long Short-Term Memory (LSTM) para cada bacia brasileira selecionada no experimento preliminar.

Macro Região	Bacia	RLM ( $t = 0$ )	RLM (15 dias)	LSTM ( $t = 0$ )	LSTM (15 dias)
Norte	Madeira	9,29	9,82	10,87	11,26
	Xingu	13,39	13,96	14,16	16,51
	Tocantins	10,80	13,85	15,80	14,35
Nordeste	São Francisco	8,37	12,60	11,82	14,21
Centro-Oeste	Serra da Mesa	12,46	17,58	16,89	17,45
	Paranaíba	12,56	14,54	15,84	15,46
Sudeste	Rio Grande	14,00	15,56	20,20	17,49
	Rio Tietê	14,80	16,04	17,02	18,89
	Parapanema	18,22	19,49	19,18	20,59
	Paraná	14,20	15,56	18,21	15,70
Sul	Itaipu	20,86	21,60	21,69	22,51
	Iguaçu	22,69	23,54	23,01	23,56
	Jacuí	24,50	26,17	29,18	27,37
	Uruguai	19,07	20,97	22,54	22,69

erro possível. Os resultados comparativos (Etapa I e II) para os modelos RLM e LSTM com previsão de 60 dias, estão disponíveis na Tabela 5.4.

Para os modelos RLM e LSTM (Etapa I) foram utilizados os preditores selecionados pelo método *stepwise*, enquanto na Etapa II para os modelos de LSTM foram utilizados os preditores selecionados pela busca exaustiva.

Na Tabela 5.4, observa-se que após a otimização dos hiperparâmetros e ajuste dos preditores de entrada dos modelos de LSTM nas bacias, os valores de MSE sofreram uma redução significativa, apresentando um melhor desempenho do modelo de LSTM quando utilizada a seleção dos índices climáticos com a técnica de busca exaustiva, exceto na bacia do Rio Madeira e do Rio Iguaçu.

**Tabela 5.4:** Erro Quadrático Médio (em inglês, mean squared error - MSE) em mm para modelos desenvolvidos por Regressão Linear Múltipla (RLM) e Long Short-Term Memory (LSTM) para cada bacia brasileira selecionada nas etapas I e II.

Macro Região	Bacia	RLM 60 dias (Etapa I)	LSTM 60 dias (Etapa I)	LSTM 60 dias (Etapa II)
Norte	Madeira	9,89	10,42	10,35
	Xingu	13,91	14,16	13,34
	Tocantins	13,42	15,80	11,29
Nordeste	São Francisco	12,48	11,82	8,79
Centro-Oeste	Serra da Mesa	17,17	16,89	13,40
	Paranaíba	14,49	15,84	13,50
Sudeste	Rio Grande	15,74	20,20	14,92
	Rio Tietê	15,91	17,02	15,03
	Parapanema	19,16	19,18	18,26
	Paraná	15,53	18,21	15,21
Sul	Itaipu	21,60	21,69	21,22
	Iguaçu	23,47	23,01	23,20
	Jacuí	26,25	29,18	23,71
	Uruguai	21,02	22,54	19,74

Para avaliar a redução do MSE, consequentemente o desempenho dos modelos e maior acurácia média das previsões, a Tabela 5.5 apresenta a porcentagem de redução do MSE em cada bacia em dois casos: *i*) entre as Etapas I e II do modelo LSTM e *ii*) entre o modelo RLM na Etapa I e modelo LSTM na Etapa II.

Conforme a Tabela 5.5, nota-se que a bacia do Madeira não apresentou redução significativa

(abaixo de 1%) quando comparada com o modelo LSTM desenvolvido na etapa anterior, além disso, quando comparado ao modelo RLM da Etapa I, o desempenho é cerca de 4,7% menor, indicando que a bacia do Rio Madeira possui um comportamento linear ou que as oscilações de precipitação nesta bacia não são capturadas, ou seja, os índices utilizados não estejam adequadamente vendo esses processos. Enquanto para a bacia do Iguaçu observa-se comportamento inverso, ou seja, melhora no desempenho em comparação com o modelo RML da Etapa I e piora em relação ao modelo LSTM da Etapa I.

Com exceção dos casos destacados acima, todos os outros modelos de LSTM desenvolvidos na Etapa II apresentam melhor desempenho (MSE abaixo do anterior). As bacias Tocantins, São Francisco e Serra da Mesa apresentaram redução do MSE entre 15% a 30%, evidenciando o comportamento não linear da precipitação nas bacias.

**Tabela 5.5:** *Redução do Erro Quadrático Médio (em inglês, mean squared error - MSE) entre LSTM nas duas etapas (Etapa I e II) e entre o LSTM (Etapa II) com o RML (Etapa I) para cada bacia brasileira selecionada.*

Macro Região	Bacia	Redução do Erro (Etapa I - LSTM)	Redução do Erro (Etapa I - RML)
Norte	Madeira	0,7%	-4,7%
	Xingu	5,8%	4,1%
	Tocantins	28,5%	15,9%
Nordeste	São Francisco	25,6%	29,6%
Centro-Oeste	Serra da Mesa	20,7%	22,0%
	Paranaíba	14,8%	7,5%
Sudeste	Rio Grande	26,1%	5,2%
	Rio Tietê	11,7%	5,5%
	Parapanema	4,8%	4,7%
	Paraná	16,5%	2,1%
Sul	Itaipu	2,2%	1,8%
	Iguaçu	-0,8%	1,2%
	Jacuí	18,7%	9,7%
	Uruguai	12,4%	6,1%

## 5.3 Conclusões

Neste capítulo foi possível desenvolver modelos empíricos estatísticos que capturam padrões lineares e não lineares dos fenômenos climáticos (preditores climáticos, isto é os índices climáticos), e transforma-los em padrões para predição da precipitação em bacias brasileiras selecionadas em estudo para o intervalo de 15 a 60 dias (3 a 12 pântadas).

O método de *stepwise* foi útil para a seleção de preditores significativos nos modelos RLM. No geral, observa-se o predomínio dos índices RMM1 e RMM2 para os períodos mais curtos de previsão (15 a 30 dias) e MEI, Niño 3.4 ou SOI para os períodos mais longos (30 a 60 dias), resultado que evidencia a influência da MJO e do ENSO na precipitação no Brasil. Enquanto os padrões AAO e PSA são selecionados em defasagens longas e curtas em diversas bacias.

Em adição, algumas bacias do Sul e Sudeste apresentam contribuição do índice PDO na defasagem mais longa, enquanto as bacias do Norte e Nordeste são sugeridos índices do Atlântico (AMO, GTA, TNA e TSA) em todos os períodos de predição.

Na etapa preliminar, mostrou-se insuficiente o método de seleção *stepwise* para o desenvolvimento de modelos de LSTM. Tal resultado evidenciou a importância de uma seleção mais robusta

dos índices, por esse motivo, adotou-se a seleção por busca exaustiva, por meio do treinamento de modelos LSTM com várias combinações possíveis de preditores até atingir os melhores conjuntos (isto é, menores MSE) em cada bacia de estudo. No geral, a busca exaustiva adicionou índices ZCAS, PSA1 e PSA2 para todas as bacias e a exclusão de índices menos expressivos como: PNA e PDO.

E por fim, os resultados comparativos entre os modelos desenvolvidos na etapa I (ou preliminar, modelo LSTM e RLM com preditores do método *stepwise*) e II (modelo LSTM com ajuste de hiperparâmetros e seleção de preditores pela busca exaustiva) sugerem melhor desempenho do modelo LSTM, quando utilizada a seleção dos índices climáticos com a técnica de busca exaustiva, exceto na bacia do Rio Madeira e do Rio Iguaçu. Destaca-se também a redução do erro de até 30% em bacias do Norte, Nordeste e Centro-Oeste (Tocantins, São Francisco e Serra da Mesa).



# Capítulo 6

## Modelos com Predição S2S

Neste capítulo, são discutidos os resultados da aplicação dos modelos utilizando como preditores os índices climáticos e adicionando a previsão extendida dos modelos disponíveis no programa *sub-seasonal to seasonal* (S2S). Os modelos de previsão de precipitação, são desenvolvidos com as mesmas técnicas de modelagem linear (RML - Regressão Linear Múltipla) e não linear (LSTM - *Long Short-Term Memory*) desenvolvidas no Capítulo anterior (Capítulo 5: Desenvolvimento de Modelos Preditivos de Precipitação nas Principais Bacias Brasileiras).

### 6.1 Dados e Metodologia

No âmbito do programa S2S foram testadas dois modelos dinâmicos de previsão de precipitação, os modelos são construídos a partir de um conjunto de sistemas de previsão global (em inglês, *global ensemble forecast system*) e os dados de saída dos modelos estão disponíveis em <http://s2sprediction.net/> com frequência de atualização variável, entre 2-3 dias e período de previsão de até 45 dias. Essas previsões de precipitação são incorporadas como preditores dos modelos empíricos RLM e LSTM com os índices climáticos, conforme a mesma metodologia desenvolvida no Capítulo 5 (Desenvolvimento de Modelos Preditivos de Precipitação nas Principais Bacias Brasileiras).

A Tabela 6.1 mostra uma breve descrição dos modelos selecionados do programa S2S (Modelos ECMWF e NCEP), o período inicial dos dados disponíveis, a referência do modelo e *hyperlink* para detalhes técnicos do modelo.

**Tabela 6.1:** Informações dos modelos dinâmicos disponíveis pelo programa S2S utilizados no estudo.

Modelo	Descrição do Modelo	Período	Referência
ECMWF	Sistema de previsão por conjunto global que simula incertezas iniciais usando vetores singulares e conjunto de assimilação de dados e incertezas do modelo devido a parametrizações físicas usando um esquema estocástico, com base em 51 membros. <a href="http://s2s.cma.cn/Models/ECMWF">http://s2s.cma.cn/Models/ECMWF</a>	1998 – atual	<a href="#">ECMWF [2020]</a>
NCEP	Sistema de previsão por conjunto global para previsões mensais e sazonais, com base em 16 membros. <a href="http://s2s.cma.cn/Models/NCEP_1">http://s2s.cma.cn/Models/NCEP_1</a>	1999 – atual	<a href="#">Saha et al. [2010]</a>

Neste capítulo, foram utilizados como preditores dos modelos RLM e LSTM a precipitação prevista dos modelos ECMWF e NCEP e os índices climáticos observados até a pêntrada em que se inicia a previsão da precipitação. A precipitação prevista dos modelos ECMWF e NCEP ajudam a corrigir a previsão e reforçam o padrão da precipitação na bacia.

No caso específico do estudo são considerados os índices climáticos e a previsão do modelo ECMWF ou NCEP de 5 dias anteriores e o próprio dia que está sendo previsto para construir o valor da previsão diária na bacia. Da mesma forma, a série de saída (precipitação prevista) pode ser simulada usando uma camada densa, totalmente conectada, com quantidade de unidades igual ao comprimento da série de saída. Neste caso, foi considerado que para cada amostra de 5 dias anteriores foi calculado o valor de previsão que corresponde a geração de 9 pântadas futuras, método que evita a perda de informação. Em estudos futuros, podem ser testadas abordagens com séries de saída de comprimento maior devido ao fato que os modelos ECMWF e NCEP no programa S2S só possuem informação de 45 dias, e o alvo é a previsão estendida até 60 dias.

A seleção dos preditores seguiu-se com configuração semelhante ao capítulo anterior (Seção 5.1: Seleção dos Preditores Climáticos - Método *Stepwise*), para o desenvolvimento preliminar optou-se por utilizar os preditores obtidos no método *stepwise* (maiores detalhes ver Anexo F). Enquanto para uma visão mais acurada do modelo LSTM, a busca exaustiva (maiores detalhes ver Anexo H).

Inicialmente as variáveis utilizadas não foram padronizadas (sem remoção do ciclo anual), pois a série temporal dos modelos ECMWF e NCEP não é contínua no tempo, ou seja, as previsões são geradas para 45 dias, o que dificulta o processo de padronização da variável. No entanto, não se observa melhora ao utilizar LSTM sobre os dados padronizados. Porém, esta não remoção do ciclo anual pode comprometer os resultados dos modelos de RLM, dificultando a captura de padrões de variabilidade na escala sub-sazonal.

Os dados obtidos dos modelos ECMWF e NCEP são previsões diárias acumuladas no período de 1999 a 2018. Optou-se por desenvolver modelos diários, porque os modelos rodam diariamente ou cada 2-3 dias e torna-se difícil coincidir uma pântada da previsão do modelo com a série de índices em pântadas. Por este motivo, e pela quantidade de amostras geradas ao trabalhar com pântadas, definiu-se para este estudo utilizar dados diários nas modelagens (RLM e LSTM). Os dados dos índices climáticos, no período de 1999 a 2018, que não possuem dados diários foram interpolados (polinômio de grau 3) (Boor [2001]). Além disso, neste experimento foi utilizado um método de extrapolação de índices para coincidir com a pântada da precipitação prevista (modelo LSTM utilizando como entrada apenas o próprio índice climático a ser previsto).

O modelo RML foi desenvolvido apenas na etapa preliminar, utilizando os índices selecionados no método *stepwise* e a precipitação prevista na bacia pelos modelos do programa S2S, ou seja, aplicou-se o modelo para comparação com a LSTM apenas para a etapa de ajuste no instante  $t = 0$ . O conjunto de dados de entrada foram divididos em dois:

1. Conjunto de treinamento (85% do total de pentadas - período de 1999 a 2015): os dados utilizados nesta etapa são utilizados para a calibração e ajuste da equação de regressão;
2. Conjunto de teste (15% do total de pentadas - período de 2016 a 2018): os dados empregados durante a fase de teste, completamente ignorado durante a fase de treinamento.

No desenvolvimento do modelo LSM foram utilizados os índices climáticos e série prevista dos modelos dinâmicos selecionados pelos dois métodos: *i*) Etapa I: método *stepwise* e *ii*) Etapa II: busca exaustiva para cada bacia brasileira e o conjunto de dados de entrada foram divididos em três:

1. Conjunto de treinamento (60% do total de pentadas - período de 1999 a 2002): os dados usados efetivamente para treinar o sistema de rede neural;
2. Conjunto de validação (25% do total de pentadas - período de 2003 a 2015): os dados usados durante a fase de treinamento para assegurar-se de que não ocorreu nenhum *overfitting*, ou seja, fenômeno que a rede neural decora os exemplos do conjunto de treinamento em vez de generalizar e aprender (Theodoridis e Koutroumbas [2009]);
3. Conjunto de teste (15% do total de pentadas - período de 2016 a 2018): os dados empregados durante a fase de teste, ou seja, dados não utilizados na fase de treinamento e validação.

A saída dos modelos é transformada em pântada (acumulado de 5 dias) e sobre a série em pântadas calcula-se o erro quadrático médio (MSE) para cada bacia, como parâmetro de avaliação do modelo, conforme a equação do capítulo anterior (Equação 5.1).

## 6.2 Resultados e Discussão

Para mensurar a contribuição da precipitação prevista pelos modelos dinâmicos do programa S2S, ECMWF e NCEP, foram desenvolvidos modelos de regressão linear múltipla (RLM) e de *Long Short-Term Memory* (LSTM) para todas as bacias selecionadas.

Como teste preliminar do modelo, foi realizado o ajuste dos modelos no instante  $t = 0$  (isto é, sem defasagem dos preditores, ocorre no mesmo instante que a precipitação na bacia), com os índices selecionados pelo método *stepwise* e a precipitação prevista do ECMWF e NCEP, para comparação inicial entre os resultados dos modelos e entre o conjunto de preditores climáticos incluindo e/ou precipitação prevista dos modelos dos S2S. Na Tabela 6.2 são apresentados os resultados deste experimento inicial.

Na Tabela 6.2, observa-se que os resultados melhoram ainda para técnicas que se veem afetadas pela parametrização (ex. RLM), evidenciando a melhoria da qualidade da previsão segundo a métrica utilizada quando são combinadas os índices climáticos com a previsão de modelos meteorológicos como ECMWF.

Neste resultado preliminar, quando se aplica a seleção pelo método *stepwise*, observa-se o favorecimento dos modelos lineares, porém essa seleção prejudica os modelos não lineares (LSTM). Por exemplo, para a bacia de Rio Madeira utilizando os índices selecionados (conjunto de todos os preditores para 15 a 60 dias na Tabela 5.1: LISAM, RMM1, PSA1, RMM2, AAO, GTA, PNA, TNA e AMO) com a precipitação prevista do ECMWF o modelo RLM apresenta  $MSE = 8,82$ , porém quando se utiliza a mesma combinação de preditores no modelo LSTM o MSE aumenta para 16,07. Quando são utilizados os índices selecionados pela busca exaustiva (conforme a Tabela 5.2) com a precipitação prevista do ECMWF, o MSE decresce para 9,02, sendo esse valor similar ao obtido pelo modelo RLM, evidenciando a seleção dos preditores pelo método de busca exaustiva.

De uma maneira geral, observa-se na Tabela 6.2 um ganho significativo na utilização da precipitação prevista para os modelos desenvolvidos por LSTM para a maioria das bacias brasileiras selecionadas. Enquanto para os modelos de RLM a adição da precipitação prevista não trouxe benefícios, podendo ter influência da não padronização das séries temporais (sem remoção do ciclo



anual) e inseridas nos modelos apenas com padronização simples (*standard score* - desvio padrão total sobre o valor médio total).

Entre os modelos dinâmicos ECMWF e NCEP são identificadas comportamentos distintos, para as bacias Xingu, São Francisco, Paranaíba, Rio Grande, Rio Tietê, Paraná e Jacuí, a precipitação prevista do NCEP apresentando maiores ganhos do que para o modelo ECMWF, enquanto para as outras bacias (Madeira, Tocantins, Serra da Mesa, Paranapanema, Itaipu, Iguaçu e Uruguai) situação se inverte, modelos ECMF apresentam resultados melhores. Entretanto, os resultados obtidos não são melhores que o LSTM ou RLM com apenas os índices climáticos, sendo justificado por duas situações: *i*) otimização de hiperparâmetros do modelo, como: número de camadas, neurônios e etc, e *ii*) extrapolação robusta dos índices climáticos.

**Tabela 6.2:** *Erro Quadrático Médio (em inglês, mean squared error - MSE) em mm para modelos desenvolvidos por Regressão Linear Múltipla (RML) e Long Short-Term Memory (LSTM) para cada bacia brasileira selecionada no experimento preliminar e comparativo com os modelos desenvolvidos com os índices climáticos e os índices climáticos + previsão dos modelos dinâmicos do programa S2S (ECMWF e NCEP).*

Macro Região	Bacia	RLM (Índices)	RLM (Índices + ECMWF)	LSTM (Índices)	LSTM (Índices + ECMWF)	LSTM (Índices + NCEP)
Norte	Madeira	9,29	8,82	10,42	9,02	11,90
	Xingu	13,39	12,62	14,16	21,60	15,55
	Tocantins	10,80	11,82	15,80	11,89	13,30
Nordeste	São Francisco	8,37	10,54	11,82	14,01	9,93
Centro-Oeste	Serra da Mesa	12,46	14,79	16,89	14,91	15,68
	Paranaíba	12,56	13,30	15,84	14,92	14,50
Sudeste	Rio Grande	14,00	14,96	20,20	18,15	15,04
	Rio Tietê	14,80	15,37	17,02	16,80	16,18
	Paranapanema	18,22	18,23	19,08	18,92	19,61
	Paraná	14,20	15,55	18,21	18,45	15,62
Sul	Itaipu	20,86	20,12	21,69	21,18	21,60
	Iguaçu	22,08	21,97	23,01	21,01	24,48
	Jacuí	24,50	25,57	29,18	25,20	25,00
	Uruguai	19,07	19,31	22,54	18,39	20,18

Para prosseguir com o desenvolvimento dos modelos preditivos optou-se por selecionar a precipitação prevista do modelo NCEP, pois a disponibilidade dos dados é mais contínua no tempo do que o ECMWF.

A Tabela 6.3 apresenta os resultados da seleção de preditores pela busca exaustiva para cada bacia para as previsões de 15 e 45 dias, e a precipitação prevista do modelo NCEP. Como resultado, é possível observar que os índices climáticos LISAM e ZCAS não aportam muita informação para períodos superiores a 3 pênadas (ou seja, acima de 15 dias).

Em trabalhos futuros serão exploradas entradas distintas de preditores do modelo, como previsão dos índices climáticos por técnicas que possam apresentar uma maior acurácia nas previsões para serem posteriormente utilizados como preditores dos modelos de previsão de precipitação na escala acima de 15 dias.

A Tabela 6.4 aponta os resultados de MSE obtidos em cada bacia brasileira para as previsões de 15 e 45 dias. Os modelos de LSTM foram desenvolvidos em uma única rodada de uma arquitetura fixa, isto é, os modelos não foram otimizados para conquistar uma arquitetura de rede com o melhor desempenho.

Os resultados apresentados na Tabela 6.4 indicam um resultado preliminar dos modelos quando

**Tabela 6.3:** Índices Climáticos selecionados através da busca exaustiva com modelos de LSTM e precipitação prevista do modelo NCEP.

Macro Região	Bacia	LSTM (15 dias)	LSTM (45 dias)
Norte	Madeira	NCEP, AMO, RMM2, LISAM RMM1, AAO, PNA	NCEP, RMM1, RMM2, AMO
	Xingu	NCEP, RMM2, MEI, GTA RMM1, PNA, PDO, AAO, LISAM	NCEP, PSA1, RMM1, MEI GTA, NIÑO3.4
	Tocantins	NCEP, RMM2, MEI, GTA RMM1, PNA, PDO, AAO	NCEP, PSA1, RMM1, MEI AAO
Nordeste	São Francisco	NCEP, ZCAS, LISAM RMM1, PSA1, AAO	NCEP, PSA2, RMM1, NAO SOI, AMO, AAO
Centro-Oeste	Serra da Mesa	NCEP, LISAM, ZCAS, RMM1 PSA1, PSA2, AAO, TSA	NCEP, RMM1, PSA1, PSA2 AAO, TSA
	Paranaíba	NCEP, LISAM, ZCAS, RMM1 PSA1, AAO, PSA2	NCEP, RMM1, PSA1, AAO PSA2
Sudeste	Rio Grande	NCEP, LISAM, ZCAS, RMM1 PSA2, RMM2, PSA1	NCEP, RMM1, PSA2, RMM2 PSA1
	Rio Tietê	NCEP, LISAM, ZCAS, RMM1 PSA2, RMM2, PSA1	NCEP, RMM1, PSA2, RMM2 PSA1
	Paranapanema	NCEP, MEI	NCEP, MEI
	Paraná	NCEP, LISAM, RMM1 RMM2, PSA1, PSA2	NCEP, RMM1, RMM2, PSA1 PSA2
Sul	Itaipu	NCEP, MEI, RMM2	NCEP, MEI, RMM2 NIÑO3.4
	Iguaçu	NCEP, MEI, RMM1, LISAM	NCEP, MEI, PNA, RMM1 NIÑO3.4, RMM2
	Jacuí	NCEP, MEI, RMM1, LISAM	NCEP, NIÑO3.4, PDO, MEI AAO, RMM1
	Uruguai	NCEP, NIÑO3.4, RMM1	NCEP, NIÑO3.4, MEI, PDO RMM1, AAO

inserido como preditor a precipitação prevista pelo modelo NCEP do programa S2S. Em trabalho futuros, o ajuste fino de hiperparâmetros, uma análise e modelagem detalha de extrapolação dos índices climáticos também devem se considerar.

**Tabela 6.4:** Erro Quadrático Médio (em inglês, mean squared error - MSE) para modelos desenvolvidos em LSTM para cada bacia brasileira selecionada.

Macro Região	Bacia	Índices + NCEP (15 dias)	Índices + NCEP (45 dias)
Norte	Madeira	11,90	11,83
	Xingu	15,55	16,22
	Tocantins	14,44	15,47
Nordeste	São Francisco	14,56	13,71
Centro-Oeste	Serra da Mesa	17,59	19,93
	Paranaíba	16,19	19,75
Sudeste	Rio Grande	17,86	18,43
	Rio Tietê	21,25	20,52
	Paranapanema	25,11	24,48
	Paraná	18,81	20,02
Sul	Itaipu	21,92	22,23
	Iguaçu	25,11	24,48
	Jacuí	24,09	25,00
	Uruguai	20,32	21,06

## 6.3 Conclusões

Diante do desafio de produzir previsões mais acuradas para a escala sub-seasonal, foram desenvolvidos modelos de previsão de precipitação com preditores climáticos de variabilidade espacial

e/ou temporal distintas, com as previsões dos modelos dinâmicos produzidos pelo NCEP e ECMWF no programa S2S.

Os resultados preliminares apontam que a precipitação prevista dos modelos dinâmicos ECMWF e NCEP ajudam a corrigir a previsão e reforçar o padrão da precipitação na bacia, visto que resultados dos MSE apontam melhor desempenho. Para as bacias Xingu, São Francisco, Paranaíba, Rio Grande, Rio Tietê, Paraná e Jacuí, a precipitação prevista do NCEP apresentando maiores ganhos do que para o modelo ECMWF, enquanto para as outras bacias (Madeira, Tocantins, Serra da Mesa, Paranapanema, Itaipu, Iguaçu e Uruguai) situação se inverte, modelos ECMWF apresenta resultados melhores.

Conclui-se que os resultados obtidos não são melhores que o LSTM ou RLM com apenas os índices climáticos, porém possui potencial de melhora, após o ajuste de hiperparâmetros do modelo, conforme apontado no capítulo anterior (Capítulo 5: Desenvolvimento de Modelos Preditivos de Precipitação nas Principais Bacias Brasileiras). O ajuste dos parâmetros podem alterar significativamente os resultados dos testes.

Além disso, para coincidir com a pântada da precipitação prevista dos modelos NCEP e ECMWF, os índices climáticos foram extrapolados utilizando modelo de LSTM e como entrada apenas o próprio índice climático a ser previsto.

Em estudos futuros, sugere-se aprimoramento dos modelos desenvolvidos, explorando novas arquiteturas através do ajuste fino dos hiperparâmetros e refinamento na extrapolação dos preditores climáticos.

Em adição, podem ser testadas abordagens com séries de saída de comprimento maior devido ao fato que os modelos ECMWF e NCEP só possuem informação de 45 dias, e o alvo é a previsão estendida de até 60 dias.

# Capítulo 7

## Conclusões

Modelagem baseada em dados, ou seja, modelos empíricos baseados no comportamento observado no passado é bastante sensível à qualidade/quantidade de dados disponíveis, para o diagnóstico e/ou prognósticos de padrões lineares e/ou não-lineares. Logo, são necessários estudos prévios para entendimento da variabilidade temporal, frequência dos padrões climáticos e as possíveis variáveis preditoras dos modelos em cada bacia brasileira.

No diagnóstico dos padrões climáticos, a análise de transformada de *wavelet* identificou a variabilidade nas escalas intrasazonal, interanual, interdecadal e multidecadal dos fenômenos estudados. Os resultados obtidos são coerentes com resultados anteriores, mas ampliam o detalhamento na análise de alguns índices que não são disponibilizados regularmente pelos órgãos cooperacionais (como o LISAM, ZCAS, PSA1 e PSA2). A periodicidade do ENSO (2 a 7 anos), PDO (20 a 30 anos), AMO (entre 30 a 60 anos), AAO, AMM, TNA e TSA (0,5 a 2 anos), QBO (2,3 anos), PSA1 (3 a 5 anos), PSA2 (1,5 ano) e os fenômenos intrasazonais, LISAM, ZCAS, RMM1 e RMM2, que variam na escala do S2S e de 4 a 8 meses. Nas análises de causalidade, sugere-se que o ENSO causa os padrões de variabilidade do Atlântico, AMO e GTA, e do Pacífico, PDO, destacando a relação não-linear ENSO  $\rightarrow$  PDO e ENSO  $\rightarrow$  GTA. Além disso, observa-se a influência PDO  $\rightarrow$  GTA e PDO  $\rightarrow$  AAO, evidenciando a transferência de energia na escala de tempo interanual do Oceano Pacífico no Atlântico.

Para identificar como os padrões climáticos interferem na escala intrasazonal da precipitação em diferentes bacias brasileiras, a dimensão do conjunto de dados foi reduzido através do uso da técnica de decomposição em componentes principais, e as componentes resultantes (modos de precipitação) correspondem em 87% da variabilidade total. No geral, os modos de precipitação das bacias são influenciados pelos padrões MJO (índices RMM1 e/ou RMM2), AAO, PSA (índices PSA1 e/ou PSA2) e SMAS (índices LISAM e/ou ZCAS). Além disso, destaca a importância dos padrões do Pacífico, ENSO e PDO, para quase todas as bacias brasileiras, exceto para noroeste da região Norte (Rio Madeira) e destaque para as bacias do Sul (Itaipu, Iguazu, Jacuí e Uruguai) e Sudeste do Brasil (Rio Grande, Rio Tietê, Paranapanema e Paraná), enquanto nas bacias do Norte (Madeira, Xingu e Tocantins), Nordeste (São Francisco) e Centro-Oeste (Serra da Mesa e Paranaíba) apresentam maior influência do Atlântico, com os padrões AMO e TNA/TSA.

Nos resultados das correlações defasadas, observa-se que o tempo de resposta da precipitação para a vazão de rios podem alterar conforme a bacia. Na bacia do Rio Madeira (Noroeste do Brasil),

o tempo de resposta varia entre 15-30 dias, ao Norte (Xingu e Tocantins) o tempo de resposta é mais lento variando entre 15-40 dias, nas bacias do Centro-Oeste (Serra Mesa e Paranaíba) Sul (Itaipu, Iguaçu, Jacuí e Uruguai) e Sudeste (Rio Grande, Rio Tietê, Paranapanema e Paraná) o tempo de resposta é mais rápido do que as regiões anteriores, entre 5-20 dias. Os resultados obtidos ressaltam a importância dos padrões de variabilidade climática de baixa frequência nas variações intrasazonais da precipitação em diversas bacias brasileiras.

Para avaliar, a hipótese dos padrões climáticos serem bons preditores de precipitação foram desenvolvidos modelos empíricos estatísticos, que evidenciem as relações lineares (modelo RLM) e não-lineares (modelo LSTM) entre os fenômenos climáticos e a precipitação nas principais bacias brasileiras na escala subsazonal a sazonal (15 a 60 dias que corresponde de 3 a 12 pântadas).

O modelo LSTM, mostrou-se superior após a seleção adequada pela busca exaustiva dos preditores e ajuste dos hiperparâmetros dos modelos desenvolvidos em quase todas as bacias, exceto na bacia do Rio Madeira influenciada por fatores locais na região dos Andes, e do Rio Iguaçu influenciado pela combinação de sistemas frontais e do fluxo de umidade oriundo da Amazônia. Os principais índices selecionados nos períodos mais curtos de previsão (15 a 30 dias) foram o RMM1/RMM2 e LISAM/ZCAS, e os índices MEI, Niño 3.4 ou SOI para os períodos mais longos (30 a 60 dias), resultado que evidencia a influência da MJO e do ENSO na precipitação no Brasil. Em adição, os padrões AAO e PSA são selecionados em desfasagens longas e curtas em diversas bacias, e nas bacias do Norte e Nordeste são sugeridos índices do Atlântico (AMO, GTA, TNA e TSA) em todos os períodos de predição.

Diante do desafio de produzir previsões mais acuradas para a escala sub-sazonal, foram desenvolvidos modelos de previsão de precipitação com preditores climáticos de variabilidade espacial e/ou temporal distintas, com as previsões dos modelos dinâmicos produzidos pelo NCEP e ECMWF no programa S2S. Os resultados apontam que a precipitação prevista dos modelos dinâmicos ECMWF e NCEP contribuem positivamente para a melhoria da qualidade da previsão e reforçar o padrão da precipitação na bacia, visto que o erro quadrático médio (MSE) é reduzido. Para as bacias Xingu, São Francisco, Paranaíba, Rio Grande, Rio Tietê, Paraná e Jacuí, a precipitação prevista do NCEP apresenta maiores ganhos do que o obtido ao se usar a precipitação prevista pelo modelo ECMWF, enquanto para as outras bacias (Madeira, Tocantins, Serra da Mesa, Paranapanema, Itaipu, Iguaçu e Uruguai) situação se inverte, modelos ECMWF apresenta resultados melhores.

É importante observar que apesar de o modelo ECMWF ser considerado melhor do que o NCEP em métricas de avaliação globais, regionalmente pode ocorrer o inverso, principalmente no caso da precipitação que é notadamente uma variável de previsibilidade muito mais baixa do que de outros campos meteorológicos como a pressão e temperatura. Os resultados obtidos nos capítulos 5 e 6 evidenciam a necessidade de se explorar conjuntos de previsões obtidos por diferentes centros meteorológicos para otimizar o desempenho da previsão da precipitação aprimorada pelo uso de técnicas estatísticas e/ou inteligência artificial.

Conclui-se que os resultados obtidos não são melhores que o LSTM ou RLM com apenas os índices climáticos, porém após o ajuste dos hiperparâmetros dos modelos os valores podem ser superiores como visto no processo de modelagem utilizando apenas os índices climáticos.

## 7.1 Considerações Finais

O trabalho desenvolvido no capítulo 3 (Padrões de Variabilidade Climática) foi apresentado no AGU 2019 Fall Meeting, em dezembro de 2019 em São Francisco, CA, USA, disponível em [Malfatti \*et al.\* \[2019\]](#).

O trabalho desenvolvido no Capítulo 4 (Identificação dos Padrões de Variabilidade Climática nas Principais Bacias Brasileiras) foi submetido para a revista *International Journal of Climatology* e aguarda aprovação para publicação, com o título *Predictability of Sub-seasonal Precipitation in Brazilian Watersheds based on Teleconnection Patterns and South American Monsoon Modes* (versão original submetida no Anexo J).

Os Capítulos 5 (Desenvolvimento de Modelos Preditivos de Precipitação nas Principais Bacias Brasileiras) e Capítulo 6 (Modelos com Predição S2S) em ajustes para publicação.

## 7.2 Sugestões para Pesquisas Futuras

Em estudos futuros, sugere-se aprimoramento dos modelos desenvolvidos, explorando novas arquiteturas através do ajuste fino dos hiperparâmetros e refinamento na extrapolação dos preditores climáticos para uso em conjunto com as séries temporais previstas dos modelos do programa S2S. Além disso, desenvolvimento de modelos específicos para cada época do ano, visto que os padrões climáticos causam a precipitação em períodos específicos.

Em adição, podem ser testadas abordagens com séries de saída de comprimento maior devido ao fato que os modelos ECMWF e NCEP só possuem informação de 45 dias, e o alvo é a previsão estendida de até 60 dias. E incluir produtos de previsão numérica de outros centros mundiais que produzam estimativas de precipitação na escala sub-sazonal.

Outro ponto que também merece atenção, é a busca de índices que sejam mais adequados para a identificação do comportamento da variabilidade intrasazonal nas bacias brasileiras. Destaca-se, por exemplo: o uso de índices como o desenvolvido por [Mayta \*et al.\* \[2020\]](#) para melhor previsão na Bacia do Rio Madeira.



# Apêndice A

## Padronização de séries temporais

A padronização de séries temporais, remove sua sazonalidade (ciclo anual), através da média e desvio padrão de cada mês correspondente conforme a equação A.1 (Wilks [2006]).

$$x^{(i,j)} = \frac{x(i,j) - \mu(j)}{S(j)} \quad (\text{A.1})$$

sendo que  $x^{(i,j)}$  é o valor padronizado da variável do  $j$ -ésimo mês do  $i$ -ésimo ano, e onde  $\mu(j)$  é a média para um mês  $j$  e pode também ser chamada de ciclo sazonal (equação A.2);  $i$  é o índice que indica o ano e  $j$  é o índice que indica o mês;  $n$  é o número de anos;  $\sigma(j)$  é o desvio padrão do mês  $j$  (equação A.3).

$$\mu(j) = \frac{1}{n} \sum_{i=1}^n x(i,j) \quad (\text{A.2})$$

$$\sigma(j) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x(i,j) - \mu(j))^2} \quad (\text{A.3})$$





# Apêndice B

## Transformada de Ondeletas

A transformada de ondeletas (ou em inglês “*wavelets*”) é utilizada para transformar uma função no domínio do tempo em uma nova função que varia nas escalas de frequência e tempo. A ondeleta complexa de Morlet é usada e a transformada é realizada no espaço de Fourier usando o método desenvolvido por [Torrence e Compo \[1998\]](#). Este método decompõe a série original em um espectrograma (tempo-frequência), que permite a localização no tempo de sinais não-estacionários, ou seja, são capturadas até pequenas variações ao longo do tempo ([Weng e Lau \[1994\]](#); [Torrence e Compo \[1998\]](#)).

O termo ondeletas refere-se a um conjunto de funções em forma de pequenas ondas geradas por translações e dilatações, e uma função ondeleta  $\psi(t)$ . Desta forma, a equação [B.1](#) mostra a função ondeleta geradora em função de  $\psi$ .

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), \quad a > 0 \quad (\text{B.1})$$

onde  $a$  é o parâmetro de dilatação;  $b$  é o parâmetro de translação;  $\psi_0(t)$  é a função da transformação, chamada de ondeleta-mãe; o fator  $1/\sqrt{a}$  é a constante de normalização da energia, que mantém a energia da ondeleta-mãe ([Weng e Lau \[1994\]](#)).

A transformada em ondeletas contínua de uma série temporal ( $f(t)$ ) é definida como a convolução de  $f(t)$  com o complexo conjugado ( $\psi^*$ ) da ondeleta-mãe dilatado e transladado, conforme a seguinte equação [B.2](#):

$$W_{(a,b)} = \langle f, \psi_{a,b}^* \rangle = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi^* \left( \frac{t-b}{a} \right) dt \quad (\text{B.2})$$

Para que uma ondeleta seja admissível as seguintes condições devem ser respeitadas: *i*) ondeleta-mãe quadraticamente integrável no tempo e espaço e *ii*) média igual a zero ([Weng e Lau \[1994\]](#)). Na literatura existem diversas funções para gerar as ondeletas, tanto contínuas quanto discretas. No grupo das ondeletas discretas, a mais conhecidas e utilizadas são as de Haar ([Morettin \[1999\]](#)) e no grupo das contínuas de Morlet ([Farge \[1994\]](#)). Para dados meteorológicos a ondeleta complexa de Morlet é a mais utilizada, sua função consiste em uma onda plana, na forma de exponencial complexa, modulada por uma curva gaussiana (equação [B.3](#)) definida por:

$$\psi(\eta) = \pi^{-1/4} e^{i\omega_0\eta} e^{-\eta^2/2} \quad (\text{B.3})$$

onde  $\omega_0$  é uma frequência adimensional, que determina a forma do sinal modulador da função ondeleta; e  $\eta$  representa cada valor da série temporal. Para satisfazer as condições de admissibilidade foi utilizado o valor de  $\omega_0 = 6$ , conforme Torrence e Compo [1998].

Após a definição da função de ondeleta um conjunto de escalas deve ser definido, como a função de Morlet é contínua, as escalas são derivadas de potência de dois (Torrence e Compo [1998]).

Além disso, Torrence e Compo [1998] define o espectro de potência (variância da série temporal) como sendo o quadrado da amplitude  $|W_{(a,b)}|^2$  e o espectro de potência global pela média temporal da potência (equação B.4), conforme:

$$\bar{W}^2(s) = \frac{1}{N} \sum_{i=0}^{N-1} |W_i(s)|^2 \quad (\text{B.4})$$

No espectro de potência também se define o cone de influência, que é a região afetada pela fronteira desse espectro em que os valores dos coeficientes wavelet podem estar afetados. Esse cone é definido pelo decaimento (*e-folding time*) do espectro wavelet em cada escala, o fator de decaimento é de  $e^{-2}$  (Torrence e Compo [1998]; Grinsted *et al.* [2004]).

Para a significância estatística de 95% de um pico no espectro de ondeletas, utiliza-se a hipótese nula em que considera o sinal como ruído vermelho com dado espectro de energia de fundo ( $P_k$ ).

$$P_k = \frac{1 - \alpha_2}{|1 - \alpha e^{-2i\pi k}|^2} \quad (\text{B.5})$$

onde  $k = 0, 1, \dots, N/2$  é o índice das bandas de frequências de Fourier ( $f_k$ ), com  $f_k = k/(ndt)$ , sendo que  $dt$  o intervalo de amostragem da série temporal e  $\alpha$  é determinado de acordo com:

$$\alpha = \frac{(\alpha_1 + \sqrt{\alpha_2})}{2} \quad (\text{B.6})$$

onde  $\alpha_1$  e  $\alpha_2$  são as autocorrelações da série temporal com defasagens  $lag_{-1}$  e  $lag_{-2}$ , respectivamente.

# Apêndice C

## Correlação de Pearson

Após a padronização dos dados, conforme o apêndice A Padronização de Séries Temporais, calcula-se o coeficiente de correlação,  $r$  de Pearson, entre as séries padronizada, a partir da (equação C.1).

$$r_{xy} = \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} (\sum_{i=1}^n x_i) (\sum_{i=1}^n y_i)}{\sqrt{\sum_{i=1}^n x_i^2 - \frac{1}{n} (\sum_{i=1}^n x_i)^2} \sqrt{\sum_{i=1}^n y_i^2 - \frac{1}{n} (\sum_{i=1}^n y_i)^2}} \quad (\text{C.1})$$

sendo que  $x_i$  e  $y_i$  são as variáveis padronizadas a serem correlacionadas.

A matriz de correlação mostra os coeficientes que medem o grau de relacionamento entre duas séries, podendo variar de +1,00 e -1,00, que indicam perfeitas correlações lineares positivas e negativas, respectivamente.

Para avaliar se os coeficientes de correlação encontrados são considerados significativos para o nível de significância 0,05 (ou seja, nível de confiança de 0,95) foram aplicados o teste de hipótese – Teste T de Student. Para tanto considera as hipóteses:

- $H_0 : \rho = 0$
- $H_1 : \rho \neq 0$

A estatística de teste é calculada a partir da equação C.2:

$$t_c = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (\text{C.2})$$

Com  $n - 2$  graus de liberdade na tabela T de Student, em que  $r$  é o valor do coeficiente de correlação sujeito a teste e  $n$  é o número de observações. Caso o valor de  $t_c$  seja superior ao valor crítico de  $t$ , deve-se rejeitar a hipótese nula. Se a hipótese nula, ao nível de significância  $\alpha$ , for rejeitada pode-se concluir que efetivamente existe uma relação significativa entre as variáveis (Devore [2014]).



# Apêndice D

## Análise de Componentes Principais

A forma de calcular as componentes principais seguem os seguintes passos (Zang e Moore [2015]): *i*) obter os dados ou as  $M$  amostras de vetores de dimensão  $n$ ; *ii*) remoção da tendência dos dados; *iii*) remoção da sazonalidade dos dados; *iv*) calcular a matriz de covariância; *v*) calcular os autovalores e autovetores da matriz de covariância; *vi*) arranjar a matriz da Transformada de Hottelling (cujas linhas são formadas a partir dos autovetores da matriz de covariância arranjados de modo que a primeira linha, o elemento  $(0,0)$ , seja o autovetor correspondente ao maior autovalor, e assim sucessivamente até que a última linha corresponda ao menor autovalor).

Primeiramente os dados foram tratados removendo a tendência linear dos dados (equação D.1), através da subtração da tendência calculada ( $y_1 = y - y(t)$ ) pelo método da regressão linear aplicado na série temporal em cada ponto observado ( $y$ ):

$$y(t) = a + bt \quad (\text{D.1})$$

onde  $y(t)$  o valor predito da série temporal;  $t$  o índice do tempo;  $b$  o coeficiente angular da reta e  $a$  o valor de  $y(t)$  quando  $b$  é zero.

Após a remoção da tendência os dados foram padronizados, removendo sua sazonalidade (ciclo anual), através da média e desvio padrão de cada mês correspondente conforme apresentado no apêndice A Padronização de Séries Temporais.

Após o tratamento dos dados a matriz de covariância ( $S$ ) é calculada, conforme a equação D.2. Utiliza-se a matriz de covariância para medir o grau de relacionamento linear entre duas ou mais variáveis aleatórias.

$$S = \frac{1}{(n-1)} X'X'^T \quad (\text{D.2})$$

onde  $X'$  é a matriz bidimensional;  $X'^T$  é a matriz transposta de  $X'$  e  $S$  é a matriz de covariância simétrica ( $n \times n$ ), com a diagonal sendo a variância dos próprios pontos.

A partir da matriz de covariância ( $S$ ), calcula-se os autovalores ( $\lambda$ ), resolvendo o determinante:

$$|S - \lambda I| = 0 \quad (\text{D.3})$$

onde  $I$  representa a matriz de identidade. Para cada autovalor existe um autovetor ( $v_{ij}^*$ ) associado,

e são obtidos através da seguinte expressão:

$$S \vec{v}_{ij} = \lambda \vec{v}_{ij} \quad (\text{D.4})$$

O autovetor com maior autovalor associado, corresponde à componente principal do conjunto de dados usado. Isso significa que essa é o relacionamento mais significativo entre as dimensões dos dados.

Como a matriz de covariância é definida positiva, devido  $X'$  ser real, todos os autovalores também são positivos e fornecem a medida da variância explicada por cada modo  $i$  a ele associado, conforme a equação D.5.

$$\text{var}(i) = \frac{100 \lambda_i}{\sum_{i=1}^p \lambda_i} \% \quad (\text{D.5})$$

Para verificar se os modos são independentes entre si, calcula-se o erro amostral do autovalor, segundo North *et al.* [1982], conforme a equação D.6.

$$\delta\lambda \cong \lambda \left( \frac{2}{N} \right)^{\frac{1}{2}} \quad (\text{D.6})$$

onde,  $N$  é o número total de autovalores.

A fim de maximizar a variância dos resultados geralmente utiliza-se o método de rotação ortogonal chamado de Varimax (Kaiser [1958]). A rotação Varimax é determinada pela escolha dos elementos da matriz de rotação para maximizar, conforme a equação D.7.

$$V^* = \frac{n \sum_{i=1}^n (v_{ij}^2)^2 - (\sum_{i=1}^n v_{ij}^2)^2}{n^2}, j = 1, \dots, r \quad (\text{D.7})$$

onde  $n$  é o número de variáveis,  $r$  é o número de CPs e  $(v_{ij}^2)$  são os autovetores.

A ACP rotacionada produz novos padrões compactos que podem ser usados para agrupar áreas homogêneas (Hannachi *et al.* [2007]). Como principais objetivos da análise de ACP são atenuar as fortes restrições (isto é, variáveis ortogonais/não-correlacionadas dos CPs), suavizar a dependência do padrão da ACP com a forma do domínio e obter estruturas simples e de fácil interpretação dos padrões resultantes.

# Apêndice E

## Causalidade (PDC e knPDC)

Na prática, a inferência da causalidade de Granger envolve o ajuste de modelos lineares vetoriais autorregressivos (em inglês, *Vector Autoregressive*, VAR) (Lütkepohl [2005]).

No domínio da frequência a Causalidade de Granger pode ser inferida a partir da Coerência Parcial Direcionada (em inglês, *Partial Directed Coherence*, PDC) (Baccalá e Sameshima [2001]) que permite detecção de conexões diretas entre as variáveis lineares.

$$\Pi_{ij}(f) = \frac{|\bar{A}_{ij}(f)|}{\sqrt{\bar{a}_j^H(f) \bar{a}_j(f)}} \quad (\text{E.1})$$

onde:  $|\bar{A}_{ij}(f)|$  é o  $i, j$ -ésimo elemento da matriz  $\bar{A}(f)$ , em que as colunas são indicadas por  $\bar{a}_1(f)$  ou seja:

$$\bar{A}(f) = [\bar{a}_1(f) \cdot \cdot \cdot \bar{a}_j(f)] \quad (\text{E.2})$$

No método desenvolvido por Massaroppe e Baccalá [2019] é introduzido um kernel no modelo VAR, resultando em um processo Autorregressivo kernelizado Vetorial (em inglês, *kernel Vector Autoregressive*, kVAR) (Massaroppe e Baccalá [2015a]; Massaroppe e Baccalá [2015b]).

O método kerneriza (*kernel trick*) (Schölkopf e Smola [2018]; Príncipe [2010]) a série não-linear, tornando-as (“quase”) lineares no espaço de características  $\mathbb{F}$ , possibilitando a aplicação de um método linear (mapeamento implícito) e fatoração espectral da matriz de covariância kernelizada, no domínio da frequência (knPDC).

Desse modo, seja  $\phi(\cdot)$  um mapeamento não-linear, induzido por um kernel de Mercer (Parzen [1959]), do espaço de entrada  $\mathbb{X}$ , para o característico ( $\phi : \mathbb{X} \subset \mathbb{R}^{N \times D} \rightarrow \mathbb{F}$ ). A kernelização do modelo autorregressivo de ordem  $p$  ( $VAR(p)$ ) pode ser feita de maneira direta (equação E.3)

$$\phi[x(n)] = \sum_{r=1}^p \phi[x(n-r)] A^{\phi H}(r) + \vartheta^{\phi}(n) \quad (\text{E.3})$$

onde:

$$\vartheta^{\phi}(n)_{n \in \mathbb{Z}} \sim i.i.d. WN(0, \sum \vartheta^{\phi}) \quad (\text{E.4})$$

Para a escolha da ordem do modelo, optou-se pela ordem que melhor se adequa aos resultados. Neste trabalho escolheu-se trabalhar com modelos de ordem 12, 24 ou 48 para que a PDC/knPDC fosse capaz de reproduzir as dinâmicas (ressonâncias) dos sinais atmosféricos analisados, que apre-



sentam ciclo anual e seus (sub) múltiplos.

Após a etapa de mapeamento kernerizado a análise da causalidade de Granger é realizada através da versão kernelizada da PDC, a partir da equação da *kn*PDC (equação E.5),

$$\kappa\eta\Pi_{ij}(f) = \frac{\frac{A_{ij}^{-\phi}(f)}{\sqrt{\sigma_{ii}^{\phi}}}}{\sqrt{\bar{a}_j^{\phi H}(f) \sum_{\vartheta\phi}^{-1} \bar{a}_j^{\phi}(f)}} \quad (\text{E.5})$$

onde:

$$\bar{A}_{ij}^{\phi}(f) = \delta_{ij} - \sum_{r=1}^p a_{ij}^{\phi}(r) e^{-12\pi fr}, \quad (j^2 = -1) \quad (\text{E.6})$$

em que as colunas são indicadas por:  $\bar{a}_j^{\phi}(f)$  ou seja,

$$\bar{A}^{\phi}(f) = [\bar{a}_1^{\phi}(f) \cdot \cdot \cdot \bar{a}_j^{\phi}(f)] \quad (\text{E.7})$$

# Apêndice F

## Método *Stepwise*

No desenvolvimento de modelos preditivos, as variáveis preditoras (variáveis independentes) podem influenciar pouco na variável dependente. O método *stepwise* é um método iterativo que adiciona e remove variáveis independentes de um modelo de regressão linear múltipla (RLM, metodologia descrita no Apêndice G) com base em sua significância estatística, podendo assim, diminuir o número de variáveis a compor a equação de regressão (Draper e Smith [1998]).

Neste caso, ao adicionar e remover variáveis, foi utilizado o teste  $F$  para o critério de seleção das variáveis com nível de significância estatística. O teste  $F$  é um teste para determinar se há uma relação linear entre a variável resposta  $Y$  e algumas das variáveis regressoras  $x_1, x_2, \dots, x_p$  contribui significativamente para o modelo. Para tanto, considera-se as hipóteses..

- $H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$
- $H_1 : \beta_j \neq 0$  para qualquer  $j = 1, \dots, p$

Se rejeitar  $H_0$ , terá ao menos uma variável explicativa  $x_1, x_2, \dots, x_p$  que contribui significativamente para o modelo. Porém, em conformidade com  $H_0$ , temos pelo Teorema de distribuição de forma quadrática que

$$\frac{SQR}{\sigma^2} X_{(p)}^2 \text{ e que } \frac{SQE}{\sigma^2} X_{(n-p-1)}^2 \quad (\text{F.1})$$

e  $SQR$  e  $SQE$  são independentes. Logo, conclui-se perante a  $H_0$ , que

$$F_0 = \frac{\frac{SQR}{p}}{\frac{SQE}{n-p-1}} = \frac{QMR}{QME} F_{(p;n-p-1)} \quad (\text{F.2})$$

Portanto, rejeita-se  $H_0$  se  $F_0 > F_{(1-\alpha;p;n-p-1)}$  e se  $p - value = P[F_{p;n-p-1} > F_0] < \alpha$ , em que  $\alpha$  é o nível de significância considerado. Geralmente adota-se  $\alpha = 5\%$ , mas para testes aplicados ao método *stepwise*, esse valor é muito rigoroso, portanto adotam-se valores de  $\alpha$  entre 15% e 20%.

Para aplicação do método *stepwise* são realizada algumas etapas (Draper e Smith [1998]):

- I. Selecionar variáveis independentes aleatórias, formando um modelo inicial qualquer;
- II. Calcular o p-valor do teste  $F$  para cada variável independente, se o  $p - value$  for menor que o nível de significância ( $\alpha_e = 0,15$ ) a variável é descartada do modelo;

- III. Se todas as variáveis independentes do modelo possuir p-valor superior ao nível de significância, adota-se um valor máximo de significância ( $\alpha_s = 0,20$ ) e removem-se termos com valores maiores que este máximo;
- IV. Os passos anteriores são repetidos até que todas as variáveis independentes sejam testadas no modelo.

# Apêndice G

## Modelo de Regressão Linear Múltipla

O Modelo de Regressão Linear Múltipla (RML) é um modelo empírico que ao ser construído leva em consideração as relações empíricas entre as variáveis observadas, sem considerar os processos físicos, sendo de fácil ajuste e validação, uso e facilidade de ajuste para diferentes casos.

Este modelo é aplicado para uma variável que se deseja estimar, cuja série histórica apresenta fortes relações com outras variáveis, sendo que essas relações se associam ao fenômeno (ou processo) que se deseja modelar.

O modelo de Regressão Linear Múltipla fornece a relação entre uma determinada variável e outras variáveis independentes, pelo método dos mínimos quadrados (Wilks [2006]). A equação de regressão (equação G.1) possui a seguinte forma:

$$y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_kX_k \quad (\text{G.1})$$

sendo  $y$  a variável dependente (neste trabalho, precipitação em bacias brasileiras a ser prevista);  $X_n$  são as variáveis independentes (variáveis escolhidas que contribuem para a variação de precipitação, neste trabalho os índices climáticos);  $b_0$  é a intersecção da reta com o eixo de  $y$ ;  $b_n$  são os coeficientes angulares obtidos da regressão;  $k$  é o número de variáveis independentes.

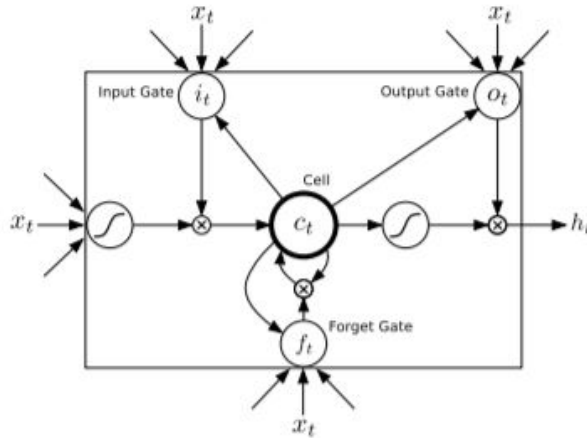


# Apêndice H

## *Long short-term memory* (LSTM)

Dentro das redes neurais recorrentes (descritas no Capítulo 2 - Revisão Bibliográfica), a LSTM (Hochreiter e Schmidhuber [1997]) é a arquitetura que se mostra mais estável e que melhor modela as dependências de longo prazo. Este tipo de arquitetura de rede foi projetado com a intenção de fornecer uma abordagem de melhor desempenho e de resolver o problema de desaparecimento do gradiente (Glorot *et al.* [2011]) que sofrem as redes neurais recorrentes (RNNs) quando lidam com grandes sequências de dados.

Para resolver o problema exposto, a arquitetura LSTM possui, além do estado oculto, outras estruturas adicionais conhecidas como portões, conforme a Figura H.1.



**Figura H.1:** Arquitetura de Long short-term memory (LSTM) .

A formulação original da LSTM apresenta três portões: *input*, *forget* e *output*. As equações H.1 a H.5 fornecem os cálculos necessários para atualizar uma unidade LSTM.

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (\text{H.1})$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (\text{H.2})$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (\text{H.3})$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \quad (\text{H.4})$$

$$h_t = o_t \circ \sigma_h(c_t) \quad (\text{H.5})$$

Como pode ser observado nas equações, o estado oculto  $h_t$  depende do estado da célula  $c_t$ , que é constantemente atualizada quando os portões input ( $i_t$ ) e forget ( $f_t$ ) são ativados (operador  $\circ$  refere-se ao produto de Hadamard) (Yang e Xu [2004]). Os portões *input* ( $i_t$ ) e *forget* ( $f_t$ ) oferecem proteção ao estado da célula  $c_t$  contra perturbações por entradas irrelevantes. Do mesmo modo, o portão de saída ( $o_t$ ) protege as outras unidades (por exemplo, o estado oculto  $h_t$ ) de perturbações por conteúdos de memória atualmente irrelevantes. Em outras palavras, se a última saída da célula for propagada para o estado final, esta será controlada pelo portão de saída  $o_t$ .

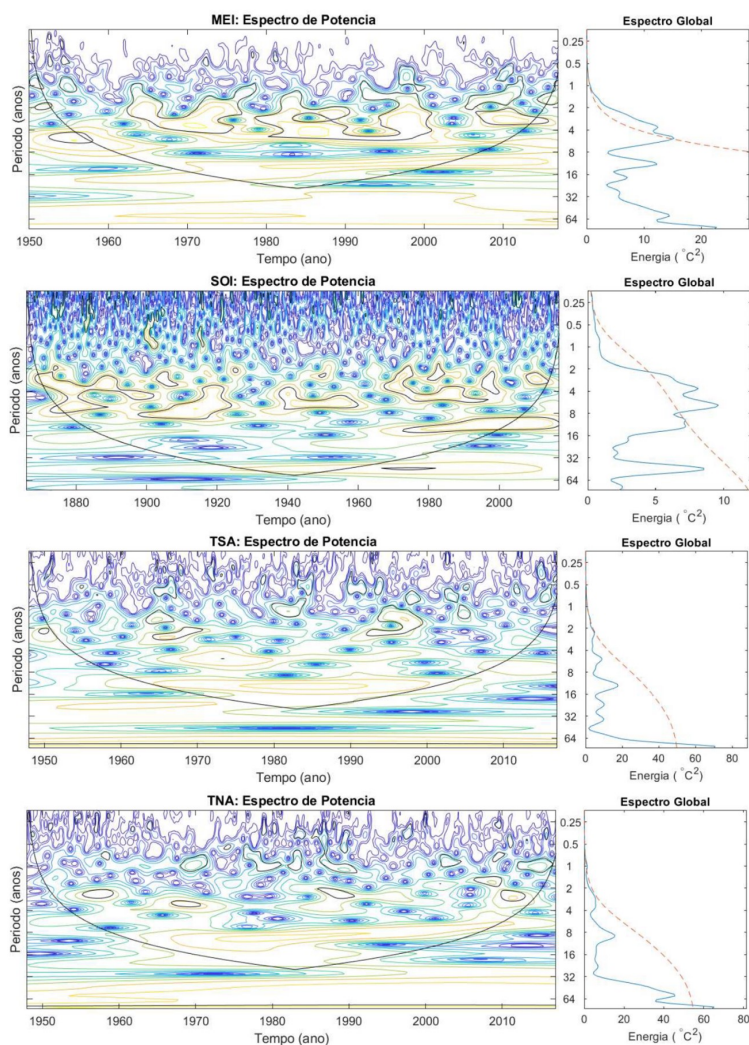
Embora as LSTMs tenham um grande sucesso na prática, não está claro o significado de seus componentes individuais. Vários autores coincidem no fato de considerar o portão *forget* o mais importante seguido do portão *input* (Bayer *et al.* [2009]; Jozefowicz *et al.* [2015]). Na maioria das aplicações de LSTMs basta inicializar os portões com pequenos pesos aleatórios para obter um bom desempenho. Por outro lado, o trabalho de Jozefowicz *et al.* [2015] mostra que quando o portão *forget* é inicializado com valor igual a 1 e o *kernel* se inicializa com uma distribuição "*glorot uniform*" a maioria das vezes resulta em desempenhos superiores da rede. Baseado nestas experiências o Keras propõe a inicialização de parâmetros seguinte para uma camada de LSTM:

```
keras.layers.LSTM(units, input_shape, activation='tanh',
recurrent_activation='sigmoid', use_bias=True,
kernel_initializer='glorot_uniform',
recurrent_initializer='orthogonal', bias_initializer='zeros',
unit_forget_bias=True, kernel_regularizer=None,
recurrent_regularizer=None, bias_regularizer=None,
activity_regularizer=None, kernel_constraint=None,
recurrent_constraint=None, bias_constraint=None, dropout=0.0,
recurrent_dropout=0.0, implementation=2, return_sequences=False,
return_state=False, go_backwards=False, stateful=False,
unroll=False)
```

As redes LSTM podem ser projetadas para análise univariada e multivariada. As variantes multivariadas também oferecem flexibilidade para aprendizado empilhado ou paralelo. A abordagem proposta neste projeto pode ser classificada como multivariada e de aprendizado empilhado. Assim, para gerar um modelo é treinada uma LSTM que aprende das séries dos índices climáticos selecionados para uma bacia específica e da previsão de modelos meteorológicos como NCEP e ECMWF que ajudam a corrigir a previsão. O parâmetro "*input shape*" serve para simular a aprendizagem multivariada e empilhada, basta especificar a quantidade de dias anteriores que devem ser considerados na previsão e a quantidade de variáveis de entrada que a LSTM deve considerar para construir o modelo.

# Apêndice I

## Ilustrações Adicionais - Capítulo 3



**Figura I.1:** *Espectro de Potência e Global da análise de ondeleta da série temporal dos índices (Morlet) a) MEI; b) SOI; c) TSA e d) TNA. O espectro de potência é delimitado pelo cone de influência, o contorno preto indica o setor do espectro de potência onde o nível de confiança de 95% é satisfeito. O espectro de potência global indica o nível de potência satisfeito para toda a série temporal, ou seja, acima da linha vermelha tracejada.*



**Tabela I.1:** Coeficiente de correlação (CC) entre os índices climáticos para o período instrumental (1980 a 2018). Os valores em negritos são considerados significativos (nível 5%) a partir do teste T de Student (Devore [2014]).

ÍNDICES	AAO	AMM	AMO	AO	GTA	LISAM	MEI	NAO	Niño3.4	PSA1	PSA2	PDO	PNA	QBO	RMM1	RMM2	SF	SOI	TNA	TSA	ZCAS
AAO	1,00	0,02	-0,29	0,06	-0,15	0,07	-0,54	0,10	-0,56	0,03	0,03	-0,42	-0,20	-0,06	0,03	0,03	0,03	0,45	-0,29	-0,14	0,19
AMM	0,02	1,00	0,66	-0,15	0,81	0,00	-0,22	-0,21	-0,22	0,02	-0,16	-0,20	0,09	-0,04	0,07	-0,01	-0,10	0,26	0,85	-0,12	-0,02
AMO	-0,29	0,66	1,00	-0,12	0,49	-0,20	-0,14	-0,27	-0,05	-0,16	-0,02	-0,06	0,10	-0,04	-0,02	-0,09	-0,18	0,14	0,85	0,32	-0,12
AO	0,06	-0,15	-0,12	1,00	-0,10	-0,01	-0,08	0,60	-0,06	0,01	0,02	-0,10	-0,24	0,04	0,12	-0,09	0,18	0,09	-0,21	-0,10	0,03
GTA	-0,15	0,81	0,49	-0,10	1,00	-0,14	0,13	-0,17	0,11	-0,07	-0,08	-0,02	0,11	-0,05	0,03	-0,05	0,03	-0,07	0,70	-0,55	0,04
LISAM	0,07	0,00	-0,20	-0,01	-0,14	1,00	0,07	0,26	0,06	0,93	-0,28	-0,12	0,18	0,15	0,07	0,11	-0,05	-0,03	0,10	0,55	0,25
MEI	-0,54	-0,22	-0,14	-0,08	0,13	0,07	1,00	0,03	0,90	0,11	-0,08	0,47	0,14	0,06	-0,02	-0,09	0,00	-0,82	-0,02	-0,20	0,14
NAO	0,10	-0,21	-0,27	0,60	-0,17	0,26	0,03	1,00	0,05	0,30	-0,04	0,00	0,08	0,07	0,15	0,04	0,15	-0,02	-0,30	-0,12	0,23
NIÑO3.4	-0,56	-0,22	-0,05	-0,06	0,11	0,06	0,90	0,05	1,00	0,09	-0,09	0,40	0,12	0,08	0,00	-0,13	-0,01	-0,77	0,06	-0,08	0,14
PSA1	0,03	0,02	-0,16	0,01	-0,07	0,93	0,11	0,30	0,09	1,00	-0,13	-0,09	0,24	0,13	0,10	0,16	0,00	-0,07	-0,06	0,03	0,41
PSA2	0,03	-0,16	-0,02	0,02	-0,08	-0,28	-0,08	-0,04	-0,09	-0,13	1,00	0,01	-0,07	-0,05	0,25	0,26	0,06	0,10	-0,14	-0,05	0,04
PDO	-0,42	-0,20	-0,06	-0,10	-0,02	-0,12	0,47	0,00	0,40	-0,09	0,01	1,00	0,16	0,02	-0,04	-0,01	-0,20	-0,37	-0,02	0,01	-0,01
PNA	-0,20	0,09	0,10	-0,24	0,11	0,18	0,14	0,08	0,12	0,24	-0,07	0,16	1,00	0,02	0,04	-0,01	-0,09	-0,13	0,15	0,03	0,16
QBO	-0,06	-0,04	-0,04	0,04	-0,05	0,15	0,06	0,07	0,08	0,13	-0,05	-0,04	0,02	1,00	-0,02	-0,06	-0,02	-0,02	-0,01	0,06	-0,05
RMM1	0,03	0,07	-0,02	0,12	0,03	0,07	-0,02	0,15	0,00	0,13	0,25	-0,04	0,04	-0,02	1,00	0,04	-0,01	0,22	0,00	-0,04	0,06
RMM2	0,03	-0,01	-0,09	-0,09	-0,05	0,11	-0,09	0,04	0,00	0,16	0,26	-0,04	0,04	0,04	0,00	1,00	0,01	-0,01	-0,10	-0,04	0,01
SF	0,03	-0,10	-0,18	0,18	0,03	-0,05	0,00	0,15	-0,01	0,00	0,06	-0,20	-0,09	-0,02	-0,01	0,01	1,00	-0,02	-0,21	-0,30	0,05
SOI	0,45	0,26	0,14	0,09	-0,07	-0,03	-0,82	-0,02	-0,77	-0,07	0,10	-0,37	-0,13	-0,02	0,22	-0,01	-0,02	1,00	0,07	0,18	-0,12
TNA	-0,29	0,85	0,85	-0,21	0,70	-0,08	-0,02	-0,30	0,06	-0,06	-0,14	-0,02	0,15	-0,01	0,00	-0,10	-0,21	0,07	1,00	0,21	-0,07
TSA	-0,14	-0,12	0,32	-0,10	-0,55	0,10	-0,20	-0,12	-0,08	0,03	-0,05	0,01	0,03	0,06	-0,04	-0,04	-0,30	0,18	0,21	1,00	-0,14
ZCAS	0,19	-0,02	-0,12	0,03	0,04	0,25	0,14	0,23	0,14	0,41	0,04	-0,01	0,16	-0,05	0,06	0,01	0,05	-0,12	-0,07	-0,14	1,00

# Apêndice J

## Paper: Predictability of Sub-seasonal Precipitation in Brazilian Watersheds based on Teleconnection Patterns and South American Monsoon Modes

### ABSTRACT

Several socio-economic activities depend on water availability, and it is essential to use reliable climate forecasts for better planning of this resource. The large territorial extension of Brazil and the high dependence on hydroelectrical power implies severe constraint on the efficient use of water management based on meteorological forecasting and knowledge of climate variability. The Brazilian watersheds are influenced by tropical and extra-tropical meteorological phenomena of different spatial and temporal scales. The main goal of this paper is to provide support for the definition of a set of predictor variables based on climate indices in the estimation of the pentad precipitation in Brazilian watersheds in the so-called sub-seasonal to seasonal (S2S) timescale, based on statistical analysis of climate indices and precipitation. Precipitation of Brazilian watersheds is influenced by the Madden-Julian Oscillation (MJO), the Antarctic Oscillation (AAO), the Pacific South American Pattern (PSA), the South American Monsoon System index (LISAM), and the South Atlantic Convergence Zone (SACZ) index. In addition, this study highlights the importance of the Pacific patterns such as the El Niño-Southern Oscillation (ENSO) and the Pacific Decadal Oscillation (PDO), which are important in most Brazilian watersheds, except those in the northwest of Brazil. The North, Northeast, and Central Brazil watersheds are more strongly influenced by the Atlantic indicators, such as the Atlantic Multidecadal Oscillation (AMO) and Tropical Northern Atlantic (TNA)/Tropical Southern Atlantic (TSA) patterns while the Pacific, ENSO and PDO standards have a greater influence on precipitation in the southern and southeastern watersheds of Brazil in the S2S timescale.

**Keywords:** streamflow; climate indices; causality; El Niño Southern Oscillation; Brazilian's watersheds.

## INTRODUCTION

Brazil has high availability of fresh water with approximately 12% of the planet's fresh water. In addition to the use for human consumption, the availability of water resources is essential for the success of economic activities such as agriculture and hydroelectric energy generation. For the water resources management, factors such as watershed physiology and interrelated climatic conditions are essential in developing accurate estimates for water availability planning (Von Sperling, 2007). In this context, accurate precipitation forecasts are an important tool for water resource management, aiding in decision making, which aims to minimize the effects of exceptional floods and water scarcity (Collischonn *et al.*, 2005).

Over the past few years, the precipitation forecast up to two weeks in advance has been operationally employed (Kumar *et al.*, 2011; Lin *et al.*, 2016). After the development of the ensemble weather forecast, it was possible to explore the uncertainties inherent to the atmospheric dynamic system, enabling performance improvement and methodological innovations (Tracton and Kalnay, 1993; Murphy, 1988). The long-term forecasting which is done for the next seasons of the year (between three to six months), has also been carried out in recent decades. The models are based on physical conservation principles as this approach tend to surpass the performance of purely statistical models (Lucio *et al.*, 2010). In this way, predictions of the El Niño and La Niña phenomena in the Pacific Ocean are now forecasted about six months in advance, allowing for better skill in evaluating the remote impact of these phenomena and others that occur in the Atlantic and Indian Ocean mainly in the equatorial region (Newman *et al.*, 2011). However, on the scale between two weeks and two months in advance, the sub-seasonal to seasonal (S2S) forecast still represents a major challenge (White *et al.*, 2017; Vitart *et al.*, 2017; Min *et al.*, 2020).

In this context, it is important to evaluate the contribution of different climatic phenomena to the precipitation forecast at the sub-seasonal scale. Thus, the aim of this article is to select potential predictors of climate indices in the forecast of precipitation in pentads in Brazilian watersheds, using the technique of causality between climate indices and precipitation.

Due to the large Brazilian's territorial extension, several mechanisms inhibit or favor precipitation. These mechanisms are of tropical and subtropical origin acting on different scales (micro, meso, synoptic, intraseasonal, interannual, interdecadal and multidecadal) and with seasonal and subseasonal variations that produce effects on the hydrological cycle in Brazilian watersheds (Satyamurty *et al.*, 1998; Grimm, 2003; Reboita *et al.*, 2010).

The watersheds in the extreme south of Brazil show less precipitation variation throughout the year, with peaks in spring and autumn, mainly caused by frontal (cold front) and mesoscale convective systems (Reboita *et al.*, 2010; Malfatti *et al.*, 2018a). In other Brazilian regions, there is a well-defined annual cycle, characterized by a dry season in winter and a rainy season in summer, typical of the South American Monsoon System (SMAS), with a strong role in the Atlantic Convergence Zone South (SACZ) over the Southeast, Central, and the northern portion of the Southern region (Grimm *et al.*, 2005), and the Intertropical Convergence Zone (ITCZ) over the northern part of the Northeast and North of Brazil (Reboita *et al.*, 2010). In addition, on the Northeast coast, the southeast trades exert an important control on precipitation, which are influenced by the relatively

cold fronts that move along the coast to the Northeast, contributing to the humidity increase in the region.

The SACZ is a meteorological phenomenon characterized by the persistent clouds band oriented in the northwest-southeast direction (NW-SE) extending from central Brazil to the central south Atlantic between October and March (Grimm *et al.*, 2005; Kodama, 1992). Intraseasonal variations can affect the position and intensity of SACZ, such as the Madden-Julian Oscillation (MJO) which is an intraseasonal phenomenon characterized by an eastward shift of a thermally direct large-scale zonal cell that causes convection variations tropical (Madden and Julian 1972; Madden and Julian, 1994). Studies indicate that when convection is over the western Pacific, there is an increase in convection over South America, favoring transient systems that contribute to the formation of SACZ, which allows the occurrence of extreme precipitation events (Kousky and Kayano, 1994; Souza and Ambrizzi, 2006; Grimm and Ambrizzi, 2009).

The ITCZ is characterized by a band of convective clouds that surround the equatorial band of the Earth's globe (Moura and Shukla, 1981; Kouadio *et al.*, 2012). The inter-hemispheric sea surface temperature gradient (SST) in the Tropical Atlantic Ocean, that is, Tropical North Atlantic (TNA) - Tropical South Atlantic (TSA), has a direct influence on the maintenance, positioning, and intensity of the ITCZ (Chiang *et al.*, 2002).

Large-scale phenomena with interannual, interdecadal and multidecadal timescale can remotely influence precipitation over the watersheds, affecting the precipitation variability in South America. The main large-scale phenomenon that can remotely influence precipitation over virtually all of Brazil is the El Niño Southern Oscillation (ENSO), which causes climatic anomalies not only in the warm region of the equatorial Pacific Ocean, but across the globe (Cai *et al.*, 2020; Souza *et al.*, 2021). In general, Brazil is affected by the ENSO phenomenon differently in two regions: the Northeast and the South (Coelho *et al.*, 2002; Grimm, 2003; Tedeschi *et al.*, 2016). In El Niño years, the southern region has its strong rainy season (spring), causing floods. On the other hand, in the same year, but at a different season (autumn), the Northeast suffers from severe droughts, due to the inhibition of the rainy season (Rodrigues *et al.*, 2011; Pinheiro *et al.*, 2021). In the La Niña years, the situation is reversed: the southern region suffers from extensive droughts, while in the northeast region there are above-average precipitation. In the Central and Southeast regions, it is not observed a well-defined precipitation anomaly pattern between the opposite phases of ENSO (Correia Filho *et al.*, 2021; Prado *et al.*, 2021). However, there is some evidence that the Southeast region experiences milder winters in terms of temperature in El Niño episodes (Grimm *et al.*, 1998). And, finally, some studies show that, on average, the North region (Amazon) has a similar behavior to the Northeast, but in the period between January and March (Foley *et al.*, 2002).

Another important phenomenon in the Pacific that remotely influences the climate in almost all of Brazil is the Pacific Decadal Oscillation (PDO), which has an interdecadal variability in the configuration of the SST in the Pacific Ocean. The positive phase is characterized by negative TSM anomalies over northern South America and positive over southeastern South American, with opposite (anomaly reversal) occurring in the negative phase (Villamayor *et al.*, 2018). In Brazil, there is a strong correlation between total rainfall and PDO in some specific regions of Brazil (southeastern Amazon, northern Northeast, Central, and South) (Kayano and Andreoli, 2004; Rebello, 2006; Molion, 2008; Sousa, 2010).

As a direct consequence of the PDO signal in the South Pacific, the PDO relationship with the South American precipitation is linked to ENSO events and the combined effect between the two oceanic events is observed. During the summer, an El Niño event is more pronounced when it is in the warm phase of the PDO. There is also a smaller number of La Niña events during the warm phase of the PDO, while in the cold phase there is a greater occurrence of intense La Niña episodes, indicating that the response of the South American climate is dependent on the PDO phase (Andreoli and Kayano, 2005). Kayano and Andreoli (2009) suggested that differences in precipitation patterns associated with ENSO may also be related to the PDO phase, which constructively influences ENSO when in the same phase (Garcia and Kayano, 2008; Kayano and Andreoli, 2009). Theoretical studies suggest that ENSO causes PDO to be related linearly and non-linearly in the decadal timescale (Ramirez *et al.*, 2017). However, the PDO influence in the S2S timescale has not been fully explored.

In the Atlantic, interannual and multidecadal variability phenomena can influence the rainfall regime, mainly in the North/Northeast and Southeast of Brazil. The Atlantic Meridional Mode (AMM), an ocean-atmosphere coupling phenomenon, with interannual variability in the tropical Atlantic, is characterized by variations in TSM and sea level pressure to the south and north of the ITCZ (Chiang and Vimont, 2004). Depending on its positive/negative phase, the AMM can prevent or favor the displacement of the ZCIT to the south, inhibiting/favoring the rains over the North and Northeast of Brazil (Wainer and Venegas, 2002).

The Atlantic Multidecadal Oscillation (AMO) is a phenomenon of low-frequency variability in the North Atlantic SST, with characteristic period ranging from 60 to 80 years. The AMO has two phases, with the positive phase (negative phase) occurring waters that are warmer (cold) than normal in the North Atlantic (Enfield *et al.*, 2001). During the South America monsoon season, there is a relationship between the AMO and the SACZ: during the negative phase (positive phase) AMO increases (reduces) the SACZ activity, providing an increase (decrease) in precipitation over the Southeast, Central, and northern part of the South region (Chiessi *et al.*, 2009). However, AMO also varies in shorter timescales and the potential impact of the S2S timescale also needs to be explored.

The relationship between the Pacific and Atlantic modes can also result in an intensification of patterns and, consequently, an alteration in the rainfall regime. ENSO impact can be intensified dependent on the AMO phase: the El Niño is more pronounced during the negative phase of the AMO, while the La Niña is more pronounced during the positive phase of the AMO (Kayano and Capistrano, 2014).

Precipitation patterns can also be influenced by the variability of the pressure belts of mid and high latitudes of the Southern Hemisphere, known as the Antarctic Oscillation (AAO) (Gong and Wang, 1999). The AAO in the positive phase prevails low pressure at high latitudes and high pressure at mid latitudes, alternating this condition in the negative phase (Gong and Wang, 1999). Carvalho *et al.* (2005) noted that the positioning of the trajectories of the cyclonic systems is influenced by the AAO phases and, consequently, the precipitation regime over South America. During the AAO negative phase, positive precipitation anomalies over the South of Brazil are observed, especially in summer and autumn. Very rainy (very dry) summers in Southeastern Brazil are associated with the intensification of PSA-type (Pacific South America) wave trains by the positive (negative) phase of

the AAO (Vasconcellos and Cavalcanti, 2010). The PSA pattern also contributes to the convection in the SACZ (Vasconcellos and Cavalcanti, 2010). In winter, this pattern can generate favorable conditions for the displacement and development of synoptic systems, such as frontal systems over South America (Cavalcanti, 2000). In addition, during MJO event, convective longitudinal anomalies emerge related to divergence at high levels in the tropical Pacific, causing teleconnection to the extratropical Pacific through changes in the PSA pattern (Grimm and Silva Dias, 1995; Mo and Nogués-Paegle, 2001). In southeastern South America, PSA patterns were observed associated with increased convective activity in different phases of MJO on the S2S scale.

Therefore, there is a wide variety of mechanisms that influence the climate, rainfall patterns and, consequently, potentially impact the river discharge in Brazil at the sub-seasonal to seasonal timescale. Depending on the system, the influence can be local or remote, direct, or indirect, whose response signals depend in many cases on the modulation of different patterns of variability.

In this present study, analyzes and discussions are carried out on the characteristics of climate variability patterns and how they influence precipitation in the main Brazilian watersheds in the S2S timescale. The following analysis defined a set of potential predictor variables supported by physical interpretations and provides the basis for the construction of empirical models of precipitation/streamflow forecast based on statistical and machine learning techniques. In addition, historical series of river flows in the main watersheds were compared with the precipitation in same watersheds, facilitating discussions on water resources planning.

## DATA AND METHODOLOGY

The climate indices data are available on websites of the National Oceanic and Atmospheric Administration (NOAA, <https://www.esrl.noaa.gov/psd/data/climateindices/list/>), European Climate Assessment & Dataset project (ECA&D, <https://www.ecad.eu/> or <https://climexp.knmi.nl>) and Bureau of Meteorology Australian Government (BOM, <http://www.bom.gov.au/>). These data sources provide monthly and/or daily data on climate indices that represent patterns of climate variability at different scales (intraseasonal, interannual, decadal or multidecadal). Table J.1 shows the indices and the references which describe how they are computed.

**Tabela J.1:** *Monthly and/or daily climate indices made available by government agencies and their respective references*

Name	Description	Reference
<b>AAO</b>	Antarctic Oscillation	Gong and Wang, 1999
<b>AMM</b>	Atlantic Meridional Mode	Chiang and Vimont, 2004
<b>AMO</b>	Atlantic Multidecadal Oscillation	Enfield <i>et al.</i> , 2001
<b>MEI</b>	Multivariate ENSO Index	Wolter and Timlin
<b>MJO</b>	Madden-Julian Oscillation (Real-time Multivariate RMM1/RMM2)	Wheeler and Hendon, 2004
<b>NAO</b>	North Atlantic Oscillation	Hurrell <i>et al.</i> , 2003
<b>PDO</b>	Pacific Decadal Oscillation	Mantua <i>et al.</i> , 1997
<b>PNA</b>	Pacific North American Pattern	Trenberth and Hurrell, 1994
<b>TNA</b>	TSM anomalies in the Tropical Northern Atlantic	Enfield <i>et al.</i> , 1999
<b>TSA</b>	TSM anomalies in the Tropical Southern Atlantic	Enfield <i>et al.</i> , 1999

Some climate indices are not available with daily resolution because they are slowly varying (such as AMO and PDO). Therefore, these indices were estimated with pentad time resolution in the 1981 to 2018 period based on a cubic interpolation (grade 3 polynomial) (Boor, 2001) of the

monthly data.

Some relevant indices associated to climate variability patterns in Brazil are not routinely available on the web. In this case, the climatic indices were computed according to the literature (Table J.2). For the PSA, the indices were generated according to the Mo (2000) reference. The PSA mode is characterized by the second (PSA1) and third (PSA2) EOF mode. In this work, the PSA time series were computed based on their reanalysis data (ERA-Interim, available from ECMWF - European Center for Medium-Range Weather Forecasts) at <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets>) of daily mean geopotential height in 500 hPa for the period 1981 to 2018 for the Southern Hemisphere (90°S – Equator).

**Tabela J.2:** *Climate indices developed and their respective references.*

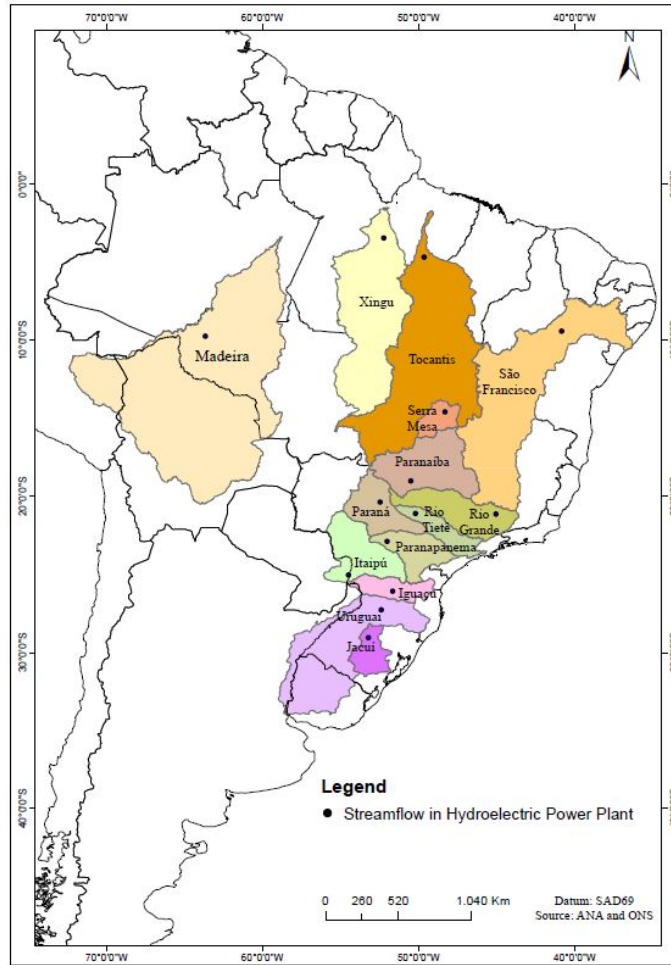
Name	Description	Referência
<b>LISAM</b>	Index of the South American Monsoon System	Silva and Carvalho, 2007
<b>PSA</b>	Pacific South American Pattern	Mo, 2000
<b>SACZ</b>	South Atlantic Convergence Zone	Silva and Carvalho, 2007

The original 0.5° x 0.5° regular resolution grid was degraded to a coarser resolution 2.5° x 2.5°, because the PSA is a large-scale pattern and for computational efficiency. The data were transformed into daily anomalies and accumulated in geopotential height pentads at 500 hPa to perform the EOF analysis (Zang and Moore, 2015). As a result, the PSA1 and PSA2 indices with pentad time resolution are obtained, represented by the time series (scores) of the second and third principal components (EOF2 and EOF3).

Likewise, the LISAM and SACZ indices were developed according to the methodology of Silva and Carvalho, based on the first (LISAM) and second (SACZ) mode of the combined EOF (Zang and Moore, 2015). In this work, were used daily data at 850 hPa of air temperature, specific humidity, and zonal/meridional components of the wind from the ERA-INTERIM reanalysis for the period 1981 to 2018.

The CHIRPS (Climate Hazards Group InfraRed Precipitation with Station Data) precipitation estimate was chosen (Funk *et al.*, 2015) for the precipitation estimate in the Brazilian watersheds with the main hydroelectric power plants (Figure J.1). CHIRPS is a quasi-global rainfall dataset with high spatial resolution (0.05° x 0.05°) and long-term temporal coverage (from 1981 to present day) which combines satellite precipitation estimates and weather station information. The data are available at different time scales (time ranges: monthly, 10 days, 5 days and daily) by the University of California at Santa Barbara (UCSB), <https://www.chc.ucsb.edu/data>. The CHIRPS data were subjected to several evaluations in Brazil. The CHIRPS precipitation dataset performs relatively well on both a regional and global scale, mainly in terms of bias and Pearson's correlation coefficient when compared to GPM (The Global Precipitation Measurement Mission) and TRMM (Tropical Rainfall Measuring Mission) (Paredes-Trejo *et al.*, 2017; Nogueira *et al.*, 2018; Costa *et al.*, 2019; Oliveira-Junior *et al.*, 2021). For this reason, the precipitation records from CHIRPS are used in the following analyses. For each watershed, the accumulated rainfall in pentads (5 days) was computed in the 1981 to 2018 period and the results were standardized by the monthly average and standard deviation, removing the seasonal cycle.

The principal component analysis (PCA) of the precipitation data was the next step for capturing the main variability modes (Zang and Moore, 2015). The orthogonal rotation method called



**Figura J.1:** Location of the main Brazilian watersheds and the streamflow points in the Hydroelectric Power Plant

Varimax (Kaiser, 1958) was applied whenever necessary, to maximize the variance of the results (*i.e.*, to obtain mutually independent modes). The first 6 principal components (PCs) that represent most of the variability of the set of input variables (87% variation from the original data) were selected. The time series resulting from the first 6 PCs (scores) were filtered (Lanczos Filter, high pass above 6 months) to eliminate low frequency variability (Thomson and Emery, 2014) and the wavelet transform (Torrence and Compo, 1998) was applied to identify the prevailing spectral peaks.

The Morlet complex wavelet (Farge, 1994) was used, with a value of  $\omega_0 = 6$  to satisfy the admissibility conditions according to Torrence and Compo (1998). After defining the wavelet function, a set of scales must be defined, as the Morlet function is continuous, the scales are derived from a power of two (Torrence and Compo, 1998). Furthermore, Torrence and Compo (1998) defines the power spectrum (time series variance) as the square of the amplitude and the global power spectrum by the time average of potency. In the power spectrum, the cone of influence is also defined, which is the region affected by the boundary of this spectrum in which the values of the wavelet coefficients can be affected. This cone is defined by the decay (e-folding time) of the wavelet spectrum at each scale, the decay factor is  $e^{-2}$  (Torrence and Compo (1998); Grinsted *et al.*, 2004). For the statistical significance of 90% of a peak in the wavelet spectrum, the null hypothesis is used, which considers



the signal as red noise with a given spectrum of background energy.

Correlation coefficient (Pearson's  $r$  coefficient) and cross-correlation (Stoica and Moses, 2005) were computed to assess the connectivity between climate indices and the precipitation modes based on the standardized climate indices timeseries. The testing for significance of the correlation using Student's  $t$ -test was determined at the 5% level. However, the correlation analysis is a simplified relationship between the indices, as the correlation does not necessarily imply causality. In this case, it is necessary to use statistical methods that identify causality, *i.e.*, the cause-and-effect relationship between the variables. Grange (1969) proposed a method that searches for causality between two variables, through linear vector autoregressive models (VAR), *i.e.*, the Granger causality inference involves the adjustment of VAR(p) models (Lutkepoh, 2005), according to the equation:

$$x(n) = \sum_{k=1}^p A_k x(n-k) + w(n), \quad N_K(0, \Sigma_w), \quad (\text{J.1})$$

Where the elements  $a_{ij,k}$  of the matrices  $A_k$  are responsible for determining the interaction of the  $k$ -th past sample of the series  $x_j(n)$  on another  $x_i(n)$  and  $\Sigma_w$  is the null-mean white process covariance matrix  $w(n) = [w_1(n) \dots w_K(n)]^T$  associated with the model residuals. In general, the models evaluate when a variable  $X$  causes, according to the Granger definition, a variable  $Y$  if, on average, the event  $Y$  is observed every time the event  $X$  occurred some period before. One of the main properties of Granger causality is that it is directional, *i.e.*, if the  $X$  causes  $Y$ , it does not imply that  $Y$  causes  $X$ .

In the frequency domain, Granger Causality can be inferred from the Partial Directed Coherence (PDC) (Baccalá and Sameshima, 2021) which allows detection of direct connections between linear variables.

$$\Pi_{ij}(f) = \frac{|\bar{A}_{ij}|(f)}{\sqrt{\bar{a}_j^H(f) \bar{a}_j(f)}} \quad (\text{J.2})$$

where:  $|\bar{A}_{ij}|(f)$  is the  $i, j$ -th element of the matrix  $\bar{A}(f)$ , where the columns are indicated by  $\bar{a}_1(f)$  that is:

$$\bar{A}(f) = [\bar{a}_1(f) \cdot \cdot \cdot \bar{a}_j(f)] \quad (\text{J.3})$$

When choosing the order of the regression model, the order that best fits the result was chosen. Tests were carried out in order 12, 24 and 48, so that the PDC could reproduce the dynamics (resonances) of the atmospheric signals analyzed, which present an annual cycle and its (sub) multiples.

To investigate the relationship between precipitation and the river discharge of the main Brazilian rivers, *i.e.*, inflow ( $m^3/s$ ) in the Hydroelectric Power Plants (HPPs), the correlation between the time series of the first six PCs of precipitation and time series of naturalized affluent flows in hydroelectric power plants in Brazil were computed. The HPPs inflow data is available at the National Electric System Operator (ONS) site (<http://www.ons.org.br>). Table J.3 provides details on the streamflow points of rivers in selected HPPs (Figure J.1).

The data available by the ONS are daily and do not contain missing data, the data are previously treated and naturalized, *i.e.*, the influent flow is restored as it would naturally occur without

**Tabela J.3:** *Main Hydroelectric Power Plants (HPP) in each Brazilian watershed. The values in bold correspond to the highest HPP in installed power (installed capacity) in Brazil.*

Macro Region	State	Watershed	Subwatershed	Hydroelectric Power Plant	Latitude	Longitude
North	RO	Madeira	Madeira	<b>Santo Antônio</b>	-8.80°	-63.95°
	PA	Xingu	Xingu	<b>Belo Monte</b>	-3.26°	-51.56°
	PA	Tocantins	Baixo Tocantins	<b>Tucuruí</b>	-3.83°	-49.65°
Northeast	BA	São Francisco	Médio São Francisco	Sobradinho	-9.43°	-40.83°
Central	GO	Serra da Mesa	Alto Tocantins	Serra da Mesa	-13.84°	-48.3°
	GO	Paranaíba	Paranaíba	<b>São Simão</b>	-19.02°	-50.5°
Southeast	RJ	Rio Grande	Grande	Funil Grande	-21.14°	-45.04°
	SP	Rio Tietê	Tietê	Nova Avanhandava	-21.12°	-50.2°
	SP	Paranapanema	Paranapanema	Rosana	-22.6°	-52.87°
	SP	Paraná	Médio Paraná	<b>Ilha Solteira</b>	-20.38°	-51.36°
South	PR	Itaipu	Baixo Paraná	<b>Itaipu</b>	-25.43°	-54.59°
	PR	Iguaçu	Iguaçu	Foz da Areia	-26.08°	-51.65°
	RS	Jacuí	Jacuí	Jacuí	-29.07°	-53.21°
	RS	Uruguai	Uruguai	Itá	-27.28°	-52.38°

the reservoirs of the National Interconnected System. For each watershed, the average affluent flow was calculated in pentads (5 days) in the period 1981 to 2018 and correlated with the precipitation modes in the same period.

## RESULTS AND DISCUSSION

To investigate the pentad precipitation variability in Brazilian watersheds, the dimension of the original data set was reduced through principal component analysis. Table J.4 shows the percentage variance (in %) and the accumulated variance for the first six main components (PCs). The first PC explains about 22.1% of the rainfall variability in the Brazilian watersheds and the first six PCs explain almost 90% of the total variance. Thus, the precipitation dataset for the fourteen Brazilian watersheds selected in this study was reduced to six PC's.

**Tabela J.4:** *Percentage of variance and accumulated variance of the first six principal components (PCs) of precipitation in the main Brazilian watersheds.*

PCs	Explained Variance (%)	Explained Variance (%)	Accumulated Variance (%)
	No rotation	Rotation	Rotation
<b>1</b>	33.2%	22.1%	22.1%
<b>2</b>	28.2%	12.1%	34.2%
<b>3</b>	10.6%	13.0%	72.0%
<b>4</b>	6.9%	7.3%	78.9%
<b>5</b>	5.6%	18.0%	84.5%
<b>6</b>	4.5%	14.5%	87.0%

The sampling error was calculated to verify the dependence between each component according to North et al. (1982) of for the first six PCs. There is a strong dependence between PCs (PC1 ↔ PC2, PC3 ↔ PC4, PC3 ↔ PC5 and PC5 ↔ PC6). The Varimax orthogonal rotation method (Kaiser, 1958) was applied to maximize the variance of the results. After rotation of the PCs, the percentage of total variance explained by the two components remains the same, but each of them explains different percentages than those explained by the components before the rotation. Note that after vector rotation by the Varimax orthogonal method (Table J.4), the components present explained variance closer to each other, presenting variability with less dependence than the method without rotation, *i.e.*, it reduces the dependence between the components.

Table J.5 highlights the significant values of CC (values in bold with 95% statistical significance) between the 6 watersheds precipitation PC's and the observed total precipitation. The watersheds

in Table J.5 are distributed according to geographic location (from North to South).

**Tabela J.5:** *Correlations between the six main components and rainfall in the main Brazilian watersheds. Values in bold correspond to a significance level of 95%, based on Student's t test.*

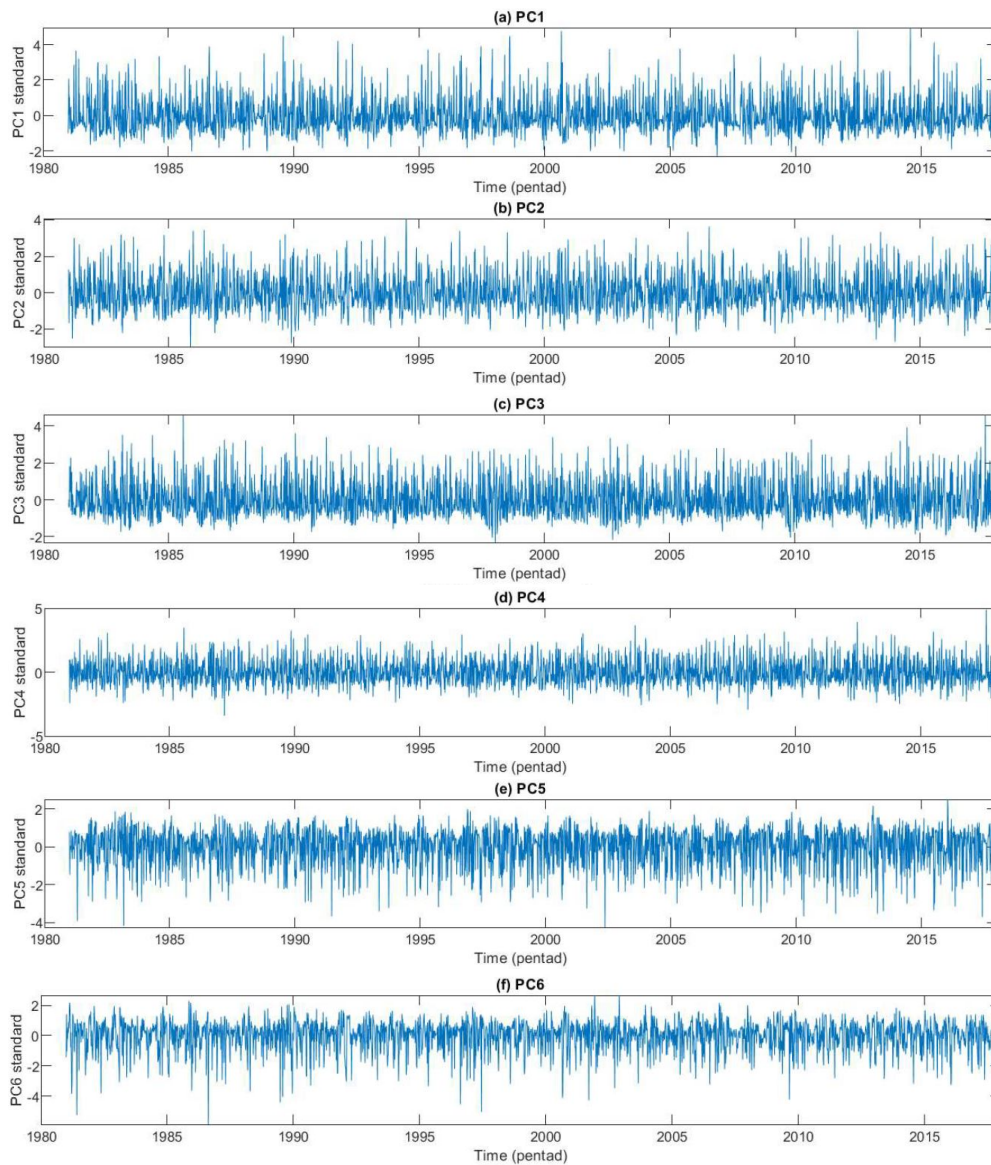
Macro Region	Watershed	PC1	PC2	PC3	PC4	PC5	PC6
North	Madeira	<b>0.21</b>	<b>0.24</b>	0.05	<b>0.99</b>	<b>-0.17</b>	<b>-0.15</b>
	Xingu	<b>0.10</b>	<b>0.96</b>	<b>-0.26</b>	<b>0.21</b>	<b>0.17</b>	<b>-0.42</b>
	Tocantins	<b>0.27</b>	<b>0.91</b>	<b>-0.30</b>	<b>0.23</b>	<b>0.14</b>	<b>-0.77</b>
Northeast	São Francisco	<b>0.23</b>	<b>0.51</b>	<b>-0.32</b>	<b>0.09</b>	<b>0.17</b>	<b>-0.89</b>
Central	Serra da Mesa	<b>0.36</b>	<b>0.60</b>	<b>-0.24</b>	<b>0.17</b>	<b>0.09</b>	<b>-0.90</b>
	Paranaíba	<b>0.80</b>	<b>0.40</b>	<b>-0.19</b>	<b>0.25</b>	<b>-0.14</b>	<b>0.71</b>
Southeast	Rio Grande	<b>0.93</b>	<b>0.20</b>	<b>-0.13</b>	<b>0.15</b>	<b>-0.33</b>	<b>-0.43</b>
	Rio Tietê	<b>0.93</b>	<b>0.11</b>	0.00	<b>0.16</b>	<b>-0.57</b>	<b>-0.18</b>
	Paranapanema	<b>0.73</b>	-0.03	<b>0.19</b>	<b>0.17</b>	<b>-0.87</b>	0.03
	Paraná	<b>0.82</b>	0.06	0.05	<b>0.34</b>	<b>-0.66</b>	-0.14
South	Itaipu	<b>0.48</b>	<b>-0.14</b>	<b>0.33</b>	<b>0.25</b>	<b>-0.95</b>	<b>0.09</b>
	Iguaçu	<b>0.30</b>	<b>-0.19</b>	<b>0.49</b>	<b>0.13</b>	<b>-0.91</b>	<b>0.18</b>
	Jacuí	-0.06	<b>-0.26</b>	<b>0.96</b>	0.03	<b>-0.34</b>	<b>0.28</b>
	Uruguai	<b>-0.07</b>	<b>-0.31</b>	<b>0.96</b>	0.06	<b>-0.41</b>	<b>0.31</b>

Table J.5 shows that the spatial distribution of PC's, the PC1 component has a strong correlation with the Southeast watersheds, PC2 with the central part of the North region, PC3 is heavily weighted to the extreme south, PC4 with the Madeira River watershed (close to the Andes), PC5 with the north of the South region, and PC6 in the watersheds of the Central and Northeast regions.

The wavelet transform was applied to the time series of the standardized and filtered time series (Figure J.2) in the intraseasonal scale (up to 6 months) of the first six PCs (Figure J.3). PC1, PC2 and PC4 are dominated by a significant peak with period of the order of 9 pentads, relatively broader in the PC4 case. PC3, PC5 and PC6 share a common significant peak at a slightly slower intraseasonal cycle (period of the order of 5 to 6 pentads). PC5 and PC6 share significant power in a band with a longer period between the intraseasonal timescale and seasonal (approximately 18 and 36 pentads). Nevertheless, PC5 and PC6 have a marginally significant power at the S2S (PC5 at 3 and 6 pentads and PC6 at approximately 6 pentads). It is noteworthy that PC1 almost shows a significant spectral peak with period of the order of 4-5 pentads (faster than the typical MJO scale). Thus, it is clear that there are 2 well defined intraseasonal peaks (8-9, 3-6 pentads) and a hybrid synoptic/intraseasonal mode influencing the precipitation in the main Brazilian watersheds (Vera *et al.*, 2006).

To show the relationships between the PC components and climate indices, Pearson's correlations were calculated in the standardized time series. Table J.6 shows the indices that showed significant correlations of 95%, from the t test of Student. In general, according to Table J.6, the indices that have the highest correlation with precipitation PCs in the main Brazilian watersheds are those that represent the interannual phenomena ENSO (MEI index), PSA and AAO, interdecadal the PDO, the multidecadal the AMO, the intraseasonal variability of the SMAS and MJO.

The lagged (1 to 12 pentads) cross-correlation between the main precipitation patterns (PC1 to PC6) and climate indices were determined (Figure J.4) to identify potential climatic predictors of precipitation in Brazilian watersheds in the subseasonal time scale. The longer lags in Figure J.4 have lighter tones while the shorter ones have stronger tones (bars from left to right). For the PC1 component (Figure J.4a), significant CC is observed for the LISAM/SACZ patterns and RMM1/RMM2, suggesting a potential precipitation predictor based on the SMAS and MJO patterns, for up to

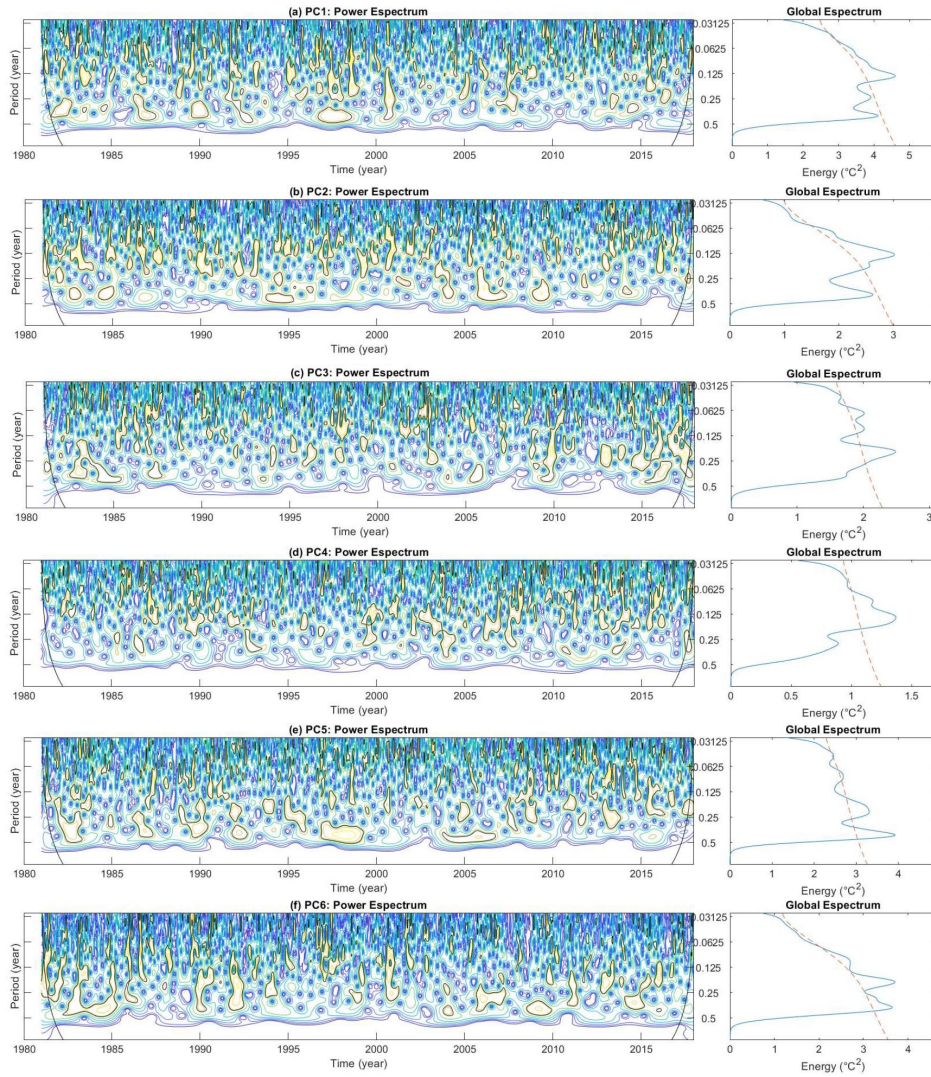


**Figura J.2:** Standardized time series of the first six principal components of precipitation.

**Tabela J.6:** Correlations between the six principal components (PCs) and climate indices. Values in bold correspond to the significance level of 95%

PCs	AAO	PSA1	PSA2	LISAM	SACZ	RMM1	RMM2	AMO	MEI	PDO	PNA	TNA
1	-0.01	-0.02	<b>-0.08</b>	<b>-0.30</b>	<b>0.13</b>	<b>-0.10</b>	0.01	-0.02	<b>0.08</b>	0.02	0.01	-0.01
2	<b>-0.09</b>	<b>0.16</b>	<b>-0.19</b>	<b>-0.32</b>	<b>0.34</b>	<b>-0.23</b>	-0.02	-0.04	<b>-0.10</b>	<b>-0.07</b>	<b>-0.07</b>	<b>-0.07</b>
3	0.03	<b>-0.07</b>	0.00	0.01	<b>-0.45</b>	<b>0.08</b>	-0.01	0.03	<b>0.14</b>	<b>0.09</b>	0.04	0.03
4	<b>-0.09</b>	0.02	<b>-0.14</b>	<b>-0.20</b>	0.04	<b>-0.15</b>	<b>0.09</b>	<b>-0.07</b>	-0.05	-0.03	-0.05	<b>-0.08</b>
5	0.01	<b>0.07</b>	0.03	<b>0.15</b>	<b>0.14</b>	0.00	-0.02	-0.01	<b>-0.12</b>	-0.06	-0.03	-0.02
6	0.01	<b>-0.07</b>	<b>0.12</b>	<b>0.43</b>	<b>-0.36</b>	<b>0.15</b>	-0.01	0.05	0.00	0.01	0.02	0.04

30 days in advance. MEI and AAO patterns are also potentially significant for all lags (1 to 12 pentads) and for longer lags (between 6 and 12 pentads), respectively. For PC2 (Figure J.4b), CC is significant for the low frequency indices and AAO, MEI, PDO, PNA and TNA at all lags. Thus, the large-scale climatic variability modes influence the subseasonal precipitation (5 to 60 days) as well as the PSA1, RMM1 and RMM2 indices, in the short to intermediate lags (up to 45 days). For PC3 (Figure J.4c), the significant CC are observed for the AAO, PSA1, MEI and PDO indices, in

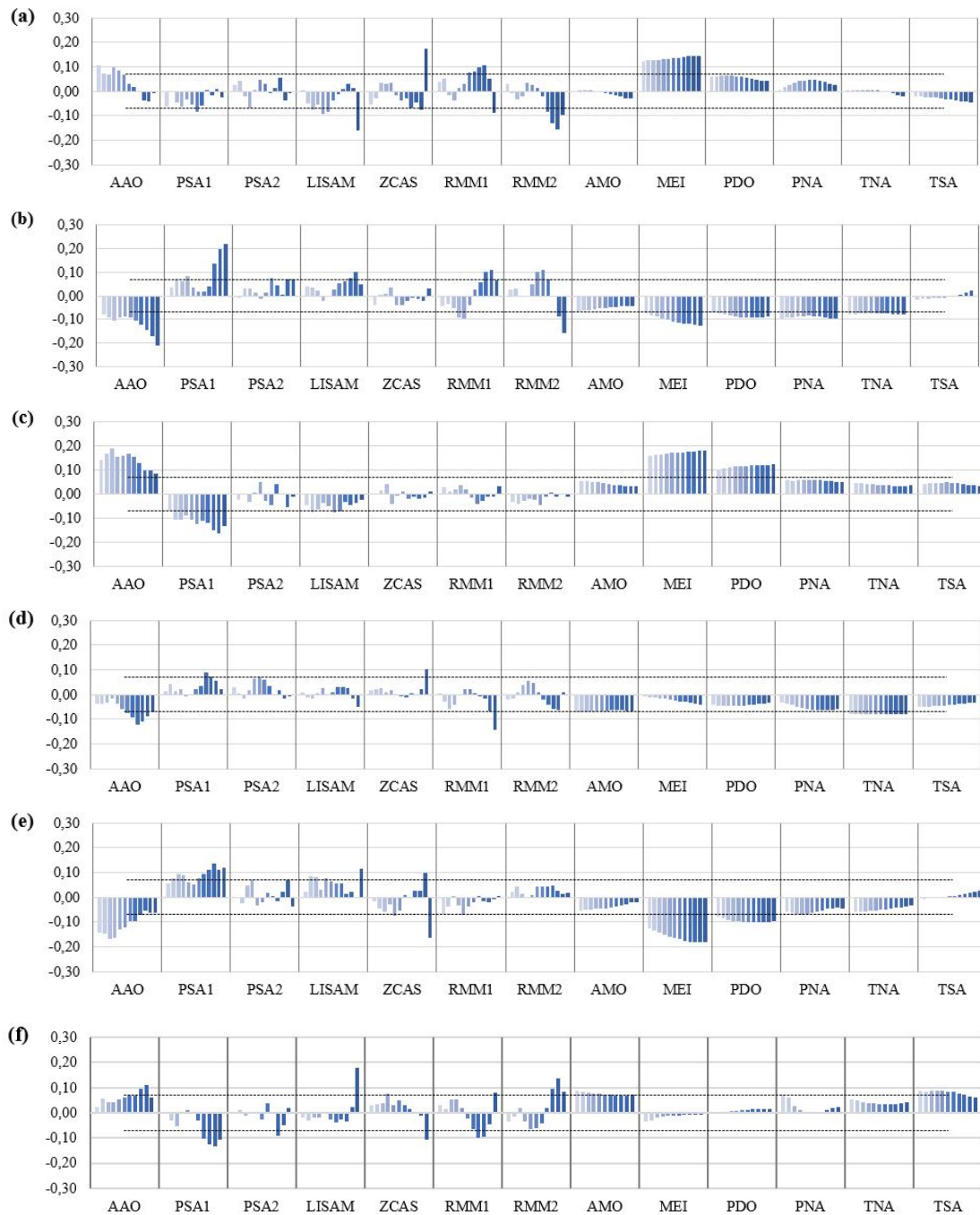


**Figura J.3:** Power and Global Spectrum of the wavelet analysis of the time series of precipitation a) PC1; b) PC2; c) PC3; d) PC4; e) PC5 e f) PC6. The power spectrum is delimited by the influence cone and the black outline indicates the sector of the power spectrum where the 90% confidence level is satisfied. The global power spectrum indicates the power level satisfied for the complete time series, that is, above the dashed red line. The period is the unit of years ( $0.03125 \approx 2$  pentads,  $0.0625 \approx 4.5$  pentads,  $0.125 \approx 9$  pentads,  $0.25 \approx 18$  pentads, and  $0.5 \approx 36$  pentads).

all lags, indicating that PC3 is a potential precipitation predictor up to 60 days. A similar pattern was also identified for the PC5 component (Figure J.4e). Finally, the PC4 (Figure J.4d) and PC6 (Figure J.4f) precipitation patterns are significantly associated to the Atlantic indices AMO, TNA or TSA (PC6) at all lags and to AAO and RMM1/RMM2 indices in the shorter lags.

Based on the initial analysis of the CC between climate indices and precipitation PCs (Table J.6), and between PCs and average precipitation in each Brazilian watershed (Table J.5), a subset of the original indices were selected for the causality analysis. Table J.7 summarizes the selected indices for each precipitation mode (PC1 to PC6).

Figure J.5 shows the PDC results for the six PCs and the selected indices according to Table J.7. The black tracing represents the (pseudo-) spectral density of the series in  $dB$ ; the red line represents the statistically significant values of the PDC; the discontinuous black line represents the threshold of the Patnaik approximation (Baccalá and Sameshima, 2021), with a significance level equal to  $\alpha$ ;



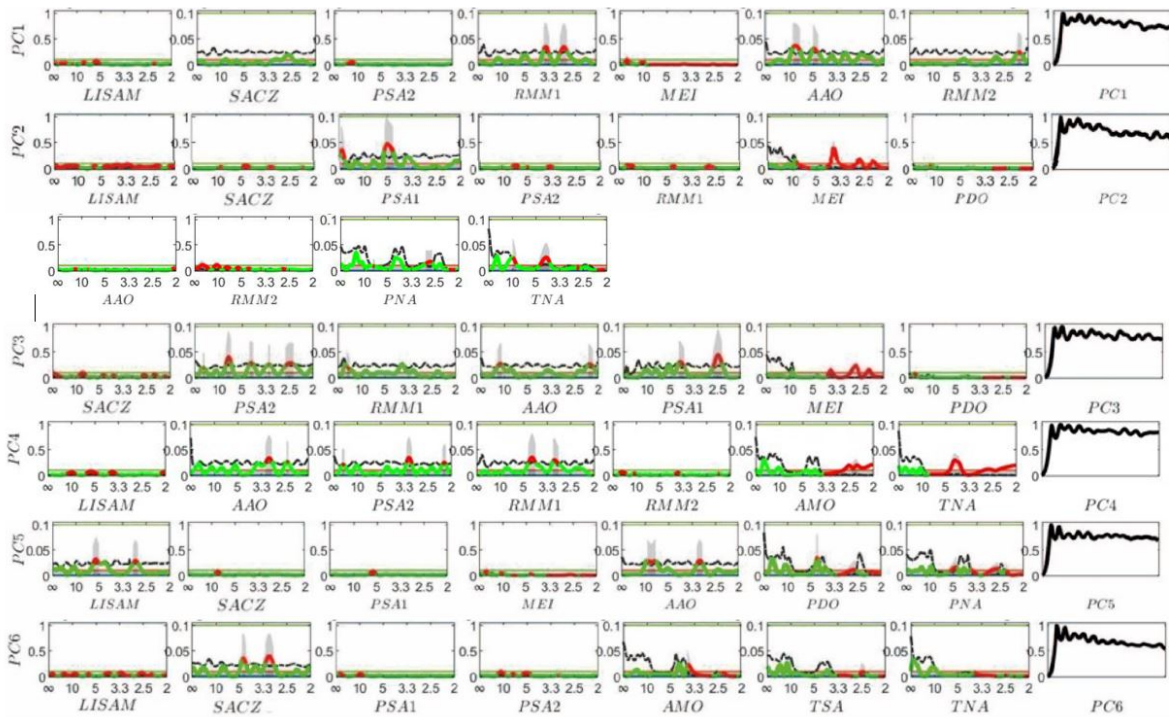
**Figure J.4:** Pentad lagged correlation (lag -12 to lag 1, from lightest to darkest tones) between precipitation modes ((a) PC1 - (f) PC6) and climate variability indices. The dashed lines represent the limit of significance from the student's *t*-test (-0.07 and +0.07).

the plotted in green the statistically non-significant values of the PDC and the values of the *x* axis in pentads. The direction of causality is observed from the *x* axis (abscissa) to the *y* axis (ordinate), in this work we chose to expose only the results where the indices cause precipitation in PCs.

As shown in Figure J.5, there are statistically significant values of causality at high frequency (intraseasonal) between all PCs and the LISAM, SACZ, MJO (RMM1 and/or RMM2), AAO and PSA indices (PSA1 and/or PSA2 indices), indicating a strong relationship between these patterns and rainfall in Brazil in all Brazilian watersheds analyzed. For the precipitation PC components

**Tabela J.7:** Selected indices for causality analysis (PDC) between precipitation modes (PCs 1 to 6) and selected climate indices with higher correlations.

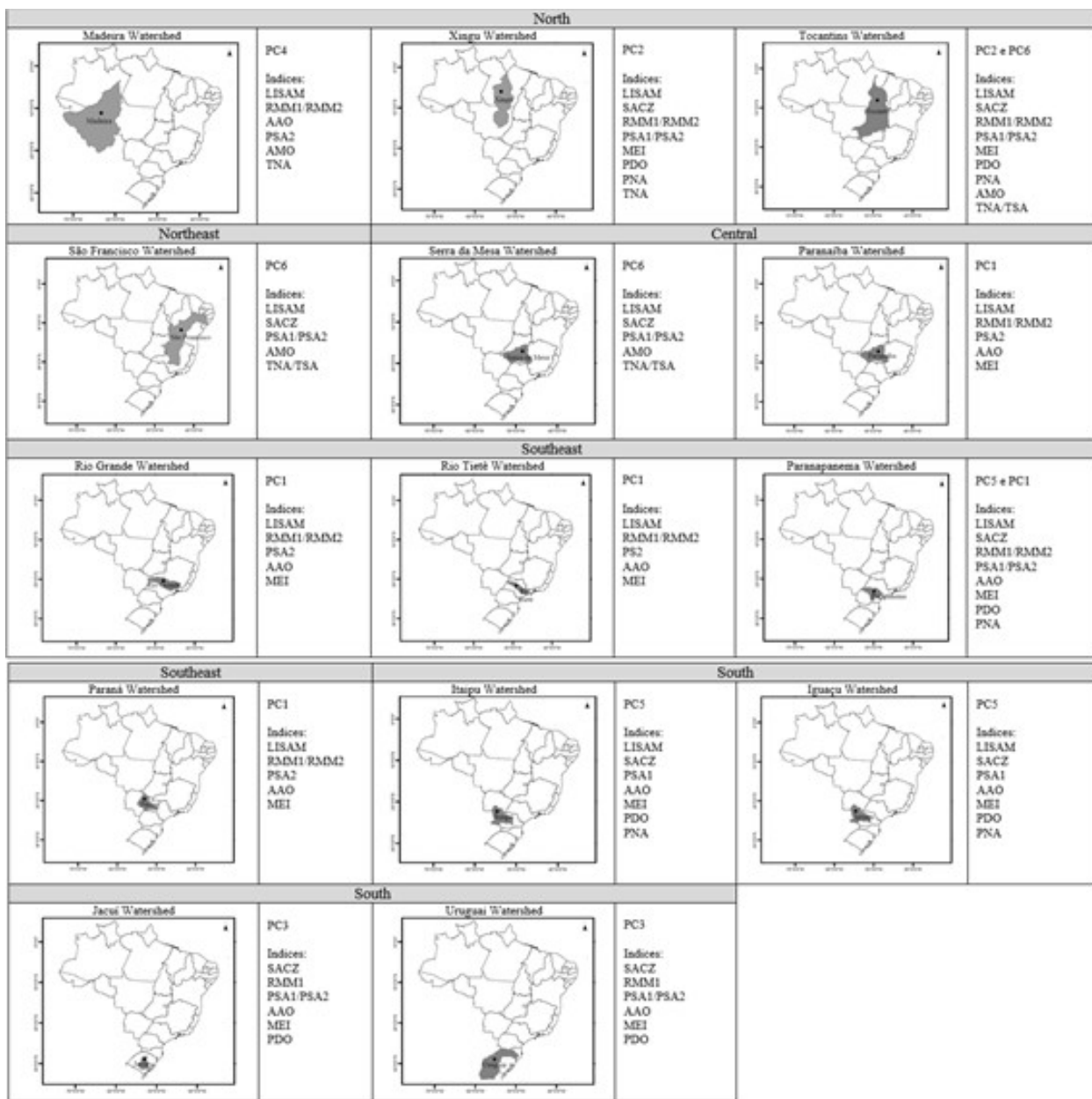
Selected indices for causality analysis													
PCs	LISAM	SACZ	AAO	PSA1	PSA2	RMM1	RMM2	MEI	PDO	AMO	PNA	TNA	TSA
<b>1</b>	x	x	x		x	x	x	x					
<b>2</b>	x	x		x	x	x	x	x	x		x	x	
<b>3</b>	x	x	x	x	x	x		x	x				
<b>4</b>	x		x		x	x	x			x		x	
<b>5</b>	x	x	x	x				x	x		x		
<b>6</b>	x	x			x					x			x



**Figura J.5:** The black line is the (pseudo-) spectral density of the series, in dB units (PC1 to PC6); the red graph represents the statistically significant PDC values; the dashed black line represents the approximation of the Patnaik threshold (Baccalá and Sameshima, 2021); the green line shows statistically non-significant PDC values and the values of the x axis in pentads. Therefore, using the significance level of 1%, the numbers represent the PDC for the set of time series of PCs x indices using an autoregressive model of order  $p = 24$ .

significant values are observed at all periods (2, 2.5, 3.3, 5, 10 pentads and above 10 pentads) for the Pacific indicators (MEI and PDO indexes). These results reinforce the Pacific Ocean SST anomalies on the Brazilian watersheds precipitation in agreement with the PC’s spectral peaks (Figure J.3). Figure J.5 shows that for the PC2, PC4 and PC6 components, there are significant spectral peaks, suggesting the influence of the Atlantic patterns, AMO, TNA and TSA, between 2 and 10 pentads. Furthermore, for the PC2 and PC5 components, the influence of the Northern Hemisphere circulation pattern (PNA) is observed in shorter periods (2 to 5 pentads). The PNA influence is mostly through the impact of the action center located in the Atlantic Ocean due to its impact in the moisture flow associated to the trades and possibly on the ITCZ positioning.

A summary of the subseasonal scale dominating precipitation patterns and the most influencing climate indices in each Brazilian watersheds in shown in Figure J.6.



**Figura J.6:** Main components (PC) with the highest correlation in each Brazilian watersheds and the respective climatic indices that cause precipitation at the intraseasonal scale.

According to Kousky and Kayano (1994), MJO influences the onset of a SACZ episode, changes in location and its intensity because of upper-level influence of the equatorial propagation of the MJO signal over South America. Another possible effect is associated to transient systems that contribute to the formation of SACZ such as the role of Rossby wave propagation triggered by changes in the South Pacific Convergence Zone induced by the MJO in the Pacific Ocean (Grimm and Silva Dias, 1995). The PSA pattern also contributes to the convection in the SACZ region (Vasconcellos and Cavalcanti, 2010), while the AAO influences the positioning of the trajectories of cyclonic systems in the subtropical and extratropical South American continent. AAO can also contribute to the intensification of Rossby wave trains of the PSA type (Vasconcellos and Cavalcanti, 2010; Machado *et al.*, 2021). The SACZ organizes convection specially in the summer months in the central-south Amazon (Serra da Mesa and Tocantins watersheds), the Central and Southeast regions



(Paranaíba, Paranapanema, Paraná, Rio Grande and Rio Tietê watersheds), south central Bahia, northern portion of the State of Paraná (Itaipu and Iguaçu watersheds) and extending towards the Southwest Atlantic Ocean (Carvalho *et al.*, 2004; Ferreira *et al.*, 2004; Marengo *et al.*, 2012).

The Madeira watershed in Northwest Brazil (Amazon catchment) is strongly correlated to PC4 (CC = 0.99) (Figure J.6), which in its turn is correlated to the SMAS, MJO, AAO, PSA and Atlantic indices (AMO and TNA). These results corroborate several studies of the Amazon Watershed that reinforce the influence of MJO, PSA, AAO (Cavalcanti and Shiumizu, 2012; Mayta *et al.*, 2020) and the Atlantic variability role in the SMAS (Jones and Carvalho, 2018; Marengo and Nobre, 2009).

In Northern Brazil, the Xingu and Tocantins catchments correlate strongly with PC2 (CC = 0.96 and CC = 0.91, respectively J.6) and moderately with the Northeast and Central watersheds (São Francisco, Serra da Mesa and Paranaíba). These results suggest, in addition to those mentioned above (LISAM/SACZ, MJO, PSA and AAO), a relationship with the Pacific (ENSO and PDO) and North Atlantic (PNA and TNA) patterns supporting previous studies on the relationship between precipitation in the North and the low frequency phenomena of the Pacific (Foley *et al.*, 2002; Molion 2008; Tedeschi *et al.*, 2016) and Atlantic (Andreoli and Kayano, 2006). In addition, the high CC (CC = -0.77) for the Tocantins catchment with PC6 (Figure J.6) and with the Northeast and Midwest watersheds (São Francisco, Serra da Mesa and Paranaíba) indicates a correlation with South Atlantic patterns (AMO and TSA) (Jones and Carvalho, 2018; Andreoli and Kayano, 2006), which, in turn, may influence the SACZ positioning, strengthening or inhibiting precipitation (Chaves, 2011).

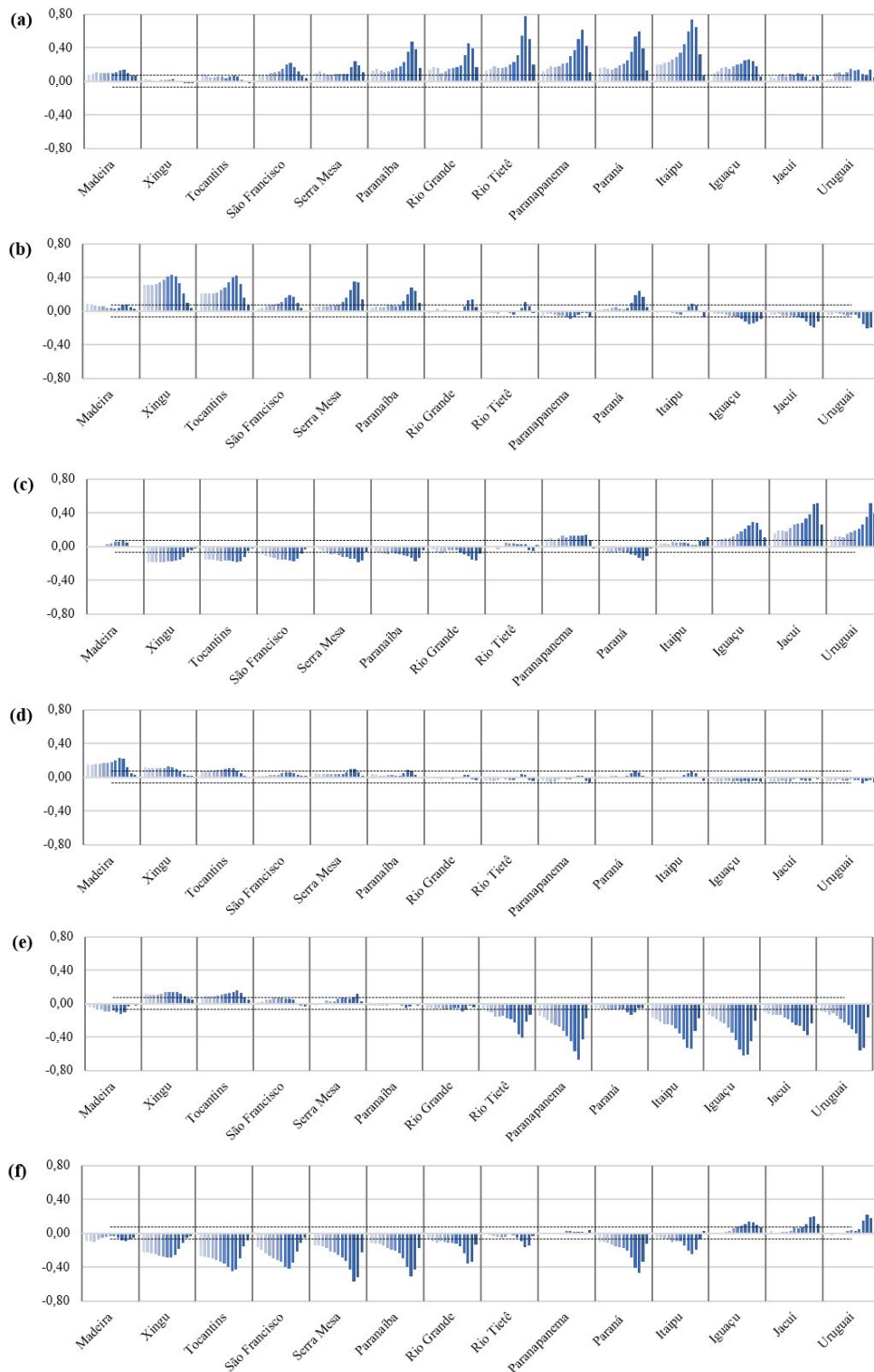
The highest correlations are observed for PC1 and PC6 (Figure 6) for the watersheds of the South region (Itaipu, Iguaçu, Jacuí and Uruguay) and Southeast (Rio Grande, Rio Tietê, Paranapanema and Paraná), indicating a strong relationship with the ENSO and PDO (Rebello, 2006; Molion, 2008; Tedeschi *et al.*, 2016). It is noteworthy that for the ENSO and PDO Pacific phenomena, opposite patterns are observed between North/Northeast and South/Southeast.

To analyze the response time of precipitation in river discharge, strategical for planning of water resources management, the CC's were calculated for naturalized affluent flows in the main rivers of each watershed, *i.e.*, at the points of the main HPP of the parents (Collischonn *et al.*, 2007; Malfatti *et al.*, 2018b). Figure J.7 shows the lagged correlations from lag -12 to lag 0 in pentads between precipitation modes (PC1 to PC6) and streamflow in the main Brazilian catchments. Longer lags have lighter tones, shorter lags have stronger tones (slashes from left to right).

In this Figure J.7, it is observed that the response time of precipitation for river flow is highly variable. In the Madeira River catchment (northwestern Brazil), the response time varies between 15-30 days (expressed by the CC values along the CP4 lags). Watersheds in the North (Xingu and Tocantins) the response time is slower, varying between 15-40 days, according to the CC values along the lags of PC2 (positive influence CC > 0) and PC6 (negative influence CC < 0).

In the Central region catchments (Serra Mesa and Paranaíba) the response time is faster than in the previous regions, between 5-20 days, according to the values observed in PC2 and PC6. Similar behavior observed in the South (Itaipu, Iguaçu, Jacuí and Uruguay) and Southeast (Rio Grande, Rio Tietê, Paranapanema and Paraná) watersheds, according to the lagged CC values with PC1 and PC5.

In addition, to the watershed precipitation, we have to consider other components of the river



**Figure J.7:** Lagged correlation (lag -12 to lag 0 in pentads, from lightest to darkest tones) between precipitation modes ((a) PC1 to (f) PC6) and streamflow in the main watersheds dashed lines represent the limit of significance from the student's *t*-test (-0.07 and +0.07).

discharge, such as the physiology (*i.e.*, the physical characteristics of the watershed, such as the area of the contributing water-catchment, its shape, soil surface conditions and subsoil geological constitution, and topographic conformation) and anthropogenic (*i.e.*, the hydraulic works built in the watershed and human occupation) (VonSperling, 2007). Nevertheless, in this context, accu-

rate precipitation forecasts helped in decision making, but factors not related to climate variability should also be considered in the strategic planning of water resources.

## CONCLUSION

In this paper, it was possible to identify the relationships between climatic variability patterns, in the intraseasonal, interannual, interdecadal and multidecadal, which interfere in the intraseasonal scale of precipitation in different Brazilian watersheds, through the correlation analysis of precipitation variability modes (PC1 to PC6, which correspond to 87% of the total variability) and climatic indices. Furthermore, it was possible to identify possible climate predictors for the main watersheds in Brazil.

In general, the watershed precipitation modes are significantly influenced by the MJO (RMM1 and/or RMM2), AAO and PSA (PSA1 and/or PSA2 indices) patterns in the SMAS (LISAM and/or SACZ indices). Furthermore, the importance of the Pacific patterns, such as ENSO and PDO, is highlighted for almost all Brazilian watersheds, except for the watershed Madeira River.

The North (Madeira, Xingu and Tocantins), Northeast (São Francisco) and Central (Serra da Mesa and Paranaíba) watersheds show a stronger Atlantic influence, with the AMO and TNA/TSA indices. The Pacific modes, ENSO and PDO, have larger influence on precipitation in the South (Itaipu, Iguaçú, Jacuí and Uruguay) and Southeast (Rio Grande, Rio Tietê, Paranapanema and Paraná) watersheds.

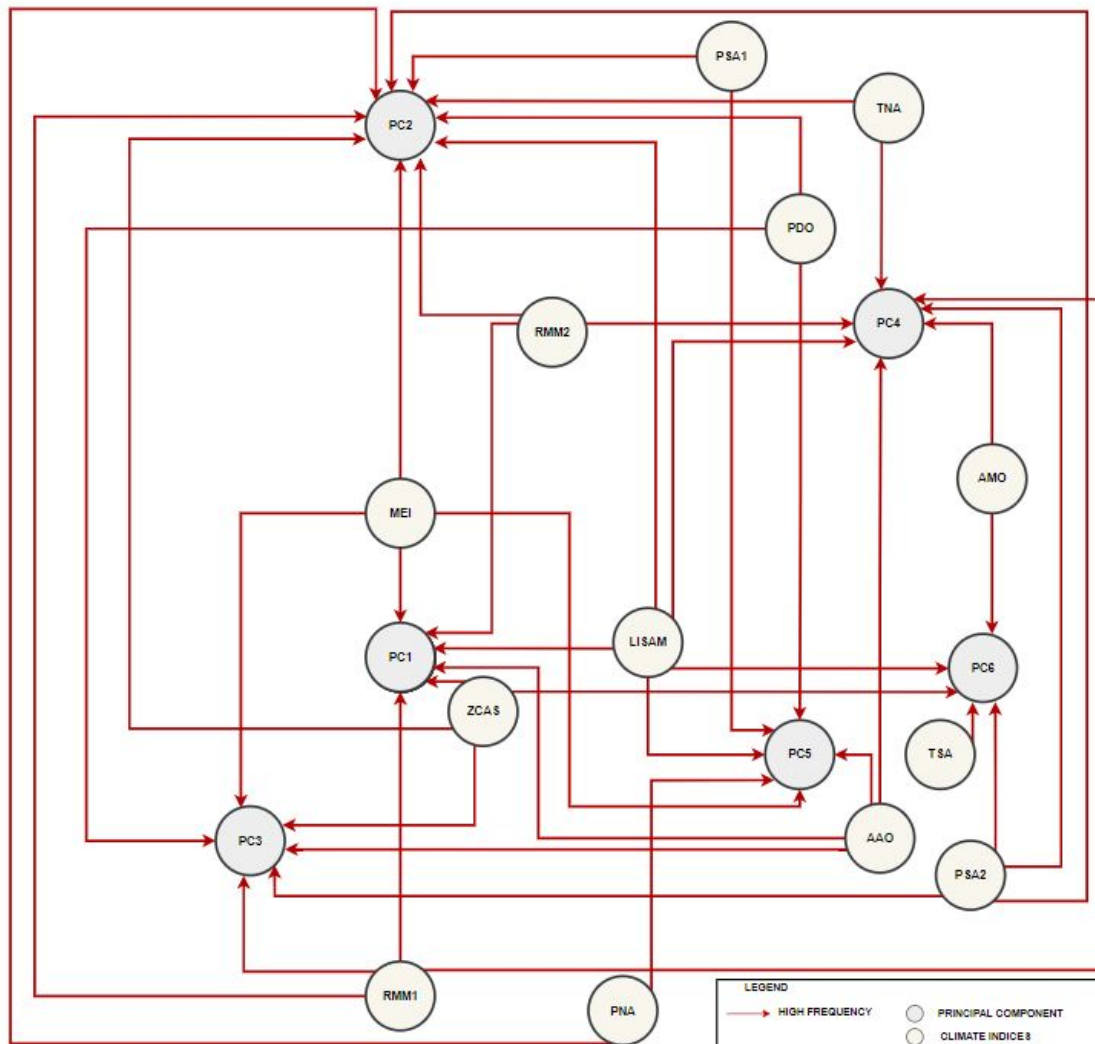
The Partial Direction Coherence (PDC) causality analysis led to the identification of the frequency ranges of the precipitation PC's that are caused, in the Granger sense, by the climate variability pattern. Figure J.8 summarizes the causal relationships between climate indices and precipitation components. Red connectors indicate high frequency causality (intraseasonal scale).

The PDC results validate the influence of several different patterns of climate variability (intraseasonal, interannual, decadal and multidecadal) on the intraseasonal time scale.

Furthermore, cross-correlation analysis indicates that the weather patterns, identified by the indices presented in Tables 1 and 2, are potential predictors for the sub-seasonal scale precipitation in the Brazilian watersheds. The LISAM and SACZ indices are potential predictors for shorter periods up to 20 days, while the RMM1/RMM2 indices are more relevant for intermediate sub-seasonal periods (up to 45 days). The AAO, MEI, PDO, AMO, PNA and TNA/TSA indices have significant correlations at all lags in the sub-seasonal scale, indicating precipitation predictors up to 60 days.

The response time of precipitation for river discharge change depending on the watershed. The response time in the Madeira River watershed (northwestern Brazil) varies between 15-30 days. The northern watersheds (Xingu and Tocantins) have a slower response time, varying between 15-40 days, as indicated by the lagged CC of PC2 (positive influence  $CC > 0$ ) and PC6 (negative influence  $CC < 0$ ).

In the Midwest watersheds (Serra Mesa and Paranaíba) the response time is faster than in the other regions (between 5-20 days) according to the lagged CC values observed in PC2 and PC6 (Figure 7). Similar behavior is observed in the South (Itaipu, Iguaçú, Jacuí and Uruguay) and Southeast (Rio Grande, Rio Tietê, Paranapanema and Paraná) catchments, according to the CC



**Figura J.8:** Causality diagram between the main components of precipitation and climate indices. Red connectors indicate high frequency causality (intra-seasonal scale).

values along the lags of PC1 and PC5.

As the response time of precipitation in river flow varies from 5 to 40 days, depending on the watershed location, precipitation modeling helps in the planning and management of water resources, anticipating its impacts between 60 to 100 days. The results obtained highlight the importance of low-frequency patterns of climate variability in controlling intraseasonal precipitation variability in most Brazilian watersheds. Thus, weather pattern indices can be used as precipitation predictors, since the indices have significant correlations in longer lags, which help to improve precipitation predictor models (ongoing research).

## ACKNOWLEDGEMENTS

The authors thank USP for its technical and financial support and CNPq for the first author's doctoral scholarship (n° 166798/2017-7).

## REFERENCE

Andreoli, R.V. and Kayano, M.T. (2005) ENSO-related rainfall anomalies in South America and associated circulation features during warm and cold Pacific Decadal Oscillation regimes. Interna-

tional Journal of Climatology, 25(15):2017–2030. <https://doi.org/10.1002/joc.1222>.

Andreoli, R.V. and Kayano, M.T. (2006) Tropical Pacific and South Atlantic effects on rainfall variability over Northeast Brazil. *International Journal of Climatology*, 26(13): 1895–1912. <https://doi.org/10.1002/joc.1341>.

Baccalá, L.A. and Sameshima, K. (2021) Partial directed coherence: twenty years on some history and an appraisal. *Biol Cybern*, 115:195–204. <https://doi.org/10.1007/s00422-021-00880-y>.

Boor, C. (2001) *A Practical Guide to Splines*. Springer-Verlag New York Inc.

Cai, W., McPhaden, M.J., Grimm, A.M., Rodrigues, R.R., Taschetto, A.S., Garreaud, R.D., Dewitte, B., Poveda, G., Ham, Y.-G., Santoso, A., Ng, B., Anderson, W., Wang, G., Geng, T., Jo, H.-S., Marengo, J.A., Alves, L.M., Osman, M., Li, S., Wu, L., Karamperidou, C., Takahashi, K. and Vera, C. Seasonal precipitation variability modes over South America associated to El Niño-Southern Oscillation (ENSO) and non-ENSO components during the 1951–2016 period. *Nature Reviews Earth & Environment*, 1(4):215–231, 2020. <https://doi.org/10.1038/s43017-020-0040-3>.

Carvalho, L.M.V., Jones, C. and Liebmann B. (2004) The South Atlantic convergence zone: Intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall. *Journal of Climate*, 17(1):88–108. [https://doi.org/10.1175/1520-0442\(2004\)017<0088:TSACZI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<0088:TSACZI>2.0.CO;2).

Carvalho, L.M.V., Charles, J. and Ambrizzi, T. (2005) Opposite phases of the Antarctic Oscillation and relationships with intraseasonal to interannual activity in the tropics during the austral summer. *Journal of Climate*, 18:702–718. <https://doi.org/10.1175/JCLI-3284.1>.

Cavalcanti, I.F.A. (2000) Teleconnection patterns orographically induced in model results and from observational data in the austral winter of the Southern hemisphere. *International Journal of Climatology*, 20(10):1191–1206. [https://doi.org/10.1002/1097-0088\(200008\)20:103.3.CO;2-7](https://doi.org/10.1002/1097-0088(200008)20:103.3.CO;2-7).

Cavalcanti, I.F.A. and Shiumizu. M.H. (2012) Climate fields over south America and variability of SACZ and PSA in hadgem2-es. *American Journal of Climate Change*, 1(3):132–144. <https://doi.org/10.4236/ajcc.2012.13011>.

Chaves, R.R. (2011) Conexões entre a TSM do atlântico sul e a convecção de verão sobre a América do Sul – Análise observacional. *Revista Brasileira de Geofísica*, 29(1):5–10. <https://doi.org/10.1590/S0102-261X2011000100001>.

Chiang, J., Kushnir, Y. and Giannini A. (2002) Deconstructing Atlantic intertropical convergence zone variability: Influence of the local cross-equatorial sea surface temperature gradient and remote forcing from the eastern equatorial pacific. *Journal of Geophysical Research*, 107(D1):ACL 3–1–ACL 3–19. <https://doi.org/10.1029/2000JD000307>.

Chiang, J.H. and Vimont, D.J. (2004) Analogous pacific and atlantic meridional modes of tropical atmosphere-ocean variability. *Journal of Climate*, 17(21):4143–4158. <https://doi.org/10.1175/JCLI4953.1>.

Chiessi, C.M., Mulitza, S., Pätzold, J., Wefer, G. and Marengo, J.A. (2009) Possible impact of the Atlantic Multidecadal Oscillation on the South American summer monsoon. *Geophysical Research Letters*, 36(21):1–5. <https://doi.org/10.1029/2009GL039914>.

Coelho, C.A.S., Uvo, C.B. and Ambrizzi, T. (2002) Exploring the impacts of the tropical Pacific SST on the precipitation patterns over South America during ENSO periods. *Theoretical and Applied Climatology*, 71(3):185–197. <https://doi.org/10.1007/s007040200004>.

Collischonn, W., Tucci, C.E.M., Clarke, R.T., Silva Dias P.L. and Oliveira. G.S. (2005) Previsão sazonal de vazão na bacia do rio Uruguai 2: Previsão climática-hidrológica. *Revista Brasileira de Recursos Hídricos*, 10(4):61–72. <https://doi.org/10.21168/rbrh.v10n4.p61-72>.

Collischonn, W., Tucci, C.E.M., Clarke, R.T., Corbo, M.D., Silva, B.C., Collischonn, B., Allasia, D. and Paz A.R. (2007) Modelo hidrológico distribuído para previsão de vazões incrementais na bacia do rio Paranaíba entre Itumbiara e São Simão. *Revista Brasileira de Recursos Hídricos*, 12(3):43–56. <https://doi.org/10.21168/rbrh.v12n3.p43-55>.

Correia Filho, W.L.F., Oliveira-Júnior, J.F. de, Silva Junior, C.A. and Santiago, D.d.B. (2021). Influence of the El Niño-Southern Oscillation (ENSO) and the synoptic systems on the rainfall variability over the Brazilian Cerrado via CHIRPS data. *Int J Climatol.*, Accepted Author Manuscript. <https://doi.org/10.1002/joc.7417>.

Costa, J.C. Pereira, G., Siqueira, M.E., Cardozo, F.S. and Silva, V.V. (2019) Validação dos dados de precipitação estimados pelo CHIRPS para o Brasil. *Revista Brasileira de Climatologia*, 24:228–243. <https://doi.org/10.5380/abclima.v24i0.60237>.

Enfield, D.B., Mestas-Nuñez, A.M. and Mayer D.A. (1999) How ubiquitous is the dipole relationship in tropical Atlantic sea surface temperatures? *Journal of Geophysical Research: Oceans*, 104(C4):7841–1848. <https://doi.org/10.1029/1998JC900109>.

Enfield, D.B., Mestas-Nuñez A.M. and Trimble, P.J. (2001) The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S. *Geophysical Research Letters*, 28(10):2077–2080. <https://doi.org/10.1029/2000GL012745>.

Farge, M. (1994) Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics*, 51:2523–2541. <https://doi.org/10.1146/annurev.fl.24.010192.002143>.

Ferreira, N.J, Sanches, M. and Silva Dias, M.A.F. Composição da zona de convergência do Atlântico Sul em períodos de El Niño e La Niña. *Revista Brasileira de Meteorologia*, 19(1):89–98.

Foley, J.A., Botta, A., Coe, M.T. and Costa, M.H. (2002) El Niño-Southern Oscillation and the climate, ecosystems and rivers of Amazonia. *Global Biogeochemical Cycle*, 16(4):79(1–20). <https://doi.org/10.1029/2002GB001872>.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. and Michaelsen, J. (2015) The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific*, 2:150066. <https://doi.org/10.1038/sdata.2015.66>.

Garcia, S.R. and Kayano, M.T. (2008) Climatological aspects of Hardley, Walker and monsoon circulations in two phases of the Pacific Decadal oscillation. *Theoretical and Applied Climatology*, 91(1-4):117–127. <https://doi.org/10.1007/s00704-007-0301-9>.

Gong, D. and Wang, S. (1999) Definition of Antarctic Oscillation index. *Geophysical Research Letters*, 26(4):459–462. <https://doi.org/10.1029/1999gl900003>.

Granger, C.W.J. (1969) Investigating causal relations by econometric models and crossspectral methods. *Econometrica*, The Econometric Society, 37(3):424–438. <https://doi.org/10.2307/1912791>.

Grimm, A.M. (2003) The El Niño impact on the summer monsoon in Brazil: Regional processes versus remote influences. *Journal of Climate*, 16(2):263–280. [https://doi.org/10.1175/1520-0442\(2003\)016<0263:TENIOT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<0263:TENIOT>2.0.CO;2).

Grimm, A.M. and Ambrizzi, T. (2009) Teleconnections into South America from the Tropics and Extratropics on Interannual and Intraseasonal Timescales. In: Vimeux, F., Sylvestre, F. and Khodri, M., Eds., *Past Climate Variability in South America and Surrounding Regions: From the Last Glacial Maximum to the Holocene*, Springer, Apeldoorn, 159-191. [http://doi.org/10.1007/978-90-481-2672-9\\_7](http://doi.org/10.1007/978-90-481-2672-9_7).

Grimm, A.M., Ferraz, S.T., and Gomes, J. (1998) Precipitation anomalies in Southern Brazil associated with El Niño and La Niña events. *Journal of Climate*, 11(11):2863–2880. [https://doi.org/10.1175/1520-0442\(1998\)011<2863:PAISBA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<2863:PAISBA>2.0.CO;2).

Grimm, A. M., & Silva Dias, P. L. (1995). Analysis of Tropical–Extratropical Interactions with Influence Functions of a Barotropic Model, *Journal of Atmospheric Sciences*, 52(20), 3538-3555. [https://doi.org/10.1175/1520-0469\(1995\)052<3538:AOTIWI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<3538:AOTIWI>2.0.CO;2).

Grinsted, A., Moore, J.C. and Jevrejeva, S. (2004). Application of cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, 11:561–566. <https://doi.org/10.5194/npg-11-561-2004>.

Hurrell, J.M., Kushnir, Y., Ottersen, G. and Visbeck, M. (2003) The North Atlantic Oscillation: climatic significance and environmental impact. American Geophysical Union.

Jones, C. and Carvalho, L.M.V. The influence of the Atlantic multidecadal oscillation on the eastern andes low-level jet and precipitation in South America. *npj Clim Atmos Sci*, 1(40). <https://doi.org/10.1038/s41612-018-0050-8>.

Kaiser H.F. (1958) The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3):187–200. <https://doi.org/10.1007/BF02289233>.

Kayano, M.T. and Andreoli, R.V. (2004) Decadal variability of northern northeast Brazil rainfall and its relation to tropical sea surface temperature and global sea level pressure anomalies. *Journal of Geophysical Research: Oceans*, 109(C11). <https://doi.org/10.1029/2004JC002429>.

Kayano, M.T. and Capistrano, V.B. (2014) How the Atlantic Multidecadal Oscillation (AMO) modifies the ENSO influence on the South American rainfall. *International Journal of Climatology*, Reading, 34(1):162–178. <https://doi.org/10.1002/joc.3674>.

Kayano, M.T. and Andreoli, R.V. (2009) Variabilidade Decenal a Multidecenal. In I.F.A. Cavalcanti, N.J. Ferreira, M.G.A.J. Silva e M.A.F. Silva Dias, editors, *Tempo e Clima no Brasil*, páginas 375–383. São paulo: Oficina de textos edição.

Kodama, Y.M. (1992) Large-scale common features of sub-tropical precipitation zones (the Baiu Frontal Zone, the SPCZ, and the SACZ). Part I: Characteristics of subtropical frontal zones. *J. Meteor. Soc.*, 70:813–835. [https://doi.org/10.2151/jmsj1965.70.4\\_813](https://doi.org/10.2151/jmsj1965.70.4_813).

Kousky, V.E. and Kayano, M.T. (1994) Principal modes of outgoing longwave radiation and 250-

mb circulation for the South American sector. *Journal of Climate*, 7:1131–1143. [https://doi.org/10.1175/1520-0442\(1994\)007<1131:PMOOLR>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<1131:PMOOLR>2.0.CO;2).

Kouadio, Y.K., Servain, J., Machado, L.A.T., and Lentinu, C.A.D. (2012) Heavy rainfall episodes in the eastern Northeast Brazil linked to large-scale ocean-atmosphere conditions in the tropical Atlantic. *Advances in Meteorology*, 2012:1–16. <https://doi.org/10.1155/2012/369567>.

Kumar, A., Chen, M. and Wang., W. (2011) An analysis of prediction skill of monthly mean climate variability. *Climate Dynamics*, 37:1119–1131. <https://doi.org/10.1007/s00382-010-0901-4>.

Li, T., Calvo, N., Yue, J., Russell III, J.M., Smith, A.K., Martin, G., Mlynczak, G., Chandran, A., Dou X. and Liu. A.Z. (2016) Southern hemisphere summer mesopause responses to El Niño–Southern oscillation. *Journal of Climate*, 29(17):6319–6328. <https://doi.org/10.1175/JCLI-D-15-0816.1>.

Lütkepohl, H. (2005) *New Introduction to Multiple Time Series Analysis*. Springer- Verlag Berlin Heidelberg.

Lúcio, P.S., Silva, F.D., Fortes L.T.G. and Santos, L.A.R. (2010) Um modelo estocástico combinado de previsão sazonal para a precipitação no Brasil. *Revista Brasileira de Meteorologia*, 25(1):70–87. <https://doi.org/10.1590/S0102-77862010000100007>.

Machado, C.B., Campos, T.L.O.B., Abou Rafee, S.A., Martins, J.A., Grimm, A.M. and Freitas, E.D. (2021) Extreme rainfall events in the macro-metropolis of São Paulo: trends and connection with climate oscillations. *Journal of Applied Meteorology and Climatology*, 60(4). <https://doi.org/10.1175/JAMC-D-20-0173.1>.

Madden, R.A. and Julian, P.R. (1972) Description of global-scale circulation cells in the tropics with a 40-50 day period. *Journal of the Atmospheric Sciences*, 29(6):1109–1123. [https://doi.org/10.1175/1520-0469\(1972\)029<1109:DOGSCC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2).

Madden, R.A. and Julian, P.R. (1994) Observations of the 40-50 day tropical oscillation – A Review. *Monthly Weather Review*, 122(5):814–837. [https://doi.org/10.1175/1520-0493\(1994\)122<0814:OOTDTP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122<0814:OOTDTP>2.0.CO;2).

Malfatti, M.G.L., Cardoso, A.O. and Hamburger, D.S. (2018a) Identificação de regiões pluviométricas homogêneas na bacia hidrográfica do rio Paraná. *Geociências*, 37(2):409–421. <https://doi.org/10.5016/geociencias.v37i2.11564>.

Malfatti, M.G.L., Cardoso, A.O. and Hamburger, D.S. (2018b) Modelo empírico linear para previsão de vazão de rios na usina hidrelétrica de Itaipu – bacia do rio Paraná. *Revista Brasileira de Meteorologia*, 33(2):257–268. <https://doi.org/10.1590/0102-7786332005>.

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M. and Francis, R.C. (1997) A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of American Meteorological Society*, 78(6):1069–1079. [https://doi.org/10.1175/1520-0477\(1997\)078<1069:APICOW>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2).

Marengo J.A. and Nobre A.C. (2009). Clima da região amazônica. In I.F.A. Cavalcanti, N.J. Ferreira, M.G.A.J. Silva e M.A.F. Silva Dias, editors, *Tempo e Clima no Brasil*. Oficina de textos: São Paulo. Marengo, J.A., Liebmann, B., Grimm, A.M., Misra, V., Silva Dias, P.L., Cavalcanti, I.F.A., Carvalho, L.M.V., Berbery, E.H., Ambrizzi, T., Vera, Saulo, C.S., A.C., Nogue Paegle, J., Zipser, E., Sethk, A. andAlvese, L.M. (2012) Review - Recent developments on the South American



monsoon system. *International Journal of Climatology*, 32:1–21.

Mayta, V.C., Silva, N.P., Ambrizzi, T., Silva Dias, P.L., and Espinoza, J.C. (2020) Assessing the skill of all-season diverse Madden–Julian oscillation indices for the intraseasonal Amazon precipitation. *Climate Dynamics*, 54:3729–3749. <https://doi.org/10.1007/s00382-020-05202-9>.

Min, Y., Ham, S., Yoo J. and Han. S. (2020) Recent progress and future prospects of subseasonal and seasonal climate predictions. *Bulletin of the American Meteorological Society*, 101(5):E640–E644. <https://doi.org/10.1175/BAMS-D-19-0300.1>.

Mo, K.C. (2020) Relationship between low-frequency variability in the Southern Hemisphere and sea surface temperature anomalies. *Journal of Climate*, 13:3599–3610. [https://doi.org/10.1175/1520-0442\(2000\)013<3599:RBLFVI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<3599:RBLFVI>2.0.CO;2).

Mo, K.C. and Nogués-Paegle, J. (2001) The Pacific-South American modes and their downstream effects. *Int. J. Climatol.*, 21, 1211– 1229. <https://doi.org/10.1002/joc.685>.

Molion, L.C.B. (2008) Perspectivas climáticas para os próximos 20 anos. *Revista Brasileira de Climatologia*, 3(1):117–128. <http://doi.org/10.5380/abclima.v3i0.25438>.

Moura, A.D. and Shukla., J. (1981) On the dynamics of droughts in Northeast Brazil: observations, theory and numerical experiments with a general circulation model. *Journal of the Atmospheric Sciences*, 38:2653–2675 [https://doi.org/10.1175/1520-0469\(1981\)038<2653:OTDODI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1981)038<2653:OTDODI>2.0.CO;2).

Murphy, M. (1988) The impact of the ensemble forecast on predictability. *Quarterly Journal of the Royal Meteorological Society*, 114(480):463–493. <https://doi.org/10.1002/qj.49711448010>.

Naumann, G., and Vargas, W.M. (2010) Joint diagnostic of the surface air temperature in southern South America and the Madden–Julian oscillation. *Weather Forecasting*, 25, 1275–1280. <https://doi.org/10.1175/2010WAF2222418.1>.

Newman, M., Alexander, M. A. and Scott, J. D. (2011) An empirical model of tropical ocean dynamics. *Climate Dynamics*, 37:1823–1841. <https://doi.org/10.1007/s00382-011-1034-0>.

Nogueira, S.M.C., Moreira, M.A. and Volparto, M.M.L. (2018) Evaluating precipitation estimates from Eta, TRMM and CHIRPS data in the South-Southeast region of Minas Gerais state - Brazil. *Remote Sensing*, 10(2):1–16. <https://doi.org/10.3390/rs10020313>.

North, G.R., Bell, T.L., Cahalan, R.F. and Moeng, F.J. Sampling Errors in the Estimation of Empirical Orthogonal Functions. *Monthly Weather Review*, 110(7):699-706. [https://doi.org/10.1175/1520-0493\(1982\)110<0699:SEITEO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2).

Oliveira-Júnior, J.F., da Silva Junior, C.A., Teodoro, P.E., Rossi, F.S., Blanco, C.J.C., Lima, M., de Gois G., Correia Filho, W.L.F., Santiago, D.B. and dos Santos Vanderley, M.H.G. (2021) Confronting CHIRPS dataset and in situ stations in the detection of wet and drought conditions in the Brazilian Midwest. *Int J Climatol.*, 41(9):4478-4493. <https://doi.org/10.1002/joc.7080>.

Paredes-Trejo, F.J., Barbosa, H.A. and Lakshimi Kumar, T.V. (2017) Validating chirps-based satellite precipitation estimates in Northeast Brazil. *Journal of Arid Environments*, 139:26–40. <https://doi.org/10.1016/j.jaridenv.2016.12.009>.

Prado, L.F., Wainer, I., Yokoyama, E., Khodri, M., and Garnier, J. (2021) Changes in summer

precipitation variability in Central Brazil over the past eight decades. *International Journal of Climatology*, 41(8):4171-4186. <https://doi.org/10.1002/joc.7065>.

Pinheiro, E., Rocha, R.P. da and Drumond, A. (2021) Assessment of 20th-century reanalysis circulation patterns associated with El Niño–Southern oscillation impacts on the tropical Atlantic and Northeastern Brazil rainy season. *Int J Climatol.*, 41(7):3824– 3840. <https://doi.org/10.1002/joc.6750>.

Ramirez, E., Silva Dias, P.L. and Raupp, C.F.M. (2018) Multiscale atmosphere–ocean interactions and the low-frequency variability in the equatorial region. *Journal of the Atmospheric Sciences*, 74(8):2503–2523. <https://doi.org/10.1175/JAS-D-15-0325.1>.

Rebello, E.R.G. (2006) A Oscilação Decadal do Pacífico e sua possível influência no regime de precipitação do Rio Grande do Sul. Em Congresso Brasileiro de Meteorologia, 14, Florianópolis. Anais... Rio de Janeiro: SBMET.

Reboita, M.S., Gan, M.A., Rocha, R.P. and Ambrizzi, T. (2010) Regimes de precipitação na América do Sul: Uma revisão bibliográfica. *Revista Brasileira de Meteorologia*, 25:185–204. <https://doi.org/10.1590/S0102-77862010000200004>.

Rodrigues, R.R., Haarsma, R.J., Campos, E.J.D. and Ambrizzi, T. (2011) The impacts of inter-El Niño variability on the tropical Atlantic and Northeast Brazil climate. *Journal of Climate*, 24(13):3402–3422. <https://doi.org/10.1175/2011JCLI3983.1>.

Satyamurty, P., Nobre, C.A. and Silva Dias, P.L. (1998) Tropics South America. In Karoly e Vicent, editors, *Meteorology of the Southern Hemisphere*, Boston.

Silva, A.E. and L.M.V Carvalho, L.M.V. (2007) Large-scale index for South America monsoon (LISAM). *Atmospheric Science Letters*, 8:51–57. <https://doi.org/10.1002/asl.150>.

Souza, E.B. and Ambrizzi, T. (2006) Modulation of the intraseasonal rainfall over tropical Brazil by the Madden-Julian oscillation. *Int. J. Climatol.*, 26:1759–1776. <https://doi.org/10.1002/joc.1331>.

Sousa, A.J.S, Ferreira, E.A., Sousa, J.R.A. and Rolim, P.A.M. (2010) Oscilação Decadal do Pacífico e sua influência na precipitação de Goiânia – GO. Em Congresso Brasileiro de Meteorologia, 16, Belém. Anais... Rio de Janeiro: SBMET.

Souza, I.P. de, Andreoli, R.V., Kayano, M.T., Vagas, F.F., Cerón, W.L., Martins, J.A., Freitas, E. and Souza, R.A.F. de. (2021) Seasonal precipitation variability modes over South America associated to El Niño–Southern Oscillation (ENSO) and non-ENSO components during the 1951–2016 period. *Int. J. Climatol.*, 41(18):4321-4338. <https://doi.org/10.1002/joc.7075>.

Stoica, P. and Moses, R. L. (2005) *Spectral Analysis of Signals*, Pearson Prentice Hall, New Jersey. Tedeschi, R.G., Grimm, A.M. and Cavalcanti, I.F. (2016) Influence of central and east ENSO on precipitation and its extreme events in South America during austral autumn and winter. *International Journal of Climatology*, 36:4797–4814. <https://doi.org/10.1002/joc.4670>.

Torrence, C. and Compo, G.P. (1998) A practical guide to wavelet analysis. *Bulletin of American Meteorological Society*, 79(1):61–78. [https://doi.org/10.1175/1520-0477\(1998\)079<0061:APGTWA>2.0.CO;2](https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2).

Thomson, R.E. and Emery, W.J. Chapter 6 - digital filters. (2014) In *Data Analysis Methods*

in *Physical Oceanography*. Elsevier, Boston. Tracton, M.S. and E. Kalnay, E. (1993) Operational ensemble prediction at the national meteorological center: Practical aspects. *Weather and Forecasting*, 8(3):379–398. [https://doi.org/10.1175/1520-0434\(1993\)008<0379:OEPATN>2.0.CO;2](https://doi.org/10.1175/1520-0434(1993)008<0379:OEPATN>2.0.CO;2).

Trenberth, K.E. and Hurrell, J.W. (1994) Decadal atmosphere-ocean variations in the Pacific. *Climate Dynamics*, 9:303–319. <https://doi.org/10.1007/BF00204745>.

Vasconcellos, F.C. and Cavalcanti, I.F.A. (2010) Extreme precipitation over Southeastern Brazil in the austral summer and relations with the southern hemisphere annular mode. *Atmospheric Science Letters*, 11(1):21–26. <https://doi.org/10.1002/asl.247>.

Villamayor, J., Ambrizzi, T., and Mohino, E. (2018). Influence of decadal sea surface temperature variability on northern Brazil rainfall in CMIP5 simulations. *Clim. Dyn.* 51, 563–579. <https://doi.org/10.1007/s00382-017-3941-1>.

Vera, C., Higgins, W., Amador, J., Ambrizzi, T., Garreaud, R., Gochis, D., Gutzler, D., Lettenmaier, D., Marengo, J., Mechoso, C. R., Nogues-Paegle, J., Dias, P. L. S., and Zhang, C. (2006). Toward a Unified View of the American Monsoon Systems, *Journal of Climate*, 19(20), 4977–5000. <https://doi.org/10.1175/JCLI3896.1>.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M., Hendon, H., Hodgson, J., Kang, H., Kumar, A., Lin, H. Liu, G. Liu, X., Malguzzi, P., Mallas, I., Manoussakis, M., Mastrangelo, D., MacLachlan, C., McLean, P., Minami, A., Mladek, R., Nakazawa, T., Najm, S., Nie, Y., Rixen, M., Robertson, A.W., Ruti, P., Sun, C., Takaya, Y., Tolstykh, M., Venuti, F., Waliser, D., Woolnough, S., Wu, T., Won, D.-J., Xiao, H., Zaripov, R. and Zhang, L. (2017) The subseasonal to seasonal (s2s) prediction project database. *Amer. Meteor. Soc.*, 98:163–173. <https://doi.org/10.1175/BAMS-D-16-0017.1>.

Von Sperling, M. (2007) Von Sperling. Estudo e modelagem da qualidade da água de rios. Editora UFMG, Belo Horizonte.

Wainer, I. and Venegas, S.A. (2002). South Atlantic multidecadal variability in the climate system model. *Journal of Climate*, 15(12):1408–1420. [https://doi.org/10.1175/1520-0442\(2002\)015<1408:SAMVIT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1408:SAMVIT>2.0.CO;2).

Wheeler, M.C. and Hendon, H.H. (2004) An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Monthly Weather Review*, 132(8):1917–1932. [https://doi.org/10.1175/1520-0493\(2004\)132<1917:AARMMI>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2).

White, C.J., Carlsen, H., Robertson, A.W., Klein, R.J.T., Lazo, J.K., Kumar, A., Vitart, F., de Perez, E.C., Ray, A.J., Murray, V., Bharwani, S., MacLeod, D., James, R., Fleming, L., Morse, A.P., Eggen, B., Graham, R., Kjellstrom, E., Becker, E., Pegion, K.V., Holbrook, N.J., McEvoy, D., Depledge, M., Perkins-Kirkpatrick, S., Brown, T.J., Street, R., Jones, L., Remenyi, T., Hodgson-Johnston, I., Buontempo, C., Lamb, R., Meinke, H., Arheimer, B. and Zebiak, S.E. (2017) Potential applications of subseasonal-to-seasonal (s2s) predictions. *Meteorol. Appl.*, 24:315–325. <http://doi.org/10.1002/met.1654>.

Wolter, K. and Timlin M.S. (1998). Measuring the strength of ENSO events: How does 1997/98 rank? *Weather*, 53(9):315–324. <https://doi.org/10.1002/j.1477-8696.1998.tb06408.x>.

Zang, Z. and J.C. Moore, J.C. (2015) Empirical orthogonal functions. In Altun Y. et al., editor, *Mathematical and Physical Fundamentals of Climate Change*. Elsevier.

# Referências Bibliográficas

- Andreoli e Kayano(2005)** R.V. Andreoli e M.T. Kayano. Enso-related rainfall anomalies in south america and associated circulation features during warm and cold pacific decadal oscillation regimes. *International Journal of Climatology*, 25(15):2017–2030. Citado na pág. 17
- Andreoli e Kayano(2006)** R.V. Andreoli e M.T. Kayano. Tropical pacific and south atlantic effects on rainfall variability over northeast brazil. *International Journal of Climatology*, 26(13): 1895–1912. Citado na pág. 50, 58
- Arias et al.(2015)** P.A. Arias, R. Fu, C. Vera e M. Rojas. A correlated shortening of the north and south american monsoon seasons in the past few decades. *Climate Dynamics*, 45(1):3183–3203. Citado na pág. 34
- Baccalá e Sameshima(2001)** A. Baccalá e K. Sameshima. Partial directed coherence: A new concept in neural structure determination. *Biological Cybernetics*, 84(6):463–474. Citado na pág. 29, 83
- Baccalá e Sameshima(2021)** L.A. Baccalá e K Sameshima. Partial directed coherence: twenty years on some history and an appraisal. *Biol Cybern*, 115:195–204. Citado na pág. 29, 41
- Baccalá et al.(2013)** L.A. Baccalá, C.S.N Brito, D.Y. Takahashi e K. Sameshima. Unified asymptotic theory for all partial directed coherence forms. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1997):20120158. Citado na pág. xiv, xv, 35, 36, 47, 48
- Baldwin et al.(2001)** M.P. Baldwin, L.J. Gray, T.J. Dunkerton, K. Hamilton, P.H. Haynes, W.J. Randel, J.R. Holton, M.J. Alexander, I. Hirota, T. Horinouchi, D.B.A. Jones, J.S. Kinnersley, C. Marquardt, K. Sato e M. Takahashi. The quasi-biennial oscillation. *Reviews of Geophysics*, 39(2):179–229. Citado na pág. 22, 26
- Barry e Chorley(2013)** R.G. Barry e R.J. Chorley. *Atmosfera, Tempo e Clima*. Bookman, 9<sup>o</sup> edição. p.512. Citado na pág. 21, 34
- Bayer et al.(2009)** J. Bayer, D. Wierstra, J. Togelius e J. Schmidhuber. Evolving memory cell structures for sequence learning. Em *Artificial Neural Networks - ICANN 2009, Anais ... Berlin, Heidelberg: Springer Berlin Heidelberg*, página 755–764. Citado na pág. 90
- Bergstra e Bengio(2012)** J. Bergstra e Y. Bengio. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13:281–305. Citado na pág. 60
- Bodnariuk et al.(2021)** N. Bodnariuk, C.G. Simionato, M. Osman e M. Saraceno. The río de la plata plume dynamics over the southwestern atlantic continental shelf and its link with the large scale atmospheric variability on interannual timescales. *Continental Shelf Research*, 212(1):1–11. Citado na pág. 20

- Bombardi e Carvalho(2011)** R.J. Bombardi e L.M.V Carvalho. The south atlantic dipole and variations in the characteristics of the south american monsoon in the wcrp-cmip3 multi-model simulations. *Climate Dynamics*, 36(11):2091–2102. Citado na pág. 19, 34
- Bombardi et al.(2014)** R.J. Bombardi, L.M.V Carvalho, C. Jones e M.S. Reboita. Precipitation over eastern south america and the south atlantic sea surface temperature during neutral enso periods. *Climate Dynamicss*, 42(5-6):1553–1568. Citado na pág. 12, 19, 33
- Boor(2001)** C. Boor. *A Practical Guide to Splines*. Springer-Verlag New York Inc., 3º edição. p.392. Citado na pág. 41, 55, 66
- Brun e Tucci(2001)** G.W. Brun e C.E.M. Tucci. Previsão em tempo real do volume afluyente ao reservatório de ernestina. *Revista Brasileira de Recursos Hídricos*, 6(2):73–79. Citado na pág. 5
- Builes-Jaramillo et al.(2018)** A. Builes-Jaramillo, A.M.T. Ramos e G. Poveda. Atmosphere-land bridge between the pacific and tropical north atlantic sst's through the amazon river basin during the 2005 and 2010 droughts. *Chaos*, 28:085702. Citado na pág. 3, 35
- Camargo e Sobel(2010)** S.J. Camargo e A.H. Sobel. Revisiting the influence of the quasi-biennial oscillation on tropical cyclone activity. *Journal of Climate*, 23(21):5810–5825. Citado na pág. 22
- Capozzoli et al.(2017)** C.R. Capozzoli, A.O. Cardoso e S.E.T. Ferraz. Padrões de variabilidade de vazão de rios nas principais bacias brasileiras e associação com Índices climáticos. *Revista Brasileira de Meteorologia*, 32(2):243–254. Citado na pág. 39
- Cardoso e Wainer(2021)** A. Cardoso e I. Wainer. Modes of variability in the tropical atlantic and its influences on the precipitation regime in brazil. Em *EGU General Assembly 2021, online Anais... EGU21-15509*. Citado na pág. 19
- Cardoso e Cataldi(2012)** A.O. Cardoso e M. Cataldi. Relações de índices climáticos e vazão de rios no brasil. Em *Congresso Brasileiro de Meteorologia, 17., 2012, Gramado. Anais ... Rio de Janeiro: SBMET*. Citado na pág. 11
- Cardoso e Silva Dias(2004)** A.O. Cardoso e P.L. Silva Dias. Identificação de trimestres extremos no regime pluviométrico do sul e sudeste do brasil e relação com anomalias da tsm. *Revista Brasileira de Meteorologia*, 19(2):149–162. Citado na pág. 18, 60
- Carvalho e Cavalcanti(2016)** L.M.V. Carvalho e I.F.A. Cavalcanti. The south american monsoon system (smas). Em L.M.V. Carvalho e c. Jones, editors, *The Monsoons and Climate Change: Observations and Modeling*, páginas 121–148. Springer climate edição. Citado na pág. 11, 12
- Carvalho et al.(2004)** L.M.V. Carvalho, C. Jones e B. Liebmann. The south atlantic convergence zone: Intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall. *Journal of Climate*, 17(1):88–108. Citado na pág. 11, 12, 49
- Carvalho et al.(2005)** L.M.V. Carvalho, J. Charles e T. Ambrizzi. Opposite phases of the antarctic oscillation and relationships with intraseasonal to interannual activity in the tropics during the austral summer. *Journal of Climate*, 18:702–718. Citado na pág. 21
- Casarin e Kousky(1986)** D.P. Casarin e V.E. Kousky. Anomalias de precipitação no sul do brasil e variações na circulação atmosférica. *Revista Brasileira. Meteorologia*, 1:83–90. Citado na pág. 13
- Cavalcanti e Shiumizu(2012)** I.F.A. Cavalcanti e M.H. Shiumizu. Climate fields over south america and variability of sacz and psa in hadgem2-es. *American Journal of Climate Change*, 1 (3):132–144. Citado na pág. 50, 58

- Cavalcanti et al.(2002)** I.F.A. Cavalcanti, J.A. Marengo, P. Satyamurty, C.A. Nobre, I. Trosnikov, J.P. Bonatti, A.O. Manzi, T. Tarasova, L.P. Pezzi, C. D’Almeida, G. Sampaio, C.C. Castro, M.B. Sanches e H. Camargo. Global climatological features in a simulation using cptec/cola agcm. *Journal of Climate*, 15:2965–2988. Citado na pág. 6
- Chaves(2011)** R.R. Chaves. Conexões entre a tsm do atlântico sul e a convecção de verão sobre a américa do sul – análise observacional. *Revista Brasileira de Geofísica*, 29(1):5–10. Citado na pág. 18, 33, 50, 60
- Chaves e Nobre(2004)** R.R. Chaves e P. Nobre. Interactions between sea surface temperature over the south atlantic ocean and the south atlantic convergence zone. *Geophysical Research Letters*, 31(L3204):1–4. Citado na pág. 12
- Chiang e Vimont(2004)** J.H. Chiang e D.J. Vimont. Analogous pacific and atlantic meridional modes of tropical atmosphere-ocean variability. *Journal of Climate*, 17(21):4143–4158. Citado na pág. 19, 26, 33
- Chiessi et al.(2009)** C.M. Chiessi, S. Mulitza, J. Pätzold, G. Wefer e J.A. Marengo. Possible impact of the atlantic multidecadal oscillation on the south american summer monsoon. *Geophysical Research Letters*, 36(21):1–5. Citado na pág. 12, 19, 34
- Chow et al.(1994)** V.T. Chow, D.R.M. Maidment, L.W. Saldarriaga e G. Juan. *Hidrología aplicada*. Santafé de Bogotá: McGraw-Hill Interamericana, 1º edição. p.584. Citado na pág. 3
- Coelho et al.(2002)** C.A.S. Coelho, C.B. UVO e T. Ambrizzi. Exploring the impacts of the tropical pacific sst on the precipitation patterns over south america during enso periods. *Theoretical and Applied Climatology*, 71(3):185–197. Citado na pág. 15
- Collischonn et al.(2005)** W. Collischonn, C.E.M. Tucci, R.T. Clarke, P.L. Silva Dias e G.S. Oliveira. Previsão sazonal de vazão na bacia do rio uruguai 2: Previsão climática-hidrológica. *Revista Brasileira de Recursos Hídricos*, 10(4):61–72. Citado na pág. 5
- Collischonn et al.(2007)** W. Collischonn, C.E.M. Tucci, R.T. Clarke, M.D. Corbo, B. C. Silva, B. Collischonn, D. Allasia e A.R. Paz. Modelo hidrológico distribuído para previsão de vazões incrementais na bacia do rio paranaíba entre itumbiara e são simão. *Revista Brasileira de Recursos Hídricos*, 12(3):43–56. Citado na pág. 5
- Correia Filho et al.(2021)** W.L.F. Correia Filho, J.F. de Oliveira-Júnior, C.A. da Silva Junior e D.d.B. Santiago. Influence of the el niño-southern oscillation (enso) and the synoptic systems on the rainfall variability over the brazilian cerrado via chirps data. *Int J Climatol.*, Accepted Author Manuscript. doi: <https://doi.org/10.1002/joc.7417>. Citado na pág. 15
- Costa et al.(2019)** J.C. Costa, G. Pereira, M.E. Siqueira, F.S. Cardozo e V.V. da Silva. Validação dos dados de precipitação estimados pelo chirps para o brasil. *Revista Brasileira de Climatologia*, 24:228–243. Citado na pág. 41
- Deser(2000)** C. Deser. On the teleconnectivity of the “artic oscillation”. *Geophysical Research Letters*, 27(6):779–782. Citado na pág. 21, 34
- Dettinger et al.(2000)** M.D. Dettinger, D.R. Cayan, G.J. McCabe e J.A. Marengo. Multiscale streamflow variability associated with el nino/southern oscillation. Em H.F. Diaz e V. Markgraf, editors, *El Nino and the Southern Oscillation - Multiscale Variability, Global and Regional Impacts*, páginas 114–147. Cambridge: University press edição. Citado na pág. 15, 16
- Devore(2014)** J.L. Devore. *Probabilidade e Estatística para Engenharia e Ciência*. Cengage Learning, 8º edição. p.693. Citado na pág. xviii, 79, 92

- Donoho e Johnstone(1994)** D.L. Donoho e I.M. Johnstone. Ideal denoising in an orthonormal basis chosen from a library of bases. *Comptes rendus de l'Académie des sciences. Série I, Mathématique*, 319(12):1317–1322. Citado na pág. [29](#)
- Draper e Smith(1998)** N.R. Draper e H. Smith. *Applied regression analysis*. Nova Jersey: Wiley, 3º edição. p.706. Citado na pág. [55](#), [85](#)
- ECMWF(2020)** ECMWF. *IFS Documentation CY47R1*. IFS Documentation. ECMWF. Citado na pág. [65](#)
- Enfield e Mayer(1997)** D.B. Enfield e D.A. Mayer. Tropical atlantic sea surface temperature variability and its relation to el niño-southern oscillation. *Journal of Geophysical Research: Oceans*, 102(1):929–945. Citado na pág. [17](#)
- Enfield et al.(1999)** D.B. Enfield, A.M. Mestas-Núñez e D.A. Mayer. How ubiquitous is the dipole relationship in tropical atlantic sea surface temperatures? *Journal of Geophysical Research: Oceans*, 104(C4):7841–1848. Citado na pág. [17](#), [26](#), [33](#)
- Enfield et al.(2001)** D.B. Enfield, A.M. Mestas-Núñez e P.J. Trimble. The atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental us. *Geophysical Research Letters*, 28(10):2077–2080. Citado na pág. [18](#), [26](#), [33](#)
- Farge(1994)** M. Farge. Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics*, 51:2523–2541. Citado na pág. [77](#)
- Ferreira et al.(2004)** N.J. Ferreira, M. Sanches e M.A.F. Silva Dias. Composição da zona de convergência do atlântico sul em períodos de el niño e la niña. *Revista Brasileira de Meteorologia*, 19(1):89–98. Citado na pág. [11](#), [49](#)
- Ferreira et al.(2012)** W.R.S. Ferreira, M.I. Vitorino, E.B. SOuza e A.M.C. Carmo. Sazonalidade da precipitação para a amazônia usando o modelo regcm3: Avaliando apenas forçante do atlântico equatorial. *Revista Brasileira de Meteorologia*, 27(4):435–446. Citado na pág. [22](#), [33](#)
- Foley et al.(2002)** J.A. Foley, A. Botta, M.T. Coe e M.H. Costa. El niño-southern oscillation and the climate, ecosystems and rivers of amazonia. *Global Biogeochemical Cycle*, 16(4):79(1–20). Citado na pág. [15](#), [50](#), [58](#)
- Freitas et al.(2009)** S.R. Freitas, K.M. Longo, M.A.F. Silva Dias, R. Chatfield, P.L. Silva Dias, P. Artaxo, M.O. Andreae, G. Grell, L.F. Rodrigues, A. Fazenda e J. Panetta. The coupled aerosol and tracer transport model to the brazilian developments on the regional atmospheric modeling system (cattbrams). part 1: Model description and evaluation. *Atmospheric Chemistry and Physics*, 9:2843–2861. Citado na pág. [6](#)
- Funk et al.(2015)** C. Funk, P. Peterson, M. Landsfeld, D. Pedreros, J. Verdin, S. Shukla, G. Husak, J. Rowland, L. Harrison, A. Hoell e J. Michaelsen. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific*, 2:150066. Citado na pág. [41](#)
- Garcia e Kayano(2008)** S.R. Garcia e M.T. Kayano. Climatological aspects of hardley, walker and monsoon circulations in two phases of the pacific decadal oscillation. *Theoretical and Applied Climatology*, 91(1-4):117–127. Citado na pág. [17](#), [34](#)
- García-Serrano et al.(2017)** J. García-Serrano, C. Cassou, H. Douville, A. Giannini e F.J. Doblaser-Reyes. Revisiting the enso teleconnection to the tropical norte atlantic. *Journal of Climate*, 30(17):6945–6957. Citado na pág. [35](#)

- Garreaud et al.(2009)** R.D. Garreaud, M. Vuille, R. Compagnucci e J. Marengo. Present-day south american climate. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 281(3-4):180–195. Citado na pág. 11, 33
- Glorot et al.(2011)** X. Glorot, A. Bordes e Y. Bengio. Deep sparse rectifier neural networks. Em *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, Anais ...Fort Lauderdale: AISTATS*, páginas 315–323. Citado na pág. 10, 89
- Gong e Wang(1999)** D. Gong e S. Wang. Definition of antarctic oscillation index. *Geophysical Research Letters*, 26(4):459–462. Citado na pág. 21, 26
- Goodwin et al.(2016)** B. P. Goodwin, E. Mosley-Thompson, A.B. Wilson, S.E. Porter e M.R. Sierra-Hernandez. Accumulation variability in the antarctic peninsula: The role of large-scale atmospheric oscillations and the interactions. *Journal of Climate*, 29(7):2579–2596. Citado na pág. 35
- Granger(1969)** C.W.J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, The Econometric Society*, 37(3):424–438. Citado na pág. 28
- Graves et al.(2013)** A.-r Graves, A. Mohamed e G. Hinton. Speech recognition with deep recurrent neural networks. Em *In Acoustics, speech and signal processing (ICASSP), 2013 Anais ... IEEE*, páginas 6645–6649. Citado na pág. 7
- Gray et al.(2010)** L.J. Gray, J. Beer, M. Geller, J.D. Haigh, M. Lockwood, K. Matthes, U. Cubasch, D. Fleitmann, G. Harrison, L. Hood, J. Luterbacher, G.A. Meehl, D. Shindell, B. Van Geel e W. White. Solar influences on climate. *Reviews of Geophysics*, 48(4):RG4001. Citado na pág. 22, 26, 29
- Grimm(2003)** A.M. Grimm. The el nino impact on the summer monsoon in brazil: Regional processes versus remote influences. *Journal of Climate*, 16(2):263–280. Citado na pág. 10, 15, 20, 33
- Grimm(2004)** A.M. Grimm. How do la niña events disturb the summer monsoon system in brazil? *Climate Dynamics*, 22(2-3):123–138. Citado na pág. 33
- Grimm(2011)** A.M. Grimm. Interannual climate variability in south america: impacts on seasonal precipitation, extreme events and possible effects of climate change. *Stochastic Environmental Research and Risk Assessment*, 25(4):537–554. Citado na pág. 11, 12
- Grimm e Ambrizzi(2009)** A.M. Grimm e T. Ambrizzi. Teleconnections into south america from the tropics and extratropics on interannual and intraseasonal timescales. Em F. Vimeux, F. Sylvestre e M. Khodri, editors, *Past climate variability in South America and surrounding regions*, páginas 159–191. Developments in paleoenvironmental research, springer netherlands edição. Citado na pág. 13
- Grimm e Tedeschi(2009)** A.M. Grimm e R.G. Tedeschi. Enso and extreme rainfall events in south america. *Journal of Climate*, 22(7):1589–1609. Citado na pág. 15
- Grimm et al.(1998)** A.M. Grimm, S.T. Ferraz e J. Gomes. Precipitation anomalies in southern brazil associated with el niño and la niña events. *Journal of Climate*, 11(11):2863–2880. Citado na pág. 15
- Grimm et al.(2005)** A.M. Grimm, C. Vera e C.R. Mechoso. The south american monsoon system. Em C.P Chang, B. Wang e N.C.G. Lau, editors, *The global monsoon system: research and forecast*, páginas 219–238. World meteorological organization, wmo/td edição. Citado na pág. 10



- Grinsted et al.(2004)** A. Grinsted, J.C. Moore e S. Jevrejeva. Application of cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, 11: 561–566. Citado na pág. [78](#)
- Guetter e Prates(2002)** A.K. Guetter e J.E. Prates. Degrau climático nas séries de vazões das bacias brasileiras. Em *Congresso Brasileiro de Meteorologia, 12, Foz do Iguaçu. Anais... Rio de Janeiro: SBMET*. Citado na pág. [7](#)
- Guilhon et al.(2003)** L.G.F. Guilhon, V.F. Rocha e J.C. Moreira. Comparação de métodos de previsão de vazões naturais afluentes a aproveitamentos hidroelétricos. *Revista Brasileira de Recursos Hídricos*, 12(3):13–20. Citado na pág. [5](#)
- Hannachi et al.(2007)** A. Hannachi, I.T. Jolliffe e D.B. Stephenson. Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, 27(9):1119–1152. Citado na pág. [82](#)
- Haykin(2009)** S. Haykin. *Neural networks and learning machines*. Person Education, 3ª edição. p.906. Citado na pág. [7](#), [8](#)
- He et al.(2015)** K. He, X. Zhang, S. Ren e J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. Em *Proceedings of the IEEE International Conference on Computer Vision, Anais ... Santiago: ICCV*, páginas 1026–1034. Citado na pág. [10](#)
- Hill et al.(2011)** K.J Hill, A.S. Taschetto e M.H. England. Sensitivity of south american summer rainfall to tropical pacific ocean sst anomalies. *Geophys. Res. Lett.*, 38(L01701):1–6. Citado na pág. [20](#)
- Hochreiter e Schmihscher(1997)** S. Hochreiter e J. Schmihscher. Long short-term memory. *Neural Computation*, 9(8):1735–1780. Citado na pág. [55](#), [56](#), [89](#)
- Hurrell et al.(2003)** J.W. Hurrell, Y. Kushnir, G. Ottersen e M. Visbeck. *The North Atlantic Oscillation: climatic significance and environmental impact*. American Geophysical Union, 1ª edição. p.279. Citado na pág. [20](#), [26](#), [34](#)
- Jacobi et al.(2015)** P.R. Jacobi, L. Giantti e Ambrizzi. Interdisciplinaridade e mudanças climáticas: caminhos de reflexão para a sustentabilidade. Em A. Philippi e V. Fernandes, editors, *Práticas da Interdisciplinaridade no Ensino e Pesquisa*, página Capítulo 13. 1. ed. são paulo: Manole edição. Citado na pág. [39](#)
- Jones e Carvalho(2002)** C. Jones e L.M.V. Carvalho. Active and break phases in the south america monsoon system. *Journal of Climate*, 15(8):905–914. Citado na pág. [11](#), [12](#)
- Jones e Carvalho(2018)** C. Jones e L.M.V. Carvalho. The influence of the atlantic multidecadal oscillation on the eastern andes low-level jet and precipitation in south america. *npj Clim Atmos Sci*, 1(40). Citado na pág. [50](#), [58](#)
- Jozefowicz et al.(2015)** R. Jozefowicz, W. Zaremba e I. Sutskever. An empirical exploration of recurrent network architectures. Em *Proceedings of the 32Nd International Conference on Machine Learning (ICML), 15 Anais ... JMLR*, página 2342–2350. Citado na pág. [90](#)
- Kaiser(1958)** H.F. Kaiser. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3):187–200. Citado na pág. [42](#), [82](#)
- Kayano e Andreoli(2009)** M.T. Kayano e R.V. Andreoli. Variabilidade decenal a multidecenal. Em I.F.A. Cavalcanti, N.J. Ferreira, M.G.A.J. Silva e M.A.F. Silva Dias, editors, *Tempo e Clima no Brasil*, páginas 375–383. São paulo: Oficina de textos edição. Citado na pág. [17](#)

- Kayano e Capistrano(2014)** M.T. Kayano e V.B. Capistrano. How the atlantic multidecadal oscillation (amo) modifies the enso influence on the south american rainfall. *International Journal of Climatology, Reading*, 34(1):162–178. Citado na pág. 19
- Kayano et al.(2009)** M.T. Kayano, C.P. Oliveira e R.V. Andreoli. Interannual relations between south american rainfall and tropical sea surface temperature anomalies before and after 1976. *International Journal of Climatology*, 29(10):1439–1448. Citado na pág. 17, 34
- Kousky e Kayano(1994)** V.E. Kousky e M.T. Kayano. Principal modes of outgoing longwave radiation and 250-mb circulation for the south american sector. *Journal of Climate*, 7:1131–1143. Citado na pág. 13, 49, 57, 58
- Kumar et al.(2011)** A. Kumar, M. Chen e W. Wang. An analysis of prediction skill of monthly mean climate variability. *Climate Dynamics*, 37:1119–1131. Citado na pág. 2
- LeCun et al.(1986)** Y. LeCun, Y. Bengio e G. Hinton. Deep learning. *Nature*, 521:436–444. Citado na pág. 9
- Li et al.(2016)** T. Li, N. Calvo, J. Yue, J.M. Russell III, A.K. Smith, G. Martin, G. Mlynchzak, A. Chandran, X. Dou e A.Z. Liu. Southern hemisphere summer mesopause responses to el niño–southern oscillation. *Journal of Climate*, 29(17):6319–6328. Citado na pág. 22
- Lihua e Zhiqiang(2017)** M. Lihua e Y. Zhiqiang. Possible solar modulation of pacific decadal oscillation. *Solar System Research*, 51(5):417–421. Citado na pág. 23, 29, 34
- Lin et al.(2016)** H. Lin, N. Gagnon, S. Beaugard, R. Muncaster, M. Markovic, B. Denis e M. Charron. Geps-based monthly prediction at the canadian meteorological centre. *Monthly Weather Review*, 144:4867–4883. Citado na pág. 2
- Lin et al.(2021)** Y.F. Lin, J.Y. Yu, C.R. Wu e F. Zheng. The footprint of the 11-year solar cycle in northeastern pacific ssts and its influence on the central pacific el niño. *Geophysical Research Letters*, 48(5):1–10. Citado na pág. 23
- Lúcio et al.(2010)** P.S. Lúcio, F.D. Silva, L.T.G Fortes e L.A.R. Santos. Um modelo estocástico combinado de previsão sazonal para a precipitação no brasil. *Revista Brasileira de Meteorologia*, 25(1):70–87. Citado na pág. 2
- Lütkepohl(2005)** H. Lütkepohl. *New Introduction to Multiple Time Series Analysis*. Springer-Verlag Berlin Heidelberg, 1º edição. p.764. Citado na pág. 83
- Machado et al.(2021)** C.B. Machado, T.L.O.B. Campos, S.A. Abou Rafee, J.A. Martins, A.M. Grimm e E.D. Freitas. Extreme rainfall events in the macro-metropolis of são paulo: trends and connection with climate oscillations. *Journal of Applied Meteorology and Climatology*, 60(4). Citado na pág. 21, 49, 57
- Madden e Julian(1972)** R.A. Madden e P.R. Julian. Description of global-scale circulation cells in the tropics with a 40-50 day period. *Journal of the Atmospheric Sciences*, 29(6):1109–1123. Citado na pág. xiii, 13, 14
- Madden e Julian(1994)** R.A. Madden e P.R. Julian. Observations of the 40-50 day tropical oscillation - a review. *Monthly Weather Review*, 122(5):814–837. Citado na pág. 13
- Madeira et al.(2018)** P.D. Madeira, C. Ramos e A. Nascimento. Influência da oscilação madden julian no comportamento da energia armazenada nos reservatórios sin. Em B. MORENO, editor, *As questões climáticas e os impactos nos negócios de energia*, páginas 9–14. Rio de janeiro: Fgv energia edição. Citado na pág. 13

- Malfatti et al.(2018a)** M.G.L. Malfatti, A.O. Cardoso e D.S. Hamburger. Identificação de regiões pluviométricas homogêneas na bacia hidrográfica do rio paran. *Geocincias*, 37(2):409–421. Citado na pg. 10, 39
- Malfatti et al.(2018b)** M.G.L. Malfatti, A.O. Cardoso e D.S. Hamburger. Modelo emprico linear para previso de vazo de rios na usina hidreltrica de itaipu – bacia do rio paran. *Revista Brasileira de Meteorologia*, 33(2):257–268. Citado na pg. 11, 57
- Malfatti et al.(2019)** M.G.L. Malfatti, L. Massaroppe e P.L. Silva Dias. Linear and nonlinear causality in monthly atmospheric and ocean time series. *Earth and Space Science Open Archive*, pgina 1. Citado na pg. 38, 73
- Mantua et al.(1997)** N.J. Mantua, S.R. Hare, Y. Zhang, J.M. Wallace e R.C. Francis. A paci-  
fic interdecadal climate oscillation with impacts on salmon production. *Bulletin of American Meteorological Society*, 78(6):1069–1079. Citado na pg. 16, 26, 33
- Marengo e Nobre(2009)** J.A. Marengo e A.C. Nobre. Clima da regio amaznica. Em I.F.A. Cavalcanti, N.J. Ferreira, M.G.A.J. Silva e M.A.F. Silva Dias, editors, *Tempo e Clima no Brasil*, pgina 197–212. So paulo: Oficina de textos edio. Citado na pg. 50
- Marengo et al.(2003)** J.A. Marengo, I.F.A. Cavalcanti, P. Satyamurty, I. Trosnikov, C.A. Nobre, J.P. Bonatti, H. Camargo, G. Sampaio, M.B. Sanches, A.O. Manzi, C.A.C. Castro, C. D’Almeida, L.P. Pezzi e L. Candido. Assessment of regional seasonal rainfall predictability using the cptec/-cola atmospheric gcm. *Climate Dynamics*, 21:459–475. Citado na pg. 6
- Marengo et al.(2012)** J.A. Marengo, B. Liebmann, A.M. Grimm, V. Misra, P.L. Silva Dias, I.F.A. Cavalcanti, L.M.V. Carvalho, E.H. Berbery, T. Ambrizzi, C.S. Vera, A.C. Saulo, J. Nogue-Paegle, E. Zipser, A. Sethk e L.M. Alvese. Review - recent developments on the south american monsoon system. *International Journal of Climatology*, 32:1–21. Citado na pg. 49
- Massaroppe e Baccal(2019)** L. Massaroppe e A. Baccal. Kernel methods for nonlinear connectivity detection. *Entropy*, 21(6):610–626. Citado na pg. 29, 83
- Massaroppe e Baccal(2015a)** L. Massaroppe e L.A. Baccal. Detecting nonlinear granger causality via the kernelization of partial directed coherence. Em *Proceedings of the 60th World Statistics Congress of the International Statistical Institute, online Anais... ISI2015. Rio de Janeiro, RJ: International Statistical Institute, The Hague, The Netherlands*. Citado na pg. 83
- Massaroppe e Baccal(2015b)** L. Massaroppe e L.A. Baccal. Kernel-nonlinear-pdc extends partial directed coherence to detecting nonlinear causal coupling. Em *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Milan: [s.n.]*, pgina 2864–2867. Citado na pg. 83
- Matthes et al.(2013)** K. Matthes, K. Kodera, R.R. Garcia, Y. Kuroda, D. Marsh e K. Labitzke. The importance of time-varying forcing for qbo modulation of the atmospheric 11 year solar cycle signal. *Journal of Geophysical Research: Atmospheres*, 118(10):4435–4447. Citado na pg. 23, 33
- Mayta et al.(2020)** V.C. Mayta, N.P. Silva, T. Ambrizzi, P.L. Silva Dias e J.C. Espinoza. Assessing the skill of all-season diverse madden–julian oscillation indices for the intraseasonal amazon precipitation. *Climate Dynamics*, 54:3729–3749. Citado na pg. 50, 58, 73
- Medina e Nunes(2011)** B.L. Medina e A.B. Nunes. Oscilao quase-bienal em fase neutra de enso e sua influncia na amrica do sul: Estudo de caso. Em *Encontro Sul-Brasileiro de Meteorologia, 4, Pelotas. Anais... Pelotas: UFPEL*. Citado na pg. 22

- Mendes e Cavalcanti(2014)** M.C.D. Mendes e I.F.A. Cavalcanti. The relationship between the antarctic oscillation and blocking events over the south pacific and atlantic oceans. *Int. J. Climatol.*, 34:529–544. Citado na pág. 21
- Min et al.(2020)** Y. Min, S. Ham, J. Yoo e S. Han. Recent progress and future prospects of subseasonal and seasonal climate predictions. *Bulletin of the American Meteorological Society*, 101(5):E640–E644. Citado na pág. 2
- Mo(2000)** K.C. Mo. Relationship between low-frequency variability in the southern hemisphere and sea surface temperature anomalies. *Journal of Climate*, 13:3599–3610. Citado na pág. 20, 21, 25, 26, 27
- Mo e Higgins(1998)** K.C. Mo e R.W. Higgins. The pacific–south american modes and tropical convection during the southern hemisphere winter. *Monthly Weather Review*, 126:1581–1596. Citado na pág. 20
- Molion(2008)** L.C.B Molion. Perspectivas climáticas para os próximos 20 anos. *Revista Brasileira de Climatologia*, 3(1):117–128. Citado na pág. 16, 50, 58
- Moquet et al.(2013)** J.S. Moquet, F.W. Cruz, J. Apaestegui, V. Novello, N. Strikis, H. Cheng, L. Edwards, J.L. Guyot e A. Sifeddine. Variabilidade multidecadal do sistema de monções de verão durante os últimos 2000 anos na região centro-oeste do brasil (goiás) com base em registros de espeleotemas. Em *Congresso da Associação Brasileira de Estudos do Quaternário, 14, Natal. Anais... São Paulo: ABEQUA*. Citado na pág. 18
- Morettin(1999)** P.A. Morettin. *Ondas e Ondaletas*. EdUSP, 2ª edição. p.320. Citado na pág. 77
- Mori e Watanabe(2008)** M. Mori e M. Watanabe. The growth and triggering mechanisms of the pna: A mjo-pna coherence. *Journal of the Meteorological Society of Japan. Ser. II*, 86(1): 213–236. Citado na pág. 34
- Murphy(1988)** J.M. Murphy. The impact of the ensemble forecast on predictability. *Quarterly Journal of the Royal Meteorological Society*, 114(480):463–493. Citado na pág. 2
- Newman et al.(2011)** M. Newman, M. A. Alexander e J. D. Scott. An empirical model of tropical ocean dynamics. *Climate Dynamics*, 37:1823–1841. Citado na pág. 2
- Nogueira et al.(2018)** S.M.C. Nogueira, M.A. Moreira e M.M.L. Volparto. Evaluating precipitation estimates from eta, trmm and chirps data in the south-southeast region of minas gerais state—brazil. *Remote Sensing*, 10(2):1–16. Citado na pág. 41
- North et al.(1982)** G.R. North, T.L. Bell, R.F. Cahalan e F.J. Moeng. Sampling errors in the estimation of empirical orthogonal functions. *Monthly Weather Review*, 110:699–706. Citado na pág. xv, 42, 43, 82
- Oliveira(2000)** G.S. Oliveira. *O El Niño e você: o fenômeno climático*. Editora Transtec. p.116. Citado na pág. 15
- Oliveira-Júnior et al.(2021)** J.F. Oliveira-Júnior, C.A. da Silva Junior, P.E. Teodoro, F.S. Rossi, C.J.C. Blanco, M. Lima, de Gois G., W.L.F. Correia Filho, D.B. Santiago e M.H.G. dos Santos Vanderley. Confronting chirps dataset and in situ stations in the detection of wet and drought conditions in the brazilian midwest. *Int J Climatol.*, Early View:1–6. Citado na pág. 41
- Palmer(1999)** T.N. Palmer. A nonlinear dynamical perspective on climate prediction. *Journal of Climate*, 12(2):575–591. Citado na pág. 3

- Paredes-Trejo et al.(2017)** F.J. Paredes-Trejo, H.A. Barbosa e T.V. Lakshimi Kumar. Validating chirps-based satellite precipitation estimates in northeast brazil. *Journal of Arid Environments*, 139:26–40. Citado na pág. 41
- Parzen(1959)** E. Parzen. *Statistical inference on time series by Hilbert space method, I*. Stanford. Citado na pág. 83
- Perez(2018)** G.M.P Perez. *Improving the quantitative precipitation forecast: a deep learning approach*. Tese de Doutorado, Dissertação (Mestrado em Meteorologia) – Instituto de Astronomia, Geofísica e Ciências Atmosféricas (IAG), Universidade São Paulo, São Paulo. Citado na pág. 10
- Pezza et al.(2007)** A.B. Pezza, I. Simmonds e J.A. Renwick. Southern hemisphere cyclones and anticyclones: Recent trends and links with decadal variability in the pacific ocean. *International Journal of Climatology*, 27:1403–1419. Citado na pág. 33, 35
- Pinheiro et al.(2021)** E. Pinheiro, R.P. da Rocha e A. Drumond. Assessment of 20th-century reanalysis circulation patterns associated with el niño–southern oscillation impacts on the tropical atlantic and northeastern brazil rainy season. *Int J Climatol.*, 41(7):3824– 3840. doi: <https://doi.org/10.1002/joc.6750>. Citado na pág. 15
- Pinto Jr(2015)** O. Pinto Jr. Thunderstorm climatology of brazil: Enso and tropical atlantic connections. *International Journal of Climatology*, 35(6):871–878. Citado na pág. 18, 34
- Ponti et al.(2016)** M. Ponti, E.S. Helou, P.J. Ferreira e N. Mascarenhas. Image restoration using gradient iteration and constraints for band extrapolation. *IEEE Journal of Selected Topics in Signal Processing*, 10(1):71–80. Citado na pág. 10
- Ponti e Paranhos da Costa(2017)** M.A. Ponti e G.B. Paranhos da Costa. Como funciona o deep learning. Em *Simpósio Brasileiro de Bancos de Dados, 32, Uberlandia. Anais ... Porto Alegre: SBC*, páginas 63–93. Citado na pág. 9
- Prado et al.(2021)** L.F. Prado, I. Wainer, E. Yokoyama, M Khodri e J. Garnier. Changes in summer precipitation variability in central brazil over the past eight decades. *International Journal of Climatology*, 1(Early View). Citado na pág. 15, 19, 60
- Príncipe(2010)** J.C. Príncipe. *Information theoretic learning: Rényi’s entropy and kernel perspectives*. New York: Springer Publishing Company, Incorporated, 1º edição. p.448. Citado na pág. 83
- Ramirez et al.(2018)** E. Ramirez, P.L. Silva Dias e C.F.M. Raupp. Multiscale atmosphere-ocean interactions and the low-frequency variability in the equatorial region. *Journal of the Atmospheric Sciences*, 74(8):2503–2523. Citado na pág. 3, 35
- Randel e Wu(2007)** W.J. Randel e F. Wu. A stratospheric ozone profile data set for 1979–2005: Variability, trends, and comparisons with column ozone data. *Journal of Geophysical Research*, 112(D6):1–12. Citado na pág. 22, 33
- Raphaldini et al.(2021)** B. Raphaldini, A.S.W. Teruya, P.L. Silva Dias, L. Massaroppe e D.Y. Takahashi. Stratospheric ozone and quasi-biennial oscillation (qbo) interaction with the tropical troposphere on intraseasonal and interannual timescales: a normal-mode perspective. *Earth Syst. Dynam.*, 12:83–101. Citado na pág. 3, 22, 33
- Raynaut(2011)** C. Raynaut. Interdisciplinaridade: mundo contemporâneo, complexidade e desafios à produção e aplicação de conhecimentos. Em J.A. Philippi e A.J. Silva Neto, editors, *Interdisciplinaridade em Ciência, Tecnologia e Inovação*, páginas 69–105. 1. ed. são paulo: Manole edição. Citado na pág. 39

- Rayner et al.(2003)** N.A. Rayner, D.E. Parker, E.B. Horton, C.K. Folland, L.V. Alexander, D.P. Rowell, E.C. Kent e A. Kaplan. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, 108 (D14):1–37. Citado na pág. 26
- Rebello(2006)** E.R.G. Rebello. A oscilação decadal do pacífico e sua possível influência no regime de precipitação do rio grande do sul. Em *Congresso Brasileiro de Meteorologia, 14, Florianópolis. Anais... Rio de Janeiro: SBMET*. Citado na pág. 16, 50
- Reboita et al.(2010)** M.S. Reboita, M.A. Gan, R.P. Rocha e T. Ambrizzi. Regimes de precipitação na américa do sul: Uma revisão bibliográfica. *Revista Brasileira de Meteorologia*, 25:185–204. Citado na pág. 10, 11, 12
- Robbins e Monro(1951)** H. Robbins e S. Monro. A stochastic approximation method. *The Annals of Mathematical Statistic*, 22(3):400–407. Citado na pág. 9
- Robertson e Mechoso(2000)** A.W. Robertson e C.R. Mechoso. Interannual and interdecadal variability of the south atlantic convergence zone. *American Meteorology Society*, 128(8):2947–2957. Citado na pág. 11
- Robertson e Mechoso(1998)** A.W. Robertson e C.R. Mechoso. Interannual and decadal cycles in river flows of southeastern south america. *Journal of Climate*, 11(10):2570–2581. Citado na pág. 11
- Rodrigues et al.(2011)** R.R. Rodrigues, R.J. Haarsma, E.J.D. Campos e T. Ambrizzi. The impacts of inter-el niño variability on the tropical atlantic and northeast brazil climate. *Journal of Climate*, 24(13):3402–3422. Citado na pág. 15
- Rodríguez-Fonseca et al.(2016)** B. Rodríguez-Fonseca, R. Suárez-Moreno, B. Ayarzagüena, J. López-Parages, I. Gómara, J. Villamayor, E. Mohino, T. Losada e A. Castaño-Tierno. A review of enso influence on the north atlantic. a non-stationary signal. *Atmosphere*, 7(7):87. Citado na pág. 35
- Ropelewski e Jones(1987)** C.F. Ropelewski e P.D. Jones. An extension of the tahiti-darwin southern oscillation index. *Monthly Weather Review*, 115:2161–2165. Citado na pág. 16, 26
- Rosenblatt(1958)** F. Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408. Citado na pág. 7
- Roy e Haigh(2011)** I. Roy e J.D. Haigh. The influence of solar variability and the quasi-biennial oscillation on lower atmospheric temperatures and sea level pressure. *Atmospheric Chemistry and Physics*, 11(22):11679–11687. Citado na pág. 23, 33
- Rumelhart et al.(1986)** D.E. Rumelhart, G.E. Hinton e R.J. Williams. Learning representations by backpropagating errors. *Nature*, 323:533. Citado na pág. 9
- Saha et al.(2010)** S. Saha, S. Moorthi, H. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, H. Liu, D. Stokes, R. Grumbine, G. Gayno, J. Wang, Y. Hou, H. Chuang, H. H. Juang, J. Sela, M. Iredell, R. Treadon, D. Kleist, P. Van Delst, D. Keyser, J. Derber, M. Ek, J. Meng, H. Wei, R. Yang, S. Lord, H. van den Dool, A. Kumar, W. Wang, C. Long, M. Chelliah, Y. Xue, B. Huang, J. Schemm, W. Ebisuzaki, R. Lin, P. Xie, M. Chen, S. Zhou, W. Higgins, C. Zou, Q. Liu, Y. Chen, Y. Han, L. Cucurull, R. W. Reynolds, G. Rutledge e M. Goldberg. The ncep climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91(8):1015–1058. Citado na pág. 65

- Santos et al.(2016a)** N.A. Santos, A.H.B. Siqueira e M.L.D. Melo. Precipitação sobre a América do sul e suas relações com a oscilação multidecadal do atlântico e a oscilação do atlântico norte. *Revista Brasileira de Climatologia*, 18:207–2018. Citado na pág. 18
- Santos et al.(2016b)** S.R.Q. Santos, C.A. Sansigolo e A.P. Santos. Dinâmica temporal em múltiplas escalas de tempo dos eventos secos e chuvosos no sudeste do Brasil. *Revista Brasileira de Geografia Física*, 9(5):1292–1300. Citado na pág. 18
- Satyamurty et al.(1998)** P. Satyamurty, C.A. Nobre e P.L. Silva Dias. Tropics south america. Em Karoly e Vicent, editors, *Meteorology of the Southern Hemisphere, Boston*, páginas 119–139. segunda edição. Citado na pág. 10
- Schossler et al.(2018)** V. Schossler, J.C. Simões, F.E. Aquino e D.R. Viana. Precipitation anomalies in the Brazilian southern coast related to the sam and ENSO climate variability modes. *Revista Brasileira de Recursos Hídricos*, 23(14):1–10. Citado na pág. 21
- Schölkopf e Smola(2018)** B. Schölkopf e A.J. Smola. *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. MIT Press, reprint edição. p.648. Citado na pág. 83
- Seluchi e Marengo(2000)** M.E. Seluchi e J.A. Marengo. Tropical-extratropical exchange of air masses during summer and winter in south america: Climatic aspects and extreme events. *International Journal of Climatology*, 20:1167–1190. Citado na pág. 12
- Servain(1991)** J. Servain. Simple climatic indices for the tropical Atlantic ocean and some applications. *J. Geophys Res.*, 96(C8):15137–15146. Citado na pág. 26
- Severo(2008)** D.L. Severo. *Padrões de circulação anômala associada a episódios úmidos e secos no Sul do Brasil*. Tese de Doutorado, Tese (Doutorado de Meteorologia) – Instituto Nacional de Pesquisas Espaciais, Ministério da Ciência e Tecnologia, São José dos Campos. Citado na pág. 20
- Shimizu e Ambrizzi(2015)** M.H. Shimizu e T. Ambrizzi. MJO influence on ENSO effects in precipitation and temperature over south america. *Theoretical and Applied Climatology*, 7(1-2):291–301. Citado na pág. 13, 33, 34
- Silva e Carvalho(2007)** A.E. Silva e L.M.V. Carvalho. Large-scale index for south america monsoon (LISAM). *Atmospheric Science Letters*, 8:51–57. Citado na pág. xiii, 12, 13, 25, 26, 27, 28
- Siqueira e Molion(2015)** A.H.B. Siqueira e L.C.B. Molion. Análises climáticas: O filtro Hodrick-Prescott aplicado aos índices atmosféricos da oscilação sul e da oscilação do atlântico norte. *Revista Brasileira de Meteorologia*, 30(3):307–318. Citado na pág. 20
- Sousa et al.(2010)** A.J.S. Sousa, E.A. Ferreira, J.R.A. Sousa e P.A.M. Rolim. Oscilação decadal do Pacífico e sua influência na precipitação de Goiânia – GO. Em *Congresso Brasileiro de Meteorologia, 16, Belém. Anais... Rio de Janeiro: SBMET*. Citado na pág. 16
- Souza e Ambrizzi(2006)** E.B. Souza e T. Ambrizzi. Modulation of the intraseasonal rainfall over tropical Brazil by the Madden-Julian oscillation. *Int. J. Climatol.*, 26:1759–1776. Citado na pág. 13
- Stoica e Moses(2005)** P. Stoica e R. Moses. *Spectral Analysis of Signals*. Upper Saddle River, NJ: Prentice Hall, 1ª edição. p.480. Citado na pág. 41
- Tedeschi et al.(2016)** R.G. Tedeschi, A.M. Grimm e I.F. Cavalcanti. Influence of central and east ENSO on precipitation and its extreme events in south america during austral autumn and winter. *International Journal of Climatology*, 36:4797–4814. Citado na pág. 15, 50, 57, 58

- Teixeira(2002)** R.F.B. Teixeira. A oscilação do atlântico norte e a qualidade da quadra chuvosa do ceará. Em *Congresso Brasileiro de Meteorologia, 12, Foz do Iguaçu. Anais... Rio de Janeiro: SBMET*. Citado na pág. 20
- Theodoridis e Koutroumbas(2009)** S. Theodoridis e K. Koutroumbas. *Pattern Recognition*. Academic Press, 4<sup>o</sup> edição. p.961. Citado na pág. 56, 67
- Thompson e Wallace(2000)** D.W.J. Thompson e J.M. Wallace. Tannular modes in the extratropical circulation. part i: Month-to-month variability. *Journal of Climate*, 13(5):1000–1016. Citado na pág. 21
- Thompson e Wallace(1998)** D.W.J. Thompson e J.M. Wallace. The arctic oscillation signature in the wintertime geopotential height and temperature fieds. *Geophysical Research Letters*, 25(9):1297–1300. Citado na pág. 21, 26
- Thomson e Emery(2014)** R.E. Thomson e W.J. Emery. Chapter 6 - digital filters. Em *Data Analysis Methods in Physical Oceanography*, páginas 593–637. Elsevier, Boston, third edition edição. doi: <https://doi.org/10.1016/B978-0-12-387782-6.00006-5>. URL <https://www.sciencedirect.com/science/article/pii/B9780123877826000065>. Citado na pág. 41
- Torrence e Compo(1998)** C. Torrence e G.P. Compo. A practical guide to wavelet analysis. *Bulletin of American Meteorological Society*, 79(1):61–78. Citado na pág. 28, 29, 41, 77, 78
- Tracton e Kalnay(1993)** M.S. Tracton e E. Kalnay. Operational ensemble prediction at the national meteorological center: Pratical aspects. *Weather and Forecasting*, 8(3):379–398. Citado na pág. 2
- Trenberth(2017)** K.E. Trenberth. *El Niño Southern Oscillation (ENSO).Reference Module in Earth Systems and Environmental Sciences*. Elsevier. Citado na pág. 15, 33
- Trenberth e Hurrell(1994)** K.E. Trenberth e J.W. Hurrell. Decadal atmosphere-ocean variations in the pacific. *Climate Dynamics*, 9:303–319. Citado na pág. 26
- Trenberth e Stepaniak(2001)** K.E. Trenberth e D.P. Stepaniak. Indices of el niño evolution. *Journal of Climate*, 14(8):1697–1701. Citado na pág. 16
- Tucci(2003)** C.E.M. Tucci. Long-term flow forecasts based on climate and hydrologic modeling: Uruguay river basin. *Water Resources Resesearch*, 39(7):1–11. Citado na pág. 5
- Tucci(2005)** C.E.M. Tucci. *Modelos Hidrológicos*. UFRGS/ABRH, 2<sup>o</sup> edição. p.608. Citado na pág. 7
- Van Loon e Meehl(2014)** H. Van Loon e G.A. Meehl. Interactions between externally forced climate signals from sunspot peaks and the internally generated pacific decadal and north atlantic oscillations. *Geophysical Research Letters*, 41(1):161–166. Citado na pág. 22, 29
- Van Loon et al.(2004)** H. Van Loon, G.A. Meehl e J.M. Arblaster. A decadal solar effect in the tropics in july–august. *Journal of Atmospheric and Solar-Terrestrial Physics*, 66(18):1767–1778. Citado na pág. 22
- Vasconcellos e Cavalcanti(2010)** F.C. Vasconcellos e I.F.A. Cavalcanti. Extreme precipitation over southeastern brazil in the austral summer and relations with the southern hemisphere annular mode. *Atmospheric Science Letters*, 11(1):21–26. Citado na pág. 21, 49, 57, 58



- Vitart et al.(2017)** F. Vitart, C. Ardilouze, A. Bonet, A. Brookshaw, M. Chen, C. Codorean, M. Déqué, L. Ferranti, E. Fucile, M. Fuentes, H. Hendon, J. Hodgson, H. Kang, A. Kumar, H. Lin, G. Liu, X. Liu, P. Malguzzi, I. Mallas, M. Manoussakis, D. Mastrangelo, C. MacLachlan, P. McLean, A. Minami, R. Mladek, T. Nakazawa, S. Najm, Y. Nie, M. Rixen, A. W. Robertson, P. Ruti, C. Sun, Y. Takaya, M. Tolstykh, F. Venuti, D. Waliser, S. Woolnough, T. Wu, D.-J. Won, H. Xiao, R. Zaripov e L. Zhang. The subseasonal to seasonal (s2s) prediction project database. *Amer. Meteor. Soc.*, 98:163–173. Citado na pág. 2
- Von Sperling(2007)** M. Von Sperling. *Estudo e modelagem da qualidade da água de rios*. Editora UFMG, 1º edição. p.588. Citado na pág. 5, 52
- Wainer e Venegas(2002)** I. Wainer e S.A. Venegas. South atlantic multidecadal variability in the climate system model. *Journal of Climate*, 15(12):1408–1420. Citado na pág. 19
- Wallace e Gutzler(1981)** L.M. Wallace e D.S. Gutzler. Teleconnections in the geopotential height field during the northern hemisphere winter. *Monthly Weather Review*, 109:785–810. Citado na pág. 21, 33
- Weng e Lau(1994)** H.Y. Weng e K.M. Lau. Wavelet, period-doubling and time-frequency localization with application to satellite data analysis. *Journal of the Atmospheric Sciences*, 51: 2523–2541. Citado na pág. 77
- Wheeler e Hendon(2004)** M.C. Wheeler e H.H. Hendon. An all-season real-time multivariate mjo index: Development of an index for monitoring and prediction. *Monthly Weather Review*, 132(8):1917–1932. Citado na pág. 14, 26
- White(2017)** C.J. et al White. Potential applications of subseasonal-to-seasonal (s2s) predictions. *Meteorol. Appl.*, 24:315–325. Citado na pág. 2
- Wilks(2006)** D.S. Wilks. *Statistical Methods in the Atmospheric Sciences*. Elsevier Academic Press Publications, 2º edição. p.634. Citado na pág. 56, 75, 87
- Wolter e Timlin(1998)** K. Wolter e M.S. Timlin. Measuring the strength of enso events: How does 1997/98 rank? *Weather*, 53(9):315–324. Citado na pág. 16, 26
- Yamakawa et al.(2016)** S. Yamakawa, M. Inoue e R. Suppiah. Relationships between solar activity and variations in sst and atmospheric circulation in the stratosphere and troposphere. *Quaternary International*, 397:289–299. Citado na pág. 23, 34
- Yang e Xu(2004)** C. Yang e C. Xu. Properties of hadamard product of inverse m-matrices. *Numerical Linear Algebra with Applications*, 11(4):343–354. Citado na pág. 90
- Yu et al.(2015)** J-Y. Yu, H. Paek, E.S. Saltzman e T. Lee. The early 1990s change in enso–psa–sam relationships and its impact on southern hemisphere climate. *Journal of Climate*, 28(23):9393–9408. Citado na pág. 33
- Yu e Zhu(2020)** T. Yu e H. Zhu. Hyper-parameter optimization: A review of algorithms and applications. *CoRR*, <http://arxiv.org/abs/12003.05689>. Citado na pág. 56
- Zang e Moore(2015)** Z. Zang e J.C. Moore. Empirical orthogonal functions. Em Altun Y. et al., editor, *Mathematical and Physical Fundamentals of Climate Change*, páginas 161–197. Elsevier edição. Citado na pág. 25, 26, 27, 41, 81
- Zhai(2017)** Q. Zhai. Evidence for the effect of sunspot activity on the el niño/southern oscillation. *New Astronomy*, 52:1–7. Citado na pág. 23, 34