



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS EXATAS E DA TERRA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS
CLIMÁTICAS

PPGCC

TESE DE DOUTORADO

**MODELOS DE DISPERSÃO PARA EXTREMOS DE PRECIPITAÇÃO. ESTUDO DE
CASO: O NORDESTE DO BRASIL.**

WASHINGTON LUIZ FÉLIX CORREIA FILHO

Natal
2014

WASHINGTON LUIZ FÉLIX CORREIA FILHO

MODELOS DE DISPERSÃO PARA EXTREMOS DE PRECIPITAÇÃO. ESTUDO DE
CASO: O NORDESTE DO BRASIL.

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciências Climáticas, da Universidade Federal do Rio Grande do Norte, como parte dos requisitos para obtenção do título de Doutor em Ciências Climáticas.

Orientador: Prof. Dr. Paulo Sérgio Lucio.

Co-orientadora: Prof. Dra. Maria Helena Constantino Spyrides.

Comissão examinadora

Prof. Dr. Francisco Javier Sigró Rodríguez (URV-Espanha).

Prof. Dr. Afrânio Márcio Corrêa Vieira (UFSCar).

Prof. Dr. David Mendes (UFRN).

Profª. Dra. Kellen Carla Lima (UFRN).

Natal

2014

MODELOS DE DISPERSÃO PARA EXTREMOS DE PRECIPITAÇÃO. ESTUDO DE
CASO: O NORDESTE DO BRASIL.

WASHINGTON LUIZ FÉLIX CORREIA FILHO

Esta Tese foi considerada adequada para a obtenção do Título de DOUTOR EM CIÊNCIAS CLIMÁTICAS, sendo aprovada em sua forma final.

Natal, 27 de Março de 2014.

Orientador: Prof. Dr. Paulo Sérgio Lucio
Universidade Federal do Rio Grande do Norte, RN.

Co-orientadora: Profa. Dra. Maria Helena Constantino Spyrides
Universidade Federal do Rio Grande do Norte, RN.

Membro Externo: Prof. Dr. Francisco Javier Sigró Rodríguez
University Rovira i Virgili, Tarragona, Espanha.

Membro Externo: Prof. Dr. Afrânio Márcio Corrêa Vieira
Universidade Federal de São Carlos, SP.

Membro Interno: Prof. Dr. David Mendes
Universidade Federal do Rio Grande do Norte, RN.

Membro Interno: Profa. Dra. Kellen Carla Lima
Universidade Federal do Rio Grande do Norte, RN.

DEDICATÓRIA

Aos meus pais Washington e Leni,
que confiaram, apoiaram e acreditaram
em todos os momentos.

minha irmã Hortência
e minha tia Creuza.

DEDICO.

AGRADECIMENTOS

Agradeço primeiramente a Deus pelas oportunidades, e também por mostrar que podemos crescer e amadurecer em momentos difíceis.

À minha família pelo apoio em todos os momentos, e também pela compreensão da ausência ao longo deste doutorado.

Ao Programa de Pós-Graduação em Ciências Climáticas (PPGCC), pela oportunidade em fazer parte deste doutorado, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão de bolsa de estudo.

Aos meus orientadores, Professores Dr. Paulo Sérgio Lucio e Dra. Maria Helena Constantino Spyrides pelo apoio, paciência, parceria e confiança em meu trabalho, pelos desafios ofertados e aperfeiçoamento profissional ao longo deste doutorado.

Aos Professores Dr. David Mendes e Dra. Rosane Chaves pelas sugestões e críticas positivas na qualificação, importantes para o desenvolvimento e melhoramento deste trabalho.

Aos professores do PPGCC, pelas disciplinas ministradas, contribuindo com o aperfeiçoamento científico e profissional ao longo deste doutorado.

Aos Professores Dra. Kellen Carla Lima, Dr. Afrânio Márcio Corrêa Vieira e Dr. Francisco Javier Sigró Rodríguez por gentilmente terem aceitado a participar da comissão examinadora de Tese.

A Mônica Cristina Damiano Mendes pelas valorosas contribuições e pelo apoio em meu trabalho de doutoramento.

Aos colegas do PPGCC pelo convívio ao longo deste doutoramento, em especial aos amigos Alexandre B. Lopo, Alexandre S. Santos, Francisco Raimundo da Silva, George U. Pedra, Jaime R. Coarati, Marcos Aurélio F. Santos, Naurinete J. C. Barreto, Ronabson C. Fernandes pela amizade, paciência, cooperação e por ter compartilhado momentos ao longo deste doutorado.

Aos meus amigos de Campina Grande, Maceió e Natal, em especial Anderson Costa, Anderson Guedes, Hydalin e Eliângela Souza, Mayara Costa, Leydson Galvêncio, Thaís Cordeiro, Jane Azevedo, Alane Saldanha, Marcelo Medeiros, Samarha Santos, Fernanda Silva, pela amizade e pelos bons momentos convividos ao longo destes últimos anos.

RESUMO

Nesta tese utilizaram-se quatro diferentes métodos com o objetivo de diagnosticar os extremos de precipitação no Nordeste do Brasil (NEB): Modelos Lineares Generalizados via regressão logística e de Poisson, análise da teoria dos valores extremos via distribuições generalizada de valores extremos (GEV) e generalizada de Pareto (GPD) e os Modelos Vetoriais Lineares Generalizados via GEV (MVLG-GEV). Os modelos de regressão logística e de Poisson foram utilizados para identificar as interações entre os extremos de precipitação e as demais variáveis com base nas razões de chances e riscos relativos. Constatou-se que a radiação de longa foi a variável indicadora para a ocorrência de extremos de precipitação no leste, norte e semiárido do NEB, e a umidade relativa foi verificada no sul do NEB. As distribuições GEV e GPD (baseados no percentil 95) mostraram que os parâmetros de localização e de escala apresentaram máximos na costa leste, norte do NEB, a GEV verificou um núcleo de máxima a oeste de Pernambuco, influenciados por sistemas meteorológicos e topografia. Os parâmetros de forma de GEV e GPD verificou na maioria das regiões que os dados foram ajustados pelas distribuições de Weibull negativa e Beta ($\xi < 0$), respectivamente. Os níveis e períodos de retorno da GEV (GPD) mostraram que no norte do Maranhão (centro da Bahia) poderá ocorrer pelo menos um evento extremo de precipitação igual ou superior a 160,9 mm/dia (192,3 mm/dia) nos próximos 30 anos. O modelo MVLG-GEV constatou que as componentes zonal e meridional do vento, evaporação e temperatura à superfície mar do Atlântico e Pacífico subsidiaram os extremos de precipitação. Os parâmetros da distribuição GEV apresentaram os seguintes resultados: a) localização (μ), o maior valor foi de $88,26 \pm 6,42$ mm no norte do Maranhão; b) escala (σ), a maioria das regiões exibiram valores positivos, exceto no sul do Maranhão; e c) forma (ξ), a maioria das regiões selecionadas foram ajustadas pela distribuição de Weibull negativa ($\xi < 0$). O sul do Maranhão e da Bahia possuem maior precisão. Quanto aos níveis de retorno, estimou-se que no centro da Bahia poderá ocorrer um evento extremo de precipitação igual ou superior a 571,2 mm/dia nos próximos 30 anos.

Palavras chave: Modelos Vetoriais Lineares Generalizados, Regressão logística, Regressão de Poisson, radiação de onda longa, evaporação.

ABSTRACT

In this thesis used four different methods in order to diagnose the precipitation extremes on Northeastern Brazil (NEB): Generalized Linear Models via logistic regression and Poisson, extreme value theory analysis via generalized extreme value (GEV) and generalized Pareto (GPD) distributions and Vectorial Generalized Linear Models via GEV (MVLG-GEV). The logistic regression and Poisson models were used to identify the interactions between the precipitation extremes and other variables based on the odds ratios and relative risks. It was found that the outgoing longwave radiation was the indicator variable for the occurrence of extreme precipitation on eastern, northern and semiarid NEB, and the relative humidity was verified on southern NEB. The GEV and GPD distributions (based on the 95th percentile) showed that the location and scale parameters were presented the maximum on the eastern and northern coast NEB, the GEV verified a maximum core on western of Pernambuco influenced by weather systems and topography. The GEV and GPD shape parameter, for most regions the data fitted by Weibull negative and Beta distributions ($\xi < 0$), respectively. The levels and return periods of GEV (GPD) on northern Maranhão (centern of Bahia) may occur at least an extreme precipitation event exceeding over of 160.9 mm /day (192.3 mm / day) on next 30 years. The MVLG-GEV model found that the zonal and meridional wind components, evaporation and Atlantic and Pacific sea surface temperature boost the precipitation extremes. The GEV parameters show the following results: a) location (μ), the highest value was 88.26 ± 6.42 mm on northern Maranhão; b) scale (σ), most regions showed positive values, except on southern of Maranhão; and c) shape (ξ), most of the selected regions were adjusted by the Weibull negative distribution ($\xi < 0$). The southern Maranhão and southern Bahia have greater accuracy. The level period, it was estimated that the centern of Bahia may occur at least an extreme precipitation event equal to or exceeding over 571.2 mm/day on next 30 years.

Keywords: VGLM, logistic regression, Poisson regression, outgoing longwave radiation, precipitation inter-annual variability.

SUMÁRIO

LISTA DE FIGURAS	ix
LISTA DE TABELAS	x
LISTA DE SIGLAS	xii
1. INTRODUÇÃO.....	17
1.1. REVISÃO DE LITERATURA	19
1.1.1. Definição de extremos no contexto climático	19
1.1.2. Fatores que subsidiam a ocorrência dos extremos de precipitação	20
1.1.3. Métodos estatísticos para diagnóstico de extremos de precipitação.....	22
CAPITULO I – Precipitation Extremes Analysis over the Brazilian Northeast via Logistic Regression	25
CAPITULO II – Analysis of Precipitation Extremes on Northeast Brazil via Poisson Regression	42
CAPITULO III – Caracterização dos extremos climáticos no Nordeste do Brasil	61
CAPITULO IV – Diagnóstico dos extremos de precipitação no Nordeste do Brasil via Modelos Vetoriais Lineares Generalizados	77
CONCLUSÕES.....	97
APÊNDICE – METODOLOGIA.....	101
REFERÊNCIA	117

LISTA DE FIGURAS

FIGURAS DO CAPITULO I

Figure 1 – Grid points distribution on NEB which was divided in four sub-regions (symbols): NNEB (circle); ENEB (cross); SANEB (square) and SNEB (triangle).	28
Figure 2 – Meteorological variables that influenced on NEB precipitation extremes by logistic regression models. Each sub-region with the highest OR values.	32
Figure 3 – ROC curve for logistic regression models. AUC values on right bottom for NEB sub-regions: NNEB (solid line), ENEB (dashed line), SANEB (dotted line) and SNEB (dashed and dotted line).....	33

FIGURAS DO CAPITULO II

Figure 1 – Grid points distribution on NEB which was divided in four sub-regions (symbols): NNEB (circle); ENEB (cross); SANEB (square) and SNEB (triangle).	46
Figure 2 – Meteorological variables that influenced on NEB precipitation extremes by Poisson regression models. Each sub-region with the highest RR values.....	49
Figure 3 – ROC curve for Poisson regression models. AUC values on right bottom for NEB sub-regions: NNEB (solid line), ENEB (dashed line), SANEB (dotted line) and SNEB (dashed and dotted line).....	51

FIGURAS DO CAPITULO III

Figura 1 – Localização dos pontos das oito regiões selecionadas para avaliação do extremos no Nordeste do Brasil.	65
Figura 2 – Diagnóstico dos extremos de precipitação no NEB a partir dos parâmetros de localização (mm) (a-b), de escala (mm) (c-d) e de forma (e-f) das distribuições GEV e GPD.	69

FIGURAS DO CAPITULO IV

Figura 1 – Localização das oito regiões selecionadas para análise de extremos no Nordeste do Brasil via modelo MVLG-GEV.	81
Figura 2 – Níveis (mm/ dia) e períodos (T em anos) de retorno estimados dos modelos MVLG-GEV gerados para as oito regiões selecionadas.	90

LISTA DE TABELAS

TABELAS DO CAPITULO I

Table 1 – Observed frequencies for precipitation extremes	30
Table 2 – Number of extreme precipitation cases (n° cases in %) of four NEB sub-regions... 31	
Table 3 – Logistic regression models for NNEB and SANEB. Coefficient regression (Coeff), Standard Error (SE), p-value, Odds Ratio (95% CI) and Goodness-of-fit (AIC test, Deviance residual and degrees of freedom (df)).....	34
Table 4 – Logistic regression models for ENEB and SNEB. Coefficient regression (Coeff), Standard Error (SE), p-value, Odds Ratio (95% CI) and Goodness-of-fit (AIC test, Deviance residual and degrees of freedom (df)).....	35

TABELAS DO CAPITULO II

Table 1 – Observed frequencies for precipitation extremes	48
Table 2 – Poisson regression models for NNEB and SANEB. Coefficient regression (Coeff), Standard Error (SE), p-value, Relative Risk (95% CI) and Goodness-of-fit (AIC test, deviance residual and degrees of freedom (df)).....	52
Table 3 – Poisson regression models for ENEB and SNEB. Coefficient regression (Coeff), Standard Error (SE), p-value, Relative Risk (95% CI) and Goodness-of-fit (AIC test, deviance residual and degrees of freedom (df)).....	53

TABELAS DO CAPITULO III

Tabela 1 – Estatística descritiva das distribuições GEV e GPD para as regiões: NMA, NCE, SMA, OPE, NAL, CBA, OBA e SBA. Parâmetros de localização (μ) (mm), de escala (σ) e forma (ξ)..	70
Tabela 2 – Níveis (mm/dia) e períodos (T em anos) de retornos estimados para as distribuições GEV e GPD.....	71

TABELAS DO CAPITULO IV

Tabela 1 – Teste da log-verossimilhança negativa, graus de liberdade (GL) e AIC dos modelos MVLG-GEV para as regiões: NMA, NCE, SMA, OPE, NAL, CBA, OBA e SBA.	87
Tabela 2 – Coeficiente de regressão (Coef), desvio padrão (DP) e valor- p para o NMA, NCE, SMA e OPE.	87
Tabela 3 – Semelhante à Tabela 2, mas para o NAL, CBA, OBA e SBA.	88
Tabela 4 – Níveis (mm/dia) e períodos (T em anos) de retorno estimados dos modelos MVLG-GEV gerados para as oito regiões selecionadas ao longo do NEB.	89

LISTA DE SIGLAS

AB – Alta da Bolívia

ACP – Análise de Componentes Principais

AIC – *Akaike Information Criterion*

AO – *Antarctic Oscillation*

AS – América do Sul

ASAS – Alta Subtropical do Atlântico Sul

ATL – *Atlantic Ocean*

AUC – *Area Under Curve*

BM – *Block Maximum*

CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CCF – *Cross Correlation Function*

CCI/CLIVAR/JCOMM – *Climate Change Initiative / Climate Variability and Predictability / Joint Technical Commission for Oceanography and Marine Meteorology*

CCM – Complexo Convectivo de Mesoescala (*Mesoscale Convective Complex*)

CI – *Confidence Interval*

COEF – Coeficiente

COEFF – *Coefficient*

CPC – *Climate Prediction Center*

CompU – *Zonal Wind Component*

CompV – *Meridional Wind component*

CNPq – Conselho Nacional de Pesquisa e Desenvolvimento Científico e Tecnológico

DF – *Degree of Freedom*

DOL – Distúrbios Ondulatórios de Leste

DP – Desvio Padrão

ECMWF - European Centre for Medium Weather Forecasts

ENOS – El Niño-Oscilação Sul

ENEB – East Northeast Brazil

ENSO – *El Niño-Southern Oscillation*

ESRL – *Earth System Research Laboratory*

ETCCDI - *Expert Team on Climate Change Detection and Indices*

EWD – *Easterly Waves Disturbances*

EV – *Evaporation*

EVT – *Extreme Value Theory*

FS – Frontal Systems

GEV – *Generalized Extreme Value Distribution*

GL – Graus de liberdade

GLM – *Generalized Linear Models*

GOF – *Goodness-of-fit*

GPD – *Generalized Pareto Distribution*

GUST – *Gust Wind*

HS – Hemisfério Sul

IC – Intervalo de Confiança

IRLS – *Iteratively Reweighted Least Squares*

ITCZ – *Intertropical Convergence Zone*

JCR – *Journal Citation Report*

LNEB – Leste do Nordeste do Brasil

MCC – *Mesoscale Convective Complex*

MCS – *Mesoscale Convective System*

MJO – *Madden-Julian Oscillation*

MLG – Modelos Lineares Generalizados

MVAG – Modelos Vetoriais Aditivos Generalizados

MVLG – Modelos Vetoriais Lineares Generalizados

MVLG-GEV – Modelos Vetoriais Lineares Generalizados via Distribuição GEV

NAO – *North Atlantic Oscillation*

NEB – *Nordeste do Brasil (Northeast Brazil)*

NINO – *Pacific Ocean*

NNEB – *Norte do Nordeste do Brasil (North NEB)*

NOAA – *National Oceanographic and Atmospheric Administration*

OMJ – *Oscilação Madden-Julian*

OLR – *Outgoing Longwave Radiation*

OR – *Odds Ratio*

PAC – *Pacific Ocean*

PCA – *Principal Component Analysis*

PPGCC – *Programa de Pós-Graduação em Ciências Climáticas*

POT – *Peaks Over Thresholds*

RC – *Razão de chance*

RH – *Relative Humidity*

ROL – *Radiação de onda longa*

ROC – *Receiver Operating Characteristics graph*

RR – *Relative Risk*

SACZ – *South Atlantic Convergence Zone*

SANEB – *Semiárido do Nordeste do Brasil*

SASA – *South America Subtropical Anticyclone*

SCM – *Sistemas Convectivos de Mesoescala*

SE – *Standard Error*

SF – *Sistemas Frontais*

SHUM – *Specific humidity*

SNEB – *Sul do Nordeste do Brasil*

SST – *Sea Surface Temperature*

TN – *Minimum Temperature*

TSM – Temperatura à Superfície do Mar

TVE – Teoria dos Valores Extremos

TX – Temperatura Máxima (*Maximum Temperature*)

VCAS – Vórtice Ciclônico de Ar Superior

UTCV – *Upper Troposphere Cyclonic Vortex*

ZCAS – Zona de Convergência do Atlântico Sul

ZCIT – Zona de Convergência Intertropical

1. INTRODUÇÃO

Nos últimos anos a comunidade internacional vem chamando a atenção acerca das mudanças climáticas e sua relação com a ocorrência de eventos extremos e seus impactos catastróficos (EASTERLING et al., 2000; MEEHL & TEBALDI, 2004; TRENBERTH, 2011; TRENBERTH, 2012; CAVALCANTI, 2012). Tais eventos estão diretamente associados à ação do homem no meio ambiente (FOLLAND et al., 2001). Estes autores relacionaram o aumento sistemático da temperatura média global com a ocorrência de eventos extremos. Segundo FOLLAND et al. (2001), não será a variabilidade da temperatura média geral que afetará a sociedade, mas as mudanças na frequência e intensidade desses eventos extremos. O aumento nos extremos de temperatura e de precipitação é resultante das mudanças climáticas e são inerentes ao desenvolvimento de eventos climáticos e aos desastres naturais (POWELL, 2010). Em particular, qualquer aumento nos extremos de precipitação poderá amplificar os episódios de secas e enchentes, pois eles geram grandes prejuízos sociais e econômicos (HUNTINGFORD et al., 2003).

Os episódios de eventos extremos de precipitação na América do Sul (AS) tornaram-se cada vez mais frequentes. Muitos destes extremos estão relacionados à variabilidade das TSM tanto da região do Atlântico (HASTENRATH, 1984; KEENLYSIDE et al., 2013) como no Pacífico (ROPELEWSKI & HALPERT, 1987; BOMBARDI et al., 2013), ambas as regiões oceânicas exercem uma forte influência no padrão de precipitação, e conseqüentemente, em seus extremos (LIEBMANN et al., 2001).

Diversos sistemas de grande escala influenciam de maneira significativa a ocorrência nos extremos de precipitação na AS, tais como: Alta da Bolívia (AB) (MARENGO et al., 1993; MARENGO et al., 2013); Jatos de Baixos Níveis na cordilheira dos Andes (BERRI & INZUNZA, 1993; POVEDA et al., 2014); Sistemas Frontais (SF) (KOUSKY, 1979); Sistemas Convectivos de Mesoescala (SCM) (MADDOX, 1980; GARAVAGLIA et al., 2013); Zona de Convergência Intertropical (ZCIT) (HASTENRATH, 1984; SOUZA & CAVALCANTI, 2009); Zona de Convergência do Atlântico Sul (ZCAS) (CARVALHO et al., 2002, CARVALHO & OYAMA, 2013); Vórtice Ciclônico da Ar Superior (KOUSKY & GAN, 1981; EVANS & BRAUN, 2012); Oscilação Madden-Julian (OMJ) (MADDEN & JULIAN, 1971; SILVA & KOUSKY, 2012); Distúrbios Ondulatórios de Leste (DOL) (KOUADIO et al., 2010; TORRES & FERREIRA, 2011).

O aumento expressivo da ocorrência de eventos extremos e de seus impactos gerados ao redor do globo, principalmente no NEB, têm motivado diversos pesquisadores a buscarem com profundidade uma melhor compreensão dos processos que desencadeiam estes eventos. No NEB, tais extremos estão relacionados à ocorrência de enchentes, decorrentes dos extremos de precipitação em curto prazo de tempo ou estiagens prolongadas, ambos ocasionam grandes danos socioeconômicos à população.

Devido à alta diversidade de informações, seja pela quantidade de variáveis meteorológicas e de sistemas atuantes responsáveis pela intensificação e ocorrência dos eventos extremos, ainda faz-se necessário entender como esses eventos ocorrem e quais as variáveis meteorológicas associadas, esta Tese tem como OBJETIVO GERAL diagnosticar e avaliar os extremos de precipitação via modelos de dispersão na Região Nordeste do Brasil.

A Tese está constituída em quatro capítulos, os capítulos I e II diagnosticaram os extremos de precipitação a partir de dois modelos de regressão pertencentes aos Modelos Lineares Generalizados (MLG): regressões Logística e de Poisson, com o objetivo de avaliar o comportamento entre os extremos de precipitação e as demais variáveis meteorológicas via razão de chances (OR) e o risco relativo (RR), respectivamente. O Capítulo III diagnosticou os extremos anuais de precipitação utilizando as distribuições generalizada de valores extremos (GEV) e generalizada de Pareto (GPD) de maneira univariada, com o objetivo de avaliar o comportamento espacial dos extremos no NEB e também seus níveis e períodos de retorno. O Capítulo IV avaliou o comportamento dos extremos anuais de precipitação no NEB, utilizando os modelos de regressão via distribuição GEV de forma múltipla com o objetivo de estimar os seguintes estes valores: 1) os coeficientes de regressão das variáveis que subsidiam estes extremos; 2) os parâmetros da distribuição GEV (localização, de escala e de forma); e 3) os níveis e períodos de retorno.

1.1. REVISÃO DE LITERATURA

1.1.1. Definição de extremos no contexto climático

O impacto das mudanças climáticas é fortemente perceptível quando associado aos eventos extremos, especialmente quando as populações vulneráveis ou com propriedades de alto valor são afetados (TRENBERTH et al., 2007).

A caracterização dos eventos extremos é facilmente reconhecida, porém sua é difícil defini-la por uma série de razões: a primeira é de que não há uma definição única para o significado da palavra “extremo”; a segunda é o conceito de extremo é relativo e depende fortemente de um contexto; e a terceira são as palavras “severo”, “raro”, “extremo” e “alto impacto” são frequentemente usadas alternadamente. (DIAZ & MURNANE, 2008)

De acordo com Diaz & Murnane (2008), as definições para os eventos extremos são:

- Severos - São eventos que criam uma enorme perda em números, como por exemplo, número de vidas, capital financeiro, ou qualidade ambiental (por exemplo perda de espécies). A severidade pode ser medida pela expectativa de perdas em longo prazo, a exposição do desastre, e a vulnerabilidade.
- Raros - São eventos que têm uma baixa probabilidade de ocorrência. Devido à raridade destes eventos, as sociedades humanas (e outros ecossistemas) muitas vezes não são bem adaptadas a eles e assim sofrem grandes quantidades de dano quando elas ocorrem. Assim, apesar de sua raridade, a grande vulnerabilidade associada a tais eventos podem muitas vezes levar a grandes perdas médias (portanto, eles são um tipo de evento grave).
- Extremos - São eventos que tem valores extremos de variáveis meteorológicas de certa importância. Os prejuízos são frequentes causados pelos valores extremos de certas variáveis meteorológicas, tais com chuva extrema em um dia, velocidade do vento forte, temperaturas altas, etc. Os extremos são geralmente definidos pelos valores máximos ou excedentes ao seu alto limiar já pré-existente, tais eventos são considerados raros, uma chuva extrema que exceda o valor do período de retorno de 100 anos, com probabilidade de apenas 0,01% de ocorrer em um ano particular.
- Alto impacto - São eventos severos que pode ser sistemas meteorológicos de curta duração (tempestades severas) ou eventos de longa duração, tais como: episódios de bloqueios que pode ser prolongado por ondas de calor ou secas.

1.1.2. Fatores que subsidiam a ocorrência dos extremos de precipitação

A ocorrência destes extremos têm fomentado diversos pesquisadores com o propósito de detectar as possíveis associações entre os extremos de precipitação aos fatores ambientais. O’Gorman & Schneider (2009) que analisaram os extremos de precipitação para cenários futuros e detectaram uma forte relação entre temperatura e a quantidade de vapor d’água em baixos níveis em regiões tropicais. Berg et al. (2013), com maior quantidade de calor liberada para a atmosfera o ciclo da água é acelerado, resultando em chuvas mais intensas em escala diária (DORE, 2005), porém em períodos mais curtos, enquanto isso favorece a maiores períodos de estiagem (GROISMAN et al., 2008).

Estes eventos extremos têm sido verificados em diversas partes do globo, cujas ocorrências estão relacionadas a vários fatores, nos quais se destacam: às alterações de sistemas meteorológicos em várias escalas (MEEHL & TEBALDI, 2004; CAVALCANTI, 2012), a temperatura do ar (DING et al., 2008), a precipitação pluvial (LIEBMANN et al., 2011; CORREIA-FILHO et al., 2014), a velocidade do vento (YAN et al., 2002; DELLAMARTA et al., 2009), a Temperatura à Superfície do Mar (UVO et al., 1998; KAYANO et al., 2013), as alterações do microclima regional provocados pelo desenvolvimento urbano desenfreado (WILLEMS et al., 2012), e finalmente, aos fatores topográficos (WOTLING et al., 2000; HOUZE, 2012).

Além destes citados acima outros fatores interferem diretamente na ocorrência destes extremos, por exemplo, a evaporação verificada por Romero et al. (1997) que investigaram as possíveis associações entre ambas e a orografia para a cidade de Catalunha – Espanha, e constataram que o efeito da orografia é responsável pela distribuição espacial da precipitação e também na formação de uma circulação ciclônica rasa sobre o oceano na costa argelina que modifica os campos de vento sobre a região da Catalunha.

Outro fator fundamental para a ocorrência dos extremos de precipitação é a velocidade do vento, ao qual exerce o papel de transportador de umidade e calor no processo inicial na formação da precipitação, e conseqüentemente, em seus extremos (GASTINEAU & SODEN, 2011). O vento em superfície explica uma pequena fração da variabilidade diária da precipitação, entretanto é determinante no processo de modulação da convecção tropical, motivados pelo aumento da evaporação, quanto maior a evaporação maior será a magnitude do vento, isto contribui para a desestabilização da camada limite e a partir do processo de

realimentação da umidade forçará a formação de convecção profunda, resultando no aumento da precipitação (BACK & BRETHERTON, 2005).

Segundo Deser et al. (2010), os principais sinais destas mudanças podem ser captados pelas flutuações das variáveis termodinâmicas, como por exemplo, a variabilidade na temperatura e na capacidade de armazenamento de água da atmosfera, do que pelas variáveis dinâmicas, como mudanças relacionadas ao padrão dos ventos. De acordo com Trenberth (2012), os extremos são avaliados de forma individual, porém em sua maioria não há associações, recorrendo-se a análise dos conglomerados de extremos na tentativa de detectar a ocorrência das anomalias da variabilidade natural. Isso ocorre especialmente em associação com o modo dominante da variabilidade natural, a mais conhecida entre estas é a El Niño - Oscilação Sul (ENOS) (TRENBERTH et al., 2002). Esta variabilidade exerce uma forte influência ao longo do globo com periodicidade que varia entre 3-7 anos, entre as fases quente e fria, originados pelas anomalias da TSM do Pacífico equatorial (DORE, 2005).

Durante a ocorrência dos ENOS, o clima sofre mudanças significativas devido às anomalias da TSM na região equatorial e central do Pacífico, resultando uma forte variabilidade nos acumulados, intensidade e frequência da precipitação (TRENBERTH et al., 2002; SANDEEP & STORDAL, 2013). Neelin et al. (2003) investigaram os efeitos do aquecimento global e do ENOS na precipitação e constataram que a forte variabilidade na precipitação, verificada na região tropical, é causada pelas mudanças no aquecimento convectivo e na dinâmica atmosférica de grande escala. Durante a fase quente do ENOS, há uma modificação na circulação da célula de Walker (oeste-leste) com o ar descendente sobre a Amazônia e o semiárido do NEB, gerando alteração na advecção de umidade da região equatorial da América do Sul, que inibe a formação de nuvens resultando em anomalias negativas de precipitação na América do Sul (GRIMM & TEDESCHI, 2009) e principalmente no NEB (RAO et al., 1995; HASTENRATH, 1984; HASTENRATH & GREISCHAR, 1993; OYAMA & NOBRE, 2004).

Uvo et al. (1998) analisaram as relações entre a precipitação mensal do NEB e as TSM do Atlântico e Pacífico tropicais durante o período chuvoso e constataram que, na região sul do NEB, o regime pluviométrico é controlado pelos sistemas frontais advindos da região sudeste, enquanto a parte norte do NEB é dominada pelas anomalias de TSM da região do Atlântico Sul, causando o deslocamento da ZCIT mais ao sul da posição normal.

Outro mecanismo que interfere diretamente nos extremos de precipitação, principalmente no NEB está relacionada à variabilidade da TSM do Atlântico (UVO et al., 1998; KAYANO et al., 2013). Dommenget & Latif (2000) diagnosticaram os efeitos desta variabilidade em escala interanual e decadal sobre o clima e constataram que a variabilidade é explicada por forçantes atmosféricas e processos dinâmicos no oceano, tais como: convecção, advecção e propagação de ondas e a variabilidade é verificada nas regiões em que os ventos alísios de nordeste e sudeste são máximos. De acordo com Xie & Carton (2004) há dois tipos de variabilidade na região do Atlântico: 1) com escala de tempo entre 8-11 anos, caracterizado pela variabilidade dependente para norte e para o sul do equador, o que provoca alteração no posicionamento da ZCIT (UVO et al., 1998; SOUZA & CAVALCANTI, 2009), 2) com escala de tempo entre 12-20 anos que consiste no dipolo (NNAMCHI et al., 2011), ambos os modos de variabilidade alteram significativamente as anomalias de precipitação na região (NOBRE & SHUKLA, 1996; GIANINNI et al., 2003).

Outros sistemas atuantes que favorecem a ocorrência dos extremos de precipitação no NEB é a Alta Subtropical do Atlântico Sul (ASAS) e DOL. Segundo Rao et al. (1993) a ASAS possui um papel fundamental no mecanismo de transporte de vapor d'água em baixos níveis ao longo do NEB, favorecendo a ocorrência de precipitações intensas.

1.1.3. Métodos estatísticos utilizados para análise de extremos

Na tentativa de compreender melhor os principais fatores e efeitos que desencadeiam a ocorrência destes extremos, tem-se aplicado diversos métodos para diagnosticá-los. Um método utilizado se baseado nos índices climáticos criados pelo CCI/CLIVAR/JCOMM e criados pelo *Expert Team on Climate Change Detection and Indices* (ETCCDI) com o objetivo de verificar as tendências dos extremos, como por exemplo, o número consecutivo de dias secos ou chuvosos (WMO, 2009; SKANSI et al., 2013).

Outro método muito utilizado para a detecção de extremos é a Teoria dos Valores Extremos (TVE), a partir das distribuições Generalizada de Valores extremos (*Generalized Extreme Value Distribution - GEV*) e Generalizada de Pareto (*Generalized Pareto Distribution - GPD*) (IDALINO et al., 2009; DAVISON & GHOLAMREZAEI 2012), extremamente eficaz, porém sua análise é realizada de forma univariada, baseada na própria série de dados, limitando a inclusão de fatores externos que possam contribuir na explicação da ocorrência de um evento extremo. Apesar de apresentar bons resultados, estas distribuições exibem uma resposta “simplificada” de causa-efeito, segundo Ding et al. (2008) verificaram

que a distribuição GPD possui a vantagem na previsão do resultados quando comparados com a distribuição GEV.

Li et al. (2005) investigaram os máximos de precipitação ao longo do sudoeste da Austrália Ocidental via picos acima dos limiares (“POT - *Peaks Over Threshold*”) da distribuição GPD e aplicaram o teste Mann-Whitney-Pettit (MWP) para detectar a variabilidade no inverno (maio-outubro) e no verão (novembro-abril). Os autores constataram pelo teste MWP, que os extremos de precipitação diária no inverno aumentaram após 1965, em contrapartida a variabilidade multidecadal, como a Oscilação Antártica (OA), os episódios de seca como visto em 1965.

Uma das alternativas que têm contribuído neste diagnóstico é a aplicação dos Modelos Lineares Generalizados (MLG), desenvolvida por Nelder e Wedderburn (1972), uma extensão dos modelos lineares clássicos que abrange uma série de modelos de regressão. O MLG é potencialmente útil e robusto, ao qual permite analisar o efeito simultâneo de diversas variáveis explicativas, o que facilita na interpretação da análise. Esses modelos são bastante difundidos em ciências climáticas, aplicados de diversas formas, tais como: modelos de ocorrência e de acumulados de precipitação (LITTLE et al., 2009), caracterização da variabilidade climática (CHANDLER, 2005), na análise de eventos extremos (YAN et al., 2002; CORREIA-FILHO et al., 2014). Chandler & Wheeler (2002) aplicaram o MLG na região oeste da Irlanda, com o objetivo de caracterizar os modelos de ocorrência e de acumulados de precipitação, e concluíram que a *North Atlantic Oscillation* (NAO) exerce uma forte influência na variabilidade das chuvas sobre a região. Ambrosino et al. (2010) aplicaram o MLG para a caracterização da variabilidade mensal de precipitação no sul da África e constataram que as variáveis que afetam a variabilidade da precipitação são: umidade relativa, El Niño da região 3.4 e a temperatura do Hemisfério Sul (HS), enquanto a orografia influencia na variabilidade regional.

Outra alternativa eficaz no diagnóstico da ocorrência dos eventos extremos seria associar o MLG via TVE (JORGENSEN et al., 2010; WANG & DEY, 2010; CALABRESE & OSMETTI, 2010), a partir das distribuições GEV e GPD, analisados pelos Modelos Vetoriais Lineares Generalizados via distribuição GEV (MVLG-GEV). O MVLG é uma extensão dos MLG tradicionais, com a vantagem em incluir a modelagem aditiva as diversas formas de análises de valores extremos, além disso, é um modelo mais robusto e ao mesmo tempo flexível, com tipos de resposta que podem ser univariada ou múltipla (YEE & WILD,

1996; YEE & STEPHENSON, 2007, VOVORAS & TSOKOS, 2009; MARAUN et al., 2011; JONES et al., 2014). Maraun et al. (2011) avaliaram a influência da circulação atmosférica em escala sinótica nos extremos de precipitação no Reino Unido via modelo MVLG-GEV. Os autores mostraram que os extremos de precipitação possuem padrões espaciais bem distintos devido aos seguintes aspectos: interação entre o fluxo de ar ocidental, variabilidade da NAO, a influência da orografia das montanhas localizadas na parte norte e sul, e o transporte de umidade dos mares circundantes do Reino Unido.

Nesta introdução foram apresentadas em caráter inicial exibiu os eventos extremos para o contexto climático, mostrando definições, tipos e suas relações entre as variáveis que possam desencadear os extremos, neste caso extremos de precipitação, e também uma breve apresentação de alguns métodos que avaliam os extremos. A seguir será apresentado o primeiro capítulo da Tese que tratará da análise de extremos de precipitação para o NEB via regressão logística.

CAPÍTULO I

Este primeiro capítulo é constituído pelo primeiro artigo da tese, intitulado “*Precipitation extremes analysis over the Brazilian Northeast via Logistic Regression*”, publicado na revista *Atmospheric and Climate Science* no ano de 2014, no volume 4, série nº1, das páginas 53-59. Esta revista está classificada pelos indicadores da CAPES *Qualis* B1 na área interdisciplinar e possui um fator de impacto JCR de 0,39. O artigo encontra-se disponível pelo link <http://dx.doi.org/10.4236/acs.2014.41007>.

PRECIPITATION EXTREMES ANALYSIS OVER THE BRAZILIAN NORTHEAST VIA LOGISTIC REGRESSION

Washington L. F. Correia-Filho¹; Paulo S. Lucio¹ e Maria H. C. Spyrides¹

¹Programa de Pós-Graduação em Ciências Climáticas, Centro de Ciências Exatas e da Terra,
Universidade Federal do Rio Grande do Norte, Natal, Brazil

ABSTRACT: This work diagnosed the precipitation extremes over the Brazilian Northeast (NEB) based on logistic regression for obtain associations between precipitation extremes and the meteorological variables by Odds Ratio (OR). It was used daily data of ten meteorological variables to the NEB (North (NNEB), East (ENEb), South (SNEB) and Semiarid (SANEB)). The OR results evidenced that the outgoing longwave radiation was the key variable on the precipitation extremes detection in three sub-regions: ENEB with 2.91 times (95% confidence interval (CI): 2.11, 4.02), NNEB with 3.63 times (95% CI: 1.93, 6.83), and SANEB with 5.40 times (95% CI: 3.04, 9.61); while on SNEB was relative humidity with 3.88 times (95% CI: 2.89, 5.20) more chance to favor the precipitation extremes. The maximum temperature, zonal wind component, evaporation, specific humidity and relative humidity also influenced on these extremes. Goodness-of-fit and Receiver Operating characteristics analysis demonstrated that all models had a good fit and good predictive capability.

Keywords: Odds Ratio, ROC Curve, Outgoing Longwave Radiation, semiarid region

RESUMO: Este trabalho diagnosticou os extremos de precipitação sobre o Nordeste Brasileiro (NEB), baseados na regressão Logística para obter associações entre extremos de precipitação e as variáveis meteorológicas pela razão de chances (RC). Utilizou-se dez variáveis meteorológicas para o NEB (Norte (NNEB), Leste (LNEB), Sul (SNEB) e Semiárido (SANEB)). Os resultados da RC evidenciaram que a radiação de onda longa foi a variável chave no diagnóstico dos extremos de precipitação em três sub-regiões: ENEB com 2,91 vezes (intervalo de confiança de 95% (IC): 2,11, 4,02), NNEB com 3,63 vezes (IC 95%: 1,93, 6,83) e SANEB com 5,40 vezes (IC 95%: 3,04, 9,61) mais chance de favorecer a extremos de precipitação, enquanto no SNEB a umidade relativa foi de 3,88 vezes (IC 95%: 2,89, 5,20). A temperatura máxima, componente do vento zonal, evaporação, umidades específica e relativa também influenciaram a ocorrência desses extremos. O ajuste de bondade e a análise da curva ROC demonstraram que todos os modelos tiveram um bom ajuste e uma boa capacidade de previsão.

Palavras-Chaves: Razão de Chance, Curva ROC, Radiação de Onda Longa, região do semiárido.

1. Introduction

The increase of extreme events in short time period became the society more vulnerable at weather and climate extremes variability, resulting in great socioeconomic losses [1]. These extremes are related at several environmental factors that favor the increase on their frequency and intensity: 1) ocean-atmospheric variables relationships, such as: air temperature [2], precipitation [3], wind speed [4] and sea surface temperature (SST) [5]; 2) regional microclimate changes due to rapid urbanization of the cities without proper urban planning [6]; and 3) orographic effects [7]. These factors when combined at atmospheric circulation or meteorological systems in several spatiotemporal scales [8] can to favor the extremes occurrence.

These extremes have motivated various researchers on seek to detect associations between precipitation extremes with environmental factors, as [9] that investigated precipitation extremes to future scenarios and detected interaction between temperature and water vapor that propel the precipitation extremes, already in tropical regions these extremes are motivated by specific humidity saturation in low levels. This temperature and specific humidity relationship also was found by [10].

On Northeast Brazil (NEB), these extremes are related at precipitation by deficit (semiarid region) or excess (capitals or coast regions) found by several researchers [11], [12], [13]. These extreme events direct or indirectly affect with great socioeconomic losses caused by floods or by prolonged droughts [14].

Several statistical methods have been implemented to extract patterns on these extremes, we aboard of these paper the generalized linear models (GLM) [15], [16]. The GLM is a flexible generalization of ordinary linear regression that allows for response variables that have other than a normal distribution. Furthermore, these models are effective and robust that it will facilitate to obtain the precipitation extremes by logistic regression models. In climate sciences, the logistic regression application is related to precipitation occurrence or amount models [17], [18], and forecast skill verification of the climate models [19]. Thus, this approach via odds ratio (OR) will aims to answer the questions:

- 1) Does SST determine the precipitation extreme occurrence?
- 2) Which is (are) variable(s) that favor the precipitation extreme occurrence?
- 3) What is the magnitude of these associations via OR on extreme intensification?
- 4) What will be the OR behavior is similar in all the NEB sub-regions?

In order to answer these questions, the goal of this article is to characterize the precipitation extremes on NEB via logistic regression models; the objective of this paper is to diagnose the precipitation extremes.

2. Material and Methods

2.1. Study Area

The NEB has 1.5×10^6 km² of area ranging between 1-18°S and 35-47°W, the region is influenced by different meteorological systems with distinct characteristics. According to [20], the NEB precipitation pattern is divided in three sub-regions: Eastern (ENE) - influenced by Mesoscale Convective Systems (MCS), Mesoscale Convective Complex (MCC) and Easterly Wave Disturbance [21], [22], [23]; Southern (SNEB) - influenced by South Atlantic Convergence Zone (SACZ), Frontal Systems (FS) and Upper Tropospheric Cyclonic Vortex (UTCV) [24], [25]; and Northern (NNEB) - influenced by Intertropical Convergence Zone (ITCZ), easterly wave disturbance [26], [22]; others meteorological

systems also occur such as: sea breeze systems on the NNEB and ENEB and the South Atlantic Subtropical Anticyclone (SASA) influences in all the sub-regions [24], [25].

Cluster analysis was performed to characterize the new NEB precipitation pattern using Euclidean distance via Ward method, resulting in four sub-regions: NNEB, ENEB, SNEB and semiarid (SANE) as the new sub-region shown on Figure 1.

2.2. Data

We use daily precipitation data from Climate Prediction Center – NOAA [27] with $1.5^\circ \times 1.5^\circ$ grid resolution, among the February-July months (rainfall period) for 1979-2011 period.

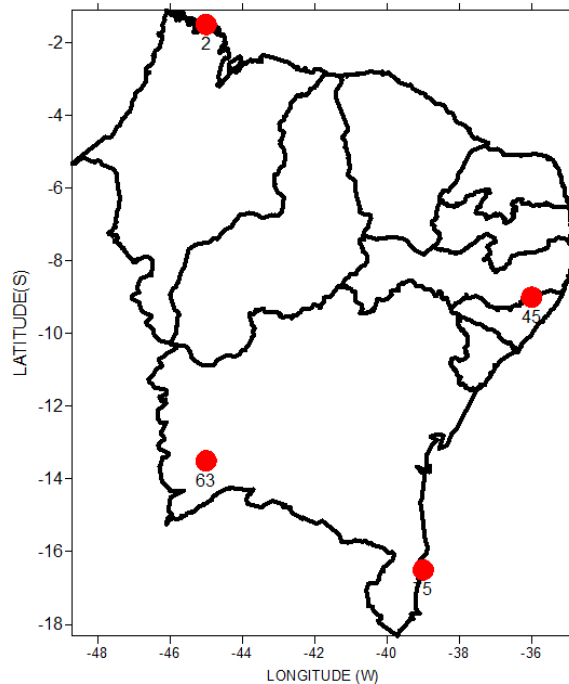


Figure 1 – Grid points distribution on NEB which was divided in four sub-regions (symbols): NNEB (circle); ENEB (cross); SANE (square) and SNEB (triangle).

On this analysis were used others variables: relative humidity (RH), minimum (TN) and maximum (TX) temperatures, evaporation (EV), zonal and meridional wind components (CompU and CompV), gust wind (GUST) and specific humidity (SHUM) provided from Era-Interim reanalyzes [28]; and interpolated satellite data of outgoing longwave radiation (OLR) provided from Earth System Research Laboratory-NOAA (ESRL-NOAA) [29]. All data are same grid resolution ($1.5^\circ \times 1.5^\circ$) shown on Figure 1. It was chosen 4 random points that

represent each sub-region: NNEB - point 02 (northern of Maranhão); SANEB - point 45 (northeastern of Alagoas); SANEB - point 63 (southwestern of Bahia); and SNEB - point 75 (southeastern of Bahia).

2.3. Principal Component Analysis for Atlantic and Pacific regions

We also use daily SST data of the two tropical regions: Atlantic Ocean (ATL) (21°S-21°N, 57°W-15°E) and Pacific Ocean (NINO), El Niño 1.2, 3, 3.4 and 4 regions (5°S-5°N, 90°W-160°E, and 10°S-0°, 80°W-90°W) provided from the Era-Interim reanalyzes. The SST in daily timescale is little used due at low degree variability, but will be important for Poisson regression build. The SST data implementation of variables explanatory follows these stages: 1) The inclusion of lags for both basins: Atlantic - 30 days [24], and Pacific - 90 days [30] due at ocean time response; 2) To calculate the anomalies for each SST regions; and 3) Principal Components Analysis (PCA) to extract the main pattern behavior.

2.4. Logistic regression models

After that PCA composition based on SST regions, we apply the cross correlation function (CCF) to identify lag of correlations between precipitation and the other variables to extract the lags. Then the logistic regression model was applied following two important criteria: 1) Given a set of independent variables, the propose is estimate the probability of precipitation extreme occurrence; and 2) the magnitude of the influence of each meteorological variable on precipitation extremes is obtained by odds ratio (OR). The logistic regression is expressed:

$$g(x) = \ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = \alpha + \beta X \quad (1)$$

α is the intercept, β is estimate coefficient, X is the explanatory variables, $g(x)$ is precipitation extreme in dichotomous form (between 0's e 1's), and π is the probability of precipitation extremes, given by:

$$\pi(x) = \frac{e^{(\alpha + \beta X)}}{1 + e^{(\alpha + \beta X)}} \quad (2)$$

2.5. Odds Ratio

Odds ratio (OR) was calculated for each variable, obtaining the association magnitude between precipitation extremes and the meteorological variables. For calculate the OR should get the odds, which is the natural measure more important in logistic regression and can be interpreted as the ratio between the odds of the precipitation extremes to occur to the odds of precipitation extremes not to occur. Both odds are dimensionless and non-negative, if the $OR < 1$ is described as exposure factor, the observed variable is not influence the precipitation extremes, while the $OR > 1$ is described as risk factor, thus the observed variable influence on the precipitation extremes. Thus, the OR depends of four probabilities that following:

$$OR = \frac{e^{\alpha} e^{\beta X} e^{\beta}}{e^{\alpha} e^{\beta X}} = e^{\beta} \quad (3)$$

Estimates of the parameters α and β area usually obtained using a statistical package, and the output for the data summarized and given in Tables 2 and 3.

These PCA obtained were transformed in time series for identify the degree of the measures of association with the precipitation extremes by relative risk. The categorization of variables included three criteria: 1) precipitation data in the upper 95th percentile ($>95p$) values were considered extreme [31], [32], [33]; 2) for OLR, OLR below 240 Wm^{-2} was considered convective cloud [34]; 3) for the other variables, the contingency tables (Table 1) was divided in quartiles, it was considered abnormal for those quartile that to obtain the higher number occurrences of precipitation extremes (Table 2).

Table 1 – Observed frequencies for precipitation extremes.

Variable	Precipitation extremes		
	Yes	No	Total
Abnormal	A	B	$n_1 = A+B$
Normal	C	D	$n_2 = C+D$
Total	A+C	B+D	$N = n_1 + n_2$

Table 2 – Number of extreme precipitation cases (n° cases in %) of four NEB sub-regions.

Variable	NNEB	ENEb	SANEb	SNEb
RH	156 (53.4)	165 (56.3)	95 (32.7)	186 (63.3)
EV	193 (66.1)	115 (39.0)	103 (34.9)	83 (28.2)
TX	163 (55.8)	131 (44.7)	212 (71.9)	86 (29.2)
TN	80 (27.3)	113 (38.6)	146 (50.3)	87 (29.6)
GUST	160 (54.8)	120 (40.9)	184 (63.4)	107 (36.4)
CompU	149 (51.0)	81 (27.6)	213 (72.2)	90 (30.6)
CompV	114 (39.0)	83 (28.1)	200 (69.0)	83 (28.2)
SHUM	88 (30.1)	129 (44.0)	227 (78.3)	153 (52.0)
OLR	278 (95.2)	147 (50.2)	276 (93.6)	166 (56.5)
ATL1	84 (28.8)	100 (33.9)	83 (28.6)	88 (29.9)
ATL2	86 (29.4)	80 (27.1)	120 (41.4)	83 (28.2)
NINO1	115 (39.4)	88 (29.8)	94 (32.4)	81 (27.5)
NINO2	138 (47.3)	90 (30.5)	174 (59.5)	83 (28.2)
PREC	292 (5.0)	295 (5.0)	290 (5.0)	294 (5.0)

2.6. Goodness-of-fit and Receiver Operating Characteristics analysis

For goodness-of-fit (GOF) analysis [35] was used two methods: Deviance residual, AIC; and p-value. Deviance Residual is the quality-of-fit statistic measure based on maximum likelihood using the sum of squared residuals in ordinary least squares. Akaike Information Criterion (AIC) is quality-of-fit measure wherein seek to select variables given a set of variables that optimize the performance of the model with the minimum AIC value. Already the p-value is other measure that indicates whether the variables contained on model has significance statistical, generally it used $p < 0.05$ value for reject the null hypothesis. To assess the accuracy of model it was used receiver operating characteristics (ROC) graph.

The ROC curve is a technique for visualizing, organizing and selecting classifiers, thus it evaluates the quality or performance of diagnostic tests [36], [37]. Generally, the ROC analysis assesses the quality of model counting of occurrence or not precipitation extremes and the exposure factor presence or absent at an extreme condition. Thus, the common measure used is Area under Curve (AUC) that interpret the average value of sensibility for all values of specificity with aim to evaluate the overall performance of a diagnostic test, ranging between 0 and 1, wherein a bigger value suggests the better overall performance of a

diagnostic test [36]. All results computation were generated on R software [38] with packages support: MASS [39], ROCR [40] and Epi [41].

3. Results and Discussions

The logistic regression model results, OR and goodness-of-fit analysis (deviance residual, AIC, p-value) are shown on Tables 3 and 4 for four NEB sub-regions (NNEB, ENEB, SANEB and SNEB) shown on Figure 2. The 5,800 daily precipitation data upper >95p were considered, corresponding 290-295 precipitation extremes cases obtained. For goodness-of-fit was verified that the SANEB (Table 2) sub-region obtained the best values of AIC with 1,386.1 and deviance residual of 1,366.1.

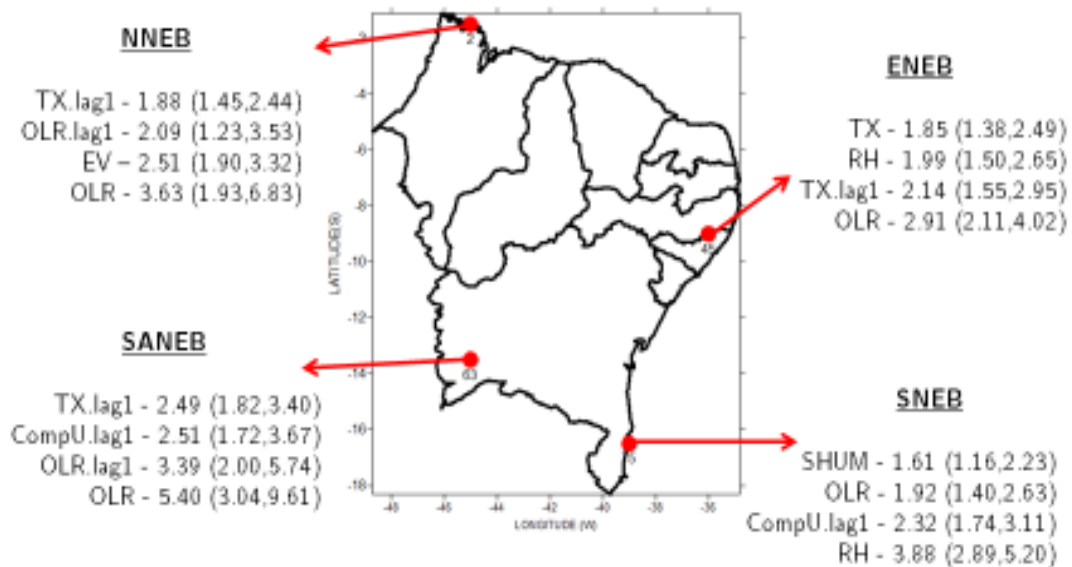


Figure 2 - Meteorological variables that influenced on NEB precipitation extremes by logistic regression models. Each sub-region with the highest OR values.

The ROC curve analysis shown on Figure 3 was observed that all models showed were above 0.80, which shows that all the models have a good predictive ability; the SANEB (Figure 3) sub-region exhibits the best AUC value of 0.935.

The measure of associations by Odds Ratio, on NNEB (Northern of Maranhão State) there is evidence that the variables that contribute on precipitation extremes are EV, OLR, TX.lag₁ and OLR.lag₁; for the ENEB (Northeastern of Alagoas State) are RH, OLR, TX and

TX.lag₁; already on SANEB (Southwestern of Bahia State) are OLR, TX.lag₁, OLR.lag₁ and OLR.lag₁; and on SNEB (Southeastern of Bahia State) are RH, OLR, SHUM and CompU.lag₁.

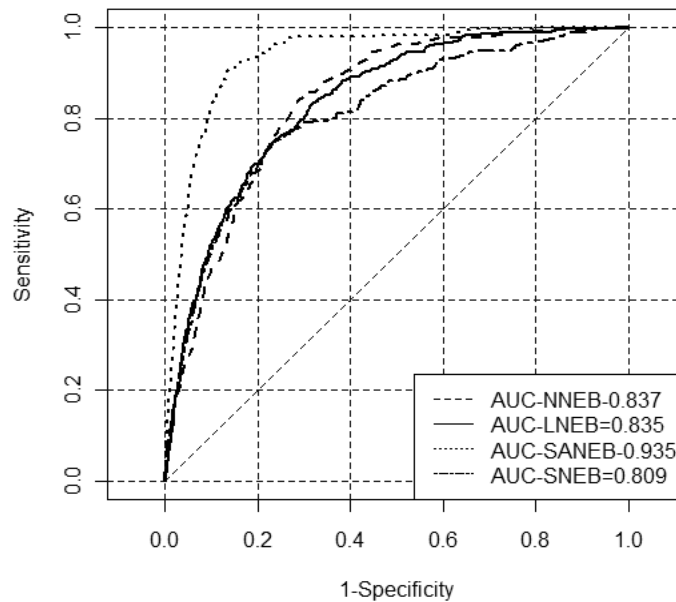


Figure 3 – ROC curve for logistic regression models. Area Under Curve (AUC) values on right bottom for NEB sub-regions: NNEB (solid line), ENEB (dashed line), SANEB (dotted line) and SNEB (dashed and dotted line).

On OR analysis was evidenced that the OLR (Tables 3 and 4) is main variable on precipitation extremes diagnose, ENEB with 2.91 times (95% confidence interval (CI): 2.11, 4.02), on NNEB with 3.63 times (95% CI: 1.93, 6.83) and SANEB with 5.40 times (95% CI: 3.04, 9.61) more chance to favor the precipitation extremes; while on SNEB the highlight is RH with 3.88 times (95% CI: 2.89, 5.20) more chance to favor the precipitation extremes.

The OR results corroborate with [42], wherein the authors detected the OLR and extreme precipitation relationships on tropics using climate indices (rain >10mm and $OLR < 180 \text{ Wm}^{-2}$), this indices indicated that association favoring the convection formation by low-level moisture convergence causing the precipitation more intense suggested by [43]. It is noted that the precipitation extremes events is well distinct for each NEB sub-region favored by several meteorological variable associated meteorological systems corroborating with [44] that describe is not only temperature has a cause-effect influence on precipitation intensity, but for a combination of different meteorological systems.

Table 3 - Logistic regression models for NNEB and SANEB. Coefficient regression (Coeff), Standard Error (SE), p-value, Odds Ratio (95% CI) and Goodness-of-fit (AIC test, Deviance residual and degrees of freedom (df)).

NNEB				SANEB			
Variable	Coeff±SE	p-value	OR (95% CI)	Variable	Coeff±SE	p-value	OR (95% CI)
Intercept	-5.80±0.29	<0.001	--	Intercept	-6.78±0.28	0.001	--
EV	0.92±0.14	<0.001	2.51 (1.90,3.32)	EV	-0.32±0.15	0.029	0.72 (0.54,0.97)
TX	0.52±0.14	<0.001	1.69 (1.30,2.21)	TX	0.64±0.16	<0.001	1.91 (1.39,2.61)
TX.lag₁	0.63±0.13	<0.001	1.88 (1.45,2.44)	TX.lag₁	0.91±0.16	<0.001	2.49 (1.82,3.40)
CompU	0.29±0.13	0.031	1.33 (1.03,1.73)	CompU	0.58±0.17	<0.001	1.80 (1.28,2.52)
CompV.lag ₁	0.58±0.14	<0.001	1.78 (1.37,2.32)	CompU.lag₁	0.92±0.19	<0.001	2.51 (1.72,3.67)
CompV.lag ₃	0.38±0.13	0.004	1.47 (1.13,1.91)	SHUM.lag ₁	0.65±0.17	<0.001	1.92 (1.36,2.70)
OLR	1.29±0.32	<0.001	3.63 (1.93,6.83)	OLR	1.69±0.29	<0.001	5.40 (3.04,9.61)
OLR.lag₁	0.74±0.27	0.006	2.09 (1.23,3.53)	OLR.lag₁	1.22±0.27	<0.001	3.39 (2.00,5.74)
ATL2	-0.27±0.14	0.049	0.76 (0.58,1.00)	NINO2	0.44±0.15	0.003	1.16 (0.97,1.85)
AIC=1, 886.2				AIC=1, 386.1			
Deviance residual=1, 866.2 d.f.=5,789				Deviance residual=1, 366.1 d.f.=5,789			

Table 4 - Logistic regression models for ENEB and SNEB. Coefficient regression (Coeff), Standard Error (SE), p-value, Odds Ratio (95% CI) and Goodness-of-fit (AIC test, Deviance residual and degrees of freedom (df)).

ENEB				SNEB			
Variable	Coeff±SE	p-value	OR (95% CI)	Variable	Coeff±SE	p-value	OR (95% CI)
Intercept	-5.17±0.17	<0.001	--	Intercept	-4.92±0.17	<0.001	--
RH	0.69±0.14	<0.001	1.99 (1.50,2.65)	RH	1.35±0.15	<0.001	3.88 (2.89,5.20)
RH.lag ₁	0.27±0.14	0.072	1.30 (0.98,1.72)	RH.lag ₁	0.33±0.15	0.025	1.39 (1.04,1.85)
EV.lag ₁	0.47±0.15	0.002	1.61 (1.19,2.18)	EV	0.34±0.16	0.034	1.41 (1.03,1.92)
TX	0.62±0.15	<0.001	1.85 (1.38,2.49)	TX	0.28±0.15	0.057	1.32 (0.98,1.73)
TN	0.36±0.14	0.007	1.37 (1.05,1.80)	TN	0.27±0.14	0.064	1.31 (0.98,1.73)
TX.lag₁	0.76±0.17	<0.001	2.14 (1.55,2.95)	SHUM.lag ₁	0.47±0.17	0.005	1.61 (1.16,2.23)
GUST	0.34±0.36	0.027	1.44 (1.04,1.98)	GUST	0.30±0.16	0.059	1.40 (1.04,1.87)
GUST.lag ₁	0.43±0.17	0.009	1.54 (1.12,2.13)	GUST.lag ₁	0.33±0.15	0.024	1.40 (1.04,1.87)
OLR	1.06±0.16	<0.001	2.91 (2.11,4.02)	OLR	0.65±0.16	<0.001	1.92 (1.40,2.63)
OLR.lag ₁	0.49±0.17	0.003	1.59 (1.15,2.20)	OLR.lag ₁	0.37±0.17	0.028	1.48 (1.07,2.04)
SHUM	0.52±0.15	<0.001	1.61 (1.19,2.18)	SHUM	0.48±0.17	0.003	1.61 (1.16,2.23)
CompU.lag ₂	0.54±0.16	<0.001	1.72 (1.26,2.33)	CompU.lag₁	0.84±0.15	<0.001	2.32 (1.74,3.11)
AIC=1, 929.5				AIC=1, 949.4			
Deviance residual= 1, 901.5		df=5,785		Deviance residual=1, 923.4		d.f.=5,786	

Analyzing the sub-regions in separate, the precipitation extremes on NNEB (Table 3) is linked at ITCZ displacement in north-south direction that transport heat and moisture into region, subsided by TX and EV combined. On SANEB (Table 3), scarcity precipitation region in NEB, it was verified that TX.lag₁, CompU, OLR and OLR.lag₁ favor precipitation extremes boosted by temperature and moisture advection associated at easterly flow forming of deep convection caused by frontal systems [20] or north axis SACZ displacement [25] that penetrating on southern NEB.

On ENEB (Table 4), this extremes are strength by RH, TX and TX.lag₁ combined with CompU.lag₁, according to [13] the intense precipitation occurrence is favored for heat and moisture transport by easterly waves disturbances and boost the MCS and MCC formation about region which are largely responsible for maintaining the precipitation regime, contributing with 50-70% of annual regime. On SNEB (Table 4), the TX.lag₁, U.lag₁ and OLR was influenced in precipitation extremes arising from different meteorological systems: breeze systems and easterly wave disturbance on eastern part, SACZ and frontal systems on southern part.

4. Conclusions

The principal conclusions obtained in this work were:

- The Atlantic and Pacific SST were not significant to support the occurrence of precipitation extremes due at low degree variability;
- OLR is the key variable for detecting the occurrence of extreme precipitation, when the $OLR < 240W/m^2$, the chances of extreme precipitation occur increases in 2.91 times (95% CI: 2.11, 4.02) on ENEB, in 3.63 times (95% CI: 1.93, 6.83) on NNEB, and in 5.40 times (95% CI: 3.04, 9.61) on SANEB;
- Relative humidity is a principal variable on SNEB for occurrence of extreme precipitation with the chance of 3.88 times (95% CI: 2.89, 5.20);
- The maximum temperature, zonal wind component, evaporation, specific humidity and relative humidity also influenced on these precipitation extremes;
- The ROC curve analysis observed that all models showed were above 0.80, which shows that all the models have a good predictive ability; the SANEB sub-region exhibits the best AUC value of 0.935.

5. Acknowledgements

The author thanks at CAPES for doctoral financial support; George Pedra and Naurinete Barreto by several contributions for this article. P. S. Lucio is sponsored by a PQ2 grant (Proc. 302493/2007-7) from CNPq (Brazil).

REFERENCES

- [1] T. R. Karl and D. R. Easterling, “Climate Extremes: Selected Review and Future Research directions,” *Climatic Change*, vol. 42, pp. 209–325, 1999.
- [2] G. C. Blain, “Modeling extreme minimum air temperature series under climate change conditions,” *Ciência Rural*, 41, pp. 1877-1883, 2011. (In Portuguese)
- [3] A. M. Grimm and R. G. Tedeschi, “ENSO and Extreme Rainfall Events in South America,” *Journal of Climate*, vol. 22, no. 7, pp. 1589–1609, 2009.
- [4] P. Friederichs, M. Göber, S. Bentzien, A. Lenz, and R. Krampitz, “A probabilistic analysis of wind gusts using extreme value statistics,” *Meteorologische Zeitschrift*, vol. 18, no. 6, pp. 615–629, 2009.
- [5] M. T. Kayano, R. V. Andreoli, R. A. F. Souza, “Relations between ENSO and the South Atlantic SST modes and their effects on the South American rainfall,” *International Journal of Climatology*, vol. 33, pp. 2008-2023, 2013.
- [6] P. Willems, K. Arnbjerg-Nielsen, J. Olsson, and V. T. V. Nguyen, “Climate change impact assessment on urban rainfall extremes and urban drainage: Methods and shortcomings,” *Atmospheric Research*, vol. 103, pp. 106–118, 2012.
- [7] R. A. Houze Jr, “Orographic effects on precipitations clouds”, *Reviews of Geophysics*, vol. 50, no. GR1001, 2012.
- [8] F. Kucharski, D. Polzin, and S. Hastenrath, “Teleconnection mechanisms of northeast Brazil droughts: modeling and empirical evidence,” *Revista Brasileira de Meteorologia*, vol. 23, no. 2, pp. 115–125, 2008. (In Portuguese)

- [9] P. A. O. Gorman and T. Schneider, “The physical basis for increases in precipitation extremes in simulations of 21st-century climate change,” *Proceedings of the National Academy of Sciences*, vol. 106, pp. 14773–14777, 2009.
- [10] P. Berg, C. Moseley, and J. O. Haerter, “Strong increase in convective precipitation in response to higher temperatures,” *Nature Geoscience*, vol. 6, no. 3, pp. 181–185, 2013.
- [11] M. D. Oyama and C. A. Nobre, “Climatic consequences of a large-scale desertification in northeast Brazil: A GCM simulation study,” *Journal of Climate*, vol. 17, no. 16, pp. 3203–3213, 2004.
- [12] B. Liebmann, G. N. Kiladis, D. Allured, C. S. Vera, C. Jones, L. M. V. Carvalho, I. Bladé, and P. L. M. Gonzáles, “Mechanisms Associated with Large Daily Rainfall Events in Northeast Brazil,” *Journal of Climate*, vol. 24, no. 2, pp. 376–396, 2011.
- [13] Y. K. Kouadio, J. Servain, L. a. T. Machado, and C. A. D. Lentini, “Heavy Rainfall Episodes in the Eastern Northeast Brazil Linked to Large-Scale Ocean-Atmosphere Conditions in the Tropical Atlantic,” *Advances in Meteorology*, vol. 2012, pp. 1–16, 2012.
- [14] S. Hastenrath, “Exploring the climate problems of Brazil’s Nordeste: a review,” *Climatic Change*, vol. 112, no. 2, pp. 243–251, 2011.
- [15] A. J. A. Nelder and R. W. M. Wedderburn, “Generalized Linear Models,” *Journal of the Royal Statistical Society. Series A (General)*, vol. 135, no. 3, pp. 370–384, 1972.
- [16] Z. Yan, S. Bate, R. E. Chandler, V. Isham, and H. Wheeler, “Changes in extreme wind speeds in NW Europe simulated by generalized linear models,” *Theoretical and Applied Climatology*, vol. 83, no. 1–4, pp. 121–137, 2005.
- [17] J. Aburrea and J. Asín, “Forecasting local daily precipitation patterns in a climate change scenario,” *Climate Research*, vol. 28, pp. 183–197, 2005.
- [18] M. K. Tippett, A. G. Barnston, and A. W. Robertson, “Estimation of seasonal precipitation tercile-based categorical probabilities from ensembles,” *Journal of climate*, vol. 20, no. 10, pp. 2210–2228, 2007.

- [19] T. M. Hamill, J. S. Whitaker, and X. Wei, "Ensemble reforecasting: Improving medium-range forecast skill using retrospective forecasts," *Monthly Weather Review*, vol. 132, no. 6, pp. 1434–1447, 2004.
- [20] V. E. Kousky, "Frontal Influences on Northeast Brazil," *Monthly Weather Review*, 107, 1140-1153, 1979.
- [21] L. C. B. Molion and S. de O. Bernardo, "Uma revisão da dinâmica das chuvas no Nordeste Brasileiro," *Revista Brasileira de Meteorologia*, vol. 17, no. 1, pp. 1–10, 2002. (In Portuguese)
- [22] R. R. Torres and N. J. Ferreira, "Case Studies of Easterly Wave Disturbances over Northeast Brazil Using the Eta Model," *Weather and Forecasting*, vol. 26, no. 2, pp. 225–235, 2011.
- [23] H.-Y. Ma, X. Ji, J. D. Neelin, and C. R. Mechoso, "Mechanisms for precipitation variability of the Eastern Brazil/SACZ convective margin," *Journal of Climate*, vol. 24, no. 13, pp. 3445–3456, 2011.
- [24] R. R. Chaves and I. F. A. Cavalcanti, "Atmospheric circulation features associated with rainfall variability over southern Northeast Brazil," *Monthly Weather Review*, vol. 129, no. 10, pp. 2614–2626, 2001.
- [25] L. M. V. Carvalho, C. Jones, and B. Liebmann, "The South Atlantic Convergence Zone: Intensity, Form, Persistence, and Relationships with Intraseasonal to Interannual Activity and Extreme Rainfall," *Journal of Climate*, vol. 17, pp. 88–108, 2004.
- [26] S. K. Mishra, V. B. Rao, and M. A. Gan, "Structure and evolution of the large-scale flow and an embedded upper-tropospheric cyclonic vortex over Northeast Brazil," *Monthly weather review*, vol. 129, no. 7, pp. 1673–1688, 2001.
- [27] M. Chen, W. Shi, P. Xie, V. Silva, V. E. Kousky, R. Wayne Higgins, and J. E. Janowiak, "Assessing objective techniques for gauge-based analyses of global daily precipitation," *Journal of Geophysical Research: Atmospheres (1984-2012)*, vol. 113, no. D4, 2008.

- [28] D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. Van De Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, and A. J. Geer, "The ERA-Interim reanalysis: configuration and performance of the data assimilation system," *Quarterly Journal of the Royal Meteorology Society*, vol.137, pp. 553–597, 2011.
- [29] B. Liebmann and C. A. Smith, "Description of a complete (interpolated) outgoing longwave radiation dataset," *Bull. Amer. Met. Soc.*, vol. 77, pp. 1275–1277, 1996.
- [30] K. E. Trenberth, "The definition of El Niño", *Bull. Amer. Met. Soc.*, vol.78, pp. 2771-2777, 1997.
- [31] K. C. Young, "A multivariate chain model for simulating climatic parameters from daily data," *Journal of Applied Meteorology*, vol. 33, no. 6, pp. 661-671, 1994.
- [32] R. P. Allan, B. J. Soden, "Atmospheric warming and the amplification of precipitation extremes". *Science*, vol. 321, no. 5895, pp. 1481-1484, 2008.
- [33] M. A. F. S. Dias et al. "Changes in extreme daily rainfall for São Paulo, Brazil.", *Climatic change*, vol. 116, no 3-4, pp. 705-722, 2013.
- [34] K. M. Lau, H. T. Wu, S. Bony, "The role of large-scale atmospheric circulation in the relationship between tropical convection and sea surface temperature." *Journal of Climate*, vol. 10, pp. 381-392, 1997.
- [35] D. W. Hosmer, S. Taber, and S. Lemeshow, "The importance of assessing the fit of logistic regression models: a case study.," *American journal of public health*, vol. 81, no. 12, pp. 1630–1635, 1991.
- [36] S. H. Park, "Receiver Operating Characteristic (ROC) Curve: Practical Review for radiologists. *Korean Journal of Radiology*," vol. 5, pp.11-18., 2004.
- [37] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [38] R Development Core Team and others, "R: A language and environment for statistical computing," 2013.

- [39] W. N. Venables, and B. D. Ripley, *Modern applied statistics with S-PLUS*, vol. 250. Springer-verlag New York, 1994.
- [40] T. Sing, O. Sander, N. Beerenwinkel, and T. Lengauer, “ROCR: visualizing classifier performance in R,” *Bioinformatics*, vol. 21, no. 20, pp. 3940–3941, 2005.
- [41] B. Carstensen, M. Plummer, E. Laara, and M. Hills, “{Epi}: A Package for Statistical Analysis in Epidemiology.” 2013.
- [42] S. Sandeep and F. Stordal, “Use of daily outgoing longwave radiation (OLR) data in detecting precipitation extremes in the tropics,” *Remote Sensing Letters*, vol. 4, no. 6, pp. 570–578, 2013.
- [43] K. E. Trenberth, A. Dai, R. M. Rasmussen, and D. B. Parsons, “The Changing Character of Precipitation,” *Bulletin of the American Meteorological Society*, vol. 84, no. 9, pp. 1205–1217, 2003.
- [44] S. C. Liu, C. Fu, C.-J. Shiu, J.-P. Chen, and F. Wu, “Temperature dependence of global precipitation extremes,” *Geophysical Research Letters*, vol. 36, no. 17, p. L17702, 2009.

CAPÍTULO II

Este segundo capítulo é constituído pelo segundo artigo da tese, intitulado “*Analysis of Precipitation extremes on Northeast Brazil via Poisson Regression*”, submetido à revista *Theoretical and Applied Climatology*, no dia 15 de fevereiro de 2014. Esta revista está classificada pelos indicadores da CAPES (*Qualis*) como uma revista **A2 na área interdisciplinar**, e possui um fator de impacto **JCR de 1,759**.

ANALYSIS OF PRECIPITATION EXTREMES ON NORTHEAST BRAZIL VIA POISSON REGRESSION

Washington L. F. Correia-Filho¹; Paulo S. Lucio¹ e Maria H. C. Spyrides¹

¹Programa de Pós-Graduação em Ciências Climáticas, Centro de Ciências Exatas e da Terra,
Universidade Federal do Rio Grande do Norte, Natal, Brazil

ABSTRACT: The increase in frequency and magnitude of climate extremes in the last decades as seen in around the globe has worried society due to great socioeconomic losses. In northeast Brazil (NEB), the climate extremes are related to absent or excessive of precipitation. This work used Poisson regression to diagnose the precipitation extremes, with the aim of obtaining associations between precipitation extremes and other variables by relative risk (RR) to four NEB sub-regions (North, East, Semiarid and South). The relative risk results evidenced that the outgoing longwave radiation is mainly responsible, with the magnitude of 2.50 times (95% confidence interval (CI): 2.09, 3.01), 5.12 times (95% CI: 2.77, 9.50) and 5.23 times (95% CI: 3.00, 9.11) of occur of extreme precipitation for East, Semiarid and North sub-regions. The result suggests the interaction of various meteorological systems such as: Intertropical Convergence Zone, South Atlantic Convergence Zone, Upper Tropospheric Cyclonic Vortex, Madden Julian Oscillation, Easterly Wave Disturbances and Frontal Systems favor the precipitation extremes over NEB. There is evidence that other variables also influence these extremes: maximum temperature, relative humidity, zonal wind components and evaporation. The results found via Poisson regression exhibited a good predictive ability verified by goodness-of-fit and ROC curve analysis.

Keywords – Madden Julian Oscillation, ROC Curve, Outgoing Longwave Radiation, Relative Risk.

RESUMO: O aumento na frequência e magnitude dos extremos climáticos nas últimas décadas ao longo do globo tem preocupado a sociedade devido a grandes perdas socioeconômicas. No Nordeste do Brasil (NEB), estes extremos estão relacionados à ausência ou excesso de precipitação. Este trabalho utilizou a regressão de Poisson para diagnosticar os extremos de precipitação, com o objetivo de obter associações entre extremos de precipitação e outras variáveis de risco relativo (RR) para quatro NEB sub-regiões (Norte, Leste, Semiárido e Sul). Os resultados do RR evidenciaram que a radiação de onda longa é o principal indicador para a ocorrência dos extremos de precipitação, com a magnitude de 2,50 vezes (intervalo de confiança (IC) de 95%: 2,09, 3,01), 5,12 (IC 95% : 2,77, 9,50) e 5,23 vezes (95% CI: 3,00, 9,11) de ocorrer estes extremos para as subregiões leste, semiárido e norte do NEB. O resultado sugere a interação de diversos sistemas meteorológicos, tais como: a Zona de Convergência Intertropical, Zona de Convergência do Atlântico Sul, Vórtice Ciclônico de Ar Superior, Distúrbios Ondulatórios de Leste, Oscilação Madden-Julian e Sistemas Frontais favorecem os extremos de precipitação no NEB. Há evidências de que outras variáveis também influenciam estes extremos: temperatura máxima, umidade relativa, componentes do vento zonal e evaporação. Os resultados obtidos pela regressão de Poisson apresentaram uma boa capacidade preditiva verificada por bondade de ajuste e análise da curva ROC.

Palavras-Chaves – Oscilação Madden-Julian, Curva ROC, Radiação de Onda Longa, Risco Relativo.

1. Introduction

The increase in frequency and magnitude of climate extremes has been evidenced in recent decades by a number of events around the globe, such as severe drought, periods of heavy rainfall and heat waves (Meehl and Tebaldi 2004; Smith 2011). Extremes events are considered "rare events" that occur in a short time period. The increase in the frequency of these events has worried society in their degree of severity, which has resulted in great

socioeconomic losses. Variability in climate extremes also interferes with various infrastructure sectors, such as food, water, energy, shelter and transportation conditioned by meteorological variable variations (Smith 2011). These events are caused by several factors that contribute to their intensification: 1) relationships between ocean and atmospheric variables such as air temperature (Blain 2011), precipitation (Liebmann et al. 2011), wind speed (Hundecha et al. 2008) and sea surface temperature (SST) (Kayano et al. 2013), associated with atmospheric circulation patterns or meteorological systems in several scales of time and space (Kucharski et al. 2008; Cavalcanti 2012); 2) regional micro-climate changes (Willems et al. 2012); and 3) orographic effects (Wotling et al. 2000; Houze 2012), when these factors are combined favor the occurrence of these events.

In the quest to understand these extremes, several researchers have seek to identify associations between precipitation extremes and the several environmental factors, Gorman and Schneider (2009) investigated precipitation extremes and found a strong relationship between temperature and water vapor, this relationship favored precipitation extremes, and in tropical regions these extremes occur due to specific humidity saturation at low levels. Berg et al. (2013) found that precipitation extremes are caused by other relationships between temperature and specific humidity.

In Northeast Brazil (NEB), these extremes are associated with precipitation via deficit (semiarid region) or excess (capitals or coast regions), as found by several researchers (Liebmann et al. 2011; Hastenrath, 2011), resulting in great socioeconomic losses caused by flash flood or by prolonged droughts (Hastenrath and Greichar 1984; Kucharski et al. 2008; Hastenrath 2011).

An alternative to diagnose the occurrence of these extremes is the generalized linear models (GLM) (Nelder and Wedderburn 1972), a potentially useful and effective method. This modeling is widespread in climate sciences and several authors have used in occurrence and accumulated precipitation research (Chandler 2005; Little et al. 2009), and also in extreme events analysis (Correia-Filho et al. 2014). These models are effective and robust which will facilitate obtaining the precipitation extremes by Poisson regression models and comparing them with the results obtained by Correia-Filho et al. (2014) using logistic regression models. In the climate sciences, the Poisson regression application is related to meteorological influences in others areas, e.g. in epidemiology (Wang et al. 2013), in ecology (Gregory et al. 2009), and also in air quality studies (To et al. 2013). Thus, this approach, via relative risk (RR), will answer the questions:

- 1) What is the magnitude of these associations via RR on extreme intensification?
- 2) Will the behavior of RR be similar in all NEB sub-regions?

In order to answer these questions, this article characterizes the precipitation extremes in NEB via Poisson regression model.

2. Materials and Methods

2.1. Study Area

NEB is 1.5×10^6 km² in area ranging between 1-18°S and 35-47°W, and the region is influenced by different meteorological systems with distinct characteristics. Using the precipitation pattern applied by Correia-Filho et al. (2014), the NEB precipitation pattern is sub-divided into four regions: Eastern (ENEB) - influenced by Mesoscale Convective Systems (MCS), Mesoscale Convective Complex (MCC) and Easterly Wave Disturbance (Torres and Ferreira 2011; Kouadio et al. 2012); Semiarid (SANEB) - influenced by MCS, MCC, Frontal Systems (FS) and Upper Tropospheric Cyclonic Vortex (UTCV) (Kousky and Gan 1981); Southern (SNEB) - influenced by South Atlantic Convergence Zone (SACZ), FS and UTCV (Chaves and Cavalcanti 2001; Carvalho et al. 2004); and Northern (NNEB) - influenced by Intertropical Convergence Zone (ITCZ) and UTCV; the sea breeze systems influence NNEB and ENEB, and the South Atlantic Subtropical Anticyclone influences all the sub-regions (Chaves and Cavalcanti 2001; Cavalcanti 2012).

2.2. Data

We use daily precipitation data from the Climate Prediction Centre-National Oceanographic Atmospheric Administration (CPC-NOAA) (Joyce et al. 2004; Chen et al. 2008) with 0.5° x 0.5° grid resolution, and resized the grid resolution to 1.5° x 1.5°, in the February-July months (the rainy period) for the 1979-2011 period. Other variables were used in this analysis: relative humidity (RH), minimum (TN) and maximum (TX) temperatures, evaporation (EV), zonal and meridional wind components (CompU and CompV), gust wind (GUST) and specific humidity (SHUM) provided from ERA-INTERIM, which reanalyzed (Dee et al. 2011); and interpolated satellite data of outgoing longwave radiation (OLR)

provided by the Earth System Research Laboratory-NOAA (ESRL-NOAA) (Liebmann and Smith 1996). All data has the same grid resolution ($1.5^\circ \times 1.5^\circ$) shown in Figure 1.

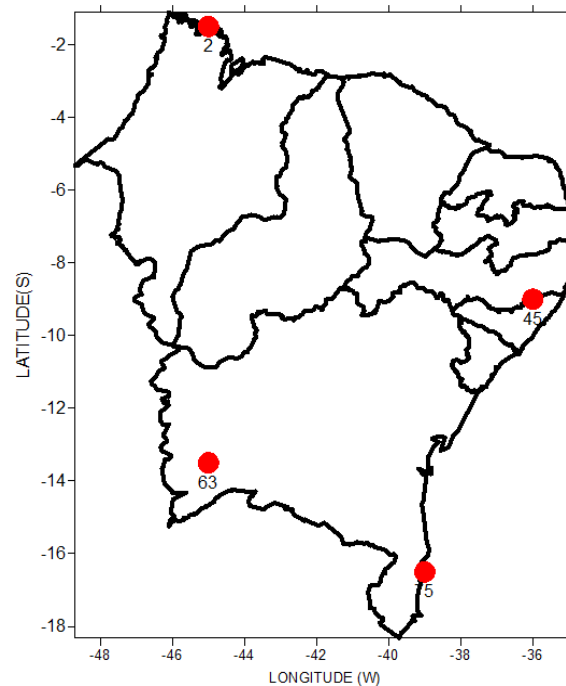


Figure 1 - Grid points distribution on NEB which was divided in four NEB sub-regions (symbols): NNEB (circle); ENEB (cross); SANEB (square) and SNEB (triangle).

It was selected four points to represent each sub-region: NNEB - Point 02 (northern of Maranhão State); SANEB - Point 45 (northeastern of Alagoas State); SANEB - Point 63 (southwestern Bahia State); and SNEB - Point 75 (southeastern Bahia State).

2.3. Principal Component Analysis for Atlantic and Pacific regions

We also used daily SST data for the two tropical regions: Atlantic Ocean (ATL) (21°S - 21°N , 57°W - 15°E) and Pacific Ocean (PAC), El Niño 1.2, 3, 3.4 and 4 regions (5°S - 5°N , 90°W - 160°E , and 10°S - 0° , 80°W - 90°W) provided from the ERA-INTERIM reanalysis. The SST on a daily timescale is little used due to low degree variability, but will be important for Poisson regression builds. The SST data implementations follow the steps applied by Correia-Filho et al. (2014), the explanatory variables follow these stages: 1) the inclusion of lags for both basins: Atlantic - 30 days, and Pacific - 90 days due to ocean time response (Chaves and Cavalcanti 2001, Molion and Bernardo 2002); 2) calculation of the anomalies for each SST regions; and 3) Principal Components Analysis (PCA) to extract the main pattern behavior.

PCA was applied to Atlantic and Pacific SST regions to extract principal signals which influence NEB rainfall patterns (Hastenrath 1984; Liebmann et al. 2011; Hastenrath 2011). These components resulted in two principal factors for both regions; the first PCA extracted the main pattern, while the second verified the seasonal variability. Atlantic SST with 66.83%, and Pacific SST with 79.14% of total variance explained. The SST results obtained by PCA were sufficient to capture the main patterns of variability for these regions. These PCA obtained were transformed in time series for identify the degree of the measures of association with the precipitation extremes by relative risk.

2.4. Poisson regression models

After the PCA composition based on SST regions, we applied the cross correlation function (CCF) to identify the lag of correlations between precipitation and the other variables. The Poisson regression model for binary data (Zou 2004) was then applied following two important criteria: 1) given a set of independent variables, the proposal is to estimate the probability of precipitation extreme occurrences; and 2) to assess the influence of each meteorological variable on precipitation extremes by relative risk (RR). The Poisson regression model with $\mu = \log[\pi(x)]$ is expressed as:

$$g(x) = \log[\pi(x)] = \alpha + \beta X \quad (1)$$

Where α is the intercept, β is estimate coefficient, $g(x)$ is count data of precipitation extreme.

To calculate the RR, all variables were categorized by three criteria: 1) precipitation data in the upper 95th percentile (>95p) values were considered extreme (Young 1994, Allan and Soden 2008, Dias et al. 2013); 2) for OLR, OLR below 240 Wm^{-2} was considered convective cloud (Lau et al. 1997); 3) for the other variables, the contingency tables (Table 1) was divided in quartiles, it was considered abnormal for those quartile that to obtain the higher number occurrences of precipitation extremes (Table 2).

Table 1 – Observed frequencies for precipitation extremes.

Variable	Precipitation extremes		
	Yes	No	Total
Abnormal	a	b	$n_1 = a+b$
Normal	c	d	$n_2 = c+d$
Total	a+c	b+d	$N = a+b+c+d$

2.5. Relative Risk (RR)

RR is a measure of association to assess the risk of an extreme precipitation event given a meteorological variable compared with of a non-extreme precipitation event (Table 1) by expression:

$$RR = \frac{\text{estimate of extreme occurrence}}{\text{estimate of non-extreme occurrence}} = \frac{a/(a+b)}{c/(c+d)} \quad (2)$$

2.6. Goodness-of-fit and ROC curve

For goodness-of-fit (GOF) analysis (Hosmer et al. 1991) three methods were used: Deviance Residual, Akaike Information Criterion (AIC) and p-value. Deviance Residual is a quality-of-fit statistical measure based on maximum likelihood using the sum of squared residuals in ordinary least squares. Akaike Information Criterion (AIC) is a quality-of-fit measure which seeks to select variables given joint variables that optimize the performance of the model with the minimum AIC value. The p-value is another measure that verifies whether the variables contained in the model have statistical significance, generally the $p < 0.05$ value is used to reject the null hypothesis. For assessing the accuracy of the model a receiver operating characteristic (ROC) graph was used. The ROC curve is a technique for visualizing, organizing and selecting classifiers, thus it evaluates the quality or performance of diagnostic tests (Park et al. 2004, Fawcett 2006).

Generally, the ROC analysis assesses the quality of a model's count of occurrences or not of precipitation extremes and the exposure factor presence or absent at an extreme condition. The most common measure used is Area under Curve (AUC) which interprets the average value of sensibility for all values of specificity with the aim to evaluate the overall performance of a diagnostic test, ranging between 0 and 1, wherein a bigger value suggests

the better overall performance of the diagnostic test (Park et al. 2004). All computation results were generated on R software (R Development Core Team 2013) with supporting packages: ggplot2 (Wickham 2013), MASS (Venables and Ripley 1994), msm (Jackson 2011), ROCR (Sing et al. 2005) and sandwich (Zeileis 2006).

3. Results and discussion

The Poisson regression model results, Relative Risk and goodness-of-fit analysis (deviance residual, AIC, p-value) are shown in Tables 3 e 4 for the four NEB sub-regions (NNEB, ENEB, SANEB and SNEB) shown on Figure 2. The 5,800 daily precipitation data points >95p were considered, corresponding to 290-295 extreme precipitation cases obtained. From the AIC test, the lowest value was obtained on NNEB from 1,473.6 (Table 2). By Residual Deviance Poisson regression, the model results ranged from 877.6 to 1,400.2, while for logistic regression verified by Correia-Filho et al. (2014) they varied from 1,366.1 to 1,923.4.

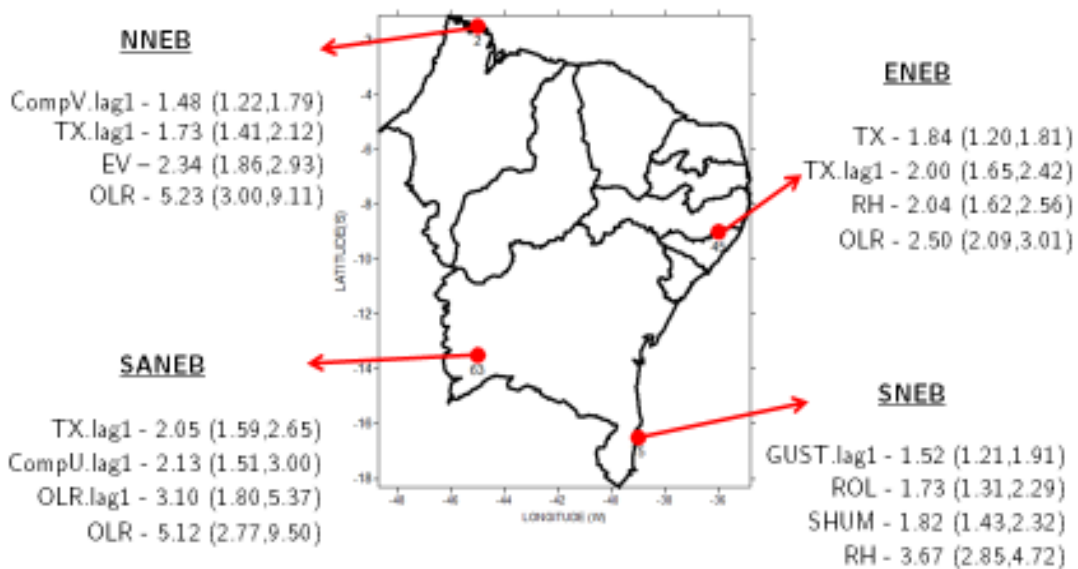


Figure 2 - Meteorological variables that influenced on NEB precipitation extremes by Poisson regression models. Each sub-region with the highest RR values.

This result shows the importance of the number of variables required for the model composition determined by AIC, selecting those variables that are most significant to the

model. The composition of the logistic regression models found by Correia-Filho et al. (2014) for the NNEB and SANEB included nine variables, and ENEB and SNEB included twelve variables, while on the Poisson regression seven, eight, nine, and eight variables were used for NNEB, SANEB, ENEB and SNEB, respectively.

Analyzing the measures of associations by RR, the variables that contributed to precipitation extreme occurrences in the NNEB (Northern of Maranhão State) were EV, TX.lag₁, CompV.lag₁ and OLR; in ENEB (Northeastern of Alagoas State) they were RH, TX, TX.lag₁ and OLR; in the SANEB (Southwestern of Bahia State) they were TX.lag₁, CompU.lag₁, OLR and SHUM.lag₁; and in the SNEB (Southeastern of Bahia State) they were RH, OLR, GUST.lag₁ and SHUM.lag₁. Based on RR results, there is evidence that the OLR is the indicator variable in three regions: ENEB, NNEB and SANEB, which demonstrates that the risk of precipitation extremes occurring is 2.50 times (95 percent confidence interval (CI): 2.09, 3.01), 5.23 times (95 percent confidence interval (CI): 3.00, 9.11), and 5.32 times (95 percent CI: 2.87, 9.86) respectively, while in SNEB the RH was the variable that increased the risk of precipitation extremes, at 3.67 times (95 percent CI: 2.85, 4.72), similar to the results found by Correia-Filho et al. (2014).

The OR results corroborate those of Sandeep and Stordal (2013) and the Correia-Filho et al. (2014) results that detected the OLR and extreme precipitation relationships in the tropics influenced by Madden Julian Oscillation (MJO) (Jones et al. 2004).

Silva and Kousky (2012) describe the MJO as an intraseasonal fluctuation that occurs in tropical regions with a typical duration of 30-60 days during wet season (3-4 months), with maximum activity between March to May (Madden and Julian 1971; Hendon and Salby 1994; Jones et al. 2004; Alves et al. 2012). The MJO favors the convective cloud formation of low-level moisture convergence causing the more intense precipitation suggested by Trenberth et al. (2003). Agreeing with Liu et al. (2009) that these precipitation extremes are not provoked only by temperature, but by a meteorological system and variables combination that provides the conditions at occurrence. The temperature also exerts a cause-effect influence on precipitation intensity.

A similar response was obtained by Correia-Filho et al. (2014), the variables that lead the occurrence of precipitation extremes in the NEB, were TX, SHUM and OLR, corroborating the results of Berg et al. (2013), which showed that an association between temperature and humidity at low levels increased these extremes.

In NNEB, the precipitation extremes are related with ITCZ position (Hastenrath and Greischar 1993; Greischar and Hastenrath 2000), which exert a great influence on NEB region originated by interhemispheric SST gradient; when this gradient is positive (negative) the displacement is more south (north) of the equator. This displacement causes effects on the meridional pressure gradient and meridional wind component (Greischar and Hastenrath 2000), transporting the heat and moisture of the ocean into the land and increasing the risk of precipitation extremes.

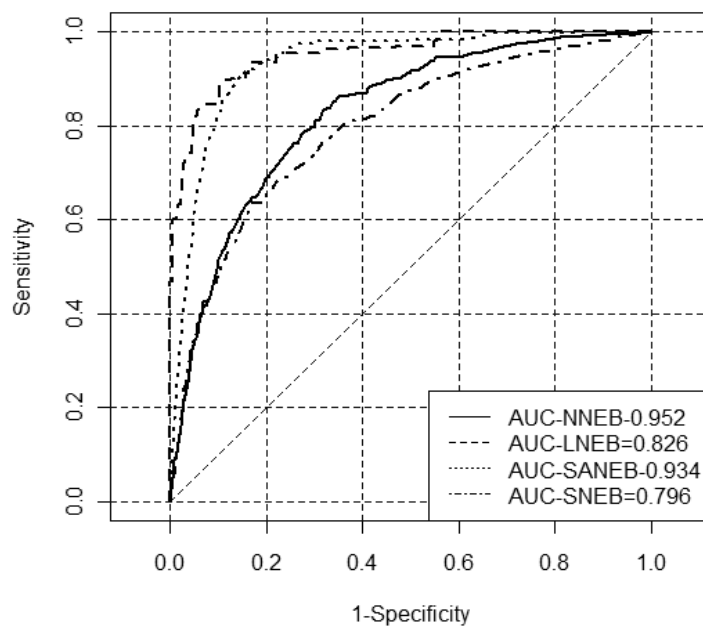


Figure 3 - ROC curve for Poisson regression models. AUC values on right bottom for NEB sub-regions: NNEB (solid line), ENEB (dashed line), SANEB (dotted line) and SNEB (dashed and dotted line).

In ENEB, these precipitation extremes are motivated by easterly wave disturbances that, if moving in a west-east direction, transporting the heat and moisture of the ocean to land, and allied the environmental conditions, favor the intensification of the MCS and MCC (Alves et al. 2001) over region (Kouadio et al. 2012).

The precipitation extremes in SANEB are boosted by MJO (Silva and Kousky 2012) associated with frontal systems and UTCV. The MJO combined with temperature and moisture advections and easterly winds, contributes to convective cloud formations.

Table 2 - Poisson regression models for NNEB and SANEB. Coefficient regression (Coeff), Standard Error (SE), p-value, Relative Risk (95% CI) and Goodness-of-fit (AIC test, deviance residual and degrees of freedom (df)).

NNEB				SANEB			
Variable	Coeff±SE	p-value	RR(95% CI)	Variable	Coeff±SE	p-value	RR(95% CI)
Intercept	-5.50±0.27	<0.001	--	Intercept	-	<0.001	--
TX	0.54±0.12	<0.001	1.73 (1.41,2.12)	TX	0.47±0.14	<0.001	1.60 (1.24,2.06)
EV	0.85±0.13	<0.001	2.34 (1.86,2.93)	TX.lag₁	0.71±0.14	<0.001	2.05 (1.59,2.65)
CompU	0.24±0.12	0.050	1.26 (1.04,1.54)	CompU	0.42±0.16	0.006	1.53 (1.19,2.03)
OLR	1.65±0.29	<0.001	5.23 (3.00,9.11)	OLR	1.63±0.29	<0.001	5.12 (2.77,9.50)
CompV.lag₁	0.39±0.13	0.002	1.48 (1.22,1.79)	OLR.lag₁	1.13±0.26	<0.001	3.10 (1.80,5.37)
CompV.lag ₃	0.28±0.13	0.020	1.32 (1.09,1.59)	CompU.lag₁	0.75±0.18	0.006	2.13 (1.51,3.00)
NINO2	0.19±0.12	0.100	1.21 (1.01,1.46)	NINO2	0.31±0.12	0.011	1.36 (1.11,1.68)
				SHUM.lag ₁	0.51±0.16	0.001	1.66 (1.24,2.22)
AIC=1,473.6				AIC=1,484.4			
Deviance residual=877.6		df=5,791		Deviance residual=888.4		df=5,790	

Table 3 - Poisson regression models for ENEB and SNEB. Coefficient regression (Coeff), Standard Error (SE), p-value, Relative Risk (95% CI) and Goodness-of-fit (AIC test, deviance residual and degrees of freedom (df)).

ENEb				SNEb			
Variable	Coeff±SE	p-value	RR(95% CI)	Variable	Coeff±SE	p-value	RR(95% CI)
Intercept	-4.65±0.13	<0.001	--	Intercept	-6.49±0.27	<0.001	--
RH	0.71±0.13	<0.001	2.04 (1.62,2.56)	TX	0.47±0.14	<0.001	1.60 (1.24,2.06)
TX.lag₁	0.70±0.15	<0.001	2.00 (1.65,2.42)	TX.lag₁	0.71±0.14	<0.001	2.05 (1.59,2.65)
TX	0.61±0.14	<0.001	1.84 (1.20,1.81)	CompU	0.42±0.16	0.006	1.53 (1.19,2.03)
TN	0.32±0.12	0.010	1.38 (1.14,1.67)	OLR	1.63±0.29	<0.001	5.12 (2.77,9.50)
GUST.lag ₁	0.32±0.14	0.019	1.38 (0.79,2.40)	OLR.lag₁	1.13±0.26	<0.001	3.10 (1.80,5.37)
OLR	0.92±0.15	<0.001	2.50 (2.09,3.01)	CompU.lag₁	0.75±0.18	0.006	2.13 (1.51,3.00)
OLR.lag ₁	0.43±0.15	0.006	1.53 (0.92,2.55)	NINO2	0.31±0.12	0.011	1.36 (1.11,1.68)
SHUM	0.43±0.14	0.002	1.53 (1.22,1.92)	SHUM.lag ₁	0.51±0.16	0.001	1.66 (1.24,2.22)
EV.lag ₁	0.39±0.14	0.006	1.48 (1.20,1.81)				
AIC=1,967.4				AIC=1,992.2			
Deviance residual=1,368.9		df=5,789		Deviance residual=1,400.2		df=5,791	

Furthermore, the penetration of frontal systems provoked by the northward displacement of the SACZ (Kousky 1979; Carvalho et al. 2004) on southern Bahia, favors the occurrence of precipitation extremes. These meteorological systems combined with the large scale circulation create an atmospheric environment favorable to the development of MCS over the region. Another system that influences the SANEB and other regions is the UTCV, characterized by subsidence motions on the central region this system, inhibiting convection when it is over the region (Kousky and Gan 1981, Chaves and Cavalcanti 2001, Cavalcanti 2012), while its edges occur intense rainfall. In SNEB, the precipitation extremes are resulting from different meteorological systems: breeze systems and easterly wave disturbance in the east; and SACZ, UTCV and frontal systems in the south.

The ROC curve for the Poisson regression shown in Figure 3 found an increase of 11.5% in the predictive capacity of the model generated for the NNEB when compared to logistic regression found by Correia-Filho et al. (2014), increasing from 0.837 to 0.952 (83.7-95.2%), while for other sub-regions there was a decrease of 0.015 (1.5%).

4. Conclusions

The initial conclusions presented in this work were:

- OLR indicated the occurrence of precipitation extremes, this result suggest the MJO occurrence over region combined with others meteorological systems such as: Intertropical Convergence Zone, South Atlantic Convergence Zone, Upper Tropospheric Cyclonic Vortex, Easterly Wave Disturbances and Frontal Systems favor the precipitation extremes over NEB;
- The highest relative risk to occur on precipitation extremes for ENEB, NNEB and SANEB were 2.50 times (95 percent confidence interval (CI): 2.09, 3.01), 5.23 times (95 percent confidence interval (CI): 3.00, 9.11), and 5.32 times (95 percent CI: 2.87, 9.86) respectively, all for OLR; wherein SNEB the highest relative risk to occur on precipitation extremes was 3.67 times (95 percent CI: 2.85, 4.72) for RH;
- PAC2 (second PCA of the Pacific SST, represents the seasonal pattern) also influences these extremes as shown by the RR results with 1.21 and 1.36 times the risk of precipitation extremes occurring in the NNEB and SANEB, respectively;

- The results found via Poisson regression exhibited similarities to the results found by Correia-Filho et al. (2014) in logistic regression; some similarities were seen in the results presented for both primary and secondary variables;
- From the AIC test, the lowest value ranged from 1,473.6 to 1,992.2. the Residual Deviance ranged from 877.6 to 1,400.2;
- The ROC curve analysis observed that all models showed were above 0.79, which shows that all the models have a good predictive ability; the SANEB sub-region exhibits the best AUC value of 0.952.

Acknowledgments

The author thanks Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES) for doctoral financial support; George Pedra and Naurinete Barreto for several contributions to the composition of this article. P. S. Lucio is sponsored by a PQ2 grant (Proc. 302493/2007-7) from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - Brazil).

References

- Allan RP, Soden BJ (2008) Atmospheric warming and the amplification of precipitation extremes. *Science*, 321(5895): 1481-1484.
- Alves JM, Teixeira RFB, Ferreira AG (2001) Um Intenso Sistema Convectivo de Mesoescala no Setor Leste do Nordeste: O Caso de 20 a 21 de maio de 1999. *Revista Brasileira de Meteorologia*, 16(1): 19-31. (In Portuguese)
- Alves JMB et al (2012) Sobre o sinal de um downscaling dinâmico às oscilações intrassazonais de precipitação no setor norte do Nordeste do Brasil. *Revista Brasileira de Meteorologia*, 27(2): 219-228. (In Portuguese)
- Berg P, Moseley C, Haerter JO (2013) Strong increase in convective precipitation in response to higher temperatures. *Nature Geoscience* 6(3): 181–185.
- Blain GC (2011) Modeling extreme minimum air temperature series under climate change conditions. *Rural Science* 41: 1877-1883.

- Carvalho LMV, Jones C, 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.
- Cavalcanti IFA (2012) Large scale and synoptic features associated with extreme precipitation over South America: A review and case studies for the first decade of the 21st century. *Atmospheric Research* 118: 27-40.
- Chandler RE (2005) On the use generalized linear models for interpreting climate variability. *Environmetrics* 16: 699-715 doi:10.1002/env.731.
- Chaves RR, Cavalcanti IFA (2001) Atmospheric circulation features associated with rainfall variability over southern Northeast Brazil. *Monthly Weather Review* 129(10): 2614–2626.
- Chen M, Shi W, Xie P, Silva VBS, Kousky VE, Wayne Higgins R, Janowiak JE (2008) Assessing objective techniques for gauge-based analyses of global daily precipitation. *Journal of Geophysical Research: Atmospheres* (1984–2012), 113(D4). doi: 10.1029/2007JD009132.
- Correia-Filho WLF, Lucio PS, Spyrides MHC (2014) Precipitation Extremes Analysis over the Brazilian Northeast via Logistic Regression, *Atmospheric and Climate Sciences*, 4(1) 2014:53-59. doi: 10.4236/acs.2014.41007.
- Dee DP et al. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* 22: 553–597.
- Dias, MAFS et al. (2013) Changes in extreme daily rainfall for São Paulo, Brazil. *Climatic change* 116 (3-4): 705-722.
- Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognition Letters* 27(8): 861–874.
- Gorman PAO, Schneider T (2009) The physical basis for increases in precipitation extremes in simulations of 21st-century climate change. *Proceedings of the National Academy of Sciences* 106: 14773-14777.
- Gregory RD et al (2009) An indicator of the impact of climatic change on European bird populations. *PLoS One* 4(3): e4678.
- Greischar L, Hastenrath S (2000) The rainy seasons of the 1990s in northeast Brazil: Real-time forecasts and verification. *Journal of Climate* 13(21): 3821-3826.

- Hastenrath S (1984) Interannual variability and annual cycle: Mechanisms of circulation and climate in the tropical Atlantic sector. *Monthly Weather Review* 112(6): 1097-1107.
- Hastenrath S (2011) Exploring the climate problems of Brazil's Nordeste: a review. *Climate Change* 112(2): 243–251.
- Hastenrath S, Greischar L (1993) Further work on the prediction of Northeast Brazil rainfall anomalies. *Journal of Climate* 6(4): 743–758.
- Hendon HH, Salby ML (1994) The life cycle of the Madden-Julian oscillation. *Journal of the Atmospheric Science* 51(15): 2225-2237.
- Hosmer DW, Taber S, Lemeshow S (1991) The importance of assessing the fit of logistic regression models: a case study. *American Journal of Public Health* 81(12) 1630–1635.
- Houze Jr RA (2012) Orographic effects on precipitating clouds. *Review of Geophysics* 50 RG1001. doi:10.1029/2011RG000365.
- Hundecha Y, St-Hilaire A, Ouarda TBMJ, El Adlouni S, Gachon P (2008) A Nonstationary Extreme Value Analysis for the Assessment of Changes in Extreme Annual Wind Speed over the Gulf of St. Lawrence, Canada. *Journal of Applied Meteorology and Climatology* 47(11): 2745–2759.
- Jackson CH (2011) Multi-state models for panel data: the msm package for R. *Journal of Statistical Software*, 38(8): 1-29.
- Jones C, Waliser DE, Lau KM, Stern W (2004) Global occurrences of extreme precipitation and the Madden-Julian Oscillation: Observations and predictability. *Journal of Climate* 17(23): 4575-4589.
- Joyce RJ, Janowiak JE, Arkin PA, Xie P (2004) CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology* 5(3): 487–503.
- Kayano, MT, Andreoli, RV, Souza, RAF (2013) Relations between ENSO and the South Atlantic SST modes and their effects on the South American rainfall. *International Journal of Climatology* 33: 2008-2023.
- Kousky VE (1979) Frontal Influences on Northeast Brazil. *Monthly Weather Review* 107: 1140-1153.

- Kousky VE, Gan MA (1981) Upper tropospheric cyclonic vortices in the tropical South Atlantic. *Tellus* 33(6): 538-551.
- Kouadio YK, Servain J, Machado LAT, Lentini CAD (2012) Heavy Rainfall Episodes in the Eastern Northeast Brazil Linked to Large-Scale Ocean-Atmosphere Conditions in the Tropical Atlantic. *Advances in Meteorology* 1–16.
- Kucharski F, Polzin D, Hastenrath S (2008) Teleconnection mechanisms of northeast Brazil droughts: modeling and empirical evidence. *Brazilian Journal of Meteorology* 23(2): 115–125.
- Lau KM, Wu HT, Bony, S (1997) The role of large-scale atmospheric circulation in the relationship between tropical convection and sea surface temperature. *Journal of Climate*, 10: 381-392.
- Liebmann B, Smith CA (1996) Description of a complete (interpolated) outgoing longwave radiation dataset. *Bulletin of the American Meteorology Society* 77: 1275–1277.
- Liebmann B, Kiladis GN, Allured D, Vera CS, Jones C, Carvalho LMV, Bladé I, Gonzáles PLM (2011) Mechanisms Associated with Large Daily Rainfall Events in Northeast Brazil. *Journal of Climate* 24(2): 376–396.
- Little M, Mcsharry PE, Taylor J (2009) Generalized Linear Models for Site-Specific Density Forecasting of U.K. Daily Rainfall. *Monthly Weather Review* (137): 1029-1045. doi:10.1175/2008MWR2614.1.
- Liu SC, Fu C, Shiu C-J, Chen J-P and Wu F (2009) Temperature dependence of global precipitation extremes. *Geophys Research Letters* 36(17): L17702.
- Madden RA, Julian PR (1971) Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific. *Journal of the Atmospheric Sciences* 28(5): 702-708.
- Meehl GA, Tebaldi C (2004) More intense, more frequent, and longer last- ing heat waves in the 21st century. *Science* 305: 994–997.
- Molion LCB, Bernardo, SO (2002) Uma Revisão da Dinâmica das Chuvas no Nordeste Brasileiro. *Revista Brasileira de Meteorologia* 17 (1):1-10. (In Portuguese)
- Nelder AJA, Wedderburn RWM (1972) Generalized Linear Models. *Journal of the Royal Statistical Society Series A (General)* 135(3): 370–384.

- Park SH, Goo JM, Jo CH (2004) Receiver operating characteristic (ROC) curve: practical review for radiologists. *Korean J of Radiol* 5(1): 11-18.
- R Development Core Team (2013) R: A language and environment for statistical computing. Vienna, Austria R Foundation for Statistical Computing.
- Sandeep S, Stordal F (2013) Use of daily outgoing longwave radiation (OLR) data in detecting precipitation extremes in the tropics. *Remote Sensing Letters* 4(6): 570–578.
- Silva VBS, Kousky VE (2012) The South American Monsoon System: Climatology and Variability, *Modern Climatology*, Dr Shih-Yu Wang (Ed.), ISBN: 978-953-51-0095-9, InTech, doi: 10.5772/38565. Available from: <http://www.intechopen.com/books/modern-climatology/the-south-american-monsoon-system-climatology-and-variability>.
- Sing T, Sander O, Beerenwinkel N, Lengauer T (2005) ROCr: visualizing classifier performance in R. *Bioinformatics* 21(20):7881
- Smith MD (2011) The ecological role of climate extremes: current understanding and future prospects. *Journal of Ecology* 99: 651–655.
- To T, Shen S, Atenafu E G, Guan J, McLimont S, Stocks B, Licskai C (2013) The air quality health index and asthma morbidity: a population-based study. *Environmental Health Perspectives* 121(1): 46-52. doi: 10.1289/ehp.1104816
- Torres RR, Ferreira NJ (2011) Case studies of easterly wave disturbances over Northeast Brazil using the Eta Model. *Weather and Forecasting* 26: 225-235.
- Trenberth KE, Dai A, Rasmussen RM, Parsons DB (2003) The Changing Character of Precipitation. *Bulletin of the American Meteorology Society* 84 (9): 1205–1217.
- Venables WN, Ripley BD (2002) *Modern applied statistics with S*. Springer.
- Wang C, Jiang B, Fan J, Wang F, Liu Q (2014) A Study of the Dengue Epidemic and Meteorological Factors in Guangzhou, China, by Using a Zero-Inflated Poisson Regression Model. *Asia-Pacific Journal of Public Health* 26(1):48-57.
- Wickham H (2009) *ggplot2: elegant graphics for data analysis*. Springer.
- Willems P, Arnbjerg-Nielsen K, Olsson J and Nguyen VTV (2012) Climate change impact assessment on urban rainfall extremes and urban drainage: Methods and shortcomings. *Atmospheric Research* 103: 106–118.

Wotling G, Bouvier C, Danloux J, Fritsch JM (2000) Regionalization of extreme precipitation distribution using the principal components of the topographical environment. *Journal of Hydrology* 233(1-4): 86–101.

Yan Z, Bate S, Chandler RE, Isham V, Wheater H (2002) An Analysis of Daily Maximum Wind Speed in Northwestern Europe Using Generalized Linear Models. *Journal of Climate* 15: 2073–2088.

Young KC (1994) A multivariate chain model for simulating climatic parameters from daily data. *Journal of Applied Meteorology* 33 (6): 661-671.

Zeileis A (2004) Econometric Computing with HC and HAC Covariance Matrix Estimators.

Zou G (2004) A Modified Poisson Regression Approach to Prospective Studies with Binary Data. *American Journal of Epidemiology* 159(7): 702–706.

CAPÍTULO III

Este capítulo está constituído pelo terceiro artigo da tese, intitulado “*Caracterização dos extremos climáticos no Nordeste do Brasil*”, que será submetido à revista *Nonlinear Processes in Geophysics*, em versão inglesa no fim de julho de 2014. Esta revista pelos indicadores da CAPES (Qualis) está classificada como uma revista **B1 na área interdisciplinar**, e possui um fator de impacto **JCR de 1,409**.

CARACTERIZAÇÃO DOS EXTREMOS CLIMÁTICOS NO NORDESTE DO BRASIL

Washington L. F. Correia-Filho¹; Paulo S. Lucio¹ e Maria H. C. Spyrides¹

¹Programa de Pós-Graduação em Ciências Climáticas, Centro de Ciências Exatas e da Terra,
Universidade Federal do Rio Grande do Norte, Natal, Brazil.

RESUMO: O impacto das mudanças climáticas tornou a ocorrência de eventos extremos mais frequentes. No Nordeste do Brasil (NEB), estes extremos estão intimamente à variabilidade da precipitação pluvial. Neste trabalho utilizaram-se as distribuições generalizadas de valores extremos (GEV) e de Pareto (GPD) (baseado no percentil 95), com o objetivo de caracterizar os extremos de precipitação no NEB e avaliar os níveis e períodos de retorno de 30,50 e 100 anos, utilizando dados diário, entre os meses de fevereiro a julho para o período de 1979-2011. Pelos parâmetros de localização e de escala, verificou-se que ambas as distribuições apresentaram configurações espaciais similares, com os máximos na costa leste do NEB (distúrbios ondulatórios de leste), no norte do Maranhão e do Ceará (Zona de convergência intertropical) e um núcleo no extremo oeste de Pernambuco (influenciado pela topografia do planalto da Borborema e da chapada do Araripe, constatado pela distribuição GEV). Pelo parâmetro de forma, constatou-se que a maioria das regiões selecionadas se ajustou pelas distribuições de Weibull negativa e beta ($\xi < 0$) para GEV e GPD, respectivamente. As estimativas dos níveis e períodos de retorno da GEV mostraram que no norte do Maranhão

poderá ocorrer pelo menos um evento extremo de precipitação no nível igual ou superior a 160,9 mm/dia em 30 anos, e de 182,1mm/dia em 100 anos, enquanto pela GPD, na região do centro da Bahia poderá ocorrerá pelo menos um evento extremo de precipitação no nível igual ou superior a 192,3 mm/dia em 30 anos, e 307,8mm /dia em 100 anos.

Palavras-chave: centro da Bahia, norte do Maranhão, oeste de Pernambuco.

ABSTRACT: The impact of climate change become more frequent occurrence of extreme events. On Northeastern Brazil (NEB), these extremes are closely to rainfall variability. This work was used the generalized extreme value (GEV) and Pareto (GPD) (based on 95th) distributions with the aim of characterizing the precipitation extremes on NEB and assess the levels and return periods for 30,50 and 100 years, using daily data between February to July months for 1979 to 2011 period. The location and scale parameters, it was found that both distributions showed the similar spatial configurations with the maximum on the east coast of the NEB (wave easterly disturbances), on northern Maranhão and Ceará States (intertropical convergence zone) and core in the extreme western of Pernambuco State (influenced by topography of the Borborema and Araripe plateaus, it verified by GEV distribution). The shape parameter, it was found that most of the selected regions are set by Weibull negative and Beta distributions ($\xi < 0$) for GEV and GPD, respectively. The levels and return periods estimated values by GEV distribution showed that on North Maranhão State may occur at least one extreme precipitation event for equal to or exceeding at level of 160.9 mm/day on next 30 years, and 182.1 mm/day on next 100 years, and while the GPD in the eastern Bahia State may occur at least an extreme precipitation event for equal to or exceeding at level of 192.3 mm/day on next 30 years, and 307.8 mm/day on next 100 years.

Keywords: centern Bahia, northern Maranhão, eastern Pernambuco.

1-INTRODUÇÃO

O aumento na frequência dos eventos extremos e de seus impactoss tem causado enormes prejuízos ao longo do globo (EASTERLING et al., 2000; MEEHL & TEBALDI,

2004; TRENBERTH, 2012; CAVALCANTI, 2012), tais eventos podem está relacionados à ação do homem no meio ambiente (FOLLAND et al., 2001).

A ocorrência destes extremos tem subsidiado diversas pesquisas na busca em diagnosticar os principais fatores que desencadeiam estes eventos (SANDEEP & STORDAL, 2013; CORREIA-FILHO et al., 2014), tais ocorrências estão relacionadas a combinação entre fatores ambientais e regionais.

Segundo Deser et al. (2010), alterações dos fatores ambientais podem ser captados pelas flutuações das variáveis termodinâmicas, como por exemplo, a variabilidade na temperatura e na capacidade de armazenamento de água da atmosfera, do que pelas variáveis dinâmicas, como mudanças relacionadas ao padrão de circulação atmosférica (MARAUN et al. 2010). De acordo com Trenberth (2012), os extremos devem ser avaliados de forma individual e em sua maioria não possuem associações, recorrendo-se a análise dos conglomerados de extremos na tentativa de detectar a ocorrência das anomalias da variabilidade natural.

Quando tratamos de extremos climáticos no Nordeste do Brasil (NEB), estes eventos estão intimamente ligados ao déficit (principalmente no semiárido nordestino) e excesso (regiões litorâneas e principalmente nas capitais) de precipitação pluvial, problemas que ocorrem com frequência na região (KURSCHASKI et al., 2008; GRIMM & TEDESCHI, 2009; LIEBMANN et al., 2011; HASTENRATH, 2011).

Diversos sistemas interferem no regime pluviométrico no NEB, principalmente em seus extremos, um destes é a variabilidade da Temperatura à Superfície do Mar (TSM) do Atlântico e Pacífico tropicais (ANDREOLI & KAYANO, 2006; GRIMM & TEDESCHI, 2009; KAYANO et al., 2013), que influencia fortemente de acordo com as fases anômalas de cada um dos oceanos que ora atuam em separado ou em conjunto. Há outros sistemas meteorológicos atuantes que favorecem a ocorrência dos extremos de precipitação no NEB, por exemplo, a Alta Subtropical do Atlântico Sul (ASAS) e os Distúrbios Ondulatórios de Leste (DOL). Segundo Rao et al. (1993) a ASAS possui um papel fundamental no mecanismo de transporte de vapor d'água em baixos níveis ao longo do NEB, favorecendo a ocorrência de precipitações intensas.

Para analisar a ocorrência destes extremos, diversas metodologias são utilizadas, entretanto duas mais conhecidas são: 1) índices climáticos, baseada em índices de extremos climáticos criados pela Organização Meteorológica Mundial (NAVEAU et al., 2005; WMO,

2009; SKANSI et al., 2013), com o propósito de verificar a tendência e sazonalidade dos extremos; 2) Teoria dos Valores Extremos (TVE) (FENG et al., 2007; BUIHAND et al., 2008; IDALINO et al., 2009; SCARROT & MACDONALD, 2012), baseadas nas distribuições generalizadas de valores extremos (GEV) e de Pareto (GPD), que são extraídas a localização, escala e forma destes eventos. A TVE é comumente aplicada em dados de temperatura e de precipitação, devido à sensibilidade de ambas para qualquer alteração sutil no clima pode ser facilmente perceptível por ambas as variáveis (SANTOS et al., 2006; IDALINO et al., 2009; SCARROT & MACDONALD, 2012). A partir da TVE para caracterizar a ocorrência dos extremos, o objetivo deste trabalho será captar sinais dos extremos de precipitação no NEB, a partir das distribuições GEV e GPD com o percentil 95, avaliando o comportamento dos níveis e períodos de retorno de 30,50 e 100 anos, utilizando dados diários de 1979-2011 para o período de fevereiro a julho (período chuvoso).

2-MATERIAL E MÉTODOS

2.1- DADOS

Neste trabalho serão utilizados dados diários de precipitação pluvial, disponibilizados pelo *Climate Prediction Center - National Oceanic and Atmospheric Administration* (CPC-NOAA) (JOYCE et al., 2004; CHEN et al., 2008), com resolução de grade de $0.5^\circ \times 0.5^\circ$ do período de fevereiro a julho de 1979-2011, e redimensionados para $1.5^\circ \times 1.5^\circ$, totalizando sete e sete pontos de grade, ao qual cobrem todo o Nordeste, abrangendo uma área entre as latitudes $0-18^\circ\text{S}$ e longitudes $35-48^\circ\text{W}$, vista na Figura 1. Destes sete e sete pontos, foram selecionados oito regiões aleatoriamente: região 1, no norte do Maranhão (NMA); região 2, no norte do Ceará (NCE); região 3, no sul do Maranhão (SMA); região 4, no oeste de Pernambuco (OPE), região 5, no nordeste de Alagoas (NAL) e regiões 6, 7 e 8 localizados no centro (CBA), oeste (OBA) e sul da Bahia (SBA), respectivamente.

2.2-TEORIA DOS VALORES EXTREMOS (TVE)

Nesta seção abordaremos a aplicação da TVE baseada nas distribuições GEV e GPD, encontrada com maiores detalhes em Coles et al. (2001).

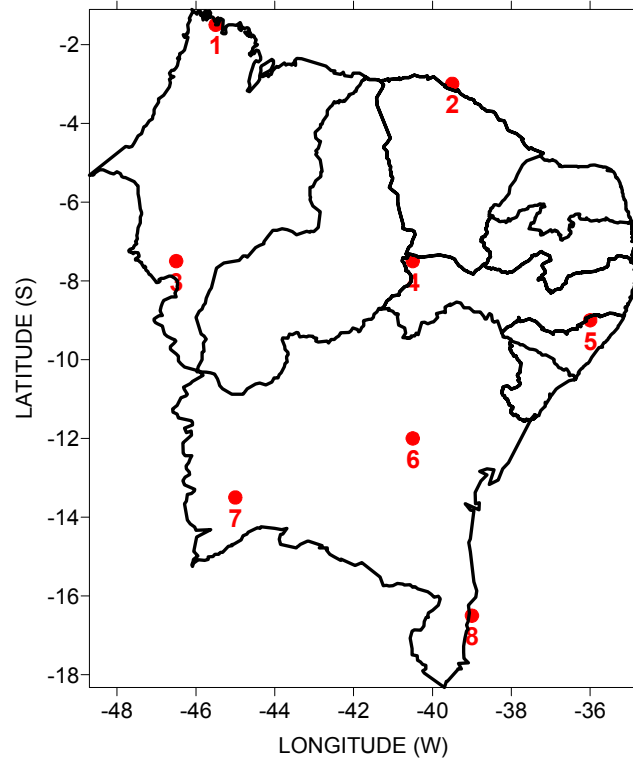


Figura 1 – Localização dos pontos das oito regiões selecionadas para avaliação do extremos no Nordeste do Brasil.

2.2.1-Distribuição Generalizada de Valores Extremos (GEV)

Este método consiste em extrair de uma série de observações independentes $X_t, t=1..n$, um valor máximo a partir de um intervalo de observações n , gerando uma série de máximos extraída destes blocos, que podem ser diário, mensal, anual. A função de distribuição acumulada da distribuição GEV é dada por:

$$F(z) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad z: 1 + \xi(x - \mu)/\sigma > 0 \quad (1)$$

Em que μ, σ e ξ são os parâmetros de localização, de escala e de forma, respectivamente.

Em caso particular, quando $\xi \rightarrow 0$ (Distribuição Gumbel), tem-se

$$G(z) = \exp \left\{ - \exp \left(- \frac{x - \mu}{\sigma} \right) \right\}, \quad -\infty < x < \infty \quad (2)$$

A obtenção do período de retorno pela seguinte forma:

$$x_T = \mu - \frac{\sigma}{\xi} \left[1 - \left\{ -\log \left(1 - \frac{1}{T} \right) \right\}^{-\xi} \right] \quad (3)$$

x_T é o nível de retorno de um evento extremo de precipitação ocorra em um dia dado um período de T anos.

2.2.2-Distribuição Generalizada de Pareto (GPD)

A distribuição Generalizada de Pareto (GPD) consiste em estabelecer um critério padrão ao conjunto de dados, sendo estes horários, diários ou mensais, sem que haja qualquer necessidade de formar blocos. Seja X_1, \dots, X_n , uma série de variáveis aleatórias independentes e identicamente distribuídas (iid), que contenha uma função de distribuição $F(x)$ desconhecida, sendo X_i as observações de valores diários de precipitação, o principal interesse é estimar o comportamento da precipitação acima do maior limiar μ (“*threshold*”) (LI et al., 2005). Neste trabalho aplicamos o método dos picos aleatórios acima dos limiares (“*Peaks Over Random Threshold*” - PORT). O PORT (SANTOS et al., 2006; SCARROT & MACDONALD, 2012) será aplicado nas séries temporais para cada ponto de grade nas variáveis de estudo, de acordo com Folland et al. (1999) a utilização dos valores dos limiares absolutos flexível (baseados em percentis) devido a heterogeneidade das áreas em estudo e também da possibilidade de processos físicos mascarar o resultado final, neste trabalho utilizamos o percentil 95. Para estimar a distribuição de excessos utilizamos a Equação:

$$F_u(y) = P\{X > u + y | X > u\} = \frac{F(u + y) - F(u)}{1 - F(u)} \quad (4)$$

Em que $0 \leq y < r_F - u$, $r_F = \inf(x: F(x) = 1) \leq \infty$, com o limite à direita $F(x)$.

A distribuição dos excessos representa a probabilidade em que um evento extremo de precipitação X_i exceda um limiar μ , ou seja, quando a GPD é parametrizada pelos parâmetros

de escala e forma nos termos dos limiares excedentes $x-\mu$, assim, representando os excedentes $x > \mu$, podemos representar a GPD desta maneira:

$$\Pr(X < x | X > u) = \begin{cases} 1 - \left[-1 + \xi \left(\frac{x - \mu}{\sigma_u} \right) \right]_+^{-1/\xi}, & \xi \neq 0 \\ 1 - \exp \left[-1 + \exp \left(\frac{x - \mu}{\sigma_u} \right) \right]_+, & \xi = 0 \end{cases} \quad (5)$$

Em que $y_+ = \max(y, 0)$. Quando o $\xi < 0$ no lado superior, então $\mu < x < \mu_-$.

De forma implícita a estimação do nível de retorno é proporcional aos limiares excedentes, quando $\varphi_u = \Pr(X > x)$, o período de retorno é calculado pela Equação:

$$F(X > x) = \varphi_u \left[1 - \Pr(X < x | X > u) \right] \quad (6)$$

Para a obtenção dos resultados dos parâmetros de localização (μ), de escala (σ) e de forma (ξ) das distribuições GEV e GPD e dos níveis de retorno para os períodos de 30, 50 e 100 anos utilizamos o software R (R DEVELOPMENT CORE TEAM, 2012) e os seguintes pacotes: “*POT*” (RIBATET, 2011), “*ismev*” (STEPHENSON, 2011).

3-RESULTADOS E DISCUSSÃO

Os resultados que serão exibidos a seguir referem-se aos parâmetros de localização, escala e de forma das distribuições GEV (Figura 1a) e GPD (Figura 1b) dos extremos de precipitação no NEB. O parâmetro de localização (μ) avalia o comportamento médio dos extremos de precipitação para GEV (Figura 1a) e GPD (Figura 1b), pode-se constatar que em ambas as distribuições às regiões de máximas estão concentradas em três regiões: 1) que se estende ao longo da costa leste do NEB, que vai do extremo leste do Rio Grande do Norte ao sul da Bahia; 2) nas regiões do NMA (região 1) e do NCE (região 2); e 3) na região do OPE (região 4). Enquanto isso, os mínimos estão concentrados na região do CBA (região 6), ao longo da região central da Bahia (verificados pela GPD). No que se refere aos valores do parâmetro de localização que estão contidos na Tabela 1, o maior foi observado na região do

NMA (região 1) com valores de 77,6mm e 42mm para as distribuições GEV e GPD, respectivamente. Já o menor foi verificada na região do NAL (região 5) com valores de 26,7mm e 13,3mm para GEV e GPD, respectivamente.

As regiões que apresentaram os maiores valores do parâmetro de localização (μ) são regiões com altos valores de precipitação, motivados pela passagem de diversos sistemas meteorológicos atuantes neste período, na parte norte do NEB é influenciada pela Zona de Convergência Inter-Tropical (ZCIT) (KAYANO et al., 2013), na parte leste (região costeira) sofre influência dos distúrbios ondulatórios de leste (DOL) (ALVES et al., 2001; TORRES & FERREIRA, 2011; KOUADIO et al., 2012). Entretanto, o núcleo de máxima observado na região do OPE (região 4) é influenciado pela influência topográfica do planalto da Borborema e da chapada do Araripe, localizada entre os estados de Pernambuco, Paraíba, Rio Grande do Norte e Ceará, também conhecida como Planalto Nordestino, em que estas regiões possuem uma elevação que varia entre 200-1000m de altitude, e que interfere substancialmente no regime pluviométrico, conseqüentemente em seus extremos.

Outro destaque se refere aos valores médios dos extremos apresentados pelo parâmetro de localização de GEV e GPD, a distribuição GPD a partir do método PORT de percentil 95 atenua os valores dos parâmetros de escala e de forma, apresentando resultado com valores muito baixos, principalmente na região do semiárido baiano com valores inferiores a 10 mm constatados na região do CBA (região 6).

Já os resultados apresentados pela GEV são similares aos que Easterling et al. (2000) apresentou quando avaliaram a tendência e variabilidade de extremos climáticos ao longo do globo, com os extremos de precipitação no NEB com valor diário de 50,8mm/dia, mas em certas regiões do NEB como o SMA (região 1), esse valor é facilmente atingido, ao qual já houve casos de extremos de precipitação diários com valores superiores a 100mm/dia, e o máximo registrado nesta região foi verificado em 2008 com o valor de 168,5mm.

Quanto aos valores deste parâmetro (Tabela 1), o maior foi verificado na região do NMA (região 1) com valores de 31,2mm e 17,4mm para as distribuições GEV e GPD, respectivamente; enquanto o menor foi constatado na região do NAL (região 5) com 7,8mm e 6,6mm para GEV e GPD, respectivamente.

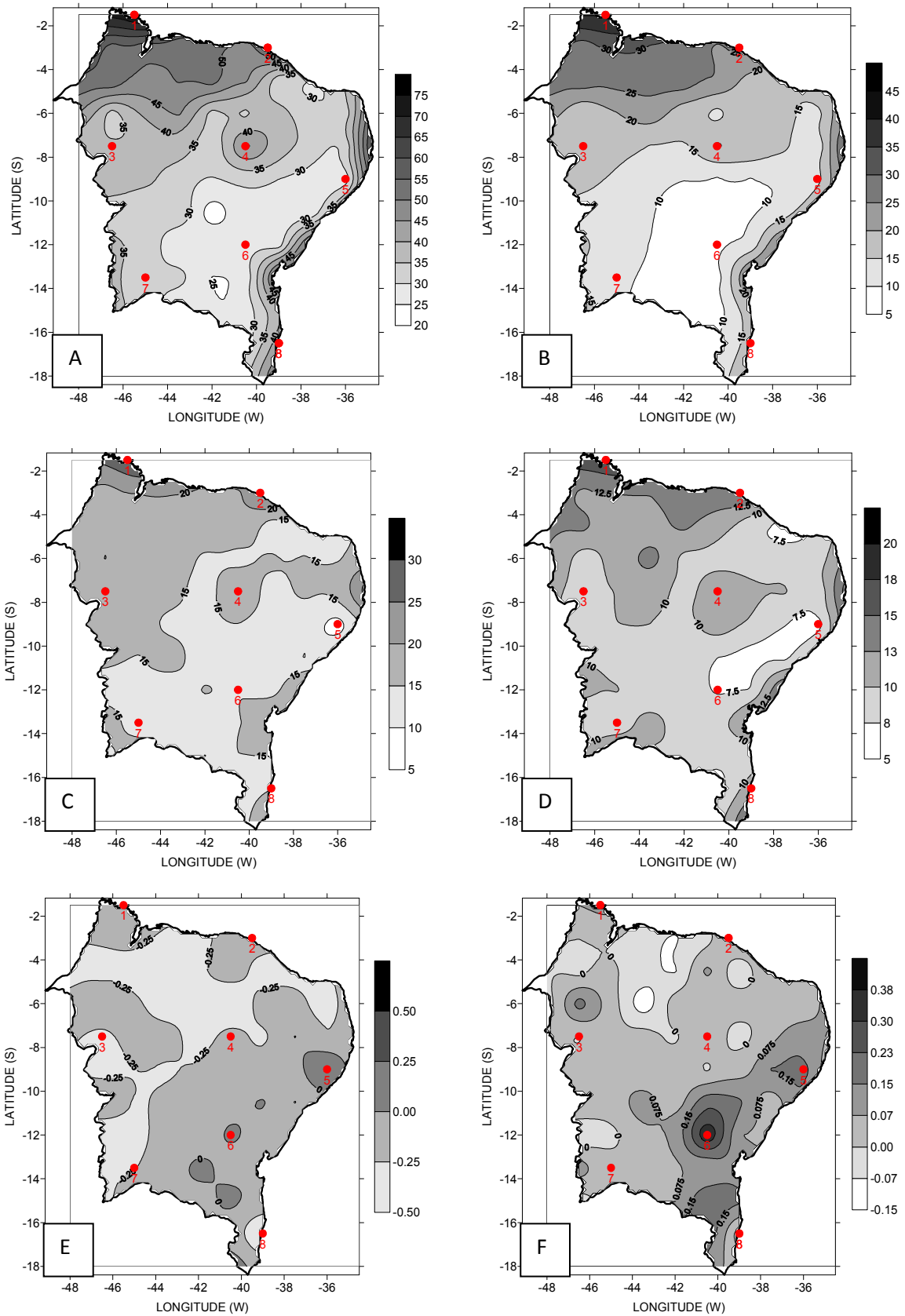


Figura 2 – Diagnóstico dos extremos de precipitação no NEB a partir dos parâmetros das distribuições GEV e GPD: de localização (mm) (a-b), de escala (mm) (c-d) e de forma (e-f).

Tabela 1 – Estatística descritiva das distribuições GEV e GPD dos extremos de precipitação para as regiões: NMA, NCE, SMA, OPE, NAL, CBA, OBA e SBA. Parâmetros de localização (μ) (mm), de escala (σ) e forma (ξ).

Localização	GEV			GPD		
	μ (mm)	σ (mm)	ξ	μ (mm)	σ (mm)	ξ
NMA (região 1)	77,6	31,2	-0,14	42,0	17,4	0,11
NCE (região 2)	59,0	28,5	-0,16	27,1	15,6	0,08
SMA (região 3)	35,6	15,2	-0,32	16,8	9,62	-0,00
OPE (região 4)	46,5	19,4	-0,17	20,8	11,8	0,07
NAL (região 5)	26,7	7,8	0,27	13,3	6,60	0,17
CBA (região 6)	28,3	15,0	0,03	9,2	7,10	0,37
OBA (região 7)	31,0	12,8	-0,26	11,5	8,84	0,03
SBA (região 8)	43,0	14,8	-0,54	17,1	11,66	-0,04

Já o parâmetro de forma (ξ) avalia o ajuste dos dados, verificando qual das famílias de distribuições existentes na GEV (Figura 1e) e GPD (Figura 1f), os dados se ajustam. Neste parâmetro é possível ver claramente o contraste entre os resultados gerados pela GEV e GPD. Pela GEV, as regiões do NAL (região 5) e CBA (região 6) os dados são ajustados pela distribuição Fréchet ($\xi > 0$) e Gumbel ($\xi = 0$), respectivamente, enquanto as demais foram ajustadas pela distribuição Weibull negativa ($\xi < 0$). No GPD, as regiões do SMA (região 3) e SBA (região 8), os dados são ajustados pelas distribuições exponencial ($\xi = 0$) e beta ($\xi < 0$), respectivamente, enquanto as demais se ajustam pela distribuição de Pareto ($\xi > 0$).

A verificação do ajuste dos dados é vital para a obtenção dos níveis e períodos de retorno estimados para 30, 50 e 100 anos, quando o parâmetro de forma é negativo o período de retorno aumenta enquanto o nível atenua até ser constante, quando o parâmetro de forma é positivo ocorre o contrário.

A partir da estimação dos níveis de e períodos de retornos para GEV e GPD (Tabela 2), constatamos que o maiores níveis de retorno foi verificados nas regiões do NMA (região 1) para GEV e do NCE (região 2) para GPD, os menores níveis foram observados no OBA

(região 7) para ambas as distribuições, na maioria das regiões selecionadas os níveis estimados pela GPD são maiores que os obtidos pela GEV.

Tabela 2 – Níveis (mm/dia) e períodos (T em anos) de retornos estimados para as distribuições GEV e GPD.

Localização	GEV			GPD		
	t=30	t=50	t=100	t=30	t=50	t=100
NMA (região 1)	160,9	170,3	182,1	88,6	99,1	115,0
NCE (região 2)	133,4	141,6	151,7	158,4	172,6	192,8
SMA (região 3)	67,1	69,6	72,3	76,4	81,1	87,6
OPE (região 4)	96,3	101,6	108,1	113,8	123,2	136,5
NAL (região 5)	70,7	81,8	99,7	88,6	99,1	114,9
CBA (região 6)	82,3	91,1	103,2	192,3	235,0	307,8
OBA (região 7)	59,7	62,2	65,2	72,4	77,8	85,3
SBA (região 8)	66,1	67,2	68,3	81,1	85,5	91,4

Pelo nível de retorno pela GEV, constata-se que poderá ocorrer pelo menos um evento extremo de precipitação de nível igual ou superior a 160,9 mm/dia em 30 anos, e de 182,1mm/dia em 100 anos. Enquanto isso nas regiões do SMA (região 3), OBA (região 7) e SBA (região 8), estes níveis permanecem praticamente estáveis, com pequenas alterações em seus extremos de precipitação.

Pela GPD, verifica-se um comportamento bem diferente do que foi visto na GEV, o maior nível de retorno ocorre na região do CBA (região 6) e que poderá ocorrer pelo menos um evento extremo de precipitação do nível igual ou superior a 192,3 mm/dia nos próximos 30 anos, e um de 307,8mm/dia nos próximos 100 anos.

4. CONCLUSÕES

Em caráter inicial, foi avaliado o comportamento dos extremos de precipitação do NEB pelas distribuições GEV e GPD para o período chuvoso de 1979 a 2011, e constatou-se que os máximos se concentraram em regiões litorâneas nas regiões do norte do Maranhão e

Ceará e na costa leste do NEB influenciadas pelos sistemas atmosféricos que atuam neste período (ZCIT e DOL), e também na região do extremo oeste de Pernambuco influenciado pela topografia, motivado pelo planalto da Borborema e da chapada do Araripe, contribuindo significativamente para o aumento dos índices pluviométricos, enquanto os mínimos foram verificados na parte central do NEB, no nordeste de Alagoas e centro da Bahia. Constatou-se pelo parâmetro de escala, regiões com características de máximas e mínimas semelhantes entre as obtidas pelo parâmetro de localização.

Quanto ao parâmetro de forma, houve um contraste entre os resultados gerados pela GEV e GPD. Verificou-se pela maioria das séries selecionadas que os dados se ajustaram pelas distribuições de Weibull negativa e beta ($\xi < 0$) para GEV e GPD, respectivamente. Os valores dos parâmetros para cada distribuição foram: de localização, entre 26,7 a 77,6mm para GEV e 9,2 a 42,0mm para GPD; de escala, entre 7,8-31,2mm para GEV e 6,6 a 17,4mm para GPD; e de forma, entre -0,54 a 0,27 para GEV e -0,04 a 0,37 para GPD.

Verificou-se uma diferença significativa nos resultados obtidos entre as distribuições utilizadas, tanto na intensidade da média dos extremos, na distribuição espacial do parâmetro de forma, mostrando indícios de que o método PORT da distribuição GPD, mesmo utilizando PORT com o percentil 95, os valores foram inferiores aos obtidos pela GEV.

As estimativas dos níveis e períodos de retorno mostraram que os valores estimados pela GPD foram superiores quando comparados a GEV, o maior valor do nível de retorno foi verificado na regiões do NMA (região 1) e NCE (região 2) para GEV e GPD, respectivamente, enquanto o menor nível foi constatado na região do OBA (região 7) para ambas as distribuições. Pela GEV conclui-se que a região do NMA (região 1) pode notificar pelo menos um evento extremo de precipitação de nível igual ou superior a 160,9 mm/dia nos próximos 30 anos, e um de 182,1mm /dia nos próximos 100 anos, enquanto pela GPD, na região do CBA (região 6) pode notificar pelo menos um evento extremo de precipitação de nível igual ou superior a 192,3 mm/dia nos próximos 30 anos e um de 307,8mm /dia para os próximos 100 anos.

AGRADECIMENTOS

O autor agradece a CAPES pela concessão da Bolsa de Doutorado. P. S. Lucio é bolsista PQ2 do CNPq (Proc: 309165/2010-5).

REFERÊNCIAS

- ALVES, J.M.B.; TEIXEIRA, R.F.B.; FERREIRA, A.G. Um Intenso Sistema Convectivo de Mesoescala no Setor Leste do Nordeste: O Caso de 20 a 21 de maio de 1999. **Revista Brasileira de Meteorologia**, v.16, n.1, p.19-31, 2001.
- ANDREOLI, R.V.; KAYANO, M. T. Tropical Pacific and South Atlantic effects on rainfall variability over Northeastern Brazil. **International Journal of Climatology**, DOI. 10.1002/JOC.1341, v. 26, n.13, p. 1895-1912, 2006.
- BUIHAND, T. A.; HAAN, L. de; ZHOU, C. On spatial extremes: With application to a rainfall problem. **Annals of Applied Statistics**, v.2, p.624–642, 2008, doi:10.1214/08-AOAS159.
- CAVALCANTI, I.F.A. Large scale and synoptic features associated with extreme precipitation over South America: A review and case studies for the first decade of the 21st century. **Atmospheric Research**, v. 118, p. 27-40, 2012.
- CHEN, M.; SHI, W.; XIE, P.; SILVA, V.B.S.; KOUSKY, V.E.; HIGGINS, R.W.; JANOWIAK, J.E. Assessing objective techniques for gauge-based analyses of global daily precipitation. **Journal of Geophysical Research**, v. 113 (D4), D04110, 2008.
- COLES, S.; BAWA, J.; TRENNER, L.; DORAZIO, P. **An introduction to statistical modeling of extreme values** (v. 208). London: Springer, 2001.
- CORREIA-FILHO, W.L.F.; LUCIO, P.S.; SPYRIDES, M.H.C. Precipitation Extremes Analysis over the Brazilian Northeast via Logistic Regression. **Atmospheric and Climate Science**, v. 4, n. 1, p. 53-59, 2014. doi: 10.4236/acs.2014.41007.
- DESER, C.; PHILLIPS, A.; BOURDETTE, V.; TENG H. Uncertainty in climate change projections: the role of internal variability. **Climate Dynamics**, 2010, doi:10.1007/s00382-010-0977-x.
- EASTERLING, D. R. et al. Observed Variability and Trends in Extreme Climate Events: A Brief Review*. **Bulletin of the American Meteorological Society**, v. 81, n. 3, p. 417-425, 2000.

GRIMM, A. M.; TEDESCHI, R. G. ENSO and Extreme Rainfall Events in South America. **Journal of Climate**, v.22, p.1589-1609, 2009, doi:10.1175/2008JCL2429.1.

FENG, S.; NADARAJAH, S.; HU, Q. Modeling annual extreme precipitation in China using the generalized extreme value distribution. **Journal of the Meteorological Society of Japan**, v.85, n. 5, p.599-613, 2007.

FOLLAND, C. K. et al. Workshop on Indices and Indicators for Climate Extremes, Asheville, NC, USA, 3–6 June 1997 Breakout Group C: Temperature Indices for Climate Extremes. **Climatic Change**, v. 42, n. 1, p. 31-43, 1999.

FOLLAND, C. K. et al., Observed climate variability and change. **Climate change: The scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change**, Cambridge University Press, p. 525-582, 2001.

HASTENRATH, S. Exploring the climate problems of Brazil's Nordeste: a review. **Climatic Change**, v.112, p.243–251, 2011 doi:10.1007/s10584-011-0227-1.

IDALINO, R. de C. de L.; OLIVEIRA, P. S. LUCIO, P. S. Modelagem de Extremos Meteorológicos via GEV e GPD - Uma análise comparativa de algumas capitais brasileiras. **Revista Brasileira de Estatística**, v.70, p.33–56, 2009.

JOYCE, R.J.; JANOWIAK, J.E.; ARKIN, P.A.; XIE, P. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. **Journal of Hydrometeorology**, v.5, n.3, p.487–503, 2004.

KAYANO, M. T.; ANDREOLI, R. V.; SOUZA, R. A. F. Relations between ENSO and the South Atlantic SST modes and their effects on the South American rainfall. **International Journal of Climatology**, v. 33, n. 8, p. 2008-2023, 2013.

KOUADIO, Y.K. et al. Heavy Rainfall Episodes in the Eastern Northeast Brazil Linked to Large-Scale Ocean-Atmosphere Conditions in the Tropical Atlantic. **Advances in Meteorology**, 2012.

KUCHARSKI, F. POLZIN, D., HASTENRATH, S. Teleconnection Mechanisms of Northeast Brazil Droughts: Modeling and Empirical Evidence. **Revista Brasileira de Meteorologia**, v.23, n. 2, p.115-125, 2008.

- LI, Y.; CAI, W.; CAMPBELL, E. P. Statistical Modeling of Extreme Rainfall in Southwest Western Australia. **Journal of Climate**, v.18, p.852–863, 2005.
- LIEBMANN, B.; KILARDIS, G. N.; ALLURED, D.; VERA, C. S.; JONES, C.; CARVALHO, L. M. V.; BLADÉ, I.; GONZÁLES, P. L. M. Mechanisms Associated with Large Daily Rainfall Events in Northeast Brazil, **Journal of Climate**, v.24, p.1589-1609, 2011, doi:10.1175/2008JCL2429.1.
- MARAUN, D.; RUST, H. W.; OSBORN, T. J. Synoptic airflow and UK daily precipitation extremes. **Extremes**, v. 13, n. 2, p. 133-153, 2010.
- MEEHL, G.A.; TEBALDI, C. More intense, more frequent, and longer last- ing heat waves in the 21st century. **Science**, v. 305, p. 994–997, 2004.
- NAVEAU, P., NOGAJ, M.; AMMANN, C.; YIOU, P.; COOLEY, D.; JOMELLI, V. Statistical methods for the analysis of climate extremes. **Comptes Rendus Geoscience**, v.337, n.1013–1022, 2005, doi:10.1016/j.crte.2005.04.015.
- R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. Vienna, Austria: **R Foundation for Statistical Computing**. ISBN 3-900051-07-0, 2012.
- RAO, V. B.; LIMA, M. C. de; FRANCHITO, S.H. Seasonal and Internannual Variations of Rainfall over Eastern Brazil. **Journal of Climate**. 6, p. 1754-1763, 1993.
- RIBATET, M. **POT: Generalized Pareto Distribution and Peaks Over Threshold**. R package version 1.1-1. <http://CRAN.R-project.org/package=POT>, 2011.
- SANDEEP, S.; STORDAL, F. Use of daily outgoing longwave radiation (OLR) data in detecting precipitation extremes in the tropics. **Remote Sensing Letters**, v.4, n.6, p.570–578, 2013.
- SANTOS, P. A.; ALVES, M. I. F.; GOMES, M. I. Peaks Over Random Threshold Methodology for tail index and high quantile Estimation. **REVSTAT**, v.4, p.227–247, 2006.
- SCARROTT, C.,MACDONALD, A. A review of Extreme Value Threshold Estimation and uncertainty quantification. **REVSTAT**, v.10, p.33–60, 2012.

SKANSI, M. M. et al. Warming and wetting signals emerging from analysis of changes in climate extreme indices over South America. **Global and Planetary Change**, v. 100, p. 295-307, 2013.

STEPHENSON, A. G. ismev: **An Introduction to Statistical Modeling of Extreme Values**. R package version 1.36. <http://CRAN.R-project.org/package=ismev>, 2011.

TORRES, R.R.; FERREIRA, N.J. Case studies of easterly wave disturbances over Northeast Brazil using the Eta Model. **Weather and Forecasting**, v.26, p.225-235, 2011.

TRENBERTH, K. E. Framing the way to relate climate extremes to climate change. **Climatic Change**, v.115, p. 283-290, 2012.

UVO, C.B. et al. The relationships between tropical Pacific and Atlantic SST and northeast Brazil monthly precipitation. **Journal of Climate**, v. 11, n. 4, 1998.

WMO. **Guidelines on Analysis of extremes in a changing climate in support of informed decisions for adaptation**, v.53, 2009.

CAPÍTULO IV

Este capítulo está constituído pelo quarto artigo da tese, intitulado “*Diagnóstico dos extremos de precipitação no Nordeste do Brasil via Modelos Vetoriais Lineares Generalizados*”, que será submetido à revista *Climate Dynamics*, em versão inglesa no fim de julho de 2014. Esta revista pelos indicadores da CAPES (*Qualis*) está classificada como uma revista **A1 na área interdisciplinar**, e possui um fator de impacto **JCR de 4,321**.

DIAGNÓSTICOS DOS EXTREMOS DE PRECIPITAÇÃO NO NORDESTE DO BRASIL VIA MODELOS VETORIAIS LINEARES GENERALIZADOS

Washington L. F. Correia-Filho¹; Paulo S. Lucio¹ e Maria H. C. Spyrides¹

¹Programa de Pós-Graduação em Ciências Climáticas, Centro de Ciências Exatas e da Terra, Universidade Federal do Rio Grande do Norte, Natal, Brazil.

RESUMO: O aumento dos eventos extremos e de seus impactos têm causado enormes prejuízos socioeconômicos. No Nordeste do Brasil (NEB), estes episódios de extremos estão relacionados à precipitação pluvial, verificados por episódios de secas e enchentes. Para diagnosticar estes extremos de precipitação no NEB foi utilizado os Modelos Vetoriais Lineares Generalizados via distribuição generalizada de valores extremos (MVLG-GEV) de forma múltipla. A partir dos parâmetros da distribuição GEV, concluiu-se: pela localização (μ), o maior valor foi verificado no norte do Maranhão com o valor de $88,26 \pm 6,42$ mm; pela escala (σ), que a maioria das regiões exibiram valores positivos, com o máximo de 7,40 no norte de Alagoas; e pela forma (ξ), que a maioria das regiões selecionadas foram ajustadas pela distribuição de Weibull negativa ($\xi < 0$). Os testes de log-verossimilhança negativa e AIC mostraram que os modelos do sul do Maranhão e da Bahia possuem maior precisão em seus resultados. As variáveis que influenciaram a ocorrência dos extremos de precipitação foram: componentes zonal e meridional do vento e a evaporação, temperatura à superfície do Atlântico e Pacífico. A partir dos níveis e períodos de retorno mostraram que os valores estimados pelo modelo MVLG-GEV, verificou-se que no centro da Bahia estimou o maior

nível de retorno, com pelo menos um evento extremo de precipitação do nível igual ou superior a 571,2 mm/dia nos próximos 30 anos, e de 875,7mm /dia em 100 anos.

Palavras-chave: Componente zonal e meridional do vento, valor- p , Modelos Vetoriais Lineares Generalizados.

ABSTRACT: The increase in extreme events and their impacts have caused great socioeconomic losses. On Northeastern Brazil (NEB), these extreme episodes are related to rainfall, its verified by droughts and floods episodes. For diagnose these precipitation extremes on NEB, it was used Vectorial Generalized Linear Models via generalized extreme value distribution (VGLM-GEV) in multiple forms. For the GEV distribution, it was concluded: location parameter (μ), the highest value was found on northern Maranhão State with the value of 88.26 ± 6.42 mm; scale parameter (σ), the most regions were showed the positive values with a maximum of 7.40 in the northern Alagoas State; and shape parameter (ξ), that the most of selected regions were adjusted by Weibull negative distribution ($\xi < 0$). Negative log-likelihood and AIC tests were showed that the southern Maranhão and southern Bahia State have a greater accuracy. The variables that influencing the occurrence of precipitation extremes were: zonal and meridional wind components and evaporation, Atlantic and Pacific sea surface temperature. The levels and return periods showed that the estimated values by VGLM-GEV models, it was found that the highest return level on centern of Bahia State at least one extreme precipitation event for equal to or exceeding at level of 571.2 mm/day on next 30 years, and 875.7 mm/day on next 100 years.

Keywords: zonal and meridional wind components; p value; Vectorial Generalized Linear Models.

1-INTRODUÇÃO

A grande preocupação da comunidade internacional acerca das mudanças climáticas nos últimos anos está relacionada à ocorrência frequente de eventos extremos e de seus impactos, que resultam em enormes prejuízos (JONES et al., 2004; SMITH, 2011; DAVISON & GHOLAMREZAEI, 2012; CAVALCANTI, 2012; SANDEEP & STORDAL, 2013).

Diversos fatores contribuem para a ocorrência destes eventos extremos, ao qual interferem tanto no aumento da frequência quanto na intensificação destes eventos, tais como: 1) a interação das variáveis ambientais, tais como: temperatura do ar (LIU et al., 2009), velocidade do vento (GASTINEAU & SODEN, 2011), Temperatura à Superfície do Mar (TSM) (KAYANO et al., 2013) e entre outras, associadas aos sistemas meteorológicos em diversas escalas de tempo e espaço (KURCHASKI et al., 2008, MARAUN et al., 2011) e 2) alterações no microclima local, provocados pelo desenvolvimento urbano (PATHIRANA et al., 2014) e 3) fatores topográficos (JOHANSSON & CHEN, 2003).

No Nordeste do Brasil (NEB), estes episódios de extremos climáticos estão relacionados à precipitação pluvial, ora resultantes de longos períodos de estiagem no semiárido nordestino como foi verificado nos anos de 2012 e 2013, ora provocados por chuvas extremas em capitais e regiões litorâneas (ALVES et al., 2001; CHAVES & CAVALCANTI, 2001; HASTENRATH, 2011; KOUADIO et al., 2012; CAVALCANTI, 2012; KAYANO et al., 2013).

Na tentativa de detectar os processos que desencadeiam estes extremos, a implementação de diversos métodos estatísticos têm contribuído de maneira substancial no aprimoramento dos resultados. Um dos métodos tradicionais para análise de extremos é a Teoria dos Valores Extremos (TVE), a partir das distribuições Generalizada de Valores Extremos (GEV) e Generalizada de Pareto (GPD) (DAVISON & GHOLAMREZAEI, 2012). Estes modelos são extremamente eficazes, porém sua análise é realizada de forma univariada baseada na própria série de dados, limitando a inclusão de fatores externos que possibilite uma melhor compreensão da ocorrência do evento extremo em questão. Uma maneira alternativa para diagnosticar a ocorrência destes extremos no NEB sob o aspecto multivariado é a utilização dos Modelos Vetoriais Lineares Generalizados, desenvolvido por Yee & Wild (1996), que incorpora diversos modelos em um só, neste trabalho abordaremos a GEV de forma múltipla ou multivariada (MVLG-GEV) (YEE & STEPHENSON, 2007; MARAUN et al., 2011; JONES et al., 2014), buscando responder os seguintes questionamentos:

- 1) Será que os resultados dos modelos MVLG estimam bem os parâmetros de localização (μ), de escala (σ) e de forma (ξ) para a distribuição GEV múltipla?
- 2) Será que os resultados estimados pelos modelos MVLG possuem uma boa capacidade preditiva quando nos referimos aos níveis e períodos de retorno?

Com o propósito em respondê-las, o objetivo deste trabalho é analisar os extremos de precipitação a partir do MVLG-GEV utilizando dados de reanálise de doze variáveis meteorológicas do período de 1979-2011.

2-MATERIAL E MÉTODOS

2.1-Área de estudo

O Nordeste Brasileiro (NEB) possui uma área de $1,5 \times 10^6$ km², compreendidas entre as latitudes 1-18°S e longitudes 35-47°W. A região é influenciada por diversos sistemas meteorológicos com características bem distintas. Utilizando o padrão de precipitação aplicado por Correia Filho et al. (2014), o NEB é subdividido em quatro sub-regiões: Leste (LNEB), influenciada por Sistemas Convectivos de Mesoescala (SCM) (MADDOX, 1980; ALVES, 2001; GARAVAGLIA et al., 2013), Distúrbios Ondulatórios de Leste (TORRES & FERREIRA, 2011; KOUADIO et al., 2012); Semiárido (SANEB), influenciada por SCM, CCM, Sistemas Frontais (SF), Vórtice Ciclônico de Ar Superior (VCAS) (KOUSKY & GAN, 1981; EVANS & BRAUN, 2012) e Oscilação Madden-Julian (OMJ) (MADDEN & JULIAN, 1971; SILVA & KOUSKY, 2012); Sul (SNEB), influenciado pela Zona de Convergência do Atlântico Sul (ZCAS), SF e o VCAS (CHAVES & CAVALCANTI, 2001; CARVALHO & OYAMA, 2013); e o Norte (NNEB) – influenciado pela Zona de Convergência Intertropical (ZCIT) e o VCAS; e o Anticiclone Subtropical do Atlântico Sul que influencia todas as sub-regiões (CAVALCANTI, 2012). Além destes sistemas meteorológicos que influenciam no regime pluviométrico, a variabilidade da TSM do Atlântico tropical e do Pacífico (TRENBERTH, 1997; SOUZA & CAVALCANTI, 2009; HASTENRATH, 2011; KAYANO et al., 2013) também interferem fortemente significativamente, tanto na frequência como na intensidade.

2.2-Dados

Neste trabalho utilizaram-se dados diários de precipitação pluvial disponibilizado pelo *Climate Prediction Center - National Oceanic and Atmospheric Administration* (CPC-NOAA) (JOYCE et al., 2004; CHEN et al., 2008), com resolução de grade de $1,5^\circ \times 1,5^\circ$ do período entre fevereiro à julho do período de 1979 a 2011. Diversas variáveis foram utilizadas nesta análise: umidade relativa (RH) (%), temperaturas mínimas (TN) e máximas (TX) ($^\circ\text{C}$), evaporação (EV) (mm), componentes zonal (U) e meridional (V) do vento (CompU e CompV), rajada máxima do vento (RAJ) (ms^{-1}) e umidade específica (SHUM) (g/kg), disponibilizados pelo ERA-INTERIM do *European Centre for Medium Weather Forecasts* (ECMWF) (DEE et al. 2011), dados interpolados de satélite de radiação de onda longa (ROL) (Wm^{-2}) do *Earth System Research Laboratory – NOAA* (ESRL-NOAA) (LIEBMANN & SMITH, 1996). Destes sete e sete pontos, foram selecionados oito regiões aleatoriamente: região 1, no norte do Maranhão (NMA); região 2, no norte do Ceará (NCE); região 3, no sul do Maranhão (SMA); região 4, oeste de Pernambuco (OPE), região 5, no nordeste de Alagoas (NAL) e regiões 6, 7 e 8 localizados no centro (CBA), oeste (OBA) e sul da Bahia (SBA), respectivamente.

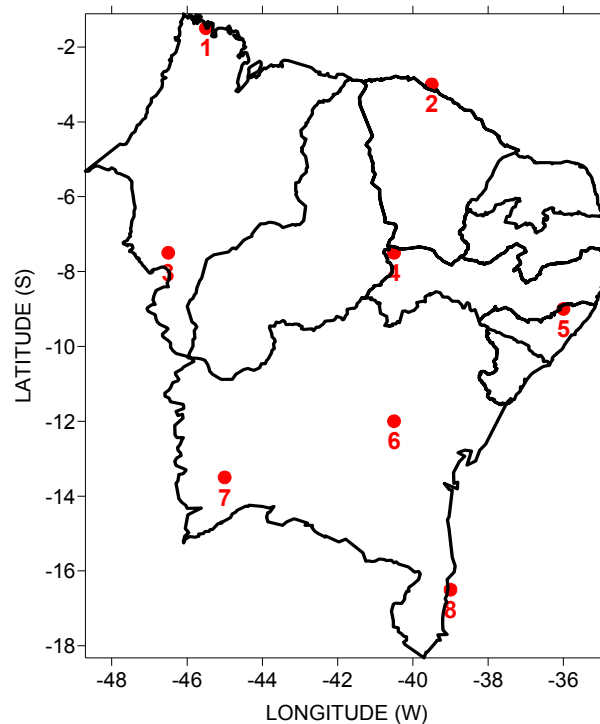


Figura 1 – Localização das oito regiões selecionadas para análise de extremos no Nordeste do Brasil via modelo MVLG-GEV.

2.3-Análise de componentes principais para as regiões do Atlântico e Pacífico

Utilizaram-se dados diários da temperatura da superfície do mar (TSM) das regiões tropicais dos Oceanos Atlântico (ATL, 21°S-21°N e 57°W-15°E) e Pacífico (PAC, região dos El Niños 1.2, 3, 3.4 e 4, entre 5°S-5°N, 90°W-160°E e 10°S-0°, 80°W-90°W) fornecido pelo ERA-INTERIM (DEE et al., 2011). A inserção dos dados de TSM seguiram os passos aplicados por Correia Filho et al. (2014), as variáveis explicativas seguiram estas etapas: 1) defasagem de 90 dias para a região do Pacífico (TRENBERTH, 1997) e de 30 dias para a região do Atlântico (CHAVES & CAVALCANTI, 2001; MOLION & BERNARDO, 2002); 2) cálculo das anomalias para cada região; e por fim, 3) análise das componentes principais (ACP) para extração do comportamento padrão destas regiões.

A aplicação das ACP na TSM das regiões do Pacífico e Atlântico teve como propósito captar os principais sinais de variabilidade, já que ambas as regiões influenciam de maneira significativa no padrão pluviométrico no NEB (SOUZA & CAVALCANTI, 2009; HASTENRATH, 2011; KAYANO et al., 2013). Pelas ACP obtidas, foram extraídas duas componentes para cada região a primeira componente exibe o comportamento padrão de cada uma das regiões de TSM e a segunda exibe o padrão de variabilidade sazonal. No Pacífico, a variância explicada obtida é de 74.91% (a 1ª ACP com 57,09% e a 2ª ACP com 17,82%), enquanto no Atlântico a variância explicada obtida é de 78.48% (a 1ª ACP com 68,02% e a 2ª ACP com 10,46%). Após a captação das ACP das regiões do Atlântico e Pacífico, realizou-se o processo de verificação individual para avaliar quais variáveis farão parte do modelo.

2.4-Escolha das variáveis

Após o procedimento de extração das ACP da TSM do Atlântico e Pacífico, todas as variáveis passaram pelo critério de seleção com o objetivo de identificar individualmente aquelas que favorecem o extremo de precipitação, este critério se baseia no *p*-valor ser inferior a 0,05, aquelas que apresentam este valor são incluídas no modelo geral. Quanto a escolha das variáveis foram em decorrência da influência que cada uma exerce sobre o processo de formação de maneira direta ou indireta dos extremos de precipitação.

Iniciamos pela temperatura, variável responsável por iniciar o processo de liberação calor para a atmosfera, quando esta é associada as umidades relativa específica (ambas quantifica a umidade na região), fornecem uma condição propícia a formação de precipitação

(O’GORMAN & SCHNEIDER, 2009; BERG et al., 2013). Já a TSM do Atlântico e Pacífico, possui uma influência significativa, além da liberação de calor, fornece também uma grande quantidade de vapor d’água dos oceanos para atmosfera (KOUADIO et al., 2012), substancial no regime pluviométrico do Norte e Leste do NEB (KAYANO et al., 2013), as TSM anômalas intensificam ou inibem a formação dos processos convectivos na precipitação.

As componente zonal e meridional do vento e a rajada, estão relacionadas aos padrões de circulação em superfície, ao qual exerce o papel de transportador de umidade e calor no processo inicial na formação da precipitação, e conseqüentemente, em seus extremos (GASTINEAU & SODEN, 2011). A evaporação é responsável pela liberação de calor para a atmosfera, que associada ao vento atuam fortemente no processo de modulação da convecção tropical, quanto maior a evaporação maior será a magnitude do vento, isto contribui para a desestabilização da camada limite e a partir do processo de realimentação da umidade forçará a formação de convecção profunda, resultando no aumento da precipitação (BACK & BRETHERTON, 2005).

A radiação de onda longa (ROL) é importante no processo de identificação de nuvens com desenvolvimento intenso, que apresentam forte convecção profunda, que indica que pode ocorrer precipitação intensa, a convecção tropical é caracterizada por ROL seja inferior à 240 Wm^{-2} (LAU et al., 1997).

2.5-Métodos Estatísticos

2.5.1-Modelos Vetoriais Lineares Generalizados (MVLG)

Para a análise de eventos extremos, tradicionalmente é realizado pela teoria de valores extremos (distribuições GEV e GPD), só que este método apenas utiliza entrada de dados univariada. Nos últimos anos, diversos métodos estatísticos surgiram como alternativa para analisar os extremos, incorporando a TVE aos modelos de regressão, incluindo variáveis explicativas no modelo, aprofundando e aprimorando a interpretação dos resultados.

Um destes modelos são os Modelos Vetoriais Lineares Generalizados (MVLG) (YEE & WILD, 1996), ao qual é a extensão dos tradicionais Modelos Lineares Generalizados (MLG) (NELDER & WEDDERBURN, 1972). Uma das vantagens deste modelo é flexibilidade nos tipos de dados de entrada e de resposta, o que tem contribuído

significativamente no aperfeiçoamento dos resultados com respostas univariada, múltipla e/ou multivariada (YEE & STEPHENSON, 2007; MARAUN et al. 2011; JONES et al., 2014).

A utilização do modelo MVLG utilizando a distribuição GEV como função de ligação, irá avaliar a influência das variáveis meteorológicas nos extremos de precipitação. Para isto, utilizamos a distribuição Generalizada de Valores Extremos (GEV) com a finalidade de modelar o comportamento dos eventos extremos de precipitação (COLES et al., 2001).

O MVLG está descrito de forma mais detalhes em Yee & Wild (1996). Supondo que a resposta observada y é um vetor q -dimensional, o MVLG é definido como um modelo para qual o distribuição condicional de Y dado um x explicativo na forma:

$$f(y | x; B, \varphi) = h \cdot (y, \eta_1, \dots, \eta_M, \varphi) \quad (1)$$

Para função conhecida $h(\cdot)$, em que $B = (\beta_1, \beta_2, \dots, \beta_M)$ é uma matriz $p \times M$ de coeficientes de regressão desconhecidos, o j -ésimo preditor linear é:

$$\eta_j = \eta_j(x) = \beta_j^T x = \sum_{k=1}^p \beta_{(j)k} x_k, \quad j = 1, \dots, M \quad (2)$$

Em que $x = (x_1, \dots, x_p)^T$, se x_1 for o intercepto então $x_1 = 1$.

2.5.2-Distribuição Generalizada de Valores Extremos (GEV)

Este método consiste em extrair de uma série de observações independentes $X_t, t=1 \dots n$, um valor máximo a partir de um intervalo de observações n , gerando uma série de máximos extraída destes blocos, que podem ser diário, mensal, anual. A função de distribuição acumulada da distribuição GEV é dada por:

$$F(z) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad z : 1 + \xi(x - \mu)/\sigma > 0 \quad (3)$$

Em que μ , σ e ξ são os parâmetros de localização, de escala e de forma, respectivamente. Em caso particular, quando $\xi \rightarrow 0$ (Distribuição Gumbel), tem-se

$$G(z) = \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}, \quad -\infty < x < \infty \quad (4)$$

Para os casos com $\xi < 0$ ou $\xi > 0$ são conhecidas como distribuições Weibull negativa e Fréchet. Para o ajuste da série de dados utilizaremos o método da máxima verossimilhança, obtida pela seguinte expressão:

$$L(\mu, \sigma, \xi) = \frac{1}{\sigma^n} \prod_{i=1}^n \left(1 + \xi \frac{x_i - \mu}{\sigma}\right)^{-(1/\xi + 1)} \times \exp\left\{-\sum_{i=1}^n \left(1 + \xi \frac{x_i - \mu}{\sigma}\right)^{-1/\xi}\right\} \quad (5)$$

Em que x_1, \dots, x_n são máximos anuais para n anos a uma dada série. Após a identificação da distribuição de ajuste do modelo, será estimado os níveis de retorno do extremo de precipitação, dada pela seguinte expressão:

$$x_T = \mu - \frac{\sigma}{\xi} \left[1 - \left\{ -\log\left(1 - \frac{1}{T}\right) \right\}^{-\xi} \right] \quad (6)$$

x_T é o nível de retorno de um evento extremo de precipitação ocorra em um dia dado um período de T anos.

Os resultados computacionais foram gerados no software R (R DEVELOPMENT CORE TEAM, 2014), pelos seguintes pacotes: “MASS” (VENABLES & RIPLEY, 2002), “VGAM” (YEE, 2013).

3-RESULTADOS E DISCUSSÕES

Os resultados exibidos a seguir se referem aos modelos gerados pelo MVLG, na verificação dos extremos de precipitação no NEB. Na Tabela 1, são apresentados a qualidade de ajuste da regressão MVLG-GEV (log-verossimilhança negativa, AIC, p-valor), nas Tabelas

2 e 3 são mostradas os coeficientes estimados dos modelos MVLG-GEV e os parâmetros de localização (μ), de escala (σ) e de forma (ξ) da distribuição GEV.

Os testes de log-verossimilhança negativa e AIC analisam a qualidade de ajuste do modelo, quanto menor for o valor de ambos os testes mostra que o modelo possui maior precisão, e verifica-se pelo primeiro teste que os valores variam entre 106,95 no NMA (região 1) a 137,84 no CBA (região 6), enquanto pelo segundo, os valores variam entre 237,56 no SMA (região 3) a 297,68 no CBA (região 6). Nota-se que em ambos os testes que no SBA (região 8) possui uma maior precisão, enquanto a menor é verificada no CBA (região 6).

Quando avaliamos os resultados dos parâmetros da distribuição GEV, o parâmetro de localização (μ) que representa a média dos extremos de precipitação, é possível verificar que o menor valor é obtido no NAL (região 5) com o valor de $17,22 \pm 7,14$ mm, enquanto o maior é verificado no NMA (região 1) com o valor de $88,26 \pm 6,42$ mm.

O parâmetro de escala (σ) que representa o desvio padrão deste extremos, verifica-se em sete das oito regiões escolhidas foram obtidos valores positivos com o máximo de 7,40 no NAL (região 5), enquanto o menor foi constatado no OPE (região 4) com o valor negativo de 3,25, este valor pode resultar em uma atenuação dos valores nos níveis de retorno. O parâmetro de forma (ξ) avalia o ajuste da distribuição de densidade da série, nota-se que a maioria dos locais estudados a série de dados são ajustados pela distribuição de Weibull negativa ($\xi < 0$), com exceção no NAL (região 5) e CBA (região 6) que suas séries são ajustadas pela distribuição de Fréchet ($\xi > 0$).

Analisando os coeficientes dos modelos MVLG para as oito regiões selecionadas (Tabelas 2 e 3), nota-se uma forte influência positiva da componente zonal do vento (CompU) na maioria das séries, apenas no SBA (região 8) ocorre o contrário. Outra variável que subsidia a ocorrência dos extremos de precipitação é a componente meridional do vento (CompV), com influência positiva no NCE (região 2) e OBA (região 7), e negativa no NMA (região 1). No SMA (região 3), NAL (região 5) e CBA (região 6), a variável motivadora para a ocorrência dos extremos de precipitação é a evaporação.

Outra observação bastante significativa são as ausências de variáveis atuantes na formação de nuvens e que favorecem a ocorrência de extremos de precipitação, como as temperaturas máximas e mínimas, umidades relativa e específica.

Tabela 1 – Teste da log-verossimilhança negativa, Graus de Liberdade (GL) e teste AIC dos modelos MVLG-GEV para as regiões: NMA, NCE, SMA, OPE, NAL, CBA, OBA e SBA.

	NMA	NCE	SMA	OPE	NAL	CBA	OBA	SBA
AIC	265,15	296,87	243,90	270,53	247,91	297,68	262,09	237,56
-Loglik	106,95	137,43	106,96	124,27	112,96	137,84	118,04	107,78
d.f.	85	85	78	85	85	85	83	85

Tabela 2 – Coeficiente de regressão (Coef), desvio padrão (DP) e valor- p para o NMA, NCE, SMA e OPE.

Variáveis	NMA		NCE		SMA		OPE	
	Coef±DP	valor- p	Coef±DP	valor- p	Coef±DP	valor- p	Coef±DP	valor- p
Intercepto (μ)	88,26±6,42	<0,001	64,34±4,30	<0,001	59,70±1,71	<0,001	43,55±3,86	<0,001
Intercepto (σ)	1,55±0,42	<0,001	2,54±0,35	<0,001	-3,25±0,53	<0,001	1,99±0,27	<0,001
Intercepto (ξ)	-2,57±0,88	0,004	-0,18±0,18	0,324	-0,72±0,25	0,005	-3,61±2,40	0,136
EV (μ)					11,70±1,79	<0,001		
EV (σ)					-3,40±0,46	<0,001		
CompU (μ)	3,08±1,40	0,031	3,35±1,00	0,001	5,53±0,44	<0,001		
CompU (σ)	-0,64±0,11	<0,001	0,15±0,09	0,089	-0,32±0,08	<0,001		
CompV (μ)	-2,16±1,07	0,046	2,74±0,47	<0,001	-0,65±0,08	<0,001	2,18±0,91	0,019
CompV (σ)	-0,09±0,06	0,132	-0,13±0,07	0,085	0,12±0,02	<0,001	-0,35±0,05	<0,001
GUST (μ)							0,91±0,45	0,048
GUST (σ)							0,17±0,03	<0,001
ATL1 (μ)	-0,84±0,14	<0,001	0,77±0,09	<0,001	-0,65±0,99	<0,001		
ATL1 (σ)	-0,05±0,01	<0,001	-0,03±0,01	0,026	0,03±0,14	0,082		
ATL2 (μ)			-0,51±0,15	0,001	0,67±0,27	<0,001	0,22±0,15	0,151
ATL2 (σ)			-0,10±0,02	<0,001	-0,02±0,20	0,301	0,11±0,01	<0,001
PAC1 (μ)	-0,54±0,09	<0,001			-0,11±0,38	0,005	0,06±0,03	0,079
PAC1 (σ)	0,03±0,05	<0,001			-0,03±0,01	<0,001	-0,02±0,00	<0,001

Tabela 3 – Semelhante à Tabela 2, mas para o NAL, CBA, OBA e SBA.

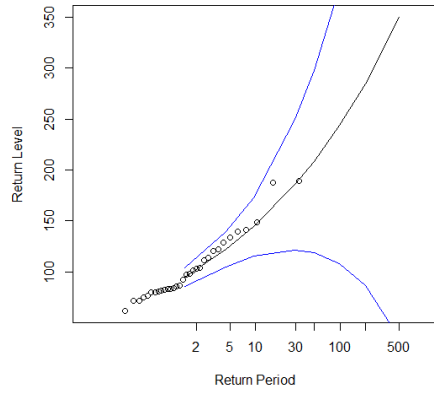
Variable	NAL		CBA		OBA		SBA	
	Coef±DP	valor- <i>p</i>	Coef±DP	valor- <i>p</i>	Coef±DP	valor- <i>p</i>	Coef±DP	valor- <i>p</i>
Intercepto (μ)	17,22±7,14	0,018	55,90±7,08	<0,001	60,65±4,33	<0,001	44,99±0,72	<0,001
Intercepto (σ)	7,40±1,70	<0,001	3,86±0,72	<0,001	3,15±0,60	<0,001	0,28±0,21	0,178
Intercepto (ξ)	0,33±0,15	0,032	0,24±0,16	0,116	-0,75±0,25	0,004	-2,59±0,98	0,010
EV (μ)	8,47±1,39	<0,001	17,16±4,49	<0,001				
EV (σ)	3,82±0,85	<0,001	0,72±0,63	0,254				
OLR (μ)	0,08±0,02	<0,001						
OLR (σ)	-0,01±0,01	0,067						
CompU (μ)	0,63±0,15	<0,001	1,15±0,63	0,071	1,28±0,40	0,002	-0,94±0,40	0,021
CompU (σ)	-0,21±0,09	0,017	0,28±0,07	<0,001	-0,17±0,08	0,030	-0,46±0,06	<0,001
CompV (μ)					5,41±0,51	<0,001	-0,50±0,18	0,006
CompV (σ)					-0,37±0,09	<0,001	0,06±0,03	0,029
GUST (μ)					-1,07±0,55	0,053		
GUST (σ)					-0,26±0,09	0,006		
ATL1 (μ)			-0,20±0,06	0,001	-0,73±0,17	<0,001	-0,11±0,06	0,093
ATL1 (σ)			-0,03±0,01	0,004	0,04±0,01	0,004	0,06±0,01	<0,001
ATL2 (μ)								
ATL2 (σ)								
PAC1 (μ)	-0,16±0,01	<0,001	0,24±0,06	<0,001				
PAC1 (σ)	0,01±0,01	0,072	-0,03±0,01	<0,001				
PAC2 (μ)					-0,65±0,14	<0,001	-0,46±0,05	<0,001
PAC2 (σ)					0,06±0,02	0,001	0,05±0,01	<0,001

Durante a composição de cada modelo MVLG-GEV para cada região, estas variáveis “ausentes” quando inseridas nos modelos, apenas contribuíam de duas formas: aumentando de maneira abrupta o desvio padrão, ou seja, aumentando significativamente a variação residual, o que comprometia o ajuste do modelo. A resposta extraída pelo modelo MVLG-GEV é bem interessante, pois constata três variáveis influenciam significativamente na ocorrência dos extremos de precipitação são: componentes zonal (CompU) e meridional (CompV) do vento e a evaporação. Outras variáveis também influenciam os extremos, de maneira coadjuvante, tais como: radiação de onda longa, rajada máxima e as ACP da TSM do Atlântico e Pacífico tropicais.

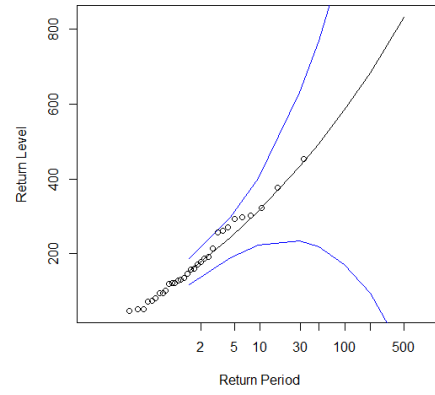
Tabela 4 – Níveis (mm/dia) e períodos (T em anos) de retorno estimados dos modelos MVLG-GEV gerados para as oito regiões selecionadas ao longo do NEB.

Localização	Período (em anos)			Localização	Período (em anos)		
	t=30	t=50	t=100		t=30	t=50	t=100
NMA (região 1)	185,7	208,7	243,7	NAL (região 5)	190,5	197,2	204,4
NCE (região 2)	432,1	493,7	543,3	CBA (região 6)	571,2	687,2	875,7
SMA (região 3)	245,5	327,4	437,5	OBA (região 7)	103,6	108,2	114,2
OPE (região 4)	155,0	177,8	214,4	SBA (região 8)	91,6	95,8	101,1

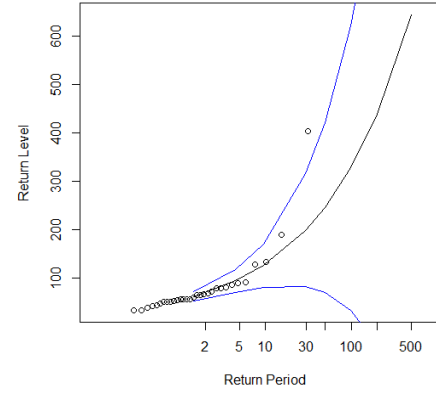
Os maiores valores dos parâmetros de localização (μ) são verificados na parte norte do NEB, no NMA (região 1), NCE (região 2) e SMA (região 3), estes máximos podem estar interligados a variabilidade da TSM do atlântico como do Pacífico, as anomalias geradas por esta variável acarreta em alterações nos campos de pressão em superfície, alterando os padrões de vento e também a circulação atmosférica, associadas ao posicionamento da ZCIT (SOUZA & CAVALCANTI, 2009). Quando estas anomalias são positivas (negativas), o seu deslocamento tende ser mais ao sul (norte) da sua posição padrão atingindo aproximadamente 3-4°S, favorecendo (inibindo) a ocorrência de extremos de precipitação no NNEB (SOUZA & CAVALCANTI, 2009; GASTINEAU & SODEN, 2011; CAVALCANTI, 2012). No SMA (região 3), pode ser constatado alguma influência de sistemas do norte do Brasil, pois trata-se de uma região limítrofe entre o norte e o NEB. Enquanto no NAL (região 5) e SBA (região 8) são influenciadas pelos DOL (TORRES & FERREIRA, 2011; KOUADIO et al., 2012), este sistema é ocasionado por ondas ou perturbações que se propagam da costa oeste africana em direção a costa leste do NEB.



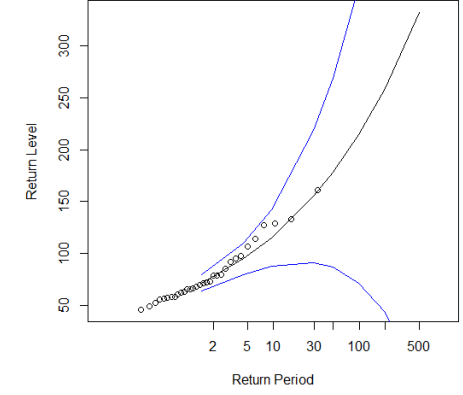
(a) NMA



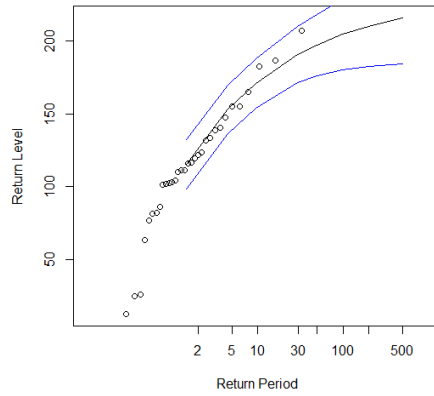
(b) NCE



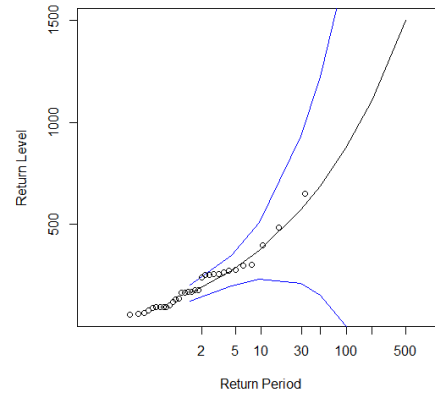
(c) SMA



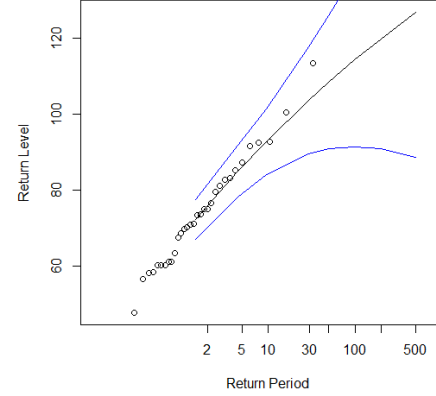
(d) OPE



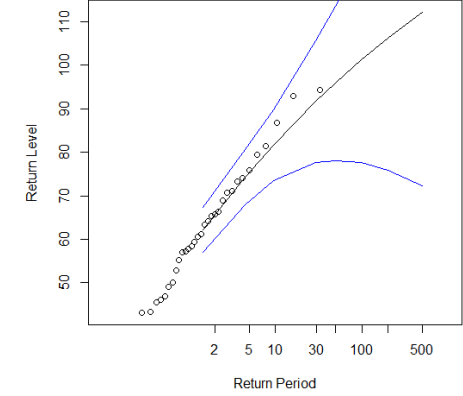
(e) NAL



(f) CBA



(g) OBA



(h) SBA

Figura 2 – Níveis (mm/ dia) e períodos (T em anos) de retorno estimados dos modelos MVLG-GEV gerados para as oito regiões selecionadas.

Este sistema é impulsionado pelo ASAS na região com ventos de sudeste, transportando umidade e calor dos oceanos para a costa, e aliadas aos processos dinâmicos e termodinâmicos da região, neste caso a circulação local (componentes zonal e meridional) aliada a evaporação (Tabela 3) propiciam a formação de SCM e CCM na costa leste do NEB, gerando chuvas intensas (ALVES et al., 2001; KOUADIO et al., 2012).

Os extremos verificados no SMA (região 3) e CBA (região 6) são influenciadas pelo VCAS, este sistema meteorológico provoca um aumento abrupto nos regimes de precipitação nas regiões afetadas, caracterizado pelos movimentos de subsidência na parte central deste sistema, inibindo a convecção no centro do sistema, enquanto suas bordas registram precipitações intensas (KOUSKY & GAN, 1981; CHAVES & CAVALCANTI, 2001; CAVALCANTI, 2012). Os extremos constatados no CBA (região 6), OBA (região 7) e SBA (região 8) sejam motivados pela incursão de sistemas frontais (SF) que se desprendem das ZCAS (KOUSKY, 1979; CHAVES & CAVALCANTI, 2001), isto ocorre devido a intensidade da ZCAS, quando muito intensas as SF se desprendem do eixo principal e se propaga até o extremo sul do NEB, provocando chuvas intensas sobre a região.

A Figura 2 apresenta os períodos e níveis de retorno dos modelos gerados para cada uma das oito regiões do NEB, e verifica-se que a maioria dos níveis de retorno possuem uma boa precisão, constata-se que CBA (região 6) estima o maior nível de retorno, pelo menos um evento extremo de precipitação de nível igual ou superior a 571,2mm/dia em 30 anos, e de 875,7mm/dia em 100 anos. Enquanto isso na região do SBA (região 8) verificou-se o menor nível, e que poderá ocorrer pelo menos um evento extremo de precipitação de nível igual ou superior a 91,6 mm/dia em 30 anos, e de 101,1mm/dia em 100 anos.

4-CONCLUSÕES

As primeiras conclusões sobre a utilização do MVLG via GEV para avaliar os extremos dos precipitação no NEB. Os resultados mostraram uma grande vantagem quanto a resposta à forma múltipla da análise de extremos, devido à inserção de variáveis explicativas que auxiliaram na interpretação da resposta, diferente da forma tradicional que é realizada de maneira univariada.

Pelo parâmetro de localização (μ) constatou-se que o menor valor foi obtido no NAL (região 5) de $17,22 \pm 7,14$ mm, enquanto o maior é verificado no NMA (região 1) com o valor

de $88,26 \pm 6,42$ mm. O parâmetro de escala (σ) verificou que com exceção do OPE (região 4), as demais regiões apresentaram valores positivos, com o máximo de 7,40 no NAL (região 5).

Pelo parâmetro de forma (ξ) verificou-se que com exceção do NAL (região 5) e CBA (região 6) que as séries são ajustadas pela distribuição de Fréchet ($\xi > 0$), as demais foram ajustadas pela distribuição de Weibull negativa ($\xi < 0$).

Os testes de log-verossimilhança negativa e AIC mostraram que o SMA (região 3) e SBA (região 8) são as regiões que possuem os modelos MVLG-GEV mais precisos com valores de 106,95 e 237,56.

As variáveis que subsidiariam a ocorrência dos extremos de precipitação via MVLG-GEV foram as componentes zonal (U), meridional (V) do vento e a evaporação, associadas às ACP do Atlântico (ATL) e Pacífico (PAC). Entretanto no SMA (região 3), NAL (região 5) e CBA (região 6), a evaporação foi a principal motivadora na ocorrência de extremos de precipitação.

Verificou-se que estas variáveis são influenciadas pela variabilidade da circulação atmosférica sobre a região, e conseqüentemente na intensidade dos sistemas atuantes como: ZCIT, no NNEB; ZCAS, SF e VCAS no SNEB, VCAS no SANEB e DOL, SCM e CCM no LNEB.

As estimativas dos níveis e períodos de retorno mostraram que os valores estimados pelo modelo MVLG-GEV para os níveis de retorno para os próximos 30 e 100 anos, e verificou-se que no CBA (região 6) pode registrar o maior nível de retorno, com pelo menos um evento extremo de precipitação de nível igual ou superior a 571,2 mm/dia nos próximos 30 anos, e de 875,7 mm /dia em 100 anos, enquanto o menor nível de retorno é constatado no SBA (região 8) poderá notificar pelo menos um evento extremo de precipitação de nível igual ou superior a 91,6 mm/dia nos próximos 30 anos e de 101,1 mm /dia em 100 anos.

AGRADECIMENTOS

O autor agradece a CAPES pela concessão da Bolsa de Doutorado. P. S. Lucio é bolsista PQ2 do CNPq (Proc: 309165/2010-5).

REFERÊNCIAS

- ALVES, J.M.B.; TEIXEIRA, R.F.B.; FERREIRA, A.G. Um Intenso Sistema Convectivo de Mesoescala no Setor Leste do Nordeste: O Caso de 20 a 21 de maio de 1999. **Revista Brasileira de Meteorologia**, v.16, n.1, p.19-31, 2001.
- BACK, L.E.; BRETHERTON, C.S. The relationship between wind speed and precipitation in the Pacific ITCZ. **Journal of climate**, v. 18, n. 20, 2005.
- BERG, P.; MOSELEY, C.; HAERTER, J.O. Strong increase in convective precipitation in response to higher temperatures, **Nature Geoscience**, v. 6, n. 3, p. 181–185, 2013.
- CARVALHO, M. A. V. de; OYAMA, M. D. Atlantic Intertropical Convergence Zone width and intensity variability: observational aspects. **Revista Brasileira de Meteorologia**, v. 28, n. 3, p. 305-316, 2013.
- CAVALCANTI, I.F.A. Large scale and synoptic features associated with extreme precipitation over South America: A review and case studies for the first decade of the 21st century. **Atmospheric Research** v. 118, p. 27-40, 2012.
- CHAVES, R.R., CAVALCANTI, I.F.A. Atmospheric Circulation features Associated with Rainfall Variability over Southern Northeast Brazil. **Monthly Weather Review**, v.129, p.2614-2626, 2001.
- CHEN, M.; SHI, W.; XIE, P.; SILVA, V.B.S.; Kousky, V.E.; HIGGINS, R.W.; JANOWIAK, J.E. Assessing objective techniques for gauge-based analyses of global daily precipitation. **Journal of Geophysical Research**, v. 113 (D4), D04110, 2008.
- COLES, S.; BAWA, J.; TRENNER, L.; DORAZIO, P. **An introduction to statistical modeling of extreme values** (v. 208). London: Springer, 2001.
- CORREIA-FILHO, W.L.F.; LUCIO, P.S.; SPYRIDES, M.H.C. Precipitation Extremes Analysis over the Brazilian Northeast via Logistic Regression. **Atmos and Clim Sci**, v. 4, n. 1, p. 53-59, 2014. doi: 10.4236/acs.2014.41007.
- DAVISON, A.C.; GHOLAMREZAEI, M.M. Geostatistics of extremes. **Proceedings of the Royal Society A**, v. 468, p.581-608, 2012, doi:10.1098/rpsa2011.0412.
- DEE et al. The ERA-Interim reanalysis: configuration and performance of the data assimilation system, **Quarterly Journal of the Royal Meteorological Society**, 137(I.656): 553–597, 2011.

EVANS, J. L.; BRAUN, A. A Climatology of Subtropical Cyclones in the South Atlantic. **Journal of Climate**, v. 25, n. 21, 2012.

GARAVAGLIA, C. R.; DOYLE, M. E.; BARROS, V. R. Statistical relationship between atmospheric circulation and extreme precipitation in La Plata Basin. **Meteorological Applications**, 2013.

GASTINEAU, G.; SODEN, B. J. Evidence for a weakening of tropical surface wind extremes in response to atmospheric warming. **Geophysical Research Letters**, v. 38, n. 9, 2011.

HASTENRATH, S. Interannual variability and annual cycle: Mechanisms of circulation and climate in the tropical Atlantic sector. **Monthly Weather Review**, v. 112, n.6, p.1097-1107, 1984.

HASTENRATH, S. Exploring the climate problems of Brazil's Nordeste: a review. **Climatic Change**, v. 112, p. 243–251, 2011, doi:10.1007/s10584-011-0227-1.

JOHANSSON, B.; CHEN, D.. The influence of wind and topography on precipitation distribution in Sweden: Statistical analysis and modelling. **International Journal of Climatology**, v. 23, n. 12, p. 1523-1535, 2003.

JONES, C.; WALISER, D.E., LAU, K.M., STERN, W. Global occurrences of extreme precipitation and the Madden-Julian Oscillation: Observations and predictability. *Journal of Climate*, v.17, n. 23, p. 4575-4589, 2004.

JONES, M. R. et al. Objective classification of extreme rainfall regions for the UK and updated estimates of trends in regional extreme rainfall. **International Journal of Climatology**, v. 34, n. 3, p. 751-765, 2014.

JØRGENSEN, B.; GOERGEBEUR, Y., MARTÍNEZ, J.R. Dispersion models for extremes. **Extremes**, v.13, n.4, p. 399-437, 2010.

JOYCE, R.J.; JANOWIAK, J.E.; ARKIN, P.A.; XIE, P. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. **Journal of Hydrometeorology**, v.5, n.3, p.487–503, 2004.

KAYANO, M.T.; ANDREOLI, R.V.; SOUZA, R. A. F.. Relations between ENSO and the South Atlantic SST modes and their effects on the South American rainfall. *International Journal of Climatology*, v.33, p.2008-2023, 2013.

- KOUADIO, Y.K. et al. Heavy Rainfall Episodes in the Eastern Northeast Brazil Linked to Large-Scale Ocean-Atmosphere Conditions in the Tropical Atlantic. **Advances in Meteorology**, 2012.
- KOUSKY, V.E.; GAN, M.A. Upper tropospheric cyclonic vortices in the tropical South Atlantic. **Tellus**, v.33, n.6, p.538-551, 1981.
- KUCHARSKI, F.; POLZIN, D.; HASTENRATH, S. Teleconnection Mechanisms of Northeast Brazil Droughts: Modeling and Empirical Evidence. **Revista Brasileira de Meteorologia**, v. 23, n.2, p.115-125, 2008.
- LAU, K. M.; WU; H. T.; BONY, S. The role of large-scale atmospheric circulation in the relationship between tropical convection and sea surface temperature. **Journal of Climate**, v. 10, p. 381-392, 1997.
- LIEBMANN, B.; SMITH, C. Description of a Complete (Interpolated) Outgoing Longwave Radiation Dataset. **Bulletin of the American Meteorological Society**, v. 77, p.1275-1277, 1996.
- LIU, S.C.; FU, C.; SHIN, C.-J.; CHEN, J.-P.; WU, F. Temperature dependence of global precipitation extremes. **Geophys Research Letters**, v.36, n. 17, L17702, 2009.
- MADDEN, R. A.; JULIAN, P. R. Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific. **Journal of the Atmospheric Sciences**, v. 28, n. 5, p. 702-708, 1971.
- MADDOX, R. A. An objective technique for separating macroscale and mesoscale features in meteorological data. **Monthly Weather Review**, v. 108, n. 8, p. 1108-1121, 1980.
- MARAUN, D.; OSBORN, T. J.; RUST, H. W. The influence of synoptic airflow on UK daily precipitation extremes. Part I: Observed spatio-temporal relationships. **Climate Dynamics**, v. 36, n. 1-2, p. 261-275, 2011.
- MOLION, L. C. B.; BERNARDO, S. de O. Uma revisão da dinâmica das chuvas no Nordeste Brasileiro. **Revista Brasileira de Meteorologia**, v. 17, n. 1, p. 1–10, 2002.
- NELDER, A.J.A.; WEDDERBURN, R.W.M. Generalized Linear Models. **Journal of the Royal Statistical Society Series A (General)**, v.135, n. 3, p. 370–384, 1972.
- O’GORMAN, P.A.; SCHNEIDER, T. The physical basis for increases in precipitation extremes in simulations of 21st-century climate change, **Proceedings of the National Academy of Sciences**, v. 106, p. 14773-14777, 2009.

PATHIRANA, A. et al. Impact of urban growth-driven landuse change on microclimate and extreme precipitation—A sensitivity study. **Atmospheric Research**, v. 138, p. 59-72, 2014.

R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. 2011, Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

SANDEEP, S.; STORDAL, F. Use of daily outgoing longwave radiation (OLR) data in detecting precipitation extremes in the tropics. **Remote Sensing Letters**, v.4, n.6, p.570–578, 2013.

SILVA, V.B.S.; KOUSKY, V.E. The South American Monsoon System: Climatology and Variability. **Modern Climatology**. Dr Shih-Yu Wang (Ed.), ISBN: 978-953-51-0095-9, 2012, InTech, doi: 10.5772/38565.

SMITH, M. D. The ecological role of climate extremes: current understanding and future prospects. *Journal of Ecology*, v.99, p. 651–655, 2011.

SOUZA, P.; CAVALCANTI, I.F.A. Atmospheric centres of action associated with the Atlantic ITCZ position. **International Journal of Climatology**, v.29, n.14, p.2091-2105, 2009.

TORRES, R. R.; FERREIRA, N.J. Case studies of easterly wave disturbances over Northeast Brazil using the Eta Model. **Weather and Forecasting**, v.26, p.225-235, 2011.

TRENBERTH, K. E. The definition of El Niño, **Bull. Amer. Met. Soc.**, v.78, p. 2771-2777, 1997.

VENABLES, W.N.; RIPLEY, B.D. **Modern applied statistics with S**. Springer, New York, 1994.

YEE T.W. **VGAM: Vector Generalized Linear and Additive Models**, R package version 0.9, 2013.

YEE, T.W.; STEPHENSON, A.G. Vector generalized linear and additive extreme value models. **Extremes**, v.10, p.1–19, 2007.

YEE, T.W.; WILD, C.J. Vector generalized additive models. **Journal of the Royal Statistical Society, Series B, Methodological**, v.58, p.481–493, 1996.

CONCLUSÕES

Esta tese abordou, em quatro capítulos, os modelos de dispersão para diagnóstico e detecção de extremos de precipitação no Nordeste do Brasil. Os dois primeiros capítulos avaliaram a magnitude das associações entre os extremos de precipitação e as variáveis meteorológicas pela razão de chance (ou *odds ratio*) via regressão logística, enquanto o segundo artigo avaliou a magnitude das associações pelo risco relativo (ou *relative risk*) via regressão de Poisson. Os resultados apresentados pelos modelos de regressão logística e de Poisson foram similares, sendo que as medidas de associações são distintas. Deste modo, os resultados apresentados a seguir das duas análises de forma simultânea, ao qual os valores entre parênteses correspondem ao risco relativo. A razão de chances (risco relativo) indicou que a radiação de onda longa é a principal variável para a detecção da ocorrência de eventos extremos de precipitação, constatada em três sub-regiões do NEB, no norte, no semiárido e leste do NEB. Verificou-se que o decaimento da radiação de onda longa (ROL), ou seja, valores de ROL inferiores a 240 W/m^2 , as chances de ocorrer um extremo de precipitação no leste do NEB aumentam aproximadamente em 3 vezes (2,5 vezes), no semiárido do NEB aumentam em 3,63 vezes (3,67 vezes), enquanto no norte do NEB aumentam em 5,4 vezes (5,23 vezes).

A radiação de onda longa está relacionada ao processo de formação e desenvolvimento de nuvens sobre estas regiões do NEB: no norte, estes processos favorecem a formação e desenvolvimento de nuvens convectivas provocadas pela ZCIT, principalmente nos extremos; no sul, este processo de desenvolvimento convectivo das nuvens é originado pelas linhas de instabilidade e sistemas frontais, que se desprendem das ZCAS e se propagam no sul da Bahia; no semiárido, é originado tanto pelos VCAS e quando combinados com a ocorrência de Oscilação Madden-Julian potencializam a ocorrência dos extremos de precipitação.

No leste do NEB, a variável indicadora para a ocorrência dos extremos de precipitação é a umidade relativa, influenciadas principalmente pelos DOL, que quando combinados às condições locais podem ocasionar a formação de sistemas e complexos convectivos de mesoescala, resultando na ocorrência de extremos que contribuem com 70% do regime diário e 60% do período mensal.

As ACP das TSM do Atlântico e do Pacífico tropicais não subsidiam a ocorrência dos extremos quando avaliados pela regressão logística, entretanto a regressão de Poisson indicou que a 2ª ACP da TSM do Pacífico tropical favoreceu a ocorrência de extremos de precipitação, com o risco de 1,21 vezes e 1,36 vezes para o norte e semiárido do NEB,

respectivamente. Esta 2ª ACP da TSM do Pacífico tropical está relacionada à variabilidade sazonal. A regressão de Poisson apresentou melhor capacidade preditiva com o menor número de variáveis.

O terceiro capítulo avaliou o comportamento de maneira univariada dos extremos anuais de precipitação no NEB utilizando a TVE via distribuições GEV e GPD (baseado no percentil 95) e avaliou os níveis e períodos de retorno de 30,50 e 100 anos.

Os resultados apontaram que os parâmetros de localização e de escala para ambas as distribuições possuem configurações espaciais similares, registrando que os máximos foram verificados na costa leste do NEB (DOL), no norte do Maranhão e do Ceará (ZCIT), na distribuição GEV verificou uma outra região de máximos localizada a oeste de Pernambuco, possivelmente favorecida pela influência topográfica do planalto da Borborema e da chapada do Araripe. Pelos resultados do parâmetro de forma é possível verificar que para a maioria das regiões selecionadas os dados foram ajustado pelas distribuições de Weibull negativa e beta ($\xi < 0$) para GEV e GPD, respectivamente. Quanto aos valores dos parâmetros da distribuição GEV e GPD obtidos para as oito regiões selecionadas foram: de localização, variaram entre 22,0 a 77,6mm para GEV e 6,5 a 42,0mm para GPD; de escala, variaram entre 7,8-31,2mm para GEV e 6,2 a 17,9mm para GPD; e de forma, variaram entre -0,5 a 0,3 para GEV e 0,1 a 0,4 para GPD.

A partir das estimativas dos níveis e períodos de retorno, concluiu-se que no norte do Maranhão registrou os maiores níveis de retorno, pelas estimativas é possível que pelo menos um evento extremo de precipitação de nível igual ou superior a 160,9 mm/dia ocorra nos próximos 30 anos, e um de 182,1mm/dia nos próximos 100 anos, já pela GPD, na região central da Bahia verificou o maior nível de retorno, pelas estimativas poderá pelo menos um evento extremo de precipitação de nível igual ou superior a 192,3 mm/dia nos próximos 30 anos, e um de 307,8mm /dia para os próximos 100 anos.

Outro detalhe bem interessante verificado neste capítulo se refere a diferença significativa nos resultados obtidos entre as distribuições, tanto na intensidade dos parâmetros de localização, de escala e de forma dos extremos, como também em sua distribuição espacial, e verificou-se que o método PORT de percentil 95 da distribuição GPD, apesar da flexibilização do limiar foram bem inferiores quando comparado a distribuição GEV.

O quarto capítulo avaliou o comportamento de forma múltipla dos extremos anuais de precipitação no NEB a partir dos modelos vetoriais lineares generalizados utilizando a

distribuição GEV (MVLG-GEV) como função de ligação, sendo este o diferencial da tese ao qual analisa estes extremos de maneira multivariada.

Os resultados apresentados pelo MVLG-GEV indicaram que os maiores valores dos parâmetros de localização foram verificados na parte norte do NEB, reafirmando o que foi constatado no capítulo III que analisou os extremos de precipitação via GEV e GPD de forma univariada, os maiores valores foram verificados no norte do Maranhão e Ceará com $88,26 \pm 6,42$ mm e de $64,34 \pm 4,30$ mm, respectivamente. Esta influência pode estar relacionada à variabilidade da TSM tanto da região do Atlântico como do Pacífico tropicais, resultantes da modificação da ZCIT, que por sua vez altera o padrão de circulação atmosférica, verificado pelas componentes zonal e meridional do vento em superfície. Quando a variabilidade da ZCIT for positiva (negativa) favorece (inibe) a ocorrência de extremos de precipitação no norte do NEB.

Os resultados do parâmetro de escala (σ) constataram valores positivo para a maioria das regiões exibiram valores positivos, com o máximo de 7,40 no norte de Alagoas. Os resultados dos parâmetros de forma mostraram que a maioria das séries os dados são ajustados pela distribuição de Weibull negativa ($\xi < 0$), entretanto para no leste de Alagoas e no centro da Bahia as séries de dados são ajustados pela distribuição de Fréchet ($\xi > 0$). Os testes de log-verossimilhança negativa e AIC mostraram que os modelos do sul do Maranhão e da Bahia possuem maior precisão com valores de 106,95 e 237,56, respectivamente. Constatou-se que as componentes zonal (U), meridional (V) do vento e a evaporação (sul do Maranhão, nordeste de Alagoas e no centro da Bahia), associadas às ACP do Atlântico (ATL) e Pacífico (PAC) subsidiam a ocorrência dos extremos de precipitação no NEB, a influência da variabilidade da TSM do Atlântico como do Pacífico, mostra indícios de uma possível atuação do ENOS. Estas variáveis estão influenciadas pela circulação atmosférica sobre a região, e conseqüentemente na intensidade dos sistemas atuantes como: ZCIT, no NNEB; ZCAS, SF e VCAS no SNEB, VCAS no SANEB e DOL, SCM e CCM no LNEB.

A partir dos níveis e períodos de retorno mostraram que os valores estimados pelo modelo MVLG-GEV, verificou-se que no centro da Bahia estimou o maior nível de retorno, as estimativas mostraram que ocorrer pelo menos um evento extremo de precipitação de nível igual ou superior a 571,2 mm/dia nos próximos 30 anos, e de 875,7mm /dia em 100 anos.

Esta tese trouxe contribuições importantes referentes a contribuição das covariáveis e o comportamento de cada uma delas na ocorrência dos extremos de precipitação no NEB,

bem como as estimativas dos modelos que contribuíram em quantificar e qualificar a magnitude por medidas de associações (razão de chances e risco relativo) entre as variáveis e da ocorrência destes extremos. Além disso, a modelagem MVLG via distribuição GEV utilizando covariáveis permitiu caracterizar e diagnosticar a influência de cada uma delas na ocorrência destes extremos. Para um maior aprofundamento na caracterização dos extremos do NEB sob o ponto de vista dos modelos aplicados nesta tese, segue abaixo algumas das perspectivas futuras:

- Melhorar a resolução espacial de $1,5^\circ \times 1,5^\circ$ para $0,5 \times 0,5^\circ$;
- Realizar a distribuição espacial da razão de chances e risco relativo no NEB, exibindo a influência de uma variável para o extremo de precipitação e em escala sazonal.
- Inserir variáveis, tais como: velocidade vertical (variável que analisa os movimentos de subsidência / ascendência na atmosfera), umidade específica em diversos níveis, componentes zonal e meridional do vento em diversos níveis para avaliar a interação destas novas variáveis com a ocorrência dos extremos de precipitação pelo modelo MVLG via distribuições GEV e GPD-Poisson;
- Avaliar os extremos de precipitação pelo modelo MVLG via distribuição GPD-Poisson como função de ligação, estabelecendo limiares distintos com o objetivo de verificar a relação das variáveis explicativas e os extremos de precipitação;
- Analisar os extremos de precipitação pelos Modelos Vetoriais Aditivos Generalizados (MVAG) via distribuição GEV e GPD-Poisson como função de ligação, para avaliar o comportamento aditivo entre as variáveis explicativas e os extremos de precipitação.

APÊNDICE I – METODOLOGIA

Neste apêndice estão descritos os procedimentos utilizados para o diagnóstico e avaliação de extremos climáticos no Nordeste do Brasil, de maneira sucinta e objetiva. Os métodos estão divididos em: 1) análise de componentes principais (ACP); 2) a análise de cluster (AA); 3) teoria clássica dos valores extremos (TVE); e por fim, 4) modelos de dispersão baseados em Modelos Lineares Generalizados (MLG) e Modelos Vetoriais Lineares Generalizados (MVLG).

Análise de Componentes Principais (ACP)

Esta técnica de análise multivariada já é bem difundida em ciências climáticas, sua aplicação tem como objetivo captar a variância total de um conjunto de dados, a partir da redução do número de variáveis, permitindo extrair padrões e processos associados às variáveis observadas (WILKS, 2005; CORRAR et al., 2007). A matriz F ($p \times n$) é constituída de p pontos de grade de precipitação e N número de observações deste respectivo ponto.

$$F = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ f_{p1} & f_{p2} & \cdots & f_{pn} \end{bmatrix} \quad (1)$$

Cada linha da matriz F representa um ponto de grade p e as colunas representam as observações no ponto de grade p . A matriz de correlação R é obtida a partir da matriz F das variáveis padronizadas pelos desvios padrões, dada por:

$$R = \frac{1}{N-1} F^T F \quad (2)$$

Sendo R uma matriz simétrica positiva de dimensão ($p \times p$) e diagonalizável por uma matriz E , denominada de autovetores. A matriz diagonal D , cujos elementos diagonais são os autovalores de R é obtida por:

$$D = E^{-1}RE \quad (3)$$

Devido à ortogonalidade dos autovalores, a matriz inversa de $E = (E^{-1})$ é igual a sua transposta (E^T). Assim as componentes principais (CP) U_1, U_2, \dots, U_b são obtidas pela combinação linear entre a matriz dos autovetores de (E^T) e a matriz de observações F , ou seja,

$$U = E^T F \quad \text{ou} \quad F = EU \quad (4)$$

Cada linha de U representa uma CP que são séries temporais associadas aos vetores. Cada CP U_1, U_2, \dots, U_p tem uma porção da variância total dos dados em ordem decrescente dos autovetores mais significativos e_p em E . Logo,

$$U = \sum_{i=1}^p e_{ip} F_p \quad (5)$$

Em que F_p é a matriz original normalizada e e_p é o autovetor. Os valores de F do n -ésimo local podem ser estimados pela seguinte expressão:

$$F_p = e_{j_1} U_1 + e_{j_2} U_2 + e_{j_3} U_3 + \dots + a_{nb} U_p \quad \text{ou} \quad F_p = \sum_{i=1}^p e_{ip} U_p \quad (6)$$

As ACP foram aplicadas aos dados diários de temperatura à superfície do Mar (TSM) nas regiões tropicais do Atlântico e Pacífico, com o propósito de captar a variabilidade espacial destas regiões (GIANNINI et al., 2000; L'HEUREUX et al., 2013) e verificar as possíveis associações com a ocorrência de extremos de precipitação.

Análise de Agrupamento (AA)

Esta técnica permite realizar uma estrutura classificatória a partir de um conjunto de indivíduos n medidos em p variáveis, ao qual possibilitem reunir indivíduos num certo número de grupos, tal que exista homogeneidade dentro dos grupos e heterogeneidade entre os grupos. A extração dos grupos é obtida por dois passos: 1) definir uma medida de distância d a fim de estabelecer similaridades entre os pares de elementos; e 2) definir o método de ligação (BURLANDO et al., 2008). A medida de distância utilizada foi a distância Euclidiana expressa dessa maneira:

$$d_{ab} = \left[\frac{1}{N_{ab}} \sum_{j=1}^{N_{ab}} (X_{aj} - X_{bj})^2 \right]^{1/2} \quad (7)$$

Em que X_{aj} é a j -ésima característica da a -ésima observação, X_{bj} é a j -ésima característica da b -ésima observação e N_{ab} representa o número de pontos de grade de precipitação que abrangem o Nordeste do Brasil, composto por setenta e sete pontos.

O método de ligação utilizado para AA é o Ward, com o objetivo de distinguir regiões homogêneas de precipitação a partir de médias climatológicas diárias no Nordeste do Brasil utilizando setenta e sete pontos, do período de 1981 a 2011 (KALKSTEIN et al., 1987; UNAL et al., 2003; SARIS et al., 2010). Esta técnica foi desenvolvida por Ward (1963) e descrita por Kalkstein (1987), e consiste na soma dos desvios quadráticos em relação à média do grupo para identificar similaridade, expressa por:

$$W = \sum_{k=1}^K \sum_{j=1}^J \sum_{i=1}^{N_k} (X_{ijk} - X_{\bullet,jk})^2 \quad (8)$$

Em que X_{ijk} é o valor da j -ésima variável para a i -ésima observações N_k de K grupos e $X_{\bullet,jk}$ é o valor médio.

Teoria dos Valores Extremos (TVE)

Neste t3pico s3o discutidos os fundamentos da Teoria de Valores Extremos (TVE), encontrados com maiores detalhes em Coles et al. (2001) e Embrechts et al. (2004). A teoria dos valores extremos 3 feita pela extra33o dos valores m3ximos de uma s3rie de dados, $M_n = \max\{X_1, \dots, X_n\}$, em que X_1, \dots, X_n 3 uma sequ3ncia de vari3veis aleat3rias independentes que possuem uma fun33o de distribu33o F , em que X representa o valor da ocorr3ncia do extremo, M_n corresponde ao m3ximo de extremos ocorridos em um per3odo de observa33es n (este per3odo pode ser di3rio, mensal, anual, etc.). A distribu33o M_n pode ser derivada para os valores de n :

$$\Pr\{M_n \leq z\} = \Pr\{X_1, \dots, X_n\} \Rightarrow \Pr\{M_n \leq z\} = \{F(z)\}^n \quad (9)$$

Em que n da distribu33o F^n tende ao infinito, para qualquer $z < z_+$, sendo z_+ o ponto superior de F^1 . Quando $F^n(z) \rightarrow 0$, a distribu33o de M_n degenera ao ponto de massa em z_+ . Este processo permite uma renormaliza33o linear da vari3vel M_n :

$$M_n^* = \frac{M_n - b_n}{a_n},$$

Com o aumento de n , quando $a_n > 0$ e b_n permanecem constantes, tem-se a seguinte express3o:

$$\Pr\{M_n - b_n/a_n \leq z\} \rightarrow G(z) \text{ como } n \rightarrow \infty$$

Sendo G , uma fun33o de distribu33o n3o-degenerada pertencente a uma das fam3lias que seguem abaixo:

$$I: G(z) = \exp \left\{ -\exp \left[-\left(\frac{z-\beta}{\alpha} \right) \right] \right\}, \quad -\infty < z < \infty;$$

$$II: G(z) = \begin{cases} 0, & z \leq \beta, \\ \exp \left\{ -\left(\frac{z-\beta}{\alpha} \right)^{-\alpha} \right\}, & z > \beta, \end{cases}$$

$$III: G(z) = \begin{cases} \exp \left\{ -\left(\frac{z-\beta}{\alpha} \right)^{-\alpha} \right\}, & z < \beta, \\ 1 & z \geq \beta, \end{cases}$$

A modelagem de extremos via distribuição GEV inicia-se a partir de uma série de observações independentes, X_1, \dots, X_n , estes dados são limitados a uma sequência ou janela de observações de comprimento n , para cada valor máximo de m gera-se uma série máxima por bloco ou período, $X_{p,1}, \dots, X_{m,n}$. Uma vez determinado o comprimento destes blocos ou período, por exemplo, o bloco máximo anual, n será o número de observações no ano para a obtenção dos máximos anuais por blocos. A escolha do tamanho do bloco ou período é essencial para a estimação dos parâmetros da distribuição GEV e também para o nível de retorno, pois se for muito pequeno poderá interferir na estimação do viés e extrapolação, e se muito grande gerariam poucos blocos de máximos, resultando no aumento da variância.

De forma geral, estas distribuições de classe são chamadas de distribuição de valores extremos dos tipos I, II e III conhecidas como famílias: Gumbel, Fréchet e Weibull negativa, respectivamente. As famílias Gumbel ($\alpha=0$), Fréchet ($\alpha>0$) e Weibull negativa ($\alpha<0$) possuem parâmetros de localização (μ), escala (β) e forma (α) distintas, com aplicações diferenciadas para cada família, exibindo diferentes comportamentos de valor extremo.

A forma simplificada desta distribuição proposto por Jenkinson (1955):

$$G(z) = \exp \left\{ -\left[1 + \xi \left(\frac{z-\mu}{\sigma} \right) \right]^{-1/\xi} \right\} \quad (10)$$

Pelo ajuste $\{z : 1 + \xi(z - \mu/\sigma) > 0\}$, os parâmetros satisfazem $-\infty < \mu < 0$ e $-\infty < \xi < \infty$, com os parâmetros de localização (μ), de escala (σ) e de forma (ξ).

A estimação do período e nível de retorno dos extremos a partir da distribuição GEV é obtida pela inversão da Equação 10,

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi} \left\{ 1 - \left[-\log(1-p) \right]^{-\xi} \right\}, \\ \mu - \sigma \log \left[-\log(1-p) \right], \end{cases} \quad (11)$$

Em que $G(z_p) = 1 - p$. z_p é o nível de retorno em relação ao período de retorno $1/p$, p é a probabilidade de ocorrência do extremo, caso os máximos por blocos seja anual, z_p será dado em anos.

Enquanto a modelagem de extremos via distribuição GPD é baseada a partir de um limiar previamente estabelecido ao conjunto de dados, sem a necessidade de formar blocos de máximos como realizado na GEV. Este critério é aplicado a toda série de dados, desde que X_1, \dots, X_n seja uma sequência de variáveis aleatórias independentes e identicamente distribuídas (iid) e que contenha uma função de distribuição marginal F , considerando eventos extremos todo aquele valor de X_i que exceda o maior limiar u . Dado um termo X_i de uma série X , com comportamento aleatório, os limiares serão obtidos pela seguinte expressão:

$$\Pr\{X > u + y \mid X > u\} = \frac{F(u + y) - F(u)}{1 - F(u)}, \quad y > 0 \quad (12)$$

Em que F é uma distribuição conhecida. Sendo X_1, \dots, X_n uma série de valores aleatórios e independentes com função de distribuição comum F , teremos algo semelhante ao $M_n = \max\{X_1, \dots, X_n\}$. Dado um termo arbitrário X_i em uma série X de n grande, teremos $\Pr = \{M_n \leq z\} \approx G(z)$, em que:

$$G(z) = \exp \left\{ - \left[1 + \xi \cdot \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}$$

Os parâmetros de localização (μ), de escala (σ) e de forma (ξ). Então, o limiar suficiente é u , a função de distribuição de $(X - u)$ é condicional $X > u$, então teremos aproximadamente:

$$H(y) = 1 - \left(1 + \frac{\xi \cdot y}{\sigma}\right)^{-1/\xi} \quad (13)$$

Definida em $\{y : y > 0 \cdot e \cdot (1 + \xi \cdot y / \sigma) > 0\}$,

$$\text{Em que } \bar{\sigma} = \sigma + \xi(u - \mu) \quad (14)$$

A distribuição GPD pode ser feita com maior precisão, justificando (13) como uma distribuição limite que u aumenta, a família de distribuições (13) é chamada de família generalizada de Pareto. Assim como na distribuição GEV, as distribuições de classe são chamadas de distribuição de valores extremos dos tipos I, II e III conhecidas como famílias só que com outras famílias exponenciais ($\sigma=0$), Pareto ($\sigma>0$) e Beta ($\sigma<0$), respectivamente.

Dado um conjunto de dados que consiste em uma sequência de medidas independentes e identicamente distribuídas (iid) X_1, \dots, X_n , os máximos são verificados a partir do limiar u previamente escolhidos $\{X_i : X_i > u\}$, em seguida identificados estes excedentes pela $x_{(1)}, \dots, x_{(k)}$, e definida pelos limiares que excederam pela $y_j = x_{(j)} - u$, para $j = 1, \dots, k$.

A escolha do limiar é semelhante ao do tamanho do bloco em GEV, implicando em um balanço entre viés e variância. Dois métodos são aplicados nesta fase: 1) análise exploratória verificando a estimação; 2) avaliação de estabilidade dos parâmetros estimados, baseados no ajuste dos modelos cruzando uma amplitude de diferentes limiares.

A estimação do nível de retorno em N anos a partir da distribuição GPD é obtida pela expressão (COLES et al., 2001),

$$x_N = u(t) + \frac{\sigma(t)}{\xi} [(Nn \zeta_u \theta)^\xi - 1], \quad (15)$$

Em que N é o período de retorno (anos), n é o número de observações de conjunto, ζ_u é a probabilidade de um excedente em μ e θ é o índice extremal. Para ambos os termos na Equação 15 o tempo t varia entre zero e o número de anos.

Outro procedimento que auxilia a escolha dos limiares é estimar o modelo a um u fixo, conhecido como picos fixos acima dos limiares (*Peaks Over Fixed Threshold - POFT*), ou então, pelos picos aleatórios acima dos limiares (*Peaks Over Random Threshold - PORT*), que são limiares flexíveis, geralmente utiliza os percentis. De acordo com Folland et al. (1999) é preferível a utilização dos percentis a valores dos limiares absolutos, o PORT, devido a heterogeneidade das áreas, e também processos físicos que podem mascarar o resultado final. A qualidade de ajuste do modelo gerada tanto pela GEV como pela GPD é avaliada pelo diagnóstico de diversos gráficos: probabilidade, quantis, densidade e de níveis de retorno.

Modelos de dispersão

Modelos Lineares Generalizados (MLG)

A representação de resultados a partir de modelos é uma forma de descrever os resultados e as relações existentes entre variáveis. Um modelo bem ajustado apresenta muitas vantagens:

- A forma estrutural do modelo descreve padrões de associações e interações;
- As inferências sobre os parâmetros do modelo ajudam a avaliar quais variáveis explicativas afetam a resposta, enquanto controlam efeitos de variáveis;
- O valor dos parâmetros estimados determina a magnitude e a importância dos efeitos;
- Os modelos permitem analisar o efeito simultâneo de várias variáveis explicativas;
- Os modelos descrevem os efeitos, sendo mais informativo do que um mero teste de associação.

Os modelos lineares generalizados (MLG) é uma extensão de modelos lineares clássicos que foi desenvolvida por Nelder e Wedderburn (1972), com a principal característica de poder tratar da mesma forma uma grande quantidade de modelos conhecidos e largamente aplicados. Assim são casos especiais de MLG:

- Modelos clássicos de Regressão com erro normal;
- Modelos clássicos de análise de variância e de covariância com erro normal;

- Modelo de análise de variância com erros aleatórios;
- Modelo log-linear aplicado à análise de tabelas de contingência;
- Modelos logit e probit para análise de proporções, dentre outros.

Um MLG é definido a partir de duas componentes mais uma função relacionando estas componentes:

1) **Componente aleatória** - Definida a partir das variáveis respostas X_1, \dots, X_n , supostamente independentes, cada uma com densidade na forma da família exponencial dada por:

$$f(y; \theta_i, \phi) = \exp[\varphi\{y\theta_i - b(\theta_i)\} + (y, \phi)] \quad (16)$$

Em que $b(\cdot)$ e $c(\cdot)$ são funções supostas conhecidas, θ é o parâmetro natural que caracteriza a densidade e ϕ^{-1} é o parâmetro de dispersão. A esperança e a variância da variável aleatória Y são dadas por:

$$E(Y_i) = \mu_i = b'(\theta_i), \text{ e } Var(Y_i) = \phi^{-1} V_i, \quad (17)$$

$V = d\mu/d\theta$ é a função de variância. Esta família exponencial representa a única fonte de variação aleatória do modelo linear generalizado.

2) **Componente sistemática** - É representada como uma função linear de um conjunto de parâmetros desconhecidos. Dessa forma temos $x_i = (x_{i1}, \dots, x_{ip})'$, que representa os valores de p variáveis explicativas, $\beta = (\beta_1, \dots, \beta_p)'$, $p < n$ é um vetor de parâmetros conhecidos a serem estimados, $X = (x'_1, x'_2, \dots, x'_n)'$ é a matriz do modelo e $\eta_i = \sum_{j=1}^p x_{ij}\beta_j = x'_i\beta$ ou $\eta = X\beta$ é o vetor de preditores lineares que corresponde a componente sistemática, e descreve a parte determinística do modelo. Em geral, a relação entre as componentes sistemática e aleatória não possuem relação aditiva.

3) **Função de ligação** - É uma função g , contínua e diferenciável, que relaciona o preditor linear (η) e o valor esperado (μ) de Y . Ou seja:

$$g(\mu_i) = \eta_i \quad (18)$$

No modelo linear clássico tem-se $\eta = \mu$ que é chamada ligação identidade. Essa ligação é adequada no sentido em que ambos η e μ podem assumir valores na reta real. A solução produzida pelos valores dos coeficientes selecionados minimiza a soma quadrática dos resíduos, sendo imparcial, atribuindo a menor variância total entre as alternativas lineares imparciais. Três características são primordiais para o MLG: 1) A relação entre cada variável explicativa e a variável resposta seja aproximadamente linear em sua estrutura; 2) Os resíduos sejam independentes com média zero e variância constante; 3) A regressão e resíduo não sejam correlacionadas.

A qualidade do ajuste dos MLG geralmente se faz a partir da análise de resíduos, com o propósito de avaliar os seguintes critérios: 1) utilização da estatística da bondade de ajuste (*goodness-of-fit*); 2) avaliação da discrepância (*leverage*); e 3) influência para pontos de dados individuais. A aplicação da análise da bondade de ajuste foi realizada pelos seguintes métodos: Estatística χ^2 de *Pearson*, *Deviance*, *Akaike Information Criterion* (AIC) e o gráfico da curva *Receiver Operating Characteristics* (ROC). A estatística de Pearson χ^2 generalizada é um método que se baseia na soma dos resíduos quadráticos de Pearson, descrita abaixo:

$$\chi^2 = \sum_{i=1}^n R_{Pearson} \Rightarrow \sum_{i=1}^n \frac{(Y - \mu)^2}{\mu} \quad (19)$$

Quando o tamanho da amostra for muito grande, então $\chi^2 \sim \chi_{n-k}^2$, em que n e k são os números da amostra e de variáveis explicativas, respectivamente. O propósito deste teste é avaliar o grau de associação entre a variável resposta com as variáveis explicativas.

A *Deviance* é definida como o dobro da diferença entre as razões de máxima log-verossimilhança entre os modelos proposto ($\tilde{\theta}_p$) *versus* saturado ou completo ($\hat{\theta}_s$) (NELDER & WEDDERBURN, 1972). Este método verifica o comportamento entre o modelo proposto comparado ao ideal, descrito como:

$$Deviance = -2(\log(p(y | \tilde{\theta}_p)) - \log(p(y | \hat{\theta}_s))) \quad (20)$$

A *Deviance Residual* é outra medida utilizada para a qualidade do ajuste, baseada na comparação entre *Deviances* dos modelos proposto (M_1) versus modelo completo ou saturado (M_2), e os parâmetros de M_1 e M_2 sejam $p < n$ e $q < p$ graus de liberdade, respectivamente, então teremos:

$$D_{Res} = \frac{D(M_1) - D(M_2)}{(p - q)} \quad (21)$$

O teste AIC consiste em pré-selecionar de um conjunto de variáveis explicativas, aquelas que melhor se ajustam ao modelo final (AKAIKE, 1974; HOSMER & LEMESHOW, 1991). Para a composição final de cada modelo, analisou-se cada uma das variáveis explicativas de maneira univariada, verificando o grau de associação entre a variável explicativa e resposta. Este grau de associação é constatado pelo nível de significância estatística do valor p que deve ser inferior a 0,05, indicando uma forte associação entre as variáveis, e assim, logo é inserida no modelo completo ou global. A aplicação do teste AIC reconstrói um novo modelo proposto a partir do modelo global, onde as variáveis explicativas que realmente interferem na variável resposta permanecem no modelo, enquanto aquelas que possuem interação ou que possuam multicolinearidade entre variáveis são excluídas. O AIC é escrito pela seguinte expressão:

$$AIC = -2L + 2k \quad (22)$$

Em que L é o máximo da log-verossimilhança e k é o número de variáveis explicativas no modelo (incluindo a constante). Quanto menor o valor de AIC significa que o modelo possui um bom ajuste.

Outro teste de qualidade é o gráfico da *receiver operating characteristics* (ROC), também conhecida como estatística C, é uma técnica bastante utilizada em modelos de regressões logística e de Poisson (quando este adota resposta dicotômica). A curva ROC é uma técnica de visualização, organização e seleção, que permite avaliar o desempenho do

modelo (PARK et al., 2004; FAWCETT, 2006). Quando obtemos duas classes de previsão (classificação binária), em que os resultados são marcados como positivo ou negativo. Há quatro possíveis resultados neste classificador binário (Tabela 1), se a previsão for positiva e a observação também é positiva, contabiliza-se como um evento positivo verdadeiro (PV), em caso contrário, a observação é negativa contabiliza-se como positivo falso (FP). Por outro lado, se a previsão é negativa e a observação é positiva contabiliza-se como negativo falso (NF), e se a previsão é negativa e a observação é negativa contabiliza-se como negativo verdadeiro (NV).

Tabela 1 – Tomada de decisão. Sensibilidade (TP/D+) e especificidade (TN/D-) (PARK et al., 2004).

Resposta observada	Resposta predita pelo modelo		
	Positivo	Negativo	Total
Positivo	PV	PF	T+
Negativo	NF	NV	T-
Total	D+	D-	N

VP – Verdadeiro positivo, FP – Falso positivo, VN – Verdadeiro negativo, FN – Falso negativo.

Na Tabela 1, a sensibilidade = $PV/PV + PF$ e especificidade = $NV/NV + NF$ pode ser calculada. Quanto maior a sensibilidade, o modelo possui uma boa capacidade preditiva. A estatística $C = \text{Área sob a curva ROC} = \hat{P} = (Y_1 > Y_2)$, em que Y_1 = probabilidade prevista para ocorrência dos eventos, Y_2 = probabilidade prevista para ausência dos eventos. O valor de C varia de 0,5 (péssima capacidade preditiva) a 1 (ótima capacidade preditiva).

Modelo de Regressão Logística

Para o modelo de regressão logística, a variável independente pode ser contínua ou discreta, mas a variável resposta é tipicamente binária (DOMINGUEZ-ALMENDROS et al., 2011). A importância do modelo logístico vem do fato de que muitos conjuntos de dados ocorrem naturalmente no padrão de um resultado binário, mesmo relacionado com uma série de variáveis independentes. Nesta Tese utilizou este modelo para avaliar a ocorrência de

extremos de precipitação relacionada aos seguintes fatores: temperatura, umidade relativa e específica, radiação de onda longa e etc., em um total de doze variáveis.

Este modelo possui dois critérios importantes: 1) Dado um conjunto de dados de variáveis independentes, estimar a probabilidade de que o evento de interesse ocorra obtido pelas tabelas de contingência (Tabela 2), ou seja, a ocorrência de extremo de precipitação; e 2) Avaliar a influência de cada variável independente em sua resposta, em forma de razão de chances, também conhecida como *Odds Ratio*.

Tabela 2 – Tabela de contingência para contabilização dos extremos de precipitação e as variáveis meteorológicas.

Variável	Extremo de Precipitação		
	Sim	Não	Total
Anormal	A	B	$n_1 = A+B$
Normal	C	D	$n_2 = C+D$
Total	A+C	B+D	$N = n_1 + n_2$

Para a probabilidade de ocorrência do extremo de precipitação (π = probabilidade de sucesso), a chance é definida:

$$\Omega = \frac{\pi}{1-\pi} \quad (23)$$

As chances são não-negativas, $\Omega > 1$, quando a ocorrência do extremo sobrepõe o não-extremo. A *odds ratio* (ou razão de chances) é definida como:

$$\theta = \frac{\Omega_1}{\Omega_2} = \frac{\pi_1 / 1 - \pi_1}{\pi_2 / 1 - \pi_2} \quad (24)$$

Em que $\theta > 1$ ($\theta < 1$) representa fator de risco (proteção ou inibição), enquanto $\theta = 1$ não há influência de proteção ou risco. A expressão da regressão logística é definida como:

$$g(x) = \ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = \alpha + \beta X \quad (25)$$

$\pi(x)$ é a probabilidade de ocorrência de extremo de precipitação dada por:

$$\pi(x) = \frac{e^{\alpha + \beta X}}{1 + e^{\alpha + \beta X}} \quad (26)$$

Para a estimação dos parâmetros β_0 e β_1 do modelo, utiliza-se o resultado da tabela de contingência (Tabela 2), e aplica-se o log-odds nos dados observados, assim teremos:

$$\beta_0 = y_0 = \log\left(\frac{C}{D}\right) \text{ e } \beta_1 = y_1 - y_0 = \log\left(\frac{A}{B}\right) - \beta_0 = \log\left(\frac{A}{B}\right) - \log\left(\frac{C}{D}\right) \quad (27)$$

O parâmetro β_1 reflete a magnitude de uma associação entre status de grupo e resposta em termos de uma componente linear do modelo.

Modelo de Regressão de Poisson

A regressão de Poisson utiliza dados de contagens, suas variáveis assumem qualquer valor inteiro positivo. Nas tabelas de contingência para a regressão de Poisson, os dados correspondem às contagens das caselas obtidas na classificação cruzada entre duas ou mais variáveis categorizadas a partir das tabelas de contingências, neste caso, a caselas corresponderá à ocorrência dos extremos e não extremos para cada variável meteorológica (Tabela 2).

A regressão de Poisson tem média positiva. Embora se possa modelar a média de uma Poisson em MLG usando a ligação identidade, é mais comum modelar o logaritmo da média. O logaritmo da média é o parâmetro natural da distribuição Poisson e a ligação logarítmica é a ligação canônica para um MLG com componente aleatória Poisson. Seja μ , o valor esperado para uma variável Poisson Y e X uma variável explicativa. O modelo log-linear tem a forma:

$$\ln(\mu) = \alpha + \beta X \quad (28)$$

Para este modelo, a média satisfaz a relação exponencial:

$$\mu = \exp(\alpha + \beta X) = e^{\alpha + \beta X} \quad (29)$$

A cada acréscimo em X tem um impacto multiplicativo de e^β sobre μ . A média de Y em x+1 é igual à média de Y em x multiplicada por e^β . Se $\beta = 0$, então, $e^\beta = e^0 = 1$ e o fator multiplicativo é 1; ou seja, a média de Y não muda em X. Se $\beta > 0$, então $e^\beta > 1$ e a média de Y cresce quando X cresce. Se $\beta < 0$, a média decresce.

O modelo linear do logaritmo das frequências usado para representar os valores esperados (média) dos l_{ij} para uma tabela 2x2 é

$$\ln(\eta_{ij}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB} \quad (30)$$

E, especificamente, conduz às expressões para os valores estimados do logaritmo das frequências para cada casela na tabela 2x2, mostrados na Tabela 3. O parâmetro μ denota o nível médio geral dos logaritmos das frequências; λ_i^A denota a influência do nível i da variável A; λ_j^B denota a influência do nível j da variável B, e; λ_{ij}^{AB} denota a influência conjunta no nível (i, j) das variáveis A e B. O modelo requer as seguintes restrições nos parâmetros:

$$\lambda_1^A = -\lambda_2^A, \quad \lambda_1^B = -\lambda_2^B, \quad \lambda_{11}^{AB} = -\lambda_{12}^{AB} = -\lambda_{21}^{AB} = \lambda_{22}^{AB}$$

Tabela 3 – Modelo log-linear para dados em uma tabela de contingência 2x2.

	B ₁	B ₂	Resposta Média
A ₁	$\mu + \lambda_1^A + \lambda_1^B + \lambda_{11}^{AB}$	$\mu + \lambda_1^A + \lambda_2^B + \lambda_{12}^{AB}$	$\mu + \lambda_1^A$
A ₂	$\mu + \lambda_2^A + \lambda_1^B + \lambda_{21}^{AB}$	$\mu + \lambda_2^A + \lambda_2^B + \lambda_{22}^{AB}$	$\mu + \lambda_2^A$
Resposta Média	$\mu + \lambda_1^B$	$\mu + \lambda_2^B$	μ

Modelos Vetoriais Lineares Generalizados

O modelo vetorial linear generalizado (MVLG) é uma extensão dos MLG, composto por uma grande quantidade de classes de modelos de regressão que são cobertos pelo MLG tradicional, com a vantagem de fornecer modelos de respostas tanto de forma univariada como também multivariada, como por exemplo, análise de sobrevivência, séries temporais, regressão quantílica e análise de valores extremos (YEE, 2004; YEE, 2010). Em ciências climáticas, a aplicação do MVLG foi aplicada aos dados de precipitação (YEE & STEPHENSON, 2007; VOVORAS & TSOKOS, 2009). A inserção da análise da TVE em modelos de regressão nos MVLG como função de ligação tem aperfeiçoado os resultados neste tipo de análise, aplicado tanto de maneira univariada (a própria variável se explica) como múltipla e/ou multivariada (conjunto de variáveis), visto em alguns trabalhos (YEE & STEPHENSON, 2007; JORGENSEN et al. 2010; WANG & DEY, 2010; CALABRESE & OSMETTI, 2010, JONES et al., 2014).

O MVLG está descrito de forma mais detalhes em Yee & Wild (1996). Supondo que a resposta observada y é um vetor q -dimensional, o MVLG é definido como um modelo para qual o distribuição condicional de Y dado um x explicativo na forma:

$$f(y | x; B, \phi) = h \cdot (y, \eta_1, \dots, \eta_M, \phi) \quad (31)$$

Para função conhecida $h(\cdot)$, em que $B = (\beta_1, \beta_2, \dots, \beta_M)$ é uma matriz $p \times M$ de coeficientes de regressão desconhecidos, o j -ésimo preditor linear é:

$$\eta_j = \eta_j(x) = \beta_j^T x = \sum_{k=1}^p \beta_{(j)k} x_k, \quad j = 1, \dots, M \quad (32)$$

Em que $x = (x_1, \dots, x_p)^T$, se x_1 for o intercepto então $x_1 = 1$. Note que todos os parâmetros são modelos como funções de x (31).

Para o modelo de valores extremos simples M é o número de parâmetros a ser estimado e q deve ser qualquer inteiro positivo. O η_j do MVLG deve ser aplicado diretamente aos parâmetros de distribuição e não apenas a uma média do MLG.

REFERÊNCIAS

- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v. AC-19, p.716–723, 1974.
- AMBROSINO, C.; CHANDLER, R. E.; TODD, M. C. Southern African Monthly Rainfall Variability: An Analysis Based on Generalized Linear Models. **Journal of Climate**, v. 24, n. 17, 2011.
- BACK, L.E.; BRETHERTON, C.S. The relationship between wind speed and precipitation in the Pacific ITCZ. **Journal of climate**, v. 18, n. 20, 2005.
- BERG, P.; MOSELEY, C.; HAERTER, J.O. Strong increase in convective precipitation in response to higher temperatures, **Nature Geoscience**, v. 6, n. 3, pp. 181–185, 2013.
- BERRI, G.J.; INZUNZA, B.J. The effect of the low-level jet on the poleward water vapor transport in the central region of South America, **Atmospheric Environment**, v.27A(3), p.335–341, 1993.
- BOMBARDI, R. J.; CARVALHO, L. M.; JONES, C.; REBOITA, M. S. Precipitation over eastern South America and the South Atlantic Sea surface temperature during neutral ENSO periods. **Climate Dynamics**, p. 1-16, 2013.
- BURLANDO, M.; ANTONELLI, M.; RATTO, C. F. Mesoscale wind climate analysis: identification of anemological regions and wind regimes. **International Journal of Climatology**, v. 28, n. 5, p. 629-641, 2008.
- CALABRESE, R.; OSMETTI, S.A. Generalized Extreme Value Regression for Binary Rare Events Data: an Application to Credit Defaults, **Working Papers 201120**, Geary Institute, University College Dublin, 2011.
- CARVALHO, M. A. V.; OYAMA, M. D. Atlantic Intertropical Convergence Zone width and intensity variability: observational aspects. **Revista Brasileira de Meteorologia**, v. 28, n. 3, p. 305-316, 2013.
- CARVALHO, L.M.V.; JONES, C.; LIEBMANN, B. The South Atlantic convergence zone: Intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall. **Journal of Climate**, v.17, n. 1, p. 88-108, 2002.

- CAVALCANTI, I.F.A. Large scale and synoptic features associated with extreme precipitation over South America: A review and case studies for the first decade of the 21st century. **Atmospheric Research** v. 118, p. 27-40, 2012.
- CHANDLER, R. E. On the use of generalized linear models for interpreting climate variability. **Environmetrics**, v. 16, n. 7, p. 699-715, 2005.
- CHANDLER, R. E.; WHEATER, H. S. Analysis of rainfall variability using generalized linear models: a case study from the west of Ireland. **Water Resources Research**, v. 38, n. 10, p. 10-1-10-10, 2002.
- CHAVES, R.R., CAVALCANTI, I.F.A. Atmospheric Circulation features Associated with Rainfall Variability over Southern Northeast Brazil. **Monthly Weather Review**, 129: 2614-2626, 2001.
- COLES, S.; BAWA, J.; TRENNER, L.; DORAZIO, P. **An introduction to statistical modeling of extreme values** (Vol. 208), 2001, London: Springer.
- CORRAR, L. J; PAULO, E.; DIAS FILHO, J. M. **Análise Multivariada - Para os Cursos de Administração, Ciências Contábeis e Economia**, Ed. Atlas, São Paulo, 1ª Edição, 344 p., 2007.
- CORREIA-FILHO, W.L.F.; LUCIO, P.S.; SPYRIDES, M.H.C. Precipitation Extremes Analysis over the Brazilian Northeast via Logistic Regression. **Atmospheric and Climate Science**, v. 4, n. 1, p. 53-59, 2014. doi: 10.4236/acs.2014.41007.
- DAVISON, A.C.; GHOLAMREZAEI, M.M. Geostatistics of extremes. **Proceedings of the Royal Society A**, v. 468, p.581-608, 2012, doi:10.1098/rpsa2011.0412.
- DELLA-MARTA, P.M. et al. The return period of wind storms over Europe. **International Journal of Climatology**, v. 29, n. 3, p. 437-459, 2009.
- DESER, C.; PHILLIPS, A.; BOURDETTE, V.; TENG H. Uncertainty in climate change projections: the role of internal variability. **Climate Dynamics**, 2010, doi:10.1007/s00382-010-0977-x.
- DIAZ, H.F.; MURNANE, R.J. (Ed.). **Climate extremes and society**. Cambridge University Press, 2008.

- DING Y.; CHENG, B.; JIANG, Z. A Newly-Discovered GPD-GEV Relationship Together with Comparing Their Models of Extreme Precipitation in Summer, **Advances in Atmospheric Sciences**, v.25, p. 507–516, 2008.
- DOMÍNGUEZ-ALMENDROS, S.; BENÍTEZ-PAREJO, N.; GONZALEZ-RAMIREZ, A. R. Logistic regression models. **Allergologia et immunopathologia**, v.39, n° 5, p.295–305, 2011.
- DOMMENGET, D.; LATIF, M. Interannual to decadal variability in the tropical Atlantic. **Journal of Climate**, v. 13, n. 4, p. 777-792, 2000.
- DORE, M. H. Climate change and changes in global precipitation patterns: what do we know? **Environment international**, v. 31, n. 8, p. 1167-1181, 2005.
- EASTERLING, D. R. et al. Observed Variability and Trends in Extreme Climate Events: A Brief Review*. **Bulletin of the American Meteorological Society**, v. 81, n. 3, p. 417-425, 2000.
- EMBRECHTS, P.; KLÜPPELBERG, C.; MIKOSCH, T. **Modelling extremal events for insurance and finance**. Berlin: Springer-Verlag, 2004.
- EVANS, J. L.; BRAUN, A. A Climatology of Subtropical Cyclones in the South Atlantic. **Journal of Climate**, v. 25, n. 21, 2012.
- FAWCETT, T. An introduction to ROC analysis. **Pattern Recognition Letters**, v.27, p.861–874, 2006.
- FOLLAND, C. K. et al., Observed climate variability and change. **Climate change: The scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change**, Eds. J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. van der Linden, X. Dai, K. Maskell, and C. I. Johnson. Cambridge University Press, p. 525-582, 2001.
- GARAVAGLIA, C. R.; DOYLE, M. E.; BARROS, V. R. Statistical relationship between atmospheric circulation and extreme precipitation in La Plata Basin. **Meteorological Applications**, 2013.
- GASTINEAU, G.; SODEN, B. J. Evidence for a weakening of tropical surface wind extremes in response to atmospheric warming. **Geophysical Research Letters**, v. 38, n. 9, 2011.
- GIANNINI, A.; KUSHNIR, Y.; CANE, M. A. Interannual variability of Caribbean rainfall, ENSO, and the Atlantic ocean. **Journal of Climate**, v. 13, n° 2, p. 297-311, 2000.

GIANNINI, A.; SARAVANAN, R.; CHANG, P. Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. **Science**, v. 302, n. 5647, p. 1027-1030, 2003.

GRIMM, A.M.; TEDESCHI, R.G. ENSO and Extreme Rainfall Events in South America, **Journal of Climate**, v. 22, n. 7, p. 1589–1609, 2009.

GROISMAN, P. Y.; KNIGHT, R.W. Prolonged Dry Episodes over the Conterminous United States: New Tendencies Emerging during the Last 40 Years. **Journal of Climate**, v.21, n.9, 2008.

HASTENRATH, S. Interannual variability and annual cycle: Mechanisms of circulation and climate in the tropical Atlantic sector. **Monthly Weather Review**, v. 112, n.6, p.1097-1107, 1984.

HASTENRATH, S; GREISCHAR, L. Further work on the prediction of Northeast Brazil rainfall anomalies. **Journal of Climate**, v. 6, n.4, p.743–758, 1993.

HOSMER, D.W., TABER, S. & LEMESHOW, S. The importance of assessing the fit of logistic regression models: a case study. **American journal of public health**, v. 81, nº12, p.1630–1635, 1991.

HOUZE Jr, R.A. Orographic effects on precipitating clouds. **Review of Geophysics**, v.50 RG1001, 2012, doi:10.1029/2011RG000365.

HUNTINGFORD, C. et al. Regional climate-model predictions of extreme rainfall for a changing climate. **Quarterly Journal of the Royal Meteorological Society**, v. 129, n. 590, p. 1607-1621, 2003.

IDALINO, R. C. L., OLIVEIRA, P. S., LUCIO, P. S., Modelagem de Extremos Meteorológicos via GEV e GPD - Uma análise comparativa de algumas capitais brasileiras. **Revista Brasileira de Estatística**, v.70, p.33–56, 2009.

JENKINSON, A. F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. **Quarterly Journal of the Royal Meteorological Society**, v. 87, p. 145-58, 1955.

JONES, M. R., et al. Objective classification of extreme rainfall regions for the UK and updated estimates of trends in regional extreme rainfall. **International Journal of Climatology**, v. 34, n. 3, p. 751-765, 2014.

- JØRGENSEN, B.; GOERGEBEUR, Y., MARTÍNEZ, J.R. Dispersion models for extremes. **Extremes**, v.13, n.4, p.399-437, 2010.
- KALKSTEIN, L. S.; TAN, G.; SKINDLOV J. An evaluation of objective clustering procedures for use in Synoptic climatological classification, **Journal of Climate and Applied Meteorology**, v. 26, nº 6, 717-730, 1987.
- KAYANO, M. T.; ANDREOLI, R. V.; SOUZA, R. A. F. Relations between ENSO and the South Atlantic SST modes and their effects on the South American rainfall. **International Journal of Climatology**, v. 33, n. 8, p. 2008-2023, 2013.
- KEENLYSIDE, N. S.; DING, H.; LATIF, M. Potential of equatorial Atlantic variability to enhance El Niño prediction. **Geophysical Research Letters**, v. 40, n. 10, p. 2278-2283, 2013.
- KOUADIO, Y.K. et al. Heavy Rainfall Episodes in the Eastern Northeast Brazil Linked to Large-Scale Ocean-Atmosphere Conditions in the Tropical Atlantic. **Advances in Meteorology**, 2012.
- KOUSKY, V. E. Frontal influences on northeast Brazil. **Monthly Weather Review**, v. 107, n. 9, p. 1140-1153, 1979.
- KOUSKY, V.E.; GAN, M.A. Upper tropospheric cyclonic vortices in the tropical South Atlantic. **Tellus**, v.33, n.6, p.538-551, 1981.
- L'HEUREUX, M. L.; COLLINS, D. C.; HU, Z-Z. Linear trends in sea surface temperature of the tropical Pacific Ocean and implications for the El Niño-Southern Oscillation. **Climate Dynamics**, v. 40, nº 5-6, p. 1223-1236, 2013.
- LI, Y.; CAI, W.; CAMPBELL, E. P. Statistical Modeling of Extreme Rainfall in Southwest Western Australia. **Journal of Climate**, v. 18, p.852–863, 2005.
- LIEBMANN, B.; JONES, C.; CARVALHO, L. Interannual variability of daily extreme precipitation events in the state of São Paulo, Brazil. **Journal of Climate**, v. 14, n. 2, 2001.
- LIEBMANN, B.; KILARDIS, G.N.; ALLURED, D.; VERA, C.S.; JONES, C.; CARVALHO, L.M.V.; BLADÉ, I., GONZÁLES, P.L.M. Mechanisms Associated with Large Daily Rainfall Events in Northeast Brazil, **Journal of Climate**, v. 24, p.1589-1609, 2011, doi:10.1175/2008JCL2429.1.

- LITTLE, M.; MCSHARRY, P.E.; TAYLOR, J. Generalized Linear Models for Site-Specific Density Forecasting of U.K. Daily Rainfall. **Monthly Weather Review**, v.137, p.1029-1045, 2009, doi:10.1175/2008MWR2614.1.
- MADDEN, R. A.; JULIAN, P. R. Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific. **Journal of the Atmospheric Sciences**, v. 28, n. 5, p. 702-708, 1971.
- MADDOX, R. A. An objective technique for separating macroscale and mesoscale features in meteorological data. **Monthly Weather Review**, v. 108, n. 8, p. 1108-1121, 1980.
- MARAUN, D.; OSBORN, T. J.; RUST, H. W. The influence of synoptic airflow on UK daily precipitation extremes. Part I: Observed spatio-temporal relationships. **Climate Dynamics**, v. 36, n. 1-2, p. 261-275, 2011.
- MARENGO, J. A.; HASTENRATH, S. Case studies of extreme climatic events in the Amazon basin. **Journal of Climate**, v. 6, n. 4, p. 617-627, 1993.
- MARENGO, J. A. et al. Two Contrasting Severe Seasonal Extremes in Tropical South America in 2012: Flood in Amazonia and Drought in Northeast Brazil. **Journal of Climate**, v. 26, n. 22, p. 9137-9154, 2013.
- MEEHL, G.A.; TEBALDI, C. More intense, more frequent, and longer lasting heat waves in the 21st century. **Science**, v. 305, p. 994-997, 2004.
- NEELIN, J. D.; CHOU, C.; SU, H. Tropical drought regions in global warming and El Niño teleconnections. **Geophysical Research Letters**, v. 30, n. 24, 2003.
- NELDER, A.J.A.; WEDDERBURN, R.W.M. Generalized Linear Models. **Journal of the Royal Statistical Society Series A (General)**, v.135, n. 3, p. 370-384, 1972.
- NNAMCHI, H. C.; LI, J.; ANYANDIKE, R.N.C. Does a dipole mode really exist in the South Atlantic Ocean? **Journal of Geophysical Research**, v.116(D15), p.D15104, 2011.
- NOBRE, P.; SHUKLA, J. Variations of sea surface temperature, wind stress, and rainfall over the tropical Atlantic and South America. **Journal of climate**, v. 9, n. 10, p. 2464-2479, 1996.
- O'GORMAN, P.A.; SCHNEIDER, T. The physical basis for increases in precipitation extremes in simulations of 21st-century climate change, **Proceedings of the National Academy of Sciences**, v. 106, p. 14773-14777, 2009.

- OYAMA, M.D.; NOBRE, C. A. Climatic consequences of a large-scale desertification in northeast Brazil: A GCM simulation study. **Journal of Climate**, v. 17, n. 16, p. 3203–3213, 2004.
- PARK, S.H., GOO, J. M., JO, C.H. 2004: Receiver Operating Characteristic (ROC) Curve: Practical Review for Reviews for Radiologists, **Korean J. Radiol.** 5, 11-18.
- POVEDA, G.; JARAMILLO, L.; VALLEJO, L. F. Seasonal precipitation patterns along pathways of South American low-level jets and aerial rivers. **Water Resources Research**, 2014.
- POWELL, E. J. **Climate Extremes in the Southeast United States: Observed Variability, Spatial Classification, and Related Planning**. 2014. Tese de Doutorado. University of Georgia, 305p.
- RAO, V. B.; LIMA, M. C. de; FRANCHITO, S.H. Seasonal and Internannual Variations of Rainfall over Eastern Brazil. **Journal of Climate**. n.6, p. 1754-1763, 1993.
- RAO, V. B.; HADA, K.; HERDIES, D.L. On the severe drought of 1993 in north-east Brazil. **International Journal of Climatology**, v. 15, n. 6, p. 697-704, 1995.
- ROMERO, R.; RAMIS, C.; ALONSO, S. Numerical simulation of an extreme rainfall event in Catalonia: Role of orography and evaporation from the sea. **Quarterly Journal of the Royal Meteorological Society**, v. 123, n. 539, p. 537-559, 1997.
- ROPELEWSKI, C. F.; HALPERT, M. S. Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. **Monthly Weather Review**, v. 115, p.1606–1626, 1987.
- SANDEEP, S.; STORDAL, F. Use of daily outgoing longwave radiation (OLR) data in detecting precipitation extremes in the tropics. **Remote Sensing Letters**, v. 4, n. 6, p. 570-578, 2013.
- SARIS, F.; HANNAH, D. M.; EASTWOOD, W. J. Spatial variability of precipitation regimes over Turkey. **Hydrological Sciences Journal–Journal des Sciences Hydrologiques**, vol. 55, n° 2, 234-249, 2010.
- SILVA, V.B.S.; KOUSKY, V.E. The South American Monsoon System: Climatology and Variability. **Modern Climatology**. Dr Shih-Yu Wang (Ed.), ISBN: 978-953-51-0095-9, 2012, InTech, doi: 10.5772/38565.

- SKANSI, M. M. et al. Warming and wetting signals emerging from analysis of changes in climate extreme indices over South America. **Global and Planetary Change**, v. 100, p. 295-307, 2013.
- SOUZA, P.; CAVALCANTI, I.F.A. Atmospheric centres of action associated with the Atlantic ITCZ position. **International Journal of Climatology**, v.29, n.14, p.2091-2105, 2009.
- TORRES, R.R.; FERREIRA, N.J. Case studies of easterly wave disturbances over Northeast Brazil using the Eta Model. **Weather and Forecasting**, v.26, p.225-235, 2011.
- TRENBERTH, K.E.; CARON, J.M.; STEPANIAK, D.P., WORLEY, S. The evolution of ENSO and global atmospheric surface temperatures. **Journal of Geophysical Research** 107:D8, 2002, doi:10.1029/2000JD000298.
- TRENBERTH, K.E. et al. Observations: Surface and atmospheric climate change. **Climate Change 2007: The Physical Science Basis**, S. Solomon et al., Eds., Cambridge University Press, p. 235–336, 2007.
- TRENBERTH, K. E. Changes in precipitation with climate change. **Climate Research**, v. 47, n. 1, p. 123, 2011.
- TRENBERTH, K. E. 2012. Framing the way to relate climate extremes to climate change. **Climatic Change**, v.115, p. 283-290, 2012.
- UNAL, Y.; KINDAP, T., KARACA, M. Redefining the climate zones of Turkey using cluster analysis. **International Journal of Climatology**, v. 23, p. 1045–1055, 2003.
- UVO, C.B. et al. The relationships between tropical Pacific and Atlantic SST and northeast Brazil monthly precipitation. **Journal of Climate**, v. 11, n. 4, 1998.
- VOVORAS D.; TSOKO, C. P. Statistical analysis and modeling of precipitation data. **Nonlinear Analysis: Theory, Methods e Applications**, v.71, p.e1169–e1177, 2009, doi:10.1016/j.na.2009.01.144.
- WANG, X.; DEY, D.K. Generalized extreme value regression for binary response data: an application to B2B electronic payments system adoption. **The Annals of Applied Statistics**, v.4, n. 4, p.2000-2023, 2010.
- WARD JR, J. H. Hierarchical grouping to optimize an objective function. **Journal of the American statistical association**, v. 58, n. 301, p. 236-244, 1963.

WILKS, D. **Statistical Methods in the Atmospheric Sciences**, Ed. Academic Press, London, 592p., 2005.

WILLEMS, P.; ARNBJERG-NIELSEN, K.; OLSSON, J.; NGUYEN, V.T.V. Climate change impact assessment on urban rainfall extremes and urban drainage: Methods and shortcomings. **Atmospheric Research**, v. 103, p. 106–118, 2012.

WMO – WORLD METEOROLOGICAL ORGANIZATION. **Guidelines on Analysis of extremes in a changing climate in support of informed decisions for adaptation**, v.53, 2009.

WOTLING, G.; BOUVIER, C.; DANLOUX, J.; FRITSCH, J.M. Regionalization of extreme precipitation distribution using the principal components of the topographical environment. **J of Hydrol**, v. 233, n.1-4, p. 86–101, 2000.

XIE, S.-P.; CARTON, James A. Tropical Atlantic variability: Patterns, mechanisms, and impacts. **Geophysical Monograph Series**, v. 147, p. 121-142, 2004.

YAN, Z. et al. An analysis of daily maximum wind speed in northwestern Europe using generalized linear models. **Journal of Climate**, v. 15, n. 15, 2002.

YEE, T.W.; WILD, C. J. Vector generalized additive models. **Journal of the Royal Statistical Society, Series B, Methodological**, v.58, p.481–493, 1996.

YEE, T.W. Quantile regression via vector generalized additive models. *Statistics in Medicine*, v. 23, 2295 – 2315, 2004.

YEE, T.W.; STEPHENSON, A.G. Vector generalized linear and additive extreme value models. **Extremes**, v.10, p. 1–19, 2007.

YEE, T.W. VGLMs and VGAMs: an overview for applications in fisheries research. **Fisheries Research**, v. 101, n. 1, p. 116-126, 2010.