



ISSN:1984-2295

Revista Brasileira de Geografia Física

Homepage: <https://periodicos.ufpe.br/revistas/rbgfe>



Identificação de preditores para as chuvas do setor leste do Nordeste do Brasil utilizando análise de correlação canônica

Geber Barbosa de Albuquerque Moura¹, JoséIVALDO Barbosa de Brito², FRANCISCO de ASSIS SALVIANO de SOUSA³, ENILSON PALMEIRA CAVALCANTI⁴, Jhon Lennon Bezerra da Silva⁵, CRISTINA RODRIGUES NASCIMENTO⁶, PABRÍCIO MARCOS OLIVEIRA LOPES⁷

¹ Doutor em Oceanografia, Professor Titular, Departamento de Agronomia, Universidade Federal Rural de Pernambuco. Av. D. Manoel de Medeiros, SN, Dois Irmãos, Recife, Pernambuco, Brasil. CEP: 52171-900. E-mail: geber@depa.ufpe.br (autor correspondente).

² Doutor em Recursos Naturais, Professor Associado, Universidade Federal de Campina Grande, Campina Grande, Paraíba, Brasil. E-mail: jivaldobrito@gmail.com.

³ Doutor em Hidráulica e Saneamento, Professor Titular, Departamento de Ciências Atmosféricas, Universidade Federal de Campina Grande, Campina Grande, Paraíba, Brasil. E-mail: francisco.salviano@ufcg.edu.br.

⁴ Doutor em Recursos Naturais, Professor Titular, Centro de Tecnologia e Recursos Naturais, Universidade Federal de Campina Grande, Campina Grande, Paraíba, Brasil. E-mail: enilson.cavalcanti@ufcg.edu.br.

⁵ Doutorando em Engenharia Agrícola, Departamento de Engenharia Agrícola, Universidade Federal Rural de Pernambuco. E-mail: jhonlennoigt@hotmail.com.

⁶ Doutora em Engenharia Agrícola, Professora Adjunta, Departamento de Agronomia, Universidade Federal Rural de Pernambuco – UFRPE, Recife, Pernambuco, Brasil. E-mail: crisrodnas@gmail.com.

⁷ Doutor em Sensoriamento Remoto, Professor Associado, Departamento de Agronomia, Universidade Federal Rural de Pernambuco – UFRPE, Recife, Pernambuco, Brasil. E-mail: pabriciope@gmail.com.

Artigo recebido em 03/02/2020 e aceito em 25/03/2020

RESUMO

O objetivo deste trabalho foi encontrar as melhores variáveis preditoras através de análise de correlação canônica nos ventos alísios, Temperatura da Superfície do Mar (TSM), Pressão atmosférica à superfície no Oceano Pacífico Equatorial e TSM no Atlântico Tropical (área do Dipolo), de forma que se possam elaborar modelos de previsão da precipitação pluvial (período chuvoso) do setor leste do Nordeste do Brasil para os quatro meses mais chuvosos dos três grupos homogêneos, com antecedência de três meses. Os grupos foram escolhidos a partir de análise de agrupamento utilizando o método hierárquico. Para estudar as correlações canônicas entre a precipitação dos grupos com os dados padronizados de TSM, vento e pressão atmosférica, as análises fundamentaram-se na série dos totais de precipitação de abril a julho e dados defasados de médias de três meses (média de Novembro a Janeiro) de TSM, vento em 850 hPa no Pacífico Equatorial e pressão da atmosfera em Tahiti e Darwin para o período de 1986 a 2017. Percebe-se que os principais preditores para os grupos homogêneos foram, por ordem de maior importância: Média de três meses de atraso do índice de ventos alísios Equatorial central (MedWC), Média da pressão atmosférica à superfície em Darwin (Mdarwin), Média do EN 34 (MEN34), Média da pressão atmosférica à superfície em Tahiti (Mtahiti) e Média de índice de ventos alísios leste (MedWE). Nota-se deste atraso que a principal influência está no Pacífico, no ENOS.

Palavras-chave: vento, TSM, precipitação.

Predictors identification for rain in the east sector of the Northeast Brazil using canonical correlation analysis

ABSTRACT

The objective of this work was to find the best predictor variables through canonical correlation analysis in trade winds, Sea Surface Temperature (SST), Atmospheric pressure at the surface in the Equatorial Pacific Ocean and SST in the Tropical Atlantic (Dipole area), that models for forecasting rainfall (rainy season) in the eastern sector of northeastern Brazil can be developed for the four rainiest months of the three homogeneous groups, three months in advance. The groups were chosen from the cluster analysis using the hierarchical method. To study the canonical correlations between the precipitation of the

groups with the standardized data of SST, wind and atmospheric pressure, the analyzes were based on the series of precipitation totals from April to July and lagged data of three-month averages (average from November to July). January) of SST, wind at 850 hPa in the Equatorial Pacific and atmospheric pressure in Tahiti and Darwin for the period from 1986 to 2017. It can be seen that the main predictors for homogeneous groups were, in order of greatest importance: Average of three months delay of the central Equatorial trade winds index (MedWC), mean of the atmospheric pressure at the surface in Darwin (Mdarwin), mean of the EN 34 (MEN34), mean of the atmospheric pressure at the surface in Tahiti (Mtahiti) and mean of the east trade winds (MedWE). It is noted from this delay that the main influence is in the Pacific, in the ENSO. Keywords: wind, SST, precipitation.

Introdução

A Região Nordeste do Brasil (NEB) enfrenta de forma intensificada problemas como a escassez de água, em decorrência principalmente das mudanças climáticas e variações meteorológicas que promovem eventos de seca constantes. Estes transtornos geram diversos conflitos nos setores de atividades essenciais na NEB, prejudicando inclusive as condições socioeconômicas local e regional. Com isso, a previsão e o monitoramento climático desta Região são essenciais para o planejamento e gerenciamento adequado dos recursos hídricos.

Prever as características futuras das variáveis meteorológicas tem sido o objetivo de muitos estudos ao longo do tempo, e vários modelos foram e têm sido utilizados na elaboração de previsões climáticas. Elaboradas com um adequado grau de precisão, estas previsões proporcionam aos tomadores de decisão e governos a possibilidade de antecipar e planejar, inteligentemente, futuras ações dirigidas a setores de atividades socioeconômicas e à sociedade como um todo.

Há evidências observacionais, teóricas e resultados de modelos de circulação geral da atmosfera (MCGA) de que as condições oceânicas e atmosféricas sobre as Bacias do Atlântico e Pacífico Tropical influenciam fortemente na variabilidade interanual do clima sobre as Américas (Hastenrath & Greischar, 1993; Aragão et al., 1994; Nobre & Shukla, 1996; Kayano e Capistrano, 2014; Builes-Jaramillo et al., 2018).

O setor leste da Região Nordeste do Brasil (NEB) devido a sua localização está submetido à influência de sistemas meteorológicos, os quais interagem entre si, conferem-lhe características próprias (Harzallah et al., 1996; Gomes et al., 2015; Moraes et al., 2015; Barbosa et al., 2016; Salgueiro et al., 2016; Dias et al., 2018; Fedorova et al., 2018; Marengo et al., 2018). A Zona de Convergência Intertropical atua no período de março a abril. A penetração de ar frio das massas de ar polar normalmente está associada à formação de frentes frias que influenciam na precipitação no período que vai de abril a agosto. As ondas de leste que são sistemas que se formam sobre o Oceano Atlântico

atuam entre maio e julho. Os sistemas de brisas marítimas afetam o regime de chuvas no Litoral do NEB, com maior intensidade no período que vai de maio a junho (Hounsou-Gbo et al., 2016).

O NEB sofre influências significativas das circulações atmosféricas e oceânicas no Atlântico e Pacífico (Alves et al., 1993; Rao et al., 1993; Hounsou-Gbo et al., 2015; Barbosa et al., 2016; Hounsou-Gbo et al., 2016; Marengo et al., 2018).

Tem-se obtido êxito considerável nas previsões dos eventos do fenômeno El Niño/Oscilação Sul (ENOS), por meio de simulações (modelos) dos fenômenos físicos da interface oceano - atmosfera. Além disso, o conhecimento sobre o Atlântico tropical (Dipolo do Atlântico), conjuntamente com as informações sobre o ENOS, nos permite elaborar uma previsão sobre a precipitação pluvial no semiárido nordestino (Kayano et al., 2011; Nóbrega et al., 2014; Capotondi et al., 2015; Andreoli et al., 2016; Hounsou-Gbo et al., 2016).

Vários modelos de previsão climática para o semiárido do Nordeste (Grimm e Tedeschi, 2009; Silva et al., 2011; Gonzalez et al., 2013), já operacionais, vêm sendo utilizados por diversos organismos regionais, nacionais e internacionais. Com esses modelos é possível elaborar prognósticos sobre a qualidade da estação chuvosa do norte do semiárido do Nordeste (período que vai de fevereiro a maio), com antecedência de um a três meses e, para os casos de anos muito chuvoso, com até seis meses de antecedência (Hounsou-Gbo et al., 2016). A confiabilidade dessas previsões pode chegar até a 80%, na dependência de definições acentuadas dos parâmetros envolvidos (Ward e Folland, 1991; Hastenrath e Greischar, 1993).

Para a faixa costeira ao leste da NEB, de climas úmidos, ainda não existem modelos operacionais para prognósticos confiáveis de previsão climática, e diante desta necessidade de prever as principais características futuras da precipitação pluviométrica no setor leste justifica-se sendo este o principal viés da presente pesquisa. Estudos recentes, também, demonstram a importância dos fenômenos oceânicos ENOS e

Dipolo do Atlântico na definição da estação chuvosa dessa região (Moura et al., 2000; Moura et al., 2009).

A precipitação pluvial é uma variável de muita importância para o desenvolvimento socioeconômico da NEB. Portanto, o conhecimento detalhado da distribuição espacial e temporal da precipitação nas micro e mesoregiões, bem como a verificação das anomalias de TSM dos oceanos como principal variável física influenciadora das condições climáticas são de grande importância para a tomada de decisões por parte dos governos e da sociedade civil.

Neste trabalho, usa-se a técnica de Análise de Correlações Canônicas (ACC), a qual permite encontrar combinações lineares, chamadas “variáveis canônicas”, a partir de dois conjuntos de dados (por exemplo, X2: preditor e X1: preditando), de forma que a correlação entre estas variáveis seja maximizada.

A ACC também é usada eficazmente para descrição de dados, verificação de modelos numéricos e construção de modelos estatísticos de previsão, proporcionando o conhecimento de quais configurações tendem a ocorrer simultaneamente entre dois ou mais campos distintos e qual o grau de conexão entre eles (L’Heureux et al., 2015; Rana et al., 2018). Ela se resume em associar índices a cada um dos conjuntos de dados (X2: variável explicativa e X1: resposta), definidos como combinações lineares dos valores em cada um dos conjuntos (Função Ortogonal Empírica - FOE), de forma a maximizar a correlação entre os dois índices. Desta forma, busca-se reter ao máximo a informação contida nas variáveis originais (Hair et al., 1998; Mingoti, 2005).

A técnica de ACC tem sido aplicada em vários estudos de modelagem estatística. Sua grande vantagem está no fato de possibilitar a busca de padrões de oscilação conjunta entre dois ou mais campos físicos. A partir daí, encontra-se um conjunto de equações prognósticas e se constrói um modelo preditor (Barnett e Preisendorfer, 1987).

Recentemente, alguns pesquisadores usaram a análise de correlação canônica para previsões e cenários climáticos e estudos diagnósticos, como por exemplo, Rana et al. (2018) para a previsão sazonal da precipitação pluvial de inverno no centro-sudoeste da Ásia, e Forootan et al. (2016) para entender a variabilidade das chuvas na Austrália.

Desta forma, o objetivo deste trabalho foi identificar variáveis preditoras através da análise de correlação canônica aplicada em regiões homogêneas de precipitação pluvial no setor leste do

Nordeste do Brasil com ventos alísios, TSM, Pressão atmosférica à superfície no Oceano Pacífico Equatorial e TSM no Atlântico Tropical (área do Dipolo), de forma que se possam elaborar modelos que auxiliem na previsão da precipitação do setor leste do Nordeste do Brasil, com antecedência de até três meses.

Material e métodos

Descrição da Área de Estudo

A sub-região denominada de setor leste do Nordeste é uma faixa ao longo do Litoral oriental, de aproximadamente 200 Km de largura e estende-se do Rio Grande do Norte ao Sul da Bahia (Figura 3). Esta sub-região pode ser dividida em três unidades geomorfológica: Litoral Oriental; Zona da Mata e parte do Agreste.

O Litoral Oriental mostra uma diversidade de formas geográficas como: Restingas, Dunas, Lagunas, Mangues, Tabuleiros sobre a Formação Barreiras e Colinas sobre a Formação Barreiras. A essas feições da geomorfologia dessa sub-região litorânea, somam-se também a embocadura do Rio São Francisco e a baía de Todos os Santos.

A faixa costeira do Nordeste do Brasil, também conhecida como Zona da Mata, estende-se do Rio Grande do Norte ao sul da Bahia. Essa sub-região do NEB apresenta clima quente e úmido com totais pluviométricos anuais elevados (1.000 a 2.400 mm/ano) e com um número médio de dias de chuva/ano variando de 150 a 210. As principais chuvas dessa região começam em março prolongando-se até julho e, em alguns anos, até agosto. Portanto, a Zona da Mata contrasta com o semiárido brasileiro (SAB) onde os totais anuais de precipitação são abaixo de 800 mm, apresentando 300 mm de total anual em Cabaceiras na Paraíba.

O agreste é a área de transição entre a Zona da Mata, região úmida e o Sertão semiárido. No agreste, os terrenos mais férteis são ocupados por minifúndios, onde predominam as culturas de subsistência e a pecuária leiteira.

Nessa sub-região se encontra as principais metrópoles regionais que são as cidades de Salvador, capital do Estado da Bahia e Recife, capital do Estado de Pernambuco.

No NEB, encontram-se regiões montanhosas com climas úmidos como é o caso da parte leste do Planalto da Borborema que separa o agreste/litoral do semiárido. A serra da Borborema é a mais importante feição geomorfológica do Nordeste oriental. A exposição de suas encostas, voltadas para leste e

sudeste, e a dimensão do seu relevo exercem papel fundamental na concentração de umidade, na formação de solos mais profundos e na existência da vegetação de floresta. Estendendo-se de Alagoas ao Rio Grande do Norte, o grande conjunto planáltico da Borborema apresenta três aspectos no seu modelado: a escarpa, as superfícies elevadas dos maciços e as superfícies baixas planas.

Dados de Pluviometria Sobre o Leste do NEB

Foram utilizados totais mensais de precipitação de 27 postos pluviométricos distribuídos espacialmente ao longo do setor leste do NEB (Zona da Mata e Agreste) (Figura 3), disponíveis no banco de dados do Instituto Nacional de Meteorologia (INMET) e da Agência Pernambucana de Águas e Climas (APAC). Estes dados compreendem um período de observação entre os anos de 1986 e 2017, com todos os postos pluviométricos apresentando um período superior a 30 anos de dados.

O período chuvoso do setor leste foi escolhido de abril a julho, de acordo com o resultado encontrado com análise de agrupamento pelo método de Ward (Ward, 1963) mostrado no Dendograma na Figura 2. Nota-se no Dendograma, que de janeiro a março ocorre à interface entre o período seco-chuvoso, e de agosto e setembro a interface entre o período chuvoso-seco. No subgrupo de outubro a dezembro tem-se a estação seca. Em janeiro em alguns anos há ocorrência de chuvas devido aos Vórtices Ciclônicos da Alta Troposfera. No período de abril a julho se encontra dois subgrupos: abril e maio onde a ZCIT é o sistema atmosférico mais atuante na ocorrência de chuvas; abril e maio onde tem-se atuação da ZCIT e Distúrbios de Leste; e junho e julho influência de Distúrbios de Leste e de Sistemas de Brisas.

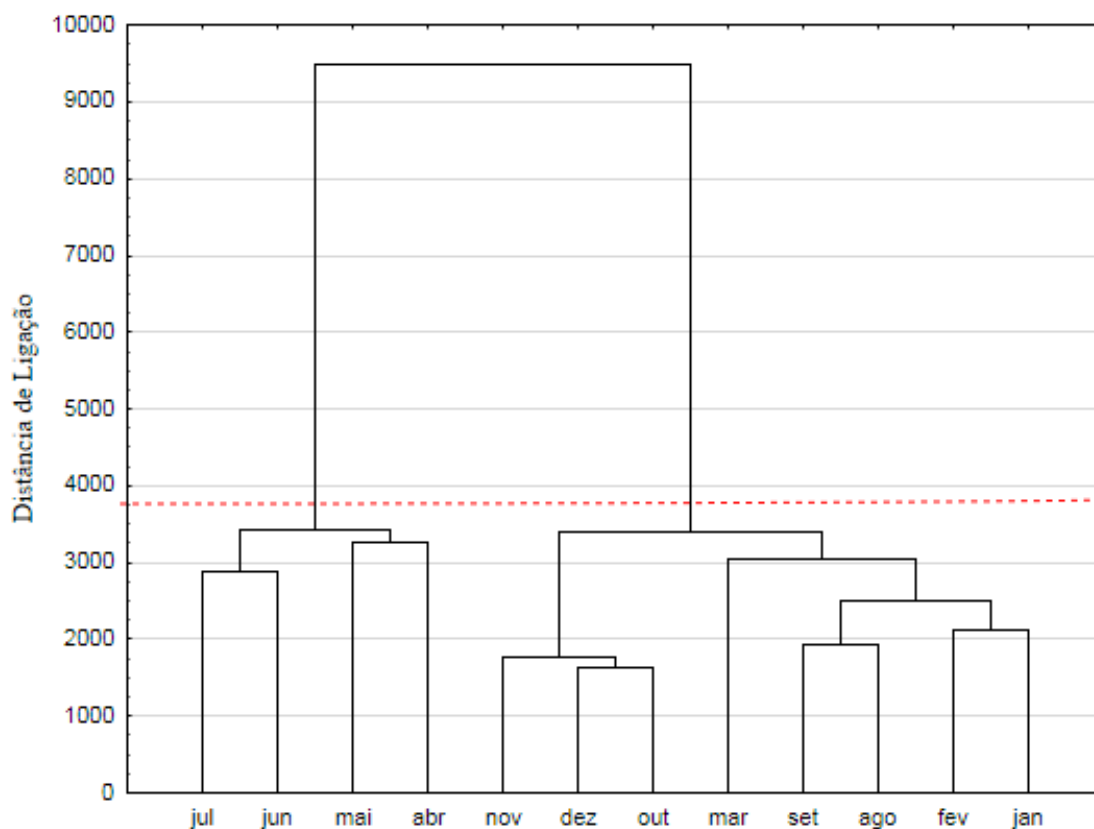


Figura 1. Dendograma resultante de Análise de Agrupamento, utilizando-se método de Ward e distância euclidiana para dados mensais de precipitação para os 27 postos, para o período de dados de 1986 a 2017.

Os grupos foram escolhidos a partir de análise de agrupamento utilizando o método hierárquico de Ward (Ward, 1963). A Figura 2 apresenta os três grupos, e a Tabela 1 mostra os

municípios dos três grupos. Têm-se que as localidades com menos chuvas se concentram no Grupo 3 onde em sua maioria fica a sotavento do Planalto da Borborema no caso de Paraíba e

Pernambuco (Figura 3). O Grupo 2 encontram-se no Vale do São Francisco em Alagoas e no Agreste da Bahia as localidades com chuvas intermediárias, parte também estão a barlavento do Planalto da

Borborema. O Grupo 1 estão as localidades que se encontram próxima ou no Litoral, e é onde acontecem as maiores chuvas.

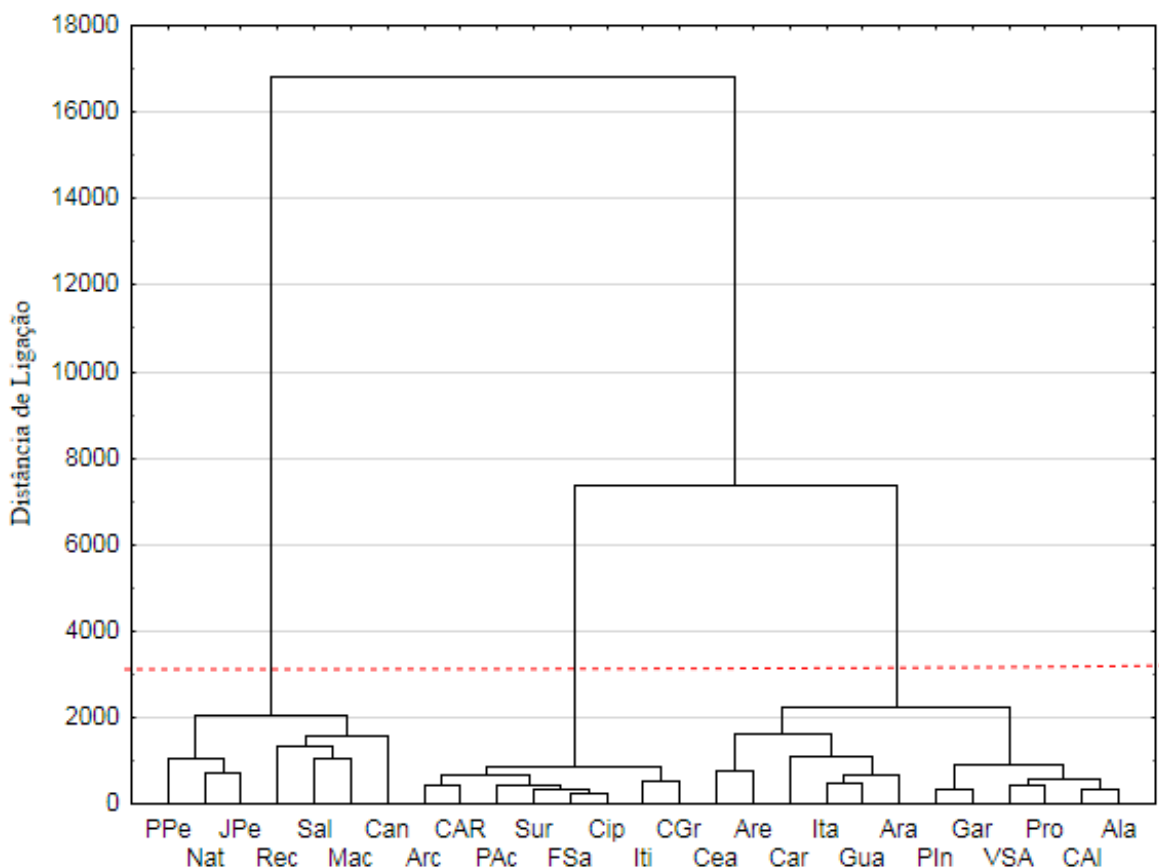


Figura 2. Dendrograma resultante da Análise de Agrupamento, utilizando-se o método de Ward e distância euclidiana para as 27 estações do setor leste do Nordeste do Brasil no período de dados de 1986 a 2017.

Tabela 1. Relação das estações dos grupos homogêneos.

Grupo	Estações pluviométricas
Grupo 1	Natal, João Pessoa, Recife, Porto de Pedras, Maceió, Salvador, Canavieiras.
Grupo 2	Ceará Mirim, Areia, Vitória de Santo Antão, Garanhuns, Palmeira dos Índios, Propriá, Aracaju, Itabaianinha, Alagoinha, Cruz das Almas, Guaratinga, Caravelas.
Grupo 3	Campina Grande, Surubim, Arcoverde, Caruaru, Pão de Açúcar, Cipó, Feira de Santana, Itiruçu.

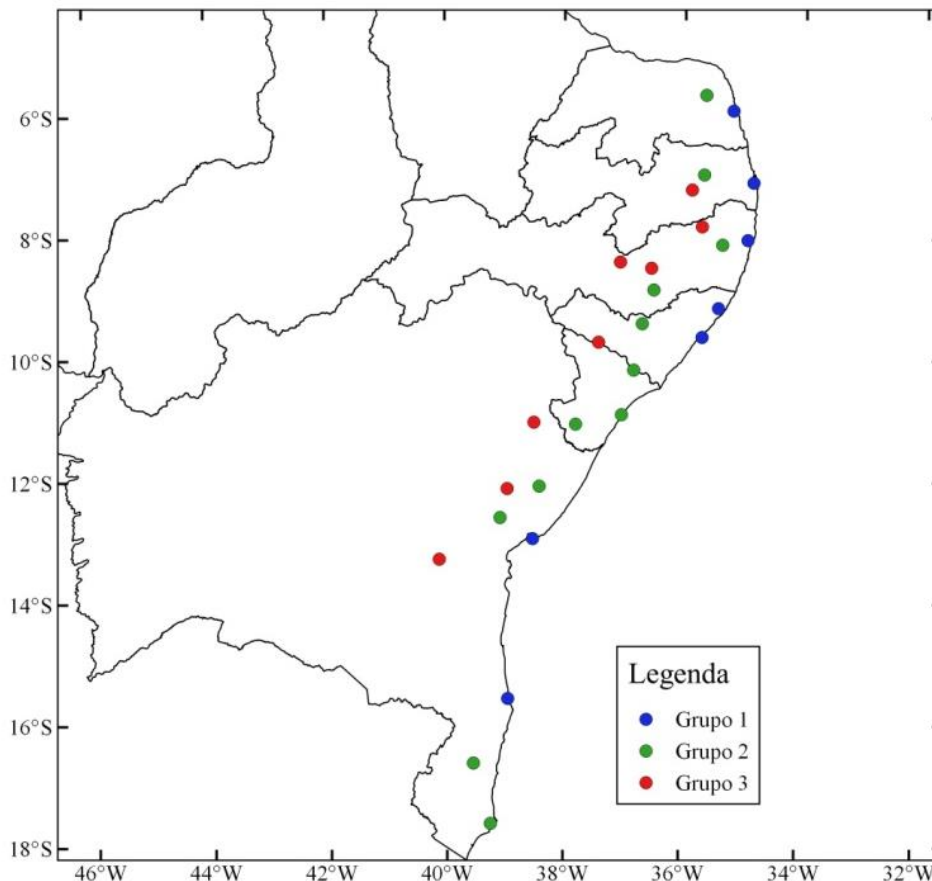


Figura 3. Localização dos três grupos homogêneos no setor leste do Nordeste do Brasil. O símbolo azul representa os postos do Grupo 1, o verde do Grupo 2, o vermelho do Grupo 3.

Análise de Agrupamento

Para a análise de agrupamento das estações, considerou-se a distância utilizada como função de grupamento para medir a similaridade ou dissimilaridade entre os indivíduos (postos), P_i e P_j como distância euclidiana, que são dadas pela Equação 1.

$$D(P_i, P_j) = \left[\sum_{q=1}^n N_i (P_{iq} - P_{jq})^2 \right]^{1/2} \quad (1)$$

em que, P_{ij} representa a precipitação do período chuvoso (abril a julho) das estações P_i e P_j , n representa o número de anos, e q representa os anos, N_i representa os indivíduos (estações) organizados em K classes.

O critério de agrupamento utilizado foi o de Ward (1963) o qual propõe que, em qualquer estágio de uma análise, a perda de informação resultante de agrupamento pode ser medida pela soma dos

quadrados dos desvios (SQD) de cada ponto com relação à média do grupo ao qual pertence. Logo para os N_i indivíduos (estações) organizados em K classes, onde X_i é o indivíduo pertencente a i -ésima classe (Equação 2).

$$SQD = \sum_{i=1}^n \left[X_i^2 - \frac{1}{n} \left(\sum X_i \right)^2 \right] \quad (2)$$

Dados de Pressão da Atmosfera, Vento em 850 hPa e Temperatura da Superfície do Mar (TSM)

Foram utilizadas temperaturas mensais de TSM na área dos Niños: Niño 1+2 (0-10°S, 90°W-80°W), Niño 3 (5°N-5°S, 150°W-90°W), Niño 3.4 (5°N-5°S, 170°W-120°W) e Niño 4 (5°N-5°S, 160°E-150°W) (Figura 4), temperaturas mensais do Atlântico Tropical Norte (ATN), na área compreendida entre 5°N-23,5°N e 15°W-57,5°W, e temperaturas mensais do Atlântico Tropical Sul (ATS), na área compreendida entre 0-20°S e 10°E-

30°E (Figura 6), de janeiro de 1986 a dezembro de 2017.

As temperaturas nas áreas do Atlântico Tropical Norte (ATN) e Atlântico Tropical Sul (ATS) correspondem à região associada ao padrão de Dipolo do Atlântico Tropical (Moura & Shukla, 1981; Servain, 1991). Os dados de índices de ventos alísios mensais em 850 hPa no Oceano Pacífico Equatorial nas seguintes áreas: Pacífico Central (5° N-5°S, 175°W-140°W), Pacífico Este (5°N-5°S, 135°W-150°W) e Pacífico Oeste (5°N-5°S, 135°E-180°W) (Figura 4) e os dados de pressão atmosférica mensais à superfície das localidades Darwin e Tahiti (Figura 5) também foram utilizados para o mesmo período. Todos os dados TSM, ventos alísios,

pressão da atmosfera foram obtidos da base de dados do Climate Prediction Center – CPC, centro pertencente ao NCEP (National Centers for Environmental Prediction). Para estudar as correlações canônicas entre a precipitação de cada grupo com os dados padronizados de TSM, vento e pressão atmosférica, as análises fundamentaram-se na série dos totais de precipitação de abril a julho e dados defasados (Tabela 2) de médias de três meses de TSM, vento (WC, WE, WW) em 850 hPa no Pacífico equatorial e pressão da atmosfera em Tahiti e Darwin para o período de 1986 a 2017.

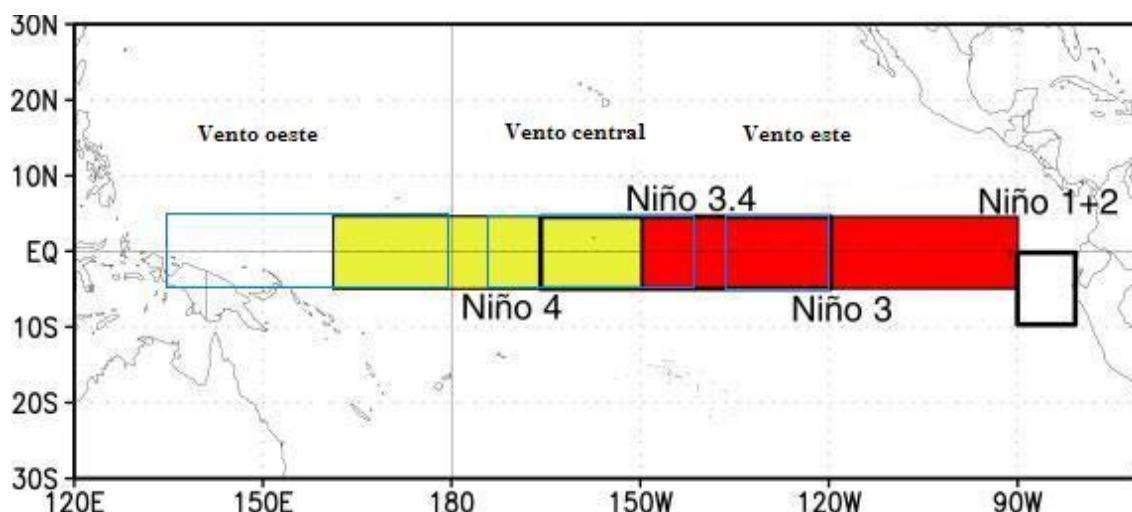


Figura 4. Localização das 4 regiões do El Niño na Bacia do Pacífico adaptada. Fonte: http://www.cpc.noaa.gov/products/analysis_monitoring/ensostuff/nino_regions.shtml.

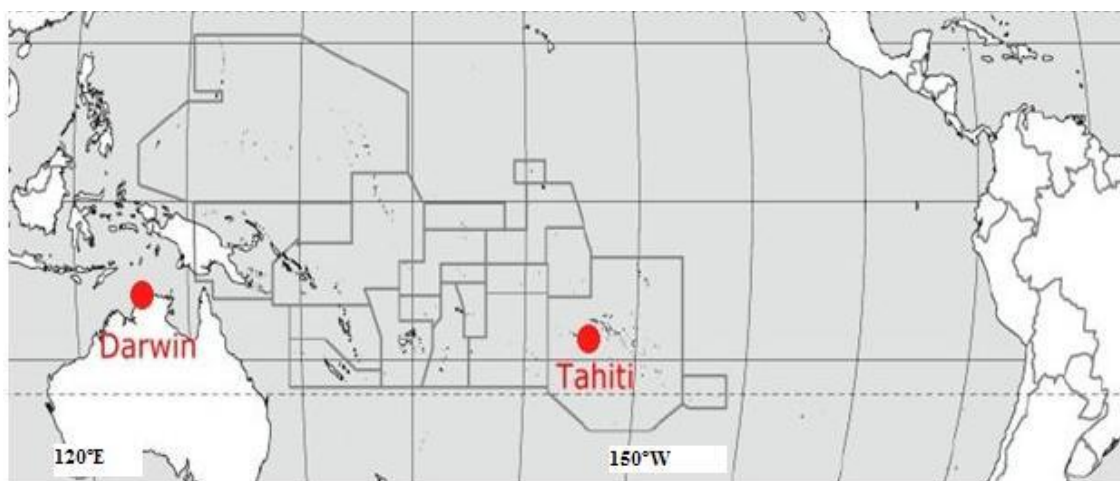


Figura 5. Localização das 4 regiões do El Niño na Bacia do Pacífico adaptada. Fonte: <https://open.uj.edu.pl/mod/page/view.php?id=1396>.

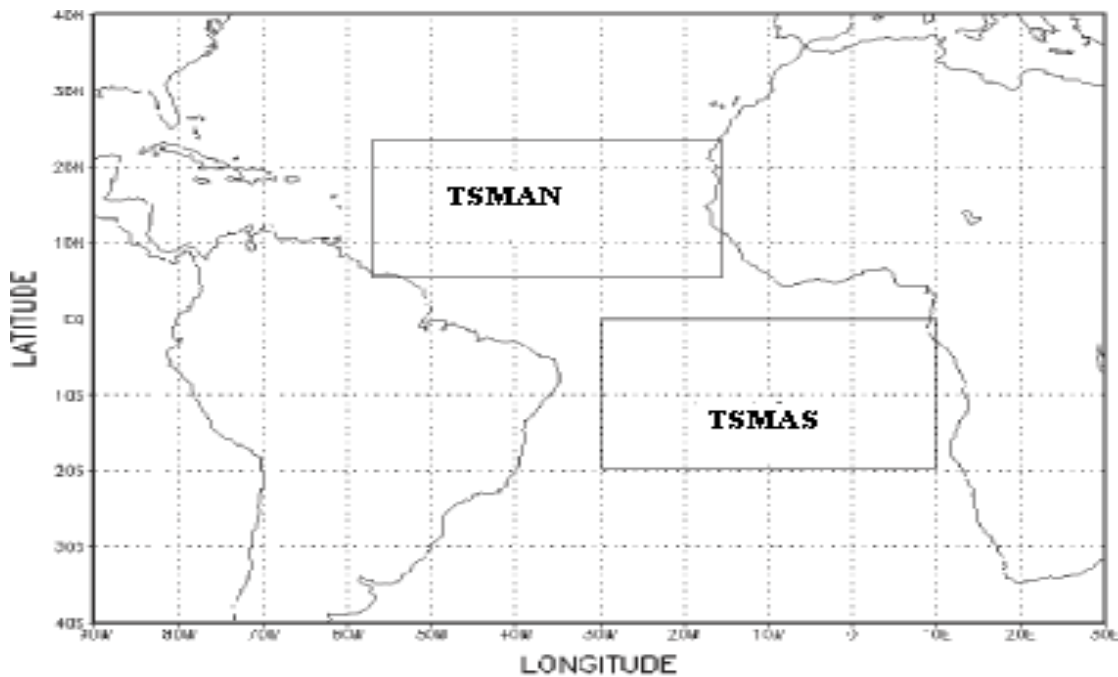


Figura 6. Localização das regiões de temperatura da superfície do mar do Oceano Atlântico.

Correlações Lineares

As correlações lineares foram feitas com os dados pluviométricos de cada grupo com os dados de TSM (Média de TSM de três meses do atlântico Norte (MtsmAN), Média de TSM de três meses do atlântico Sul (MtsmAS), Média do EN12 de três meses (MEN12), Média do EN3 de três meses (MEN3), Média do EN4 de três meses (MEN4), Média do EN34 de três meses (MEN34)), Vento em 850 hPa (Média do índice de ventos alísios central Pacífico Equatorial de três meses (MedWC), Média do índice de ventos alísios leste Pacífico Equatorial de três meses (MedWE), Média do índice de ventos alísios oeste Pacífico Equatorial de três meses

(MedWW) e Pressão da atmosfera (Média da pressão atmosférica à superfície em Darwin de três meses (Mdarwin), Média da pressão atmosférica à superfície em Tahiti de três meses (Mtahiti)).

As correlações foram feitas entre as precipitações médias anuais do período chuvoso de abril a julho com as variáveis defasada de três meses (Novembro a Janeiro) de tsm, vento e pressão.

A partir dessas correlações lineares, identificaram-se as variáveis de TSM, vento e pressão de correlações mais significativas ($> |0,4|$), com relação à precipitação e significativas em $p < 0,05$ com teste estatístico de distribuição t de Student.

Tabela 2. Atrasos de El Niño área 1 e 2 (EN12), El Niño área 3 (EN3), El Niño área 3.4 (EN34), El Niño área 4 (EN4), Pressão à superfície em Darwin (Darwin), Pressão à superfície em Tahiti (Tahiti), Temperatura da superfície do mar no Atlântico Norte (TsmAN), Temperatura da superfície do mar no Atlântico Sul (TsmAS), índice de ventos alísios área central do Pacífico Equatorial (WC), índice de ventos alísios área leste do Pacífico Equatorial (WE) e índice de ventos alísios área oeste do Pacífico Equatorial (WW) com relação à média dos quatro meses de Precipitação.

Atrasos (lags)	Período chuvoso para os três grupos	Médias de EN12, EN3, EN34, EN4, Darwin, Tahiti, TsmAS, TsmAN, WC, WE e WW normalizadas
Atraso	Abril a Julho	Novembro a Janeiro

Análise de Correlação Canônica (ACC)

A Análise Correlação Canônica caracteriza-se por avaliar relações entre dois grupos influenciados, no mínimo, por dois caracteres. De

maneira genérica, considera-se que o primeiro grupo é estabelecido por p caracteres e o segundo por q. O número de correlações canônicas é igual ao menor número de caracteres que constitui um dos grupos (p

ou q) e sua magnitude sempre decresce com a ordem em que são estimadas. Entretanto, o primeiro coeficiente é sempre maior ou igual, em valor absoluto, a qualquer coeficiente de correlação simples ou múltipla, entre os caracteres do primeiro e do segundo grupo.

A ACC é baseada na determinação de variáveis canônicas ortogonais, por isso as variáveis em cada grupo devem ser linearmente independentes. Se esta condição não for satisfeita de início, será necessário descartar as variáveis que são combinações lineares das demais (variáveis redundantes).

Variável Canônica e Correlação Canônica

Sejam dois grupos de variáveis X1 e X2, definidos como sendo que: X1 = (x1, x2, ..., xp) é o vetor de medidas de p características que constituem o grupo I, e X2 = (x1, x2, ..., xp) é o vetor das medidas de q características que constituem o grupo II.

Considerando X1 e X2 juntamente, temos o vetor X((p+q)x1) e a matriz de covariância Σ conforme apresentados nas Equações 3 e 4.

$$X_{((p+q) \times 1)} = \begin{bmatrix} X_1^{(1)} \\ X_2^{(1)} \\ \vdots \\ X_p^{(1)} \\ X_1^{(2)} \\ X_2^{(2)} \\ \vdots \\ X_q^{(2)} \end{bmatrix} = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix} \quad (3)$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \vdots & \Sigma_{12} \\ \dots & \dots & \dots \\ \Sigma_{21} & \vdots & \Sigma_{22} \\ \dots & \dots & \dots \end{bmatrix} \quad (4)$$

Assim, os pq elementos de Σ₁₂ medem a associação entre os dois grupos. O objetivo é concentrar a atenção em algumas poucas combinações lineares de variáveis pertencentes a X¹ e X², ao invés de se trabalhar com todas as covariâncias contidas em Σ₁₂. Adotando-se a notação utilizada por Anderson (1984), sejam U e V,

combinações lineares das variáveis X¹ e X², respectivamente definidas pelas Equações (5), (6), (7), (8) e (9).

$$U = a'X^1 \quad (U = a_1x^1 + a_2x^1 + \dots + a_px^1) \quad (5)$$

$$V = b'X^2 \quad (V = b_1x^2 + b_2x^2 + \dots + b_qx^2) \quad (6)$$

em que, a' = [a₁, a₂, ..., a_p] é o vetor 1 x p de pesos das características do grupo I; e b' = [b₁, b₂, ..., b_q] é o vetor 1 x q de pesos das características do grupo II.

$$\text{Var}(U) = a'.\text{Cov}(X^1).a = a'.\Sigma_{11}.a \quad (7)$$

$$\text{Var}(V) = b'.\text{Cov}(X^2).b = b'.\Sigma_{22}.b \quad (8)$$

$$\text{Cov}(U, V) = a'.\text{Cov}(X^1, X^2).b = a'.\Sigma_{12}.b \quad (9)$$

Define-se como a primeira correlação canônica (r₁) aquela que maximiza a relação entre U₁ e V₁. As funções U₁ e V₁ constituem o primeiro par canônico associado àquela correlação canônica que é expressa pela Equação 10.

$$r_1 = \text{Cor}(U_1, V_1) = \frac{\text{Cov}(U, V)}{\sqrt{\text{Var}(U)\text{Var}(V)}} = \sqrt{\lambda_1} \quad (10)$$

O primeiro par de variáveis canônicas é o par de combinações linear U₁, V₁ tendo variância igual a 1, que maximiza a correlação. O k par de variáveis canônicas é o par de combinações linear U_k, V_k tendo variância igual a 1, que maximiza a correlação entre todas escolhas que não são correlacionadas com os pares de variáveis canônicas anteriores (k-1). Aqui, λ_k são os autovalores. Para os casos em que se utilizam variáveis padronizadas, têm-se Σ₁₁ = R₁₁, Σ₂₂ = R₂₂ e Σ₁₂ = R₁₂, em que R representa uma matriz de correlações.

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \quad (11)$$

O primeiro passo é a determinação dos autovalores que podem ser obtidos pela equação |R₁₁⁻¹ R₁₂ R₂₂⁻¹ R₂₁ - λI|=0. Na sequência calculam-se os autovetores (a e b) que seguem diretamente da equação do autovalor R₁₁⁻¹ R₁₂ R₂₂⁻¹ R₂₁a = λa e R₁₁⁻¹ R₁₂ R₂₂⁻¹ R₂₁b = λb. A primeira correlação canônica entre a combinação linear das características dos grupos I e II é dada por r₁ = √λ₁ em que λ₁ é o maior autovalor da matriz R₁₁⁻¹ R₁₂ R₂₂⁻¹ R₂₁ que é quadrada

e não simétrica. As demais correlações e fatores canônicos foram estimados utilizando-se os autovalores e os autovetores das expressões descritas, de ordem correspondente a p ou q-ésima correlação estimada. O critério adotado para verificar a significância da hipótese de que todas as possíveis correlações canônicas são nulas ($H_0: r_1=r_2... r_s=0$), com s = menor (p ou q) e n=número de observações, foi avaliado pelo teste aproximado de χ^2 (Trugilho et al., 2003), dado pela Equação 12.

$$\chi^2 = - [n - 0,5 (p + q + 3)] \log_e \left[\prod_{i=1}^s (1 - r_i^2) \right] \quad (12)$$

A estatística está associada a pq graus de liberdade. Se a hipótese é rejeitada, testa-se a hipótese $H_0: r_k > 0$ e $r_{k+1} = r_{k+2} = ... r_s = 0$, por meio da Equação 13.

$$\chi^2 = - [n - 0,5 (p + q + 3)] \log_e \left[\prod_{i=k+1}^s (1 - r_i^2) \right] \quad (13)$$

que está associado a χ^2 com (p-k)(p-k) graus de liberdade. Somente as raízes estatisticamente significativas foram usadas para os subseqüentes estudos sobre pares canônicos.

Interpretação das Variáveis Estatísticas Canônicas

Os coeficientes de correlação canônica (r) são as raízes quadradas dos autovalores $r_k = \sqrt{\lambda_k}$ e podem ser interpretados como coeficientes de correlação. Sendo as correlações associadas as variáveis canônicas, estas são chamadas de correlações canônicas. Como índice para controle do grau de correlação canônica entre os dois grupos de variáveis é usual adotar o maior coeficiente de correlação. Entretanto, as outras variáveis canônicas também podem ser significativas e fornecer interpretações.

Os autovalores (λ) podem ser interpretados como a parcela de variância compartilhada entre as respectivas variáveis canônicas.

Aos coeficientes de cada função canônica denominam-se coeficientes canônicos (ou canonical loading) mede a correlação linear simples entre uma variável original observada no grupo I ou II e o par canônico. Esse coeficiente reflete a variância que a variável observada compartilha com o par ou variável canônica. Quanto maior o coeficiente

canônico, maior é a importância da variável no grupo.

O quadrado das cargas de correlação canônica pode ser usado para estimar a parcela de variância compartilhada pelas variáveis canônicas em cada grupo. Se o grupo I tem p variáveis e o grupo II q variáveis, r^2 é a correlação canônica ao quadrado; multiplicando esse valor pela parcela da variância extraída (equações abaixo), obtêm-se a estimativa chamada “redundância”, ou seja, quão superpostas são as variáveis dos grupos.

$$\text{Redundância}_{\text{grupo I}} = [\sum (\text{cargas}_{\text{grupo I}}^2) / p] * r^2$$

$$\text{Redundância}_{\text{grupo II}} = [\sum (\text{cargas}_{\text{grupo II}}^2) / q] * r^2$$

A redundância será útil para avaliar a significância prática das raízes canônicas. Com amostras grandes, correlações canônicas próximas de $r = 0,30$ podem ser estatisticamente significativas. O quadrado deste coeficiente ($r^2 = 0,09$) usado na expressão de redundância indicaria que as raízes canônicas são responsáveis somente por um pouco de variabilidade nas variáveis. A avaliação da significância é subjetiva por natureza. Mas, para avaliar a participação de uma raiz canônica na variância (ou, nas variáveis), é importante a medida de redundância, ou seja, quanto de variabilidade em um grupo é explicado pelo outro.

Quando identificado um outliers, este foi checado e eliminado caso verificado comprometimento dos dados. Os outliers representam valores esporádicos que não representam à base de dados como um todo, mas sim um evento que pode ter sido realmente medido, porém único, ou uma contaminação na medida.

Os valores faltantes (missing) no banco de dados foram estimados. Segundo pode-se estimar até no máximo 25% de valores faltantes de uma variável, sem correr o risco de alterações significativas na tendência média dos dados. Os valores faltantes de precipitação foram estimados por regressão linear (múltipla ou simples) com estações dentro do grupo homogêneo. Todas as variáveis com valores estimados apresentaram um modelo estatístico adequado, ou seja, correlação significativa com as variáveis de referência e distribuição normal e constante dos resíduos. As variáveis que possuíam até 15% dos valores faltantes foram estimadas.

Neste trabalho se utilizou de três métodos para interpretação das variáveis canônicas que são: (1) pesos canônicos (coeficientes padronizados), (2) cargas canônicas (correlações estruturais) e (3) cargas cruzadas canônicas. Para as correlações

canônicas e lineares foi utilizado o software STATÍSTICA 12.0.

Resultados e Discussão

Variabilidade Pluviométrica dos Grupos Homogêneos do Setor Leste do NEB

Na Figura 7 apresenta-se uma comparação entre as séries temporais de precipitação dos três grupos de estações do setor leste do NEB no período chuvoso (abril a julho). Percebe-se que a precipitação diminui do grupo 1 para o grupo 3, ou seja, do Litoral em direção para o Agreste. Na primeira coluna da Tabela 1 vê-se que a média estacional é de 1045,0 mm no Grupo 1, 566,5 mm no Grupo 2 e 316,7 mm no 3.

Detalhadamente pode-se observar ainda na Figura 7 que o grupo 1 apresenta uma maior variabilidade, o que também é evidenciado na Tabela 3 com os desvio-padrões de 214,19, 106,74 e 74,50 mm dos grupos 1, 2 e 3, respectivamente. Isto implica que sistemas atmosféricos atuam com diferentes intensidades em cada grupo e que sistemas

como Brisas atuam mais no grupo 1. As linhas de tendência mostram que houve um leve acréscimo de precipitação no grupo 1. Nos grupos 2 e 3 praticamente não houve mudanças. Hounsou-Gbo et al. (2016) destacaram como principais razões adicionais para a variabilidade sazonal das chuvas no NEB às relacionadas a convecção profunda, linhas de instabilidade atmosférica, ocorrências de brisa e ondas atmosféricas do leste.

Na Figura 8 encontra-se a média mensal dos três grupos homogêneos mostrando que os quatro meses mais chuvosos ocorrem de abril a julho com valores mais alto do litoral para o interior nordestino.

Os sistemas como ondas de leste, brisas, Zona de Convergência Intertropical (ZCIT) e frente fria atuam nas localidades destes grupos com intensidades diferentes, principalmente devido a orografia. O grupo 1 está próximo ao oceano Atlântico, sendo influenciado por ele, com brisas, o grupo 2 está a barlavento do planalto da Borborema, já o grupo 3 está a sota-vento, um setor com subsidência de vento.

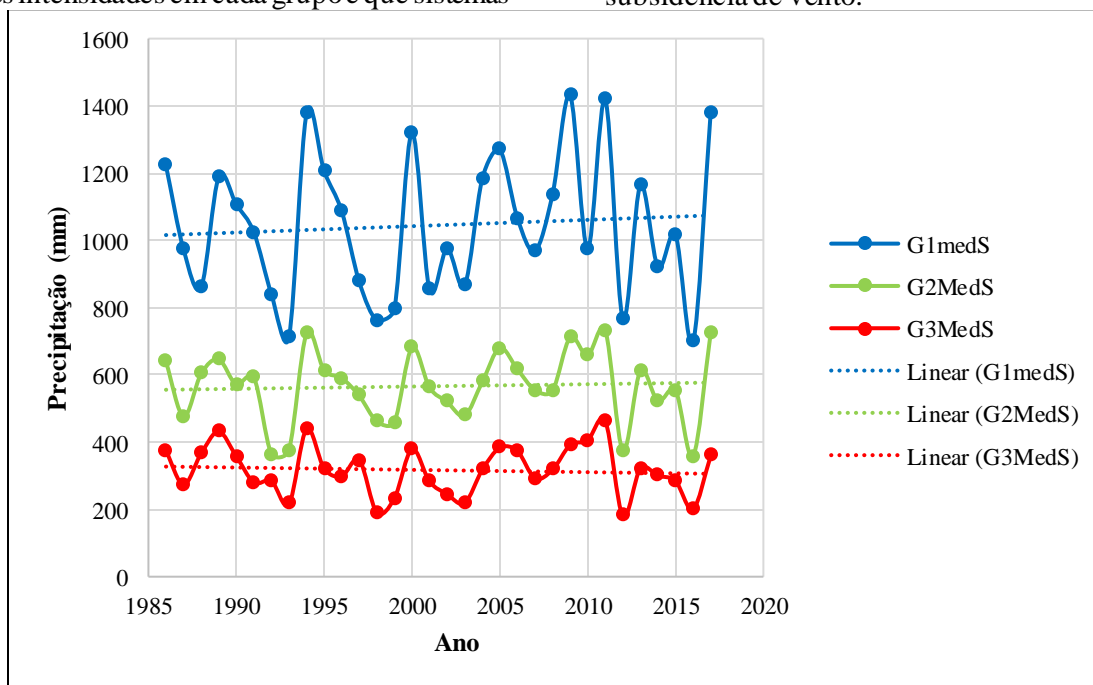


Figura 7. Séries temporais da precipitação para os 3 grupos do Setor Leste do Nordeste do Brasil no período chuvoso (abril a julho) de 1986 a 2017.

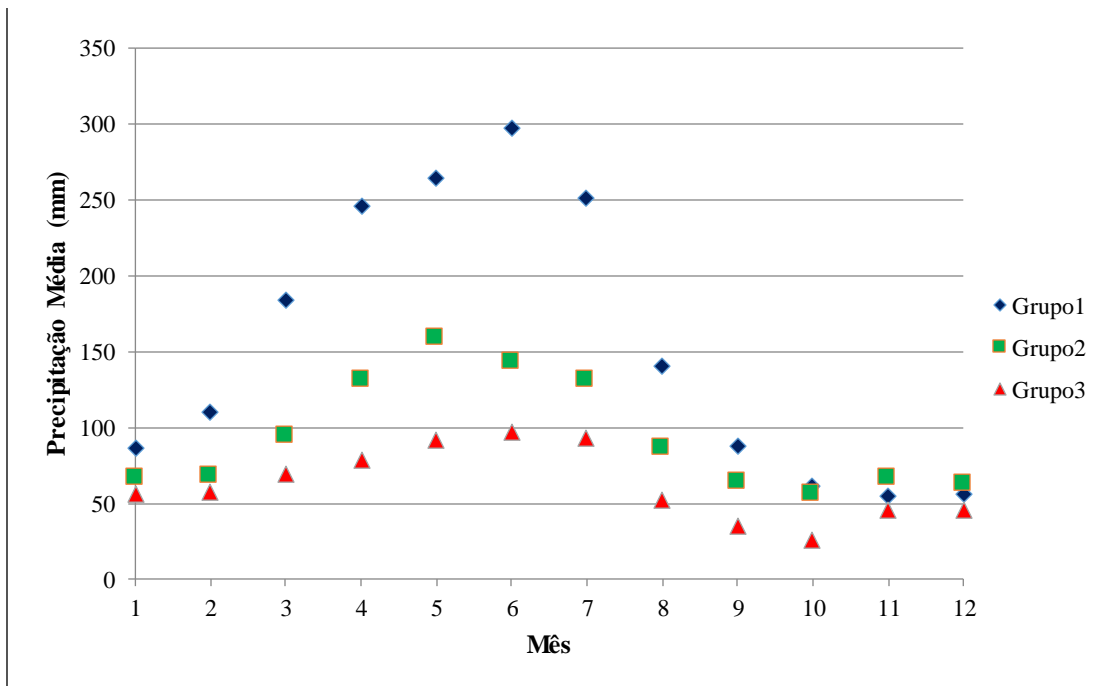


Figura 8. Séries temporais da precipitação para os 3 grupos do Setor Leste do Nordeste do Brasil, médias mensais de 1986 a 2017.

Tabela 3. Média, desvio padrão e correlações lineares de G1medS, G2MedS, G3MedS, MtsmAN, MtsmAS, Mdarwin, Mtahiti, Men12, MEN3, MEN4, MEN34, MedWE, MedWW, MedWC.

	Média	Desvio	G1medS	G2MedS	G3MedS	MtsmAN	MtsmAS	Mdarwin	Mtahiti	Men12	MEN3	MEN4	MEN34	MedWE	MedWW	MedWC
G1medS	1045,0	214,19	1,00	0,89	0,79	0,15	0,33	-0,33	0,35	-0,35	-0,42	-0,32	-0,42	0,37	0,28	0,45
G2MedS	566,47	106,73		1,00	0,87	0,24	0,45	-0,40	0,31	-0,33	-0,38	-0,25	-0,37	0,40	0,31	0,48
G3MedS	316,68	74,50			1,00	0,05	0,32	-0,37	0,30	-0,40	-0,41	-0,28	-0,37	0,48	0,20	0,43
MtsmAN	26,48	0,35				1,00	0,37	-0,10	0,09	0,11	0,11	0,11	0,05	-0,27	0,42	0,15
MtsmAS	25,65	0,33					1,00	0,15	-0,11	0,53	0,44	0,20	0,32	-0,27	0,09	-0,14
Mdarwin	6,89	1,03						1,00	-0,75	0,74	0,84	0,82	0,88	-0,55	-0,80	-0,87
Mtahiti	11,09	1,20							1,00	-0,60	-0,72	-0,82	-0,79	0,49	0,71	0,84
Men12	24,51	0,86								1,00	0,92	0,62	0,82	-0,77	-0,39	-0,79
MEN3	25,76	1,04									1,00	0,84	0,97	-0,76	-0,54	-0,84
MEN4	28,31	0,83										1,00	0,93	-0,53	-0,72	-0,76
MEN34	26,62	1,13											1,00	-0,70	-0,68	-0,86
MedWE	9,62	1,51												1,00	0,18	0,69
MedWW	1,69	2,26													1,00	0,73
MedWC	8,96	2,51														1,00

*Correlações com fonte em negrito são significativas em $p < 0,0500$

Significância Estatística e Prática Para Atraso de 3 Meses

O primeiro teste de significância estatística foi feito para as correlações canônicas de cada uma das três funções canônicas. A única correlação canônica estatisticamente significativa é a primeira (Tabela 4). Além de testes de cada função canônica separadamente, testes multivariados das funções simultaneamente também foram executados. A estatística de teste foi o lambda de Wilks. Na Tabela 2 também, encontra-se detalhada as estatísticas de teste multivariado, as quais indicam que as funções canônicas, tomadas coletivamente, são estatisticamente significantes ao nível aproximado de

0,01, só que individualmente as correlações das funções 2 e 3 com baixa significância, com 0,17 e 0,43, respectivamente. A estatística de teste lambda de Wilks, mostra 0,12 para função 1, enquanto que para as funções 2 e 3, foram encontrados os valores 0,50 e 0,76, respectivamente. Quanto mais próximo de zero o teste lambda de Wilks for, implica que os dois conjuntos de dados são bem correlacionados. Além da significância estatística, as correlações canônicas tinham tamanho suficiente para serem consideradas com significância prática. Na Tabela 4 apresenta-se a análise de correlação canônica, teste multivariado de significância.

Tabela 4. Análise de correlação canônica, teste multivariado de significância.

Função canônica	R	R ²	χ^2	df	p	Lambda Wilks
1	0,87	0,76	54,00	21,00	0,00	0,12
2	0,58	0,34	17,64	12,00	0,13	0,50
3	0,49	0,24	7,10	5,00	0,21	0,76

Análise de Redundância Para Atraso de 3 Meses

O índice de redundância foi calculado para as variáveis estatísticas independente e dependente da primeira função e estão apresentados na Tabela 5. Como pode ser visto o índice de redundância para a variável estatística dependente foi moderado (0,67), pois demonstra que 67% da variação do grupo dependente pode ser explicado pelas variáveis independentes escolhidas. Contudo a variável estatística independente tem um índice de redundância consideravelmente menor (0,14), apesar que, nesse caso, como há uma clara fronteira entre as variáveis dependentes e independentes, esse valor menor pode não ser considerado inesperado ou problemático. A baixa redundância da variável estatística independente resulta da variância compartilhada relativamente baixa na variável estatística independente (0,14), e não do R² canônico.

A partir da análise de redundância e dos testes de significância estatística, a primeira função é aceita.

A análise de redundância para a segunda e terceira função produz resultados bem diferentes (Tabela 6). Primeiro, os R² canônicos são menores 0,3436 e 0,1920, para segunda e terceira função, respectivamente. Além disso, os conjuntos de variáveis têm baixa variância compartilhada na segunda e terceira função (0,0635 e 0,0555 para as variáveis estatísticas dependentes e 0,0583 e 0,01558 para as variáveis estatísticas independentes). A combinação delas com a raiz canônica no índice de redundância produz valores de 0,0218 e 0,0107 para as variáveis estatísticas dependentes e 0,0200 e 0,0299 para a variáveis estatísticas independentes. Logo, a segunda e terceira função não são estatisticamente significantes.

Tabela 5. Cálculo dos índices de redundância para a primeira função canônica.

Variáveis	Carga canônica	Carga canônica quadrada	Carga canônica quadrada média	R ² canônico	Índice de redundância*
Dependentes					
G1MEDS	0,8919	0,7955	-	-	-
G2MEDS	0,9904	0,9612	-	-	-
G3MEDS	0,9310	0,8668	-	-	-

Soma da carga canônica quadrada	-	2,6431	0,8810	0,76	0,6696
Independentes					
MtasmAS	0,2154	0,0464	-	-	-
Mdarwin	0,4896	0,2397	-	-	-
Mtahiti	-0,4621	0,2135	-	-	-
MEN3	0,3628	0,1316	-	-	-
MEN34	-0,4626	0,2140	-	-	-
MedWE	-0,4433	0,1965	-	-	-
MedWC	0,5003	0,2503	-	-	-
Soma da carga canônica quadrada	-	1,2921	0,1846	0,76	0,1402

*Carga quadrada média vezes o R² canônico

Tabela 6. Análise de redundância de variáveis estatísticas dependentes e independentes para ambas as funções canônicas.

Variância padronizada das variáveis dependentes explicada					
Função canônica	Por sua própria variável estatística canônica (variância compartilhada)			Pela variável estatística canônica oposta (redundância)	
	Proporção	Proporção cumulativa	R²	Proporção	Proporção cumulativa
1	0,8810	0,8810	0,7600	0,6696	0,6696
2	0,0635	0,9445	0,3436	0,0218	0,6914
3	0,0555	1,0000	0,1920	0,0107	0,7021

Variância padronizada das variáveis independentes explicada					
Função canônica	Por sua própria variável estatística Canônica (variância compartilhada)			Pela variável estatística canônica Oposta (redundância)	
	Proporção	Proporção cumulativa	R²	Proporção	Proporção cumulativa
1	0,1846	0,1846	0,7600	0,1402	0,1402
2	0,0583	0,2429	0,3436	0,0200	0,1602
3	0,0155	0,2584	0,1920	0,0299	0,1901

Interpretação das Variáveis Estatísticas Canônicas Para Atraso de 3 Meses: Pesos Canônicos

Na Tabela 7 estão apresentados os pesos canônicos padronizados para cada variável estatística canônica: variáveis dependentes e independentes. A magnitude dos pesos representa sua contribuição relativa para com a variável estatística. Com base no tamanho dos pesos, a ordem de contribuição de variáveis independentes para a primeira variável estatística é MEN34 (valor negativo), MedWE,

Mdarwin, Mtahiti, MedWC, MtasmAS e MEN34 (negativo) com valor muito pequeno e a ordem de variáveis dependentes na primeira variável estatística é G2MEDS, G3MEDS e G1MEDS. Observando as funções 2 e 3, percebe-se que estas não seguem as mesmas ordens. Como os pesos canônicos são geralmente variáveis, particularmente em casos de multicolinearidade, devido ao fato de que seu cálculo apenas aperfeiçoa a correlação canônica.

Tabela 7. Pesos canônicos para as duas funções canônicas.

Coefficientes canônicos padronizados para as variáveis independentes	Pesos canônicos		
	Função 1	Função 2	Função 3
G1MEDS	0,0338	-2,1669	-0,2231
G2MEDS	0,7184	1,9647	-1,7118

Coeficientes canônicos padronizados para as variáveis dependentes	G3MEDS	0,2775	-0,0142	2,0348
		Função 1	Função 2	Função 3
MtasmAS	0,0236	0,2287	-0,6491	
Mdarwin	1,1374	-0,2711	-0,0480	
Mtahiti	0,1420	-1,9186	-0,3395	
MEN3	-0,0012	-0,3909	0,4561	
MEN34	-3,0072	1,4311	-2,0468	
MedWE	1,9469	0,9751	2,9001	
MedWC	-0,0477	0,8652	0,3593	

O par para a primeira função das variáveis canônica gerou os seguintes modelos:

$$U1 = 0,0338 G1MEDS + 0,7184 G2MEDS + 0,2775 G3MEDS$$

$$V1 = 0,0236 MtasmAS + 1,1374 Mdarwin + 0,1420 Mtahiti - 0,0012 MEN3 - 3,0072 MEN34 + 1,9469 MedWE - 0,0477 MedWC$$

Onde as variáveis estão padronizadas.

Interpretação das Variáveis Estatísticas Canônicas Para Atraso de 3 Meses: Cargas canônicas

As cargas canônicas para as variáveis estatísticas dependente e independente para as três funções canônicas estão apresentadas na Tabela 8.

Na primeira variável estatística dependente, as variáveis têm cargas que excedem 0,89, o que resulta em variância compartilhada elevada (0,88). Isso indica um alto grau de intercorrelação entre as três variáveis e sugere que estas medidas são representativas.

A carga canônica da primeira função estatística independente tem um padrão diferente, em que as cargas variam de 0,2154 a 0,5003, e três variáveis independentes com cargas negativas variando de -0,4626 a -0,4433. As cinco variáveis com as mais altas cargas na variável estatística independente são MedWC (0,5003), Mdarwin (0,4896), MEN34 (-0,4626), Mtahiti (-0,4621), e MedWE (-0,4433). As variáveis estatísticas em correlação canônica são extraídas para maximizar objetivos preditivos, portanto, ela corresponde mais aos resultados de outras técnicas de dependência. Supõe-se, que a primeira função canônica se aproxima dos resultados de regressão múltipla, e a variável estatística independente representa o conjunto de variáveis que melhor prevê as três medidas dependentes, principalmente a G2MEDS.

É possível notar, que os melhores preditores são os campos de ventos no Pacífico central e este;

pressão atmosférica em Darwin e Tahiti; e os campos de TSM nas áreas do El Niño 3.4. Rao et al. (1993) mostraram a existência de uma relação linear entre as séries dos desvios normalizados de chuvas no leste do NEB e o Índice de Oscilação Sul (IOS). Em geral, os coeficientes de correlação foram negativos e pequenos em magnitude para o período de abril a julho. Estes resultados concordaram com estudos anteriores de Ropelewski e Halpert (1986). Hounsou-Gbo et al. (2016) reforçam que o monitoramento operacional preciso das condições da superfície, principalmente TSM e vento, nas áreas oceânicas mais sensíveis relacionadas à precipitação no NEB possui relevância para a compreensão e previsão de possíveis eventos importantes durante a próxima estação chuvosa nesta região.

Os baixos valores de redundância das variáveis estatísticas são exibidos nas cargas substancialmente menores para ambas as variáveis estatísticas na segunda e terceira função. Logo, a interpretabilidade mais baixa refletida nas cargas menores, juntamente com os baixos valores de redundância, reforça a baixa significância prática da segunda função e terceira função.

Interpretação das Variáveis Estatísticas Canônicas Para Atraso de 3 Meses: Cargas cruzadas canônicas

Na Tabela 8 também estão expressas as cargas cruzadas para as três funções canônicas. Ao analisar a primeira função canônica, percebemos que ambas as variáveis dependentes (G1MEDS, G2MEDS e G3MEDS) exibem correlações fortes com a variável estatística canônica independente (função 1): 0,7722, 0,8575 e 0,8060, respectivamente. Isso reflete uma alta a moderada variância compartilhada entre essas três variáveis. Elevando ao quadrado esses termos, descobrimos o percentual da variância para cada variável explicada pela função 1. Os resultados mostram que 60% da variância em G1MEDS, 73% variância em G2MEDS e 65% da variância em G3MEDS que são explicados pela função 1. Olhando as cargas cruzadas das

variáveis independentes, vemos que as correlações são menores que 0,5 e que as variáveis mais importantes são: MedWC (0,4332), Mdarwin (0,4239), MEN34 (-0,4005), Mtahiti (-0,4001) e MedWE2 (-0,3838), que têm aproximadamente 19, 18, 16, 16 e 15%, respectivamente, das variâncias que são explicadas pela variável estatística dependente. As outras variáveis têm correlações

menores que 0,4, e isto implicam em baixa variância, e, portanto, baixa explicação da variável dependente. A última questão de interpretação foi examinar os sinais das cargas cruzadas. As variáveis independentes MtsmAS, Mdarwin, MedWC e MEN3 têm uma relação positiva, direta, já as variáveis Mtahiti3, MedWE e MEN34 têm relação inversa com a função canônica dependente.

Tabela 8. Estrutura canônica para as três funções canônicas.

		Cargas canônicas		
Correlações entre as variáveis dependentes e suas variáveis estatísticas canônicas		Função 1	Função 2	Função 3
	G1MEDS	0,8919	-0,4347	-0,1247
	G2MEDS	0,9904	0,0293	-0,1349
	G3MEDS	0,9310	-0,0230	0,3643
Correlações entre as variáveis independentes e suas variáveis estatísticas canônicas		Função 1	Função 2	Função 3
	MtsmAS	0,2154	0,2421	-0,7663
	Mdarwin	0,4896	0,2758	-0,4481
	Mtahiti	-0,4621	-0,1098	0,0207
	MEN3	0,3628	-0,2806	0,0015
	MEN34	-0,4626	0,2987	-0,2105
	MedWE	-0,4433	0,3045	-0,0431
	MedWC	0,5003	-0,0293	0,5056
		Cargas cruzadas canônicas		
Correlações entre as variáveis dependentes e variáveis estatísticas canônicas independentes		Função 1	Função 2	Função 3
	G1MEDS	0,7722	-0,2548	-0,0546
	G2MEDS	0,8575	0,0172	-0,0591
	G3MEDS	0,8060	-0,0135	0,1597
Correlações entre as variáveis independentes e variáveis estatísticas canônicas dependentes		Função 1	Função 2	Função 3
	MtsmAS	0,1865	0,1419	-0,3358
	Mdarwin	0,4239	0,1616	-0,1964
	Mtahiti	-0,4001	-0,0644	0,0091
	MEN3	0,3141	-0,1645	0,0007
	MEN34	-0,4005	0,1751	-0,0922
	MedWE	-0,3838	0,1785	-0,0189
	MedWC	0,4332	-0,0172	0,2216

Interpretação das Variáveis Estatísticas Canônicas Para Atraso de 3 Meses: Validação e diagnóstico

A validação das análises de correlação canônica foi feita através de análise de sensibilidade do conjunto de variáveis independentes. Na Tabela 9 pode-se observar o resultado da análise de sensibilidade, na qual as cargas canônicas são examinadas quanto à estabilidade quando variáveis individuais independentes são eliminadas da análise. Como se pode perceber, as cargas canônicas mostraram certa instabilidade e inconsistência

principalmente na variável MEN3. As variáveis que foram eliminadas foram Mdarwin ou MEN3. As correlações canônicas gerais permaneceram estáveis quando se retirou a variável Mdarwin, contudo quando, retirou-se MEN3 houve um decréscimo na correlação.

Cargas canônicas, como pesos, podem estar sujeitas a considerável variabilidade de uma amostra para outra. Essa variabilidade sugere que as cargas, e, portanto, as relações, podem ser específicas da amostra, resultantes de fatores casuais extrínsecos.

Tabela 9. Análise de sensibilidade dos resultados de correlação canônica consequentes da remoção de uma variável independente.

	Resultados após eliminação de		
	Variável completa	Mdarwin	MEN3
Correlação canônica (R)	0,87	0,8665	0,7848
Raiz canônica (R ²)	0,76	0,7508	0,6159
Variável estatística dependente			
Cargas Canônicas			
G1medS	0,8919	0,8664	0,8457
G2MedS	0,9904	0,9928	0,9937
G3MedS	0,9310	0,9208	0,8244
Variância compartilhada	0,8810	0,8614	0,7941
Redundância	0,6695	0,6468	0,4891
Variável estatística independente			
Cargas Canônicas			
MtasmAS	0,2154	0,5037	0,5999
Mdarwin	0,4896	-	-0,5064
Mtahiti	-0,4621	0,3496	0,3658
MEN3	0,3628	0,3503	-
MEN34	-0,4626	-0,4471	-0,4452
MedWE	-0,4433	-0,4292	0,4737
MedWC	0,5003	0,4928	0,6058
Variância compartilhada	0,1846	0,2180	0,2566
Redundância	0,1402	0,1637	0,1581

Conclusões

Há poucas pesquisas que trata de previsão climática para o setor leste do NEB, portanto, este trabalho contribui com o entendimento de como preditores nos oceanos Pacífico e Atlântico influenciam as chuvas deste setor no decorrer de três meses de atrasos.

A precipitação diminui do Litoral para o interior devido principalmente aos sistemas de brisas e distúrbios de leste que atingem o Litoral e Zona da Mata. Também existe uma diminuição da precipitação ao oeste do Planalto da Borborema onde se encontrou as menores quantidades de chuvas, que se deve a subsidência dos ventos alísios.

O Grupo 2 de precipitação pluvial que se encontra a montante do Planalto da Borborema tem as melhores respostas quanto aos preditores nos três atrasos.

No atraso os principais preditores para os grupos homogêneos são, por ordem de maior importância decrescente: MedWC, Mdarwin, MEN34, Mtahiti e MedWE. Percebe-se deste atraso que a principal influência está no Pacífico, no ENOS.

Agradecimentos

Ao Instituto Nacional de Meteorologia (INMET) e Agência Pernambucana de Águas e

Climas (APAC) pela disponibilidade dos dados meteorológicos de precipitação pluviométrica. A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 pela bolsa de doutorado.

Referências

- Alves, J.M.B., Repelli, C.A., Mello, N.S., 1993. A pré-estação chuvosa do setor norte do Nordeste Brasileiro e sua relação com a temperatura dos oceanos adjacentes. *Revista Brasileira de Meteorologia* 8, 22-30. Disponível: http://www.rbmet.org.br/port/revista/revista_artigo.php?id_artigo=419
- Anderson, T.W., 1984. *An Introduction to Multivariate Statistical Analysis*. 2th edition. John Wiley. New York. 439-449.
- Andreoli, R.V., Oliveira, S.S., Kayano, M.T., Viegas, J., Souza, R.A.F., Candido, L.A., 2016. The influence of different el Niño types on the south American rainfall. *International Journal of Climatology* 37, 1374-1390. <https://doi.org/10.1002/joc.4783>
- Aragão, J.O.R., Roucou, P., Harzallah, A., Fontaine, B; Janicot, S., 1994. Variabilité atmosphérique sur le Nordeste brésilien dans le modèle de circulation générale du LMD (1970-1988).

- Publications de l'Association Internationale de Climatologie 7, 432-438.
- Barbosa, H.A., Lakshmi Kumar, T., Paredes, F., Elliott, S., Ayuga, J.G., 2018. Assessment of Caatinga response to drought using Meteorological SEVIRI Normalized Difference Vegetation Index (2008-2016). *ISPRS Journal of Photogrammetry and Remote Sensing* 148, 235-252. <https://doi.org/10.1016/j.isprsjprs.2018.12.014>
- Barnett, T.P., Preisendorfer, R., 1987. Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Monthly Weather Review* 115, 1825-1849. [https://doi.org/10.1175/1520-0493\(1987\)115<1825:OALOMA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<1825:OALOMA>2.0.CO;2)
- Builes-Jaramillo, A., Marwan, N., Poveda, G., Kurths, J., 2018. Nonlinear interactions between the Amazon River basin and the Tropical North Atlantic at interannual timescales. *Climate Dynamics* 50, 2951-2969. <https://doi.org/10.1007/s00382-017-3785-8>
- Capotondi, A., Wittenberg, A.T., Newman, M., Di Lorenzo, E., Yu, J., et al., ... Yeh, S.W., 2015. Understanding ENSO diversity. *American Meteorological Society* 921-938. <https://doi.org/10.1175/BAMS-D-13-00117.1>
- Dias, F.J.S., Castro, B.M., Lacerda, L.D., 2018. Tidal and low-frequency currents off the Jaguaribe River estuary (4° S, 37° 4' W), Northeastern Brazil. *Ocean Dynamics* 68, 967-985. <https://doi.org/10.1007/s10236-018-1172-6>
- Fedorova, N., Levit, V., Campos, A.M.V., 2018. Brazilian Northeast Jet Stream: association with synoptic-scale systems. *Meteorological Applications* 25, 261-268. <https://doi.org/10.1002/met.1693>
- Forootan, E., Khandu, Awange, J.L., Schumacher, M., Anyah, R.O., Van Dijk, A.I.J.M., Kusche, J., 2016. Quantifying the impacts of ENSO and IOD on rain gauge and remotely sensed precipitation products over Australia. *Remote Sensing of Environment* 172, 50-66. <https://doi.org/10.1016/j.rse.2015.10.027>
- Gomes, H.B., Ambrizzi, T., Herdies, D.L., Hodges, K., Pontes da Silva, B.F., 2015. Easterly wave Disturbances over northeast Brazil: an observational analysis. *Advances in Meteorology* 2015, 1-20. <https://doi.org/10.1155/2015/176238>
- Gonzalez, R.A., Andreoli, R.V., Candido, L.A., Kayano, M.T., Souza, R.A.F., 2013. A influência do evento El Niño – Oscilação Sul e Atlântico Equatorial na precipitação sobre as regiões norte e nordeste da América do Sul. *Acta Amazônica* 43, 469-480. <http://dx.doi.org/10.1590/S0044-59672013000400009>
- Grimm, A.M., Tedeschi, R.G., 2009. ENSO and Extreme Rainfall Events in South America. *Journal of Climate* 22, 1589-1609. <https://doi.org/10.1175/2008JCLI2429.1>
- Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., 1998. *Multivariate data analysis*. Upper Saddle River, NJ: Prentice hall. 5, 207-219.
- Harzallah, A., Aragão, J.O.R., Sadourny, R., 1996. Interannual rainfall variability in Northeast Brazil: Observation and model simulation. *International Journal of Climatology* 16, 861-878. [https://doi.org/10.1002/\(SICI\)1097-0088\(199608\)16:8<861::AID-JOC59>3.0.CO;2-D](https://doi.org/10.1002/(SICI)1097-0088(199608)16:8<861::AID-JOC59>3.0.CO;2-D)
- Hastenrath, S., Greischar, L., 1993. Further work on the prediction of Northeast Brazil rainfall anomalies. *Journal of Climate* 6, 743-758. [https://doi.org/10.1175/1520-0442\(1993\)006<0743:FWOTPO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006<0743:FWOTPO>2.0.CO;2)
- Hounsou-Gbo G.A., Araujo, M., Bourlès, B., Veleda, D., Servain, J., 2015. Tropical Atlantic Contributions to Strong Rainfall Variability Along the Northeast Brazilian Coast. *Advances in Meteorology* 2015, 1-13. <http://dx.doi.org/10.1155/2015/902084>
- Hounsou-Gbo. G.A., Servain, J., Araujo, M., Martins, E.S., Bourles, B., Canaix, G., 2016. Oceanic Indices for Forecasting Seasonal Rainfall over the Northern Part of Brazilian Northeast. *Am J Clim Change* 5, 261-274. <http://dx.doi.org/10.4236/ajcc.2016.52022>
- Kayano, M.T., Andreoli, R.V., Souza, R.A.F., 2011. Evolving anomalous SST patterns leading to ENSO extremes: relations between the tropical Pacific and Atlantic Oceans and the influence on the South American rainfall. *International Journal of Climatology* 31, 1119-1134. <https://doi.org/10.1002/joc.2135>
- Kayano, M.T., Capistrano, V.B., 2014. How the Atlantic multidecadal oscillation (AMO) modifies the ENSO influence on the South American rainfall. *International Journal of Climatology* 34, 162-178. <https://doi.org/10.1002/joc.3674>
- L'Heureux, M.L., Tippett, M.K., Barnston, A.G., 2015. Characterizing ENSO coupled variability and its impact on North American seasonal precipitation and temperature. *Journal Climate* 28, 4231-4245. <https://doi.org/10.1175/JCLI-D-14-00508.1>

- Marengo, J.A., Alves, L.M., Alvala, R., Cunha, A.P., Brito, S., Moraes, O.L., 2018. Climatic characteristics of the 2010-2016 drought in the semiarid Northeast Brazil region. *Anais da Academia Brasileira de Ciências* 90, 1973-1985. <http://dx.doi.org/10.1590/0001-3765201720170206>
- Mingoti, S.A., 2005. Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada. Belo Horizonte: Editora UFMG. 297 p.
- Moraes, M.D.C., Oliveira, F.P., Coutinho, M.D., 2015. One case of simulation of Upper Tropospheric Cyclonic vortex in the Brazil Northeast – impact in the convection parameterization. *Journal of Hyperspectral Remote Sensing* 1, 27-44. <https://doi.org/10.29150/jhrs.v5.1.p027-044>
- Moura, A.D., 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)
- Moura, G.B.A., Aragão, J.O.R., Lacerda, F.F., Passavante, J.Z.O., 2000. Relação entre a precipitação no setor leste do Nordeste do Brasil e a temperatura da superfície nos oceanos Atlântico (área do Dipolo) e Pacífico. *Revista Brasileira de Engenharia Agrícola e Ambiental* 4, 247-251. <http://dx.doi.org/10.1590/S1415-43662000000200019>
- Moura, G.B.A., Aragão, J.O.R., Melo, J.S.P., Silva, A.P.N., Giongo, P.R., Lacerda, F.F., 2009. Relação entre a precipitação do leste do Nordeste do Brasil e a temperatura dos oceanos. *Revista Brasileira de Engenharia Agrícola e Ambiental* 14, 462-469. <http://dx.doi.org/10.1590/S1415-43662009000400014>
- Nobre, P., Shukla, J., 1996. Variations of sea surface temperature, wind stress and rainfall over the tropical Atlantic and South America. *Journal of Climate* 9, 2464-2479. [https://doi.org/10.1175/1520-0442\(1996\)009<2464:VOSSTW>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<2464:VOSSTW>2.0.CO;2)
- Nóbrega, J.N., Santos, C.A.C., Gomes, O.M.G., Bezerra, B.G., Brito, J.I.B., 2014. Eventos extremos de precipitação nas mesorregiões da Paraíba e suas relações com a tsm dos oceanos tropicais. *Revista Brasileira de Meteorologia* 29, 197-208. <http://dx.doi.org/10.1590/S0102-77862014000200005>
- Rana, S., Renwick, J., McGregor, J., Ankita Singh, A., 2018. Seasonal Prediction of Winter Precipitation Anomalies over Central Southwest Asia: A Canonical Correlation Analysis Approach. *Journal of Climate* 31, 727-741. <https://doi.org/10.1175/JCLI-D-17-0131.1>
- Rao, V.B.; Lima, M.C., Franchito, S.H., 1993. Seasonal and interannual variations of rainfall over eastern Northeast Brazil. *Journal of Climate* 6, 1754-1763. [https://doi.org/10.1175/1520-0442\(1993\)006<1754:SAIVOR>2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006<1754:SAIVOR>2.0.CO;2)
- Ropelewski, C.F., Halpert, M.S., 1986. North American precipitation and temperature patterns associated with the El Niño Southern Oscillation (ENSO). *Monthly Weather Review* 114, 2352-2362. [https://doi.org/10.1175/1520-0493\(1986\)114<2352:NAPATP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1986)114<2352:NAPATP>2.0.CO;2)
- Salgueiro, J.H.P.B., Montenegro, S.M.G.L., Pinto, E.J.A., Silva, B.B., Souza, W.M., Oliveira, L.M.M., 2016. Influence of oceanic-atmospheric interactions on extreme events of daily rainfall in the Sub-basin 39 located in Northeastern Brazil. *Revista Brasileira de Recursos Hídricos* 21, 685-693. <http://dx.doi.org/10.1590/2318-0331.011616023>
- Servain, J., 1991. Simple Climatic Indices for the Tropical Atlantic Ocean and some applications. *Journal of Geophysical Research* 96, 137-146. <https://doi.org/10.1029/91JC01046>
- Silva, M.T., Costa, S.C.F.E., Gomes Filho, M.F., Lucena, D.B., 2011. Estudo da Temperatura da Superfície do Mar para os Oceanos Atlântico e Pacífico Utilizando a Técnica de Análises de Componente Principal e de Agrupamento. *Revista Brasileira de Geografia Física* 4, 264-277. <https://doi.org/10.26848/rbgf.v4.2.p264-277>
- Trugilho, P.F., Lima, J.T., Mori, F.A., 2003. Correlação canônica das características químicas e físicas da madeira de clones de *Eucalyptus grandis* e *Eucalyptus saligna*. *Cerne* 9, 81-91. Disponível: <<https://www.redalyc.org/articulo.oa?id=74409108>>.
- Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. *Journal American Statistical Association* 58, 236-244. <http://dx.doi.org/10.1080/01621459.1963.10500845>
- Ward, M.N., Folland, C.K., 1991. Prediction of seasonal rainfall in the north nordeste of Brazil using eigenvectors of sea-surface temperature. *International Journal of Climatology* 11, 711-743. <https://doi.org/10.1002/joc.3370110703>.