



UNIVERSIDADE ESTADUAL DO CEARÁ
JOSÉ GOMES RIBEIRO FILHO

**APLICAÇÕES DA TÉCNICA DE ANÁLISE DE
CORRELAÇÕES CANÔNICAS NA PREVISÃO
CLIMÁTICA SAZONAL NO CEARÁ**

Fortaleza, Ceará
Dezembro 2009

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

JOSÉ GOMES RIBEIRO FILHO

**APLICAÇÕES DA TÉCNICA DE ANÁLISE DE
CORRELAÇÕES CANÔNICAS NA PREVISÃO
CLIMÁTICA SAZONAL NO CEARÁ**

Dissertação apresentada ao curso de
Mestrado em Ciências Físicas Aplicadas da
Universidade Estadual do Ceará como
requisito parcial para obtenção do título de
Mestre em Ciências Físicas Aplicadas.

Área de concentração: Meteorologia de Meso e Grande Escalas
Subárea: Climatologia

Orientador: Prof. Dr. Alexandre Araújo Costa

Fortaleza, Ceará
Dezembro 2009

UNIVERSIDADE ESTADUAL DO CEARÁ - UECE
MESTRADO EM CIÊNCIAS FÍSICAS APLICADAS

Título do Trabalho: Aplicações da Técnica de Análise de Correlações Canônicas na
Previsão Climática Sazonal no Ceará

Autor: José Gomes Ribeiro Filho

Dissertação em: 18/12/2009

Conceito obtido: _____

Nota: _____

BANCA EXAMINADORA

Prof. Dr. Alexandre Araújo Costa - UECE (Orientador)

Prof. Dr. Dirceu Silveira Reis Jr. - FUNCEME

Dr. José Maria Brabo Alves - FUNCEME

*À minha mãe Janete Lira e aos meus
filhos, Gabriel e Caio*

*"A Educação é o nosso passaporte
para o futuro, pois o amanhã
pertence às pessoas que se preparam
hoje." (MALCOLM X)*

AGRADECIMENTOS

À Deus nosso criador, por todas as graças alcançadas em minha vida;

Ao Professor Alexandre Araújo Costa, pela oportunidade, incentivo e por propiciar todo o suporte necessário para a realização deste trabalho;

À Fundação Cearense de Meteorologia e Recursos Hídricos (FUNCEME), pela cessão dos dados de precipitação observada e do *downscaling* dinâmico;

Aos pesquisadores da FUNCEME em especial a Mosefran Firmino, Francisco Venícius Fernandes Barros, Otacílio Leandro de Menezes Neto e Paulo Henrique Santiago de Maria pelo grande apoio;

A todos os companheiros do MCFA, inclusive os que já seguiram por novos caminhos, com os quais aprendi muitas coisas;

Ao Professor Antônio Carlos Santana dos Santos, que muito me auxiliou em diversas atividades;

Ao Mestrado em Ciências Físicas Aplicadas (MCFA), em especial a todos os professores e demais funcionários;

À minha família, pela tranquilidade para poder fazer o meu melhor;

Em especial, à minha namorada Samara, pelo amor, companheirismo e incentivo;

Ao meu filho Caio Ribeiro de Brito por me fazer rir nos momentos mais tensos de conclusão deste trabalho;

Por fim, agradeço a todos que diretamente ou indiretamente contribuíram para a realização deste trabalho. Muito obrigado!

RESUMO

A previsão da precipitação sobre o Nordeste Brasileiro e o Estado do Ceará, bem como de sua distribuição espacial e temporal durante a estação chuvosa, é importantíssima para os recursos hídricos e a agricultura locais, pois auxilia na orientação para a distribuição de sementes (no início do plantio) e para o gerenciamento dos reservatórios. Dada a crescente demanda dos usuários por informações sobre condições climáticas extremas (secas e enchentes), o presente trabalho tem como objetivos: 1) aplicar um modelo empírico de correlações canônicas para a previsão sazonal de anomalias de precipitação, que poderá ser utilizado por diversos setores da sociedade, a exemplo destes dois setores e 2) utilizar o modelo de correlações canônicas como corretor dos resultados de um modelo dinâmico utilizado na Fundação Cearense de Meteorologia e Recursos Hídricos na previsão climática sazonal e para agregar à informação gerada por esse modelo a informação da TSM propriamente dita compondo um modelo de previsão híbrido dinâmico-empírico. Os resultados sugerem que o uso de ACC na previsão climática sazonal proporciona bons resultados tanto diretamente como modelo empírico quanto associada a modelos dinâmicos.

ABSTRACT

Forecasting the precipitation over Northeast Brazil and the state of Ceará, as well its time and space distribution during the rainy season, is very important for local water resources and agriculture, since it provides information for seed distribution (at the beginning of the plantation) and reservoir management. Given the growing demand from users for information on extreme climate conditions (droughts and floods), this work aims at 1) implementing an empirical model of canonical correlations for prediction of seasonal precipitation anomalies, which could be used by various sectors of society, like the ones already mentioned and 2) use canonical correlations a model output correction of dynamical downscaling results from a mesoscale model used for seasonal prediction by the Ceará state Foundation for Meteorology and Water Resources and to join the data from the dynamical model and the SSTs as predictors in a hybrid, dynamical-empirical model. The results suggest that the use of CCA in seasonal climate prediction produces good results both as an empirical model and associated with dynamical models.

SUMÁRIO

1 INTRODUÇÃO.....	17
2 REVISÃO DE LITERATURA	21
2.1 PREVISÃO CLIMÁTICA	21
2.2 MODELOS EMPÍRICOS	29
2.2.1 REGRESSÃO MÚLTIPLA.....	30
2.2.2 ACP	31
2.2.3 ACC.....	32
2.3 MODELOS DINÂMICOS.....	37
2.3.1 MODELOS GLOBAIS.....	38
2.3.2 MODELOS REGIONAIS.....	39
2.4 DOWNSCALING.....	39
2.5 ERROS E TÉCNICAS ESTATÍSTICAS DE CORREÇÃO	41
2.6 PREVISÃO HÍBRIDA.....	44
3 MATERIAL E MÉTODOS.....	47
3.1 DADOS OBSERVADOS	47
3.1.1 TSM	47
3.1.2 PRECIPITAÇÃO.....	49
3.2 DOWNSCALING DINÂMICO DA FUNCEME.....	50
3.2.1 RSM.....	51
3.2.2. RAMS.....	52

3.3 CPT	54
3.4 ESTRATÉGIAS DE USO DA ACC	55
3.4.1 PREVISÃO EMPÍRICA.....	55
3.4.2 CORREÇÃO ESTATÍSTICA DOS RESULTADOS DE DOWNSCALING DINÂMICO	57
3.4.3 PREVISÃO HÍBRIDA.....	59
3.5 MÉTRICAS E DESTREZA DAS PREVISÕES	61
3.5.1 COEFICIENTE DE CORRELAÇÃO DE PEARSON	61
3.5.2 COEFICIENTE DE CORRELAÇÃO DE SPEARMAN.....	63
3.5.3 ÁREAS ROC	64
4.1 PREVISÃO EMPÍRICA PARA AS UNIDADES DE PLANEJAMENTO (BACIAS)	70
4.1.1 CONFIGURAÇÃO DO CPT	70
4.1.2 MODOS DE VARIABILIDADE	71
4.1.3 DESTREZA DO MODELO	77
4.2 CORREÇÃO DA MODELAGEM NUMÉRICA	84
4.3 PREVISÃO HÍBRIDA.....	91
4.3.1 PEARSON	91
4.3.2 SPEARMAN.....	94
4.3.3 ÁREAS ROC - ABOVE	98
4.3.3 ÁREAS ROC - BELOW	101
5 CONCLUSÕES E PERSPECTIVAS.....	105
REFERÊNCIAS BIBLIOGRÁFICAS	107

ANEXO120

LISTA DE SIGLAS E ABREVIATURAS

AMJ – Abril-Maio-Junho

ACC – Análise de Correlações Canônicas

ACP – Análise de Componente Principal

AUC – Area Under Curve

CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico

CP – Correlação de Pearson

CPT – Climate Predictability Tool

COADS – Comprehensive Ocean-Atmosphere Data Set

CRU – climate research unit

CS – Correlação de Spearman

DJF – Dezembro-Janeiro-Fevereiro

DVS – Decomposição do Valor Singular

ENOS – El Niño Oscilação Sul

FINEP – Financiadora de Estudos e Projetos

FMA – Fevereiro-Março-Abril

FOE – Funções Ortogonais Empíricas

FP – Falsos Positivos

FUNCEME – Fundação Cearense de Meteorologia e Recursos Hídricos

HCM – Hybrid Coupled Models

INMET – Instituto Nacional de Meteorologia

INPE – Instituto Nacional de Pesquisas Espaciais

IRI – International Research Institute for Climate and Society

ISAN – ISentropic ANalysis

JFM – Janeiro-Fevereiro-Março

MAM – Março-Abril-Maio

MCG – Modelo de Circulação Geral

MCGA – Modelo de Circulação Geral Atmosférico

MCGO – Modelo de Circulação Geral Oceânico

MOS – Model Output Statistics
NCEP – National Centers for Environmental Prediction
NDJ – Novembro-Dezembro-Janeiro
NE – Nordeste
NEB – Nordeste Brasileiro
NOAA – National Oceanic and Atmospheric Administration
RAMS – Regional Atmospheric Modeling System
RCP – Regressão por Componentes Principais
RSM – Regional Spectral Modeling
RLM – Regressão Linear Múltipla
ROC – Receiver Operating Characteristic Curve
SIMOC – Sistema de Modelagem dos Oceanos
TSM – Temperatura da Superfície do Mar
VP – Verdadeiros Positivos
ZCIT – Zona de Convergência Inter-Tropical

LISTA DE FIGURAS

Figura 3.1 Regiões do Oceano Pacífico onde é feito o monitoramento da TSM.....	48
Figura 3.2 Regiões de influencia da TSM na precipitação no Norte e Nordeste do Brasil.....	49
Figura 3.3 a)Regiões Pluvimétricamente Homôgeneas e b)Unidades de Planejamento (Bacias).....	50
Figura 3.4 Diagrama esquemático do Sistema de Previsão Climática Regional da FUNCEME.....	51
Figura 3.5 Grade do RSM-97/FUNCEME.....	52
Figura 3.6 Grade horizontal utilizada no modelo RAMS.	53
Figura 3.7 Esquema ilustrativo das etapas utilizadas para a previsão da precipitação, utilizando ACC.....	57
Figura 3.8 Esquema ilustrativo das etapas utilizadas para a correção estatística dos resultados de previsão da precipitação	59
Figura 3.9 Esquema ilustrativo das etapas utilizadas para a correção estatística dos resultados de previsão da precipitação	60
Figura 3.10 Curvas ROC de diferentes classificadores.....	67
Figura 4.1 a) Carga canônica da anomalia de TSM em outubro, b) Carga canônica da precipitação em DJF e c) Escores temporais para o primeiro modo dominante de correlações canônicas: TSM em vermelho e precipitação em verde.....	73
Figura 4.2 Como na Figura 4.1, mas para a TSM em novembro e a precipitação em JFM..	74
Figura 4.3 Como na Figura 4.1, mas para a TSM em dezembro e a precipitação em FMA	75
Figura 4.4 Como na Figura 4.1, mas para a TSM em janeiro e a precipitação em MAM	76
Figura 4.5 Como na Figura 4.1, mas para a TSM em fevereiro e a precipitação em AMJ	77

Figura 4.6 Coeficientes de correlação de Pearson entre as séries observadas de precipitação e as séries previstas pelo CPT no período de 1981 a 2005 para as bacias no Estado do Ceará para os trimestres: (a) DJF, (b) JFM, (c) FMA, (d) MAM e (e) AMJ	78
Figura 4.7 Como na Figura 4.6, mas para o coeficiente de correlação de Spearman	79
Figura 4.8 Como na Figura 4.6, mas para as áreas ROC - <i>Above</i>	80
Figura 4.9 Como na Figura 4.6, mas para as áreas ROC - <i>Below</i>	81
Figura 4.10 Coeficientes de correlação de Pearson entre os resultados não corrigidos do modelo regional e a precipitação observada (esquerda) e a correlação de Pearson entre os resultados corrigidos através do CPT e a precipitação observada (direita) para os trimestres: (a) DJF, (b) JFM, (c) FMA, (d) MAM e (e) AMJ	85
Figura 4.11 Como na Figura 4.10, mas para o coeficiente de correlação de Spearman	86
Figura 4.12 Como na Figura 4.10, mas para as áreas ROC - <i>Above</i>	87
Figura 4.13 Como na Figura 4.10, mas para as áreas ROC - <i>Below</i>	89
Figura 4.14 Coeficientes de correlação de Pearson entre a precipitação observada e o resultado do modelo híbrido (esquerda-acima) com a correlação de Pearson entre a precipitação observada e os resultados, via CPT, da previsão empírica pura (direita-acima), previsão numérica pura (esquerda-abixo) e a previsão numérica corrigida (direita-abaixo) para os trimestres: (a) DJF, (b) JFM, (c) FMA, (d) MAM e (e) AMJ.....	92
Figura 4.15 Como na Figura 4.14, mas para o coeficiente de correlação de Spearman	95
Figura 4.16 Como na Figura 4.14, mas para as áreas ROC - <i>Above</i>	98
Figura 4.17 Como na Figura 4.14, mas para as áreas ROC - <i>Below</i>	101

LISTA DE TABELAS

Tabela 3.1 Análise qualitativa do grau de correlação entre duas variáveis.....	63
Tabela 3.2 Tabela de Contingência (matriz de confusão).	66
Tabela 4.1 Posições dos domínios oceânicos que foram usados para a previsão empírica.	71

1 INTRODUÇÃO

O Brasil, por ser um país de grande extensão territorial, possui diferentes regimes de precipitação. De norte a sul, nota-se uma grande diversidade de climas com diferentes características regionais. Na região Norte do País, constata-se um clima equatorial chuvoso, com chuvas de primavera, verão e outono, sendo que na parte noroeste da região praticamente não se nota estação seca. As regiões Sudeste e Centro-Oeste sofrem influência tanto de sistemas tropicais quanto de latitudes médias, com estação seca bem definida no inverno e estação chuvosa no verão com chuvas convectivas. O período chuvoso no norte do NEB, por sua vez, é dividido em chuvas da pré-estação, tipicamente nos meses de dezembro e janeiro; a estação chuvosa propriamente dita, de fevereiro a maio; e a pós-estação, que compreende o final de maio e o mês de junho. Normalmente, nos meses de dezembro, janeiro e início de fevereiro, as chuvas ocorrem devido à formação de vórtices ciclônicos de ar superior e também à influência de frentes frias que vêm até a Bahia e eventualmente induzem chuvas no Ceará. Na estação chuvosa, o principal sistema que atua é a Zona de Convergência Intertropical; na pós-estação, as Ondas de Leste. Tradicionalmente, considera-se que a pré-estação e a pós-estação chuvosa têm baixa previsibilidade climática, comparadas à estação chuvosa propriamente dita.

Nesse contexto, a porção semi-árida da região Nordeste do Brasil é particularmente vulnerável às variações climáticas; a periodicidade das secas compromete principalmente a agricultura de subsistência que, na maior parte, se constitui de plantios em sistema de sequeiro (agricultura que não utiliza irrigação). De acordo com Silva (1992), o regime de precipitação pluvial nessa extensa área depende de vários sistemas atmosféricos que atuam na costa leste do Nordeste, sendo mais frequentes os seguintes: zona de convergência intertropical, brisas marítimas e terrestres, frentes frias, ciclones na alta troposfera e ondas de leste. O desenvolvimento do semi-árido do Nordeste do Brasil, principalmente no

Estado do Ceará, é fortemente dependente da precipitação pluviométrica, e, conseqüentemente, as suas variações provocam prejuízos econômicos e sociais à população do Estado. O Ceará tem como características climáticas marcantes, as irregularidades, tanto espacial quanto temporal, do seu regime de chuvas.

Segundo Diaz et al (1998) a variabilidade climática de determinadas regiões da América do Sul está relacionada em parte à variabilidade interanual da Temperatura da Superfície do Mar (TSM) nos Oceanos Pacífico e Atlântico, podendo influenciar diretamente na produtividade agrícola anual, causando perdas ou ganhos em diferentes regiões. Por exemplo, o comportamento da TSM na região denominada Niño 3, situada entre as latitudes 5°N–5°S e longitudes 150°–90°W, costuma influenciar expressivamente o comportamento da precipitação nas regiões sul e nordeste do Brasil, a qual, por sua vez, é determinante da produtividade agrícola nessas regiões, para culturas não irrigadas.

Muitas culturas são normalmente praticadas em regime de sequeiro, o que as faz depender exclusivamente da quantidade, da distribuição e da intensidade das chuvas. Os cultivos irrigados ainda são uma minoria, principalmente no que diz respeito a grandes áreas. Por esse motivo, a atividade agrícola torna-se exclusivamente sazonal, sendo praticada principalmente no período das chuvas. O milho, por exemplo, é plantado na maioria das áreas, no período chuvoso. Assim, conhecer o número de dias secos consecutivos é muito importante na determinação da época de plantio. Consideram-se como dias secos aqueles em que a precipitação é inferior a 5 mm. Muitos pesquisadores mostraram que as máximas produtividades ocorrem quando o consumo de água durante todo o ciclo está entre 500 e 800 mm e que a cultura exige um mínimo de 350-500 mm para que produza sem a necessidade de irrigação.

Especialistas acreditam que, para viabilizar a agricultura de sequeiro e a economia do Nordeste, deverão ser introduzidas culturas mais resistentes às longas estiagens, tais como o sorgo, milheto, soja, algodão, gergelim e mamona, entre outras. Essas culturas são consideradas mais resistentes às variações climáticas do semi-árido, além de apresentarem baixa requisição hídrica. As

informações a respeito das características e duração dos períodos chuvosos e menos chuvosos possibilitam que muitas atividades agrícolas, como a irrigação e o preparo do solo, possam ser planejadas e realizadas de forma racional, diminuindo os custos da produção.

Além da agricultura, os recursos hídricos são outro setor fortemente afetado pelas variações climáticas. Como a água é o insumo básico da sobrevivência de todas as espécies e indicador do desenvolvimento de uma região, ela necessita de especial atenção no seu manejo objetivando sua conservação em qualidade e quantidade. Isso é alcançado por meio da gestão dos recursos hídricos, que se refere aos procedimentos relativos à tentativa de equacionar e resolver as questões da água, otimizando seu uso.

Como exposto anteriormente, a precipitação pluvial é uma variável de muita importância para o desenvolvimento sócio-econômico do Ceará. Portanto, o conhecimento detalhado da distribuição espacial e temporal da precipitação nas micro e mesorregiões do Estado, bem como a verificação das anomalias de TSM dos oceanos como principal variável física influenciadora das condições climáticas, é de grande importância para a tomada de decisões por parte dos governos e da sociedade civil.

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

A técnica de ACC tem sido aplicada em vários estudos de modelagem estatística. Sua grande vantagem está no fato de a mesma possibilitar a busca de padrões de oscilação conjunta entre dois ou mais campos físicos. A partir daí, pode-se encontrar um conjunto de equações prognósticas e construir um modelo preditor.

Assim, o principal objetivo deste trabalho é desenvolver um modelo de previsão climatológica para a precipitação utilizando como variável preditora

a TSM, de forma que se possam prever, com antecedência de até três meses, eventos de precipitação abaixo ou acima da climatologia.

Pretende-se ainda testar o desempenho do modelo estatístico utilizando o método de validação cruzada.

2 REVISÃO DE LITERATURA

2.1 PREVISÃO CLIMÁTICA

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

As previsões climáticas são feitas para um período de, no mínimo, um mês, mas geralmente se estendem por períodos de três meses, a fim de suavizar a variabilidade intra-sazonal. Durante esses intervalos de tempo, anomalias de TSM, persistidas por vários meses, podem ter uma forte influência sobre as anomalias atmosféricas sazonais, especialmente nos trópicos. Técnicas, usando padrões de escala global de dados históricos de TSM, têm sido usadas com sucesso, desde 1986, nos serviços meteorológicos para produzir previsões experimentais da precipitação sazonal na Região Nordeste do Brasil, assim como em outras regiões. Nas últimas décadas, tem havido um interesse crescente na previsão sazonal, como já fora antecipado por Palmer e Anderson (1994).

Têm-se realizado previsões climáticas tanto com modelos dinâmicos (STERN e MIYAKODA 1995; KUMAR et al., 1996; BRANKOVIC e PALMER 1997) como com modelos estatísticos (BARNETT e PREISENDORFER 1987; BARNSTON, 1994; WARD e FOLLAND, 1991; VAUTARD et al., 1996). Os modelos dinâmicos são representações matemáticas dos processos físicos que visam simular a distribuição e evolução de umidade e energia sobre todo o globo. Para isso, eles resolvem aproximações numéricas das equações básicas do fluido

atmosférico, que representam a conservação do momentum, da massa, da energia e da água, com parametrizações físicas para processos como transferência radiativa, formação de nuvens e convecção, transporte turbulento e interação superfície-atmosfera. Por outro lado, os modelos estatísticos de previsão do clima estabelecem relações estatisticamente significativas entre variáveis meteorológicas locais em áreas continentais com preditores locais e remotos (preditores locais se referem a variáveis observadas na mesma localidade do preditando e podem incluir, por exemplo, o estado antecedente da variável a ser prevista, enquanto preditores remotos normalmente correspondem às variáveis meteorológicas em áreas oceânicas ou em outras áreas que não sejam do local da previsão).

Em modelos estatísticos de previsão climática, vários podem ser os preditores utilizados: a altitude geopotencial, a radiação de onda longa, a chuva da pré-estação no Nordeste, ventos na baixa e alta troposfera e, o mais comum, a TSM. Por ocuparem três quartos (71%) da superfície do globo e 81% da superfície no Hemisfério Sul, os oceanos exercem uma influência muito grande nas características e comportamento das variáveis meteorológicas nas relações Oceano-Atmosfera.

Diversos pesquisadores realizaram trabalhos sobre previsão climática para diversas regiões do globo, inclusive a América do Sul e o Brasil, apontando a TSM como um dos principais indicadores das variações de precipitação e temperatura, sendo, portanto, um importante preditor para essas variáveis.

Repelli e Alves (1996) apresentaram um modelo estatístico para prognosticar a variabilidade sazonal da precipitação na Região semi-árida do Nordeste do Brasil. O modelo foi baseado na técnica estatística chamada de análise de correlações canônicas (ACC) e usou o campo de Temperatura da Superfície do Mar (TSM) como preditor. A influência do campo de TSM, em janeiro, na precipitação de fevereiro a maio sobre a Região foi discutida. Em geral, o modelo foi capaz de prever a tendência da estação chuvosa dentro de três categorias (acima, abaixo ou em torno da média) para todos os casos considerados.

Ao estudar os impactos da TSM de sete sub-áreas, localizadas entre 90° W e 160° E, 5° N e 5° S, do Pacífico Tropical, sobre os padrões de precipitação sazonal (dezembro a maio) na América do Sul, em períodos de ocorrência dos fenômenos El Niño e La Niña, Ambrizzi et al. (2002) demonstraram, através de correlação linear simples, que as anomalias positivas de TSM nestas regiões influenciam positivamente e de modo significativo a precipitação no extremo norte da América do Sul ($r > 0,5$), no norte de Roraima (Brasil), noroeste da Guiana e nordeste da Venezuela, e, negativamente o Nordeste do Brasil ($r < -0,4$), destacando a região entre 90° W e 110° W, 5° N e 5° S, como a que apresentou os maiores valores de correlação entre a precipitação e a TSM.

Pesquisando as relações da precipitação pluvial sobre o Uruguai e a TSM durante o El Niño-Oscilação Sul (ENOS), Pisciottano et al. (1994), chegaram a resultados que indicam que durante eventos de El Niño, a precipitação tende a anomalias positivas de novembro a janeiro. Em contrapartida, em anos de alto Índice de Oscilação Sul, a anomalia tende a ser negativa de outubro a dezembro. Para a parte sul do Uruguai, as influências sobre a precipitação, durante eventos El Niño, são relativamente fracas.

Através da Análise de Correlações Canônicas, Barnston e Smith (1996) usaram a TSM para a previsão global da temperatura e precipitação em áreas afetadas pelo ENOS, como o norte e o leste da Austrália, partes da África, América do Norte e Sul. Foi verificada uma capacidade satisfatória de previsão do modelo para ambas as variáveis. Para a precipitação na América do Sul, os resultados mostraram que as áreas oceânicas de maior importância são as regiões dos ENOS, Índico e Atlântico Sul. Em condições de Atlântico Sul quente, Atlântico Norte frio (dipolo negativo) e Pacífico equatorial frio, o trimestre de janeiro a março mostrou relações com o aumento na precipitação no Nordeste do Brasil e Paraguai.

Grimm et al (1998) pesquisaram o impacto de eventos El Niño e La Niña sobre a precipitação pluvial no Sul do Brasil através de dados mensais de precipitação em 250 estações meteorológicas. Os resultados registram que toda a

Região Sul do Brasil, durante eventos de El Niño, apresenta consistentemente anomalias positivas de precipitação pluvial na primavera, com um pronunciado pico em novembro. Em oposição, durante eventos de La Niña, a primavera apresenta forte anomalia negativa de precipitação em toda a região, também com magnitude máxima em novembro.

Procurando detectar as zonas de TSM de todo Hemisfério Sul que possam influenciar na formação das precipitações no Sul do Brasil, Khan et al (1998) utilizaram taxas de precipitações pluviais mensais de 13 estações meteorológicas dos Estados do Rio Grande do Sul e Santa Catarina, no período de 1910-1995, dados fornecidos pelo 8º Distrito de Meteorologia (INMET); também utilizaram médias mensais de TSM que foram compostas de três diferentes fontes, fornecidas pelo INPE (Instituto Nacional de Pesquisas Espaciais). Os resultados obtidos apontam que a Zona da Corrente Circumpolar Antártico é a que causa o maior impacto para as precipitações pluviais no extremo Sul do Brasil. Outras zonas de correlação substantiva com a precipitação também foram evidenciadas fora da região do El Niño, como a Zona de Convergência do Índico Sul.

Correlações entre anomalias de precipitação e a TSM no Uruguai e Sul do Brasil foram investigadas por Diaz et al. (1998) através de séries mensais de precipitação pluvial de 40 estações meteorológicas, no período de 1917 a 1980, utilizando técnicas estatísticas multivariadas. Concluiu-se que a variabilidade dessa precipitação no local de estudo está ligada a anomalias de TSM dos Oceanos Pacífico e Atlântico. Os resultados confirmam a relação entre o ENOS e as anomalias de precipitação durante o final da primavera e início do verão, o final do outono e início do inverno. Foi também verificada a ligação entre as anomalias TSM do Oceano Atlântico com as anomalias de precipitação durante Outubro-Dezembro e Abril-Julho.

Montecinos et al. (2000) empregaram a TSM do Pacífico Tropical como a variável preditora para a previsão sazonal da precipitação na América do Sul Subtropical. No estudo foram utilizados totais mensais de precipitação pluvial de 106 estações meteorológicas localizadas nos seguintes países:

Argentina, Chile, Brasil, Uruguai e Paraguai, no período de 1946 – 1986. A técnica utilizada para avaliar a previsão da precipitação pluvial sazonal na América do Sul Subtropical foi a Análise de Correlações Canônicas, frequentemente usada por diversos autores. Os resultados apontam que a precipitação na América do Sul Subtropical tem um nível significativo de previsibilidade na maioria dos casos, usando-se como preditor a TSM do Pacífico Tropical. Isto é altamente indicativo de que a persistência de anomalias de TSM, no Pacífico Tropical, é a fonte principal de previsibilidade de chuva em escala sazonal.

Utilizando 25 anos de dados de TSM (1948-72) sobre o norte e sul do Atlântico, Moura e Shukla (1981) calcularam a correlação entre anomalias de TSM de Março-Abril-Maio de Fortaleza e Quixeramobim. Encontraram a ocorrência simultânea de anomalias de TSM quentes ao norte e anomalias de TSM frias ao sul do Atlântico equatorial relacionada a secas sobre o NEB.

Diniz (2002) realizou trabalhos envolvendo modelos de previsão climática como o que prevêem as temperaturas mínima e máxima em regiões homogêneas do Rio Grande do Sul, usando como variável preditora a TSM. Marques (2005) utilizou-se dos mesmos preditores para desenvolver um modelo estatístico de previsão climatológica da precipitação também para o Estado do Rio Grande do Sul. Seus resultados indicaram que a TSM do Oceano Pacífico, com defasagem de dois e três meses, apresenta maior influência na precipitação da região noroeste comparada à região sudeste do Estado e que a TSM do Oceano Atlântico, com defasagem de zero a um mês, apresenta relações mais evidentes na região sudeste que no noroeste do Estado. Para estas mesmas regiões, os meses de inverno apresentam os melhores ajustes nos modelos lineares derivados da TSM.

Desde meados de 1997 o Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) adaptou e operacionalizou o Sistema de Modelagem Estatística dos Oceanos – SIMOC (REPELLI E NOBRE, 2004) inicialmente desenvolvido para previsões das anomalias de TSM sobre a Bacia do Oceano Atlântico Tropical e posteriormente adaptado para prever anomalias sazonais de

precipitação (PEZZI et al, 2000). O SIMOC se baseia na técnica estatística de Análise de Correlações Canônicas (ACC) cujos preditores são os campos de TSM em áreas sobre o Pacífico e o Atlântico Tropical, e o preditando é a precipitação sobre a Região Sul do Brasil. Foi observado que o uso dos dois oceanos simultaneamente como preditores melhorou sensivelmente a precisão das previsões, quando comparados com os casos dos oceanos individuais.

Com o objetivo de apresentarem um estudo avaliando a possibilidade de utilizar informações de TSM sobre os oceanos Atlântico e Pacífico para desenvolver um modelo estocástico prognóstico de vazão no Rio Paraná, Cardoso et al. (2004) utilizaram médias mensais de TSM observadas e ajustadas pela Interpolação Ótima de Reynolds. Também foram usados dados mensais de vazão sobre o posto Posadas, no Rio Paraná. Consideraram o período de 1950 a 1988 para ajuste no modelo e o período de 1989 a 2001 para a avaliação do desempenho deste. Através da técnica de Análise de Componentes Principais (ACP), foram obtidos modos de variabilidade de TSM, que foram utilizados como preditores de vazão em um modelo de previsão. Aplicou-se, então, a análise de regressão linear múltipla e empregou-se a técnica dos mínimos quadrados para se obter a equação de regressão. Os resultados indicaram que a TSM sobre os Oceanos Atlântico e Pacífico são covariáveis da vazão no Rio Paraná, sendo evidenciadas as áreas do Oceano Pacífico Tropical e o oeste do Oceano Atlântico Sul. O modelo apresentou uma baixa destreza para defasagem superior a dois meses, mas o mesmo pode servir como mais um indicador de mudanças na vazão.

Utilizando desvios normalizados dos dados de 193 estações pluviométricas, Minuzzi e Sedyama (2004) analisaram a relação mensal e sazonal existentes entre o Padrão do Dipolo do Atlântico e a precipitação durante o período chuvoso da Região Sudeste do Brasil. Com os resultados, percebeu-se que a grande variabilidade espacial e temporal da precipitação na Região Sudeste está relacionada ao Padrão de dipolo positivo do Atlântico e que influências do dipolo durante a sua fase negativa são menos marcantes.

Muitos pesquisadores têm desenvolvido estudos no Nordeste Brasileiro com o objetivo de relacionar as anomalias de TSM com a distribuição da precipitação nesse local. Dentre eles, Menezes (1998) avaliou a importância da TSM dos Oceanos Atlântico e Pacífico no estudo prognóstico do período chuvoso do semi-árido nordestino, que compreende fevereiro, março, abril e maio. Foram utilizados totais mensais de precipitação de 113 postos pluviométricos, distribuídos na região, durante o período de 1970-1988 e dados de TSM para o mesmo período. Nesse trabalho foram utilizadas duas técnicas estatísticas: a ACP e também a regressão linear múltipla. O modelo de regressão mostrou um ajuste relativamente satisfatório, explicando cerca de 50% da variabilidade da precipitação pluvial no norte do Nordeste. O modelo mostrou-se sensível às variações das TSM do Atlântico e Pacífico Tropicais.

Foram analisadas por Uvo et al. (1998) as relações entre as anomalias de TSM do Oceano Pacífico e Atlântico Tropical com as precipitações mensais na estação chuvosa no Nordeste do Brasil, utilizando dados mensais de anomalias de precipitação pluvial, normalizadas pelo desvio padrão e dados de TSM sobre o Oceano Pacífico Tropical e Atlântico Tropical para o período 1946 a 1985. Concluíram que alguns meses apresentam maior correlação da precipitação com a TSM do que outros e os que apresentam uma correlação menor sempre mostram, de alguma forma, uma relação significativa com algum oceano, seja com o Pacífico ou com o Atlântico.

A relação entre a TSM e a temperatura no Nordeste Brasileiro foi estudada por Silva et al. (2006), com o propósito de estabelecer um modelo preditivo da temperatura do ar em função das anomalias de TSM e das coordenadas geográficas, capaz de reconstruir as séries temporais de temperatura. Foram utilizados dois modelos: um de regressão e outro de Análise de Componentes Principais, com dados mensais de temperatura obtidos de 69 estações no Nordeste do Brasil e dados de anomalias de TSM do Atlântico Norte e Sul, Trópicos Globais e Oceano Pacífico Tropical, no período de 1961 a 1990. O modelo preditivo apresentado pode ser usado com razoável nível de precisão para reconstrução de séries temporais de temperatura do ar. As anomalias de

TSM do Oceano Atlântico Norte e Sul e o Oceano Global Tropical foram melhores preditores da temperatura do ar no Nordeste que as anomalias de TSM do Oceano Pacífico Tropical.

Finalmente, utilizando-se de um modelo simples de correlações canônicas através da ferramenta *Climate Predictability Tool* (CPT), Costa (2007) mostrou resultados preliminares para a previsão de precipitações sobre o Estado do Ceará usando como variável preditora a anomalia de temperatura da superfície do mar (TSM), para a qual identificou um modo canônico claramente dominante, associando anomalias negativas da variável preditora sobre o Pacífico Equatorial e o Atlântico Tropical Norte e anomalias positivas de TSM sobre o Atlântico Tropical Sul com anomalias positivas de precipitação sobre todas as regiões pluviometricamente homogêneas no Estado do Ceará. A partir dos resultados, atestou que o modelo CPT pode ser operacionalizado como um modelo de previsão, assim como uma ferramenta de *downscaling* estatístico e correção de resultados de modelos dinâmicos globais e regionais, objetos do presente trabalho.

Apesar da influência do El Niño, a importância do Atlântico tropical em afetar a variabilidade climática regional tem sido reconhecida desde a década de 1970. Ward e Folland (1991) verificaram que as chuvas no Nordeste do Brasil em MAM são mais fortemente relacionadas com a TSM no mês anterior do Pacífico Tropical (negativamente) e Atlântico Tropical (positivamente para o Sul do Atlântico Tropical e negativamente para o Norte do Atlântico Tropical).

O posicionamento da zona de Convergência Inter-Tropical (ZCIT) está estreitamente relacionada à TSM. Mudanças relativamente pequenas na TSM do Atlântico tropical, em particular mudanças no gradiente meridional da TSM, podem mudar a posição da ZCIT, causando impactos diretamente no clima local. Portanto, as anomalias de TSM no Oceano Atlântico Tropical, através da variação média da posição da ZCIT, levam à variabilidade interanual da precipitação sobre o Norte e Nordeste do Brasil (HASTENRATH e LAMB, 1977; MARKHAM e MCLAIN, 1977; MOURA e SHUKLA, 1981; HASTENRATH, 1984).

Com isso, a previsibilidade de anomalias pluviométricas sazonais sobre o Nordeste se deve ao caráter predominante do efeito das TSM no comportamento dos fenômenos causadores de chuva sobre a região. Assim, a confiabilidade de uma previsão de anomalias sazonais de precipitação sobre o Nordeste será tanto maior quanto melhores forem as previsões das anomalias de TSM sobre os oceanos tropicais (WARD e FOLLAND, 1991).

2.2 MODELOS EMPÍRICOS

O uso de ferramentas estatísticas no prognóstico climático, em complemento às previsões fornecidas pelos modelos dinâmicos, apresenta-se como uma possibilidade bastante promissora, especialmente em regiões geográficas onde os modelos dinâmicos disponíveis ainda apresentam destreza insuficiente ou quando os recursos computacionais necessários para a implementação de modelos dinâmicos não se encontram disponíveis.

Os modelos empíricos não se baseiam nas equações físicas que regem o escoamento do fluido atmosférico, mas apenas estabelecem relações empíricas dentro de um conjunto de dados experimentais conhecidos. Eles ajustam um conjunto de equações simultâneas sobre as séries do passado de modo completamente experimental.

Apresentam as seguintes características:

- São modelos essencialmente descritivos;
- Enfatizam os relacionamentos entre variáveis;
- Requerem número reduzido de variáveis;
- Utilizam modelos matemáticos simples;
- São eficientes em predições;
- São sempre dependentes do tamanho da série de observações.

Para se ter informações acerca de um grupo de variáveis ou às vezes do conjunto total dos dados de uma região, é frequente recorrer-se à análise multivariada. Este conjunto de técnicas estatísticas é também utilizado para

reduzir ao máximo o número de variáveis envolvidas em um problema com uma pequena perda de informações. Seus principais objetivos são:

- Reduzir a dimensão de interpretação de uma matriz de dados;
- Investigar o comportamento espacial e temporal das variáveis consideradas;
- Obter grupos homogêneos das variáveis.

Na análise multivariada, existem várias técnicas de manejo de dados muito úteis para trabalhos em climatologia como, por exemplo, a Regressão Linear Múltipla (RLM), a Análise de Componentes Principais (ACP), a Análise de Correlação Canônica (ACC) entre outras.

2.2.1 REGRESSÃO MÚLTIPLA

Existem diversas maneiras de analisar dois conjuntos de dados. Dentre os modelos mais comuns de análise temos a regressão múltipla. Na regressão múltipla, uma variável é explicada por uma combinação linear de outras variáveis. A variável a ser explicada é chamada de dependente (y) e as variáveis explicativas são chamadas de independentes (x). A descrição mais detalhada da análise de regressão pode ser encontrada em Apte (1990).

A equação básica de uma regressão múltipla pode ser expressa por:

$$y_i = \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i = x_i' \beta + \varepsilon_i \quad i = 1, \dots, n$$

ou vetorialmente

$$y = X \beta + \varepsilon,$$

onde

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} x_{11} & \dots & x_{1p} \\ x_{21} & \dots & x_{2p} \\ \vdots & \dots & \vdots \\ x_{n1} & \dots & x_{np} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{pmatrix}.$$

com:

y_i é chamado de variável dependente ou variável da resposta.

x_i são chamados variáveis explanatórias ou variáveis independentes.

β é um vetor p-dimensional do parâmetro. Seus elementos são chamados.

ε_i é chamado o termo do erro, termo de distúrbio, ou ruído.

O modelo de regressão é chamado de simples quando a relação causal envolve apenas duas variáveis. O modelo de regressão é múltiplo quando o comportamento de y é explicado por mais de uma variável independente $x_1, x_2, \dots, x_n$. Os modelos acima (simples ou multivariados) simulam relacionamentos entre as variáveis do tipo linear (equação da reta ou do plano) ou não linear (equação exponencial, geométrica, etc.)

Na intenção de determinar a relação entre essas variáveis é possível realizar previsões futuras sobre fenômenos que as envolvem. Neste caso extrapolam-se para o futuro as relações de já observadas no passado entre as variáveis. Por exemplo: quão intensa é a precipitação do trimestre FMA (y) para as 8 Regiões Pluviometricamente Homogêneas no Estado do Ceará, quando se tem a TSM de Dezembro (x) em áreas sobre o Pacífico e o Atlântico Tropical.

Um exemplo de aplicação de um modelo de regressão linear múltipla para a previsão de precipitação sobre o Estado do Ceará foi proposto por Xavier et al. (2000).

2.2.2 ACP

A Análise de Componentes Principais (ACP) é uma técnica estatística que transforma um conjunto de p variáveis em um conjunto com número menor (k) de variáveis não correlacionadas.

Estas variáveis explicam uma parcela substancial das informações do conjunto original. As p variáveis $Y_1, Y_2, \dots, Y_p$, são denominadas componentes principais, de modo que Y_1 é aquela que explica a maior parcela da variabilidade total dos dados, Y_2 explica a segunda maior parcela e assim sucessivamente. (BARROSO 2003).

Os objetivos da Análise de Componentes Principais são:

- 1) a redução da dimensão dos dados originais;
- 2) facilitar a interpretação das análises realizadas.

Segundo Barroso (2003), esta análise resume o padrão de correlação entre as variáveis.

As componentes principais são combinações lineares das variáveis originais e podem ser interpretadas com base nas relações entre as variáveis originais e as componentes principais expressas por meio dos coeficientes das combinações lineares que levam às componentes principais. As correlações são medidas de cada contribuição individual de cada variável e não da contribuição multivariada das outras variáveis, mas os coeficientes são medidas das contribuições multivariadas. Normalmente utilizam-se das K primeiras componentes principais para uma análise, mas quanto menor for a parte desprezada, melhor o ajuste. Quando se tem de 80% até 90% do total da variância da amostra, para um grande conjunto de p variáveis, pode-se utilizar 1, 2 ou 3 componentes. Estas componentes, que são em número muito menor, podem representar as p variáveis originais com pouca perda de informação.

2.2.3 ACC

A Análise de Correlações Canônicas (ACC) é um procedimento estatístico multivariado que permite o exame da estrutura de relações existentes entre dois grupos ou conjuntos de variáveis (X e Y) (ABREU e VETTER, 1978).

A correlação canônica pode ser vista como uma extensão da regressão múltipla. Na correlação canônica existem duas ou mais variáveis dependentes.

Foi introduzida por Hotelling em 1936, sendo muito pouco usada no passado. Glahn (1968) foi quem primeiramente mostrou a aplicação de ACC em campos meteorológicos. Mais detalhes sobre a ACC é fornecido em Barnett e Preisendorfer (1987), Graham et al. (1987a, 1987b), Barnston e Ropelewski (1992), Barnston (1994), e suas referências. Breve tratamentos matemáticos são

encontrados no apêndice do Barnett e Preisendorfer (1987) e na seção 3b de Graham et al. (1987a).

A ACC, como o próprio nome diz, refere-se às correlações entre variáveis canônicas, ou seja, entre combinações lineares de variáveis, de tal modo que a correlação entre essas combinações é máxima. Com isso, não existe nenhuma outra combinação linear de variáveis cuja correlação seja maior que essa. Na correlação canônica não existe, a princípio, diferença entre variável independente e dependente; existem somente dois conjuntos de variáveis em que se busca a máxima correlação entre ambos.

Sejam dois grupos de variáveis X e Y , definidos como sendo:

$X' = [x_1, x_2, \dots, x_p]$ é o vetor de medidas de p características que constituem as variáveis independentes, e

$Y' = [y_1, y_2, \dots, y_q]$ é o vetor das medidas de q características que constituem as variáveis dependentes.

O problema estatístico consiste em estimar a máxima correlação entre combinações lineares de características das variáveis independentes e das variáveis dependentes, bem como estimar os respectivos coeficientes de ponderação das características em cada combinação linear. Sendo X_1 e Y_1 uma das combinações lineares das variáveis independentes e dependentes, respectivamente, tem-se:

$$X_1 = a_1x_1 + a_2x_2 + \dots + a_px_p \text{ e}$$

$$Y_1 = b_1y_1 + b_2y_2 + \dots + b_qy_q ,$$

em que:

$a' = [a_1, a_2, \dots, a_p]$ é o vetor $1 \times p$ de pesos das características de X e

$b' = [b_1, b_2, \dots, b_q]$ é o vetor $1 \times q$ de pesos das características de Y .

Uma maneira de expressar a correlação canônica pode ser determinando uma combinação linear entre x e y ,

$$U = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$$

$$V = b_1 y_1 + b_2 y_2 + \dots + b_n y_n$$

tal que a correlação $\text{Corr}(U, V)$ seja maximizada.

Considere X como sendo uma matriz $n \times p$ e Y como a matriz $n \times q$. Considere também a matriz $\text{Cov}(X, Y)$. Podemos separar esta matriz em quatro partes:

$$\text{Cov}(X, Y) = \begin{bmatrix} \sum_{p \times p}^{11} & \sum_{p \times q}^{12} \\ \sum_{q \times p}^{21} & \sum_{q \times q}^{22} \end{bmatrix}$$

As covariâncias entre variáveis de diferentes conjuntos, uma variável de X e outra de Y estarão contidas em (12) ou (21). Analisar as covariâncias em (12) ou (21) pode ser extremamente trabalhoso, ainda mais se p e q forem grandes. Porém, o principal objetivo da correlação canônica é resumir as associações entre X e Y em função de algumas poucas correlações escolhidas, ao invés das $p \times q$ correlações.

Combinação linear é uma forma simples de resumir um conjunto de variáveis, então seja:

$$U = a'X$$

$$V = b'Y$$

$$\text{Var}(U) = a' \text{Cov}(X) a = a' \sum_{11} a$$

$$\text{Var}(V) = b' \text{Cov}(Y) b = b' \sum_{22} b$$

$$\text{Cov}(U, V) = a' \text{Cov}(X, Y) b = a' \sum_{12} b$$

O que a correlação canônica procura é determinar os vetores a e b tais que:

$$\text{Corr}(U, V) = \frac{a' \sum_{12} b}{\sqrt{a' \sum_{11} a} \sqrt{b' \sum_{22} b}}$$

seja a maior possível.

Existirão mínimo $(p, q)^{-1}$ pares de variáveis canônicas independentes do par de correlação máxima, que irão expressar a variância total dos dois grupos de variáveis.

Para calcular este máximo, $\max_{a, b}$,

$$\text{Corr}(U, V) = \rho_1^*$$

restrita pela combinação linear (primeiro par de variáveis canônicas):

$$U_1 = \underbrace{e_1' \sum_{11}^{-1/2}}_{a_1'} X \text{ e } V_1 = \underbrace{f_1' \sum_{22}^{-1/2}}_{b_1'} Y$$

Neste caso, ρ_1^{*2} é o autovalor de $\sum_{11}^{-1/2} \sum_{12} \sum_{22}^{-1} \sum_{21} \sum_{11}^{-1/2}$ e e_1 é seu respectivo autovetor.

Supondo 'A' uma matriz quadrada $k \times k$ e 'v' um vetor $k \times 1$, pode-se mostrar que a equação

$$A\mathbf{v} = \lambda\mathbf{v}$$

tem k soluções, sendo λ um escalar. Cada solução é dada por um par formado por um escalar λ_i e um vetor $\mathbf{v}_i$. O escalar é chamado de autovalor de A e o vetor de autovetor de A .

A prova para o resultado acima pode ser encontrada em Johnson e Wichern(1992).

O k -ésimo par de variáveis canônicas pode ser descrito como:

$$U = e_k' \sum_{11}^{-1/2} X \text{ e } V = f_k' \sum_{22}^{-1/2} Y$$

maximizando,

$$\text{Corr}(U_k, V_k) = \rho_k^*$$

A técnica da análise de correlação canônica é muito usada em estudos exploratórios. Um pesquisador pode ter um conjunto grande de variáveis, mas pode estar interessado em estudar somente umas poucas combinações lineares de variáveis desse conjunto. Poderá, então, estudar aquelas combinações lineares cuja correlação é mais elevada.

A ACC é frequentemente utilizada como uma técnica de previsão para determinar a melhor combinação linear de dois conjuntos de dados (preditor, por exemplo, TSM, e o preditando, precipitação observada) que são altamente correlacionados. A ACC está no topo da hierarquia de modelagem de regressão (BARNETT e PREISENDORFER, 1987). Devido aos campos do preditor e do preditando conterem um grande número de variáveis altamente correlacionadas e poucas observações, recomenda-se uma pré-filtragem para um espaço menor (BARNSTON, 1994) usando funções ortogonais empíricas (análise de FOE, equivalente à ACP) (JACKSON, 1991). As séries de dados do preditor e do

preditando são primeiramente padronizadas, resultando em matrizes de correlação em que a ACP é executada, de forma a maximizar a correlação entre os dois índices. Desta forma, busca-se reter ao máximo a informação contida nas variáveis originais. A padronização garante que todos os pontos de grade e os índices de precipitação têm igual oportunidade de participar no processo de previsão (JACKSON, 1991; BARNSTON, 1994). A ACP é realizada separadamente para cada um dos campos do preditor e preditando.

O conjunto de pesos para os componentes dos preditores está relacionado com o conjunto análogo de pesos para os componentes do preditando. Estes conjuntos de ponderação, chamados padrões de carregamento, ilustram a associação entre os padrões do preditor e do preditando e, portanto, pode ser usado como orientação para os processos físicos subjacentes, bem como para fins descritivos.

A ACC também é usada eficazmente para descrição de dados, verificação de modelos numéricos e construção de modelos estatísticos de previsão, proporcionando o conhecimento de quais configurações tendem a ocorrer simultaneamente entre dois ou mais campos distintos e qual o grau de conexão entre eles.

Para o nosso objetivo, ela se resume em associar índices a cada um dos conjuntos de dados (X: variável explicativa e Y: resposta), definidos como combinações lineares dos valores em cada um dos conjuntos (Função Ortogonal Empírica - FOE), de forma a maximizar a correlação entre os dois índices. Desta forma, busca-se reter ao máximo a informação contida nas variáveis originais.

Neste trabalho, quando da construção do modelo empírico de previsão, ela será utilizada para identificar os campos de TSM (X) mais altamente correlacionados com padrões de anomalia de precipitação observada (Y). Ao ser usada como ferramenta de correção do *downscaling* dinâmico, a variável explicativa passa a ser a precipitação simulada pelo modelo dinâmico, sendo que a resposta continua a ser a precipitação observada.

Como uma ferramenta de previsão, o ACC tem várias vantagens sobre outras técnicas convencionais, tais como a regressão linear múltipla. Como um

modelo linear, no entanto, ACC não explica completamente as relações não-lineares.

Aplicações do método de ACC podem ser encontradas em outros trabalhos tais como Studzinski (1995), Diaz et al. (1998) Barnston e Smith (1996), Pezzi et al. (1998) e Repelli e Nobre (2000).

2.3 MODELOS DINÂMICOS

Os modelos de previsão numérica (modelos dinâmicos) são modelos computacionais desenvolvidos para simular o comportamento da atmosfera. Eles resolvem um sistema complexo de equações matemáticas baseadas em leis físicas de modo a prever o estado futuro da atmosfera partindo de condições iniciais específicas. Suas equações matemáticas são as formas mais completas e complexas das equações dinâmicas e termodinâmicas. Em adição, processos físicos, como a radiação, nuvens, precipitação e trocas à superfície são representadas no modelo utilizando esquemas de parametrizações.

O modelo dinâmico divide o planeta em várias camadas verticais que representam níveis da atmosfera, e divide a superfície do planeta numa malha de caixas horizontais separadas por linhas similares às linhas de latitude e longitude. Com isso o planeta é coberto por uma malha tri-dimensional.

Para fazer uma previsão, os modelos partem de um conjunto de condições iniciais. Então, suas equações são integradas no tempo para prever o estado da atmosfera num tempo futuro. Pequenos erros nas condições iniciais se amplificam a medida que a simulação avança no tempo, de modo que a diferença entre as condições atmosféricas prevista e as condições reais da atmosfera aumenta ao longo da previsão. Isto torna inviáveis previsões de tempo com mais do que duas semanas de antecedência, no entanto, as previsões climáticas seguem outro princípio: estatísticas de conjuntos (*ensembles*) de simulações forçadas por condições de contorno (como a TSM).

Os modelos de previsão numérica são classificados segundo a sua escala espacial em:

- *Global*: com resolução da ordem de várias dezenas a poucas centenas de quilômetros, visa identificar o comportamento geral da atmosfera sobre uma área extensa; estes modelos são apenas capazes de identificar fenômenos meteorológicos de grande escala (ditos sinóticos);
- *Regional ou local*: com resolução espacial típica de quilômetros a dezenas de quilômetros, visa identificar com maior detalhe o comportamento da atmosfera sobre uma região específica, sendo apropriados para identificar fenômenos meteorológicos de menor escala (ditos de mesoescala), como brisas, interação entre o escoamento atmosférico, topografia, etc.

2.3.1 MODELOS GLOBAIS

São modelos que se destinam a descrever movimentos atmosféricos sobre o planeta como um todo. São também denominados Modelo de Circulação Geral Atmosférico (MCGA) ou Modelo de Circulação Geral (MCG). Podem ser acoplados a modelos oceânicos e/ou a modelos de outros componentes do clima.

Os MCGs podem ser configurados para previsão de tempo ou para estudos climáticos. Para a previsão de tempo, o modelo é executado por um tempo físico de alguns dias (entre uma e duas semanas) e tem em vista representar a formação e a evolução de sistemas de tempo de grande escala da forma mais realista possível. Já para previsão climática, o modelo não se interessa em sistemas de tempo individuais, mas nas propriedades estatísticas da circulação atmosférica em escalas de tempo de meses ou mesmo anos.

Como variáveis de saída, eles fornecem, entre outras, temperatura do ar, velocidade do vento em várias direções, cobertura de nuvens, precipitação, umidade do solo, cobertura de neve, radiação solar, pressão e temperatura na superfície do mar.

A baixa resolução desses modelos deve-se ao poder computacional exigido para a execução dos mesmos, pois quanto maior a resolução espacial requerida, maior será o tempo de computação necessário.

2.3.2 MODELOS REGIONAIS

Os modelos regionais são aqueles cujo espaçamento de grade é menor do que aquele usado em modelos globais, o que permite a representação de fenômenos de mesoescala.

Normalmente, tais modelos se aplicam a estudos ou previsões de tempo em escala regional e local. Como sua grade não cobre toda a extensão da superfície terrestre (como os modelos globais), mas somente a atmosfera sobre uma dada área, são também chamados de modelos de área limitada.

Esses modelos permitem a representação de fenômenos que dificilmente são simulados em modelos globais, como a evolução de sistemas de tempo de mesoescala, a influência da topografia, etc. Isso se dá graças à melhor resolução dos modelos de mesoescala, cujo espaçamento de grade varia entre menos de uma a algumas dezenas de quilômetros.

Do ponto de vista de suas aplicações, recentemente os modelos de área limitada têm não somente servido para fins de previsão de tempo (como os modelos ETA, RAMS, MM5, RSM, etc., em diversos centros de meteorologia do Brasil e do exterior), mas também têm encontrado uso em estudos climáticos, através da técnica conhecida como *downscaling* dinâmico.

2.4 DOWNSCALING

Como ferramentas para previsões sazonais, os MCGs conseguem representar bem apenas os padrões climáticos de grande escala. Entretanto, devido à baixa resolução desses modelos (resolução espacial da ordem de 100 a

200 km), os mesmos ainda não proporcionam, por exemplo, um prognóstico mais detalhado da distribuição espacial da chuva na escala intrasazonal (controlada por fatores com escala tipicamente bem menor), além da própria variabilidade temporal dessa distribuição de precipitação.

Assim, com o objetivo de solucionar o problema da diferença entre a escala para a qual são realizadas as previsões por MCGs e a escala da região onde se deseja empregá-las, tem-se usado técnicas de transferência da informação meteorológica para as escalas menores (técnicas de *downscaling*) desenvolvidas pela necessidade de se ter uma informação mais detalhada no tempo e no espaço dos produtos de previsão numéricas de tempo e clima.

Basicamente o *downscaling* trata do uso dos resultados de uma modelagem de grande escala para alimentar um modelo de menor escala, seja um modelo numérico regional de área limitada, com uma resolução da ordem de dezenas de quilômetros (DICKINSON et al., 1989), seja um modelo empírico.

As técnicas de *downscaling* podem ser classificadas em duas categorias: *downscaling* dinâmico e *downscaling* estatístico.

A primeira delas envolve o uso de modelos numéricos de escalas global e regional, e requer informação climática (de superfície) detalhada, assim como uma alta disponibilidade computacional. Por outro lado, as empíricas empregam métodos estatísticos, sendo que os mais empregados são os métodos de regressão (lineares e não lineares), geradores de tempo (análises de componentes principais, análises de correlação canônica, redes neurais artificiais) e os estocásticos (teoria fuzzy – conjuntos nebulosos, algoritmos genéticos, etc).

Na técnica de *downscaling* dinâmico, utiliza-se um modelo meteorológico de resolução mais alta (regional), com espaçamento de grade da ordem de algumas dezenas de quilômetros, alimentado nas fronteiras pelas condições produzidas pelo modelo global (ou até por dados observados). Por representar com maior detalhe os processos que agem regionalmente, acredita-se que os modelos regionais podem produzir previsões de tempo e clima com maior precisão.

Como o principal argumento contra a prática *downscaling* dinâmico é o custo computacional necessário para executar as simulações históricas das previsões, a existência de uma boa rede de observação é crucial para a validação dos dados a partir de modelos de alta resolução.

Durante muito tempo, diversas abordagens menos complexas e menos onerosas como o *downscaling* estatístico davam resultados sistematicamente comparáveis ou mesmo superiores aos do *downscaling* dinâmico. No entanto, os modelos dinâmicos ganharam espaço e as técnicas estatísticas, ao invés de se contraporem a estes, têm passado a cumprir um papel complementar, seja como modelos de previsão autônomos, seja como corretores dos modelos dinâmicos.

A vantagem do uso do *downscaling* estatístico é que ele oferece uma solução imediata a um baixo custo computacional e, conseqüentemente, pode ser rapidamente implementado em centros operacionais.

2.5 ERROS E TÉCNICAS ESTATÍSTICAS DE CORREÇÃO

Os modelos climáticos regionais são ferramentas importantes tanto para o melhor entendimento dos fenômenos meteorológicos, como no auxílio da previsão de tempo e clima. Porém apresentam erros que podem ou não ser sistemáticos e que, eventualmente, podem ser relativamente altos. Esses erros se dão não somente pelo fato de que os modelos numéricos não conseguem reproduzir a grande complexidade dos processos que compõem a evolução das condições atmosféricas, mas também são inerentes à falta de representatividade espacial e temporal dos dados observados que os alimentam. O conhecimento do desempenho dos modelos, juntamente com seus erros sistemáticos, é de suma importância para que os meteorologistas possam elaborar previsões de tempo mais confiáveis.

Em geral, os modelos climáticos regionais possuem uma boa capacidade de prever anomalias ao redor de um valor médio, porém apresentam erros relativamente altos, que podem e devem ser removidos para que estes

resultados possam ser empregados de maneira apropriada na previsão climática (WANG et al., 2004; WOOD et al., 2002; LEUNG et al., 1999).

Existem dois tipos de erros atribuídos às medições: os erros estatísticos e os erros sistemáticos. O erro estatístico (ou aleatório) está relacionado com a precisão do aparelho e com as condições externas aleatórias das medidas. São intervenções caóticas, imprevisíveis, devido a numerosas causas inconstantes, totalmente ignoradas ou mal conhecidas; elas comprometem as medições desviando-as igualmente para valores maiores ou menores que o valor verdadeiro. No entanto, os erros estatísticos podem ser reduzidos, eliminando ou reduzindo os fatores aleatórios que interferem no processo de medição. Por sua vez, o erro sistemático é sempre o mesmo nos resultados, ou seja, na presença de erros sistemáticos a diferença entre o valor medido e o valor real é sempre a mesma. São interferências que desviam as medidas do valor verdadeiro sempre no mesmo sentido, produzindo medições sempre maiores ou sempre menores que o valor verdadeiro. Os erros sistemáticos podem ser eliminados ou minimizados através de alterações no procedimento experimental. Caso não seja possível eliminar os erros sistemáticos, deve-se realizar uma estimativa da sua interferência na medida do valor real. Possíveis causas para ocorrência de erros sistemáticos nos modelos são mencionadas por Steyn e Mckendry (1988).

Os tipos de erros mais comuns na precipitação prevista se referem à distribuição espacial e temporal dos eventos e a intensidade da precipitação.

Diversos erros podem corrigidos por meio de métodos estatísticos que tentam realizar uma análise dos erros com base em uma série histórica de dados observados e modelados. Steed, R. C. e Mass (2004) mostraram e compararam vários métodos para a remoção de erros de viés para modelos de previsão climática. Cada um destes métodos leva a definição estatística de viés como sendo um erro médio de um parâmetro. Os métodos comparados se utilizaram de diferentes técnicas de aplicação de remoção do viés. A descrição de um procedimento para a determinação e correção de erros de fase e intensidade com o objetivo de melhorar o desempenho dos modelos de previsão climática é vista

com detalhes em Lee e Wilson (2007), e sua aplicabilidade é mostrada por Lee e Zawadzki (2006). Stephenson (2000) investigou diversas maneiras de quantificar a habilidade das previsões de eventos climáticos usando *downscalings* diferentes.

Muitos outros métodos de correção têm sido propostos para sistemas de previsão climáticos diferentes (por exemplo, Zeng et al., 1994; Feddersen et al., 1999; Zhao et al., 1999; Wang et al., 2000; Ward e Navarra, 1997; Mason et al., 1999; Wang et al., 2004; Wodd et al., 2002), e tem sido verificado que tais métodos são eficientes para melhorar a habilidade da previsão, principalmente com base em uma série histórica longa de dados observados e modelados.

A análise dos erros geralmente se baseia no *Model Output Statistics – MOS* (GLAHN e LOWRY, 1972). O MOS é uma técnica estatística que tem o objetivo de identificar qual é a porção da variabilidade do modelo que está relacionada à variabilidade da atmosfera real e estabelecer uma inter-relação entre estes dois universos de modo a aumentar a precisão do modelo. A idéia básica por trás do MOS para previsão sazonal é recalibrar os resultados do modelo a partir de inter-relações estatísticas obtidas através de análise multivariada entre os resultados de uma rodada do modelo e séries históricas de dados observacionais.

No MOS, desenvolvem-se relações entre previsões do modelo numérico como preditores e os dados de observações (temperatura máxima e mínima, precipitação e outros) como preditantes. Como as equações do MOS são obtidas com a utilização das previsões do modelo numérico como preditores, os erros sistemáticos que são inerentes nos modelos de previsões numéricas, podem ser levados em conta.

A primeira utilização desta técnica foi para correção de variáveis como temperatura máxima e vento em superfície (GLAHN e LOWRY, 1972). Desde então os centros operacionais de previsão vem desenvolvendo e utilizando relações estatísticas entre os resultados do modelo e dados observados para otimizar o *skill* na previsão numérica de tempo e clima.

Métodos de correção estatística podem compensar as deficiências do modelo através do preenchimento de detalhes de grande escala com padrões

encontrados na natureza. Tais métodos têm sido utilizados para corrigir as anomalias de precipitação do modelo simulado (SMITH e LIVEZEY, 1999; FEDDERSEN et al., 1999) e as previsões sazonais (MO e STRAUS, 2002).

Os erros do modelo podem ser corrigidos através da realização da ACC. Neste caso, o modelo previsto é o preditor e o observado é o preditando. Os modos da ACC em seguida, são determinados de acordo com a covariância entre as precipitações observadas e as precipitações previstas, resultado do modelo numérico (LIVEZEY e SMITH, 1999).

2.6 PREVISÃO HÍBRIDA

A principal base de qualquer modelo empírico de previsão climática sazonal para os trópicos está apoiada no princípio de que, nessa região, as condições de contorno na superfície (principalmente a temperatura da superfície do mar - TSM) condicionam as características estatísticas significativas do clima. Portanto, estas condições podem, em princípio, ser usadas como preditores climáticos. Montar um modelo linear de previsão consiste em construir uma equação de regressão linear múltipla com fortes relações estatísticas. Daí a importância em escolher um bom conjunto de preditores. Conforme discutido anteriormente, a TSM apresenta uma relação significativa com a precipitação sobre a região NE do Brasil. O fato de a TSM influenciar a precipitação com uma determinada defasagem torna possível utilizá-la como um preditor da precipitação.

É importante salientar, no entanto, que diversos aspectos da não-linearidade do clima não podem ser representados por modelos estatísticos simples, particularmente aqueles baseados em relações lineares, como os que utilizam regressões lineares múltiplas, análise de componentes principais ou análise de correlações canônicas.

Nesse sentido, o uso da técnica de *downscaling* dinâmico é capaz de produzir previsões sazonais com elevado grau de acerto. Sun et al. (2005)

mostram que o *Regional Spectral Modeling* (RSM) produz previsões sazonais com elevados índices de destreza, assim como o modelo RAMS, que mostra um bom desempenho na previsão de precipitação de ocorrência de chuva no NE do Brasil (segundo Alves et al (2007) o modelo RAMS apresentou valores relativamente altos de *Skill* em áreas do Ceará e Rio Grande do Norte com valores entre 0,5 e 0,6 com pequenas áreas excedendo acima de 0,7). É importante ressaltar, no entanto, que, em determinadas regiões as previsões dinâmicas superam claramente as técnicas estatísticas, ao passo que, em outras, estas últimas se mostram superiores.

Estes fatos, o de que a TSM é um bom preditor para a precipitação sobre o NEB, o de que os modelos dinâmicos produzem previsões com alta destreza sobre a região e o de que modelos empíricos e dinâmicos podem ter habilidades de previsão complementares, serviram de base para o desenvolvimento de um modelo “híbrido” neste trabalho. O comportamento ideal de um modelo híbrido seria um desempenho melhor do que os dois modelos originais.

A idéia de combinar o conhecimento estatístico com a modelagem dinâmica não é nova: A formulação de um modelo híbrido simples foi proposto por Madden (1981).

Uma experiência de combinação de modelos estatísticos (obtidos da análise estatística de dados observados) com modelos dinâmicos (obtidos das leis físicas) surgiu em resposta ao papel fundamental desempenhado pelo oceano no fenômeno ENSO. Tal experiência constituiu os chamados modelos acoplados híbridos (*hybrid coupled models*- HCM), um termo aparentemente proposto por Neelin (1989, 1990). Tais modelos são compostos de um modelo de circulação geral oceânico (MCGO) acoplado a um modelo atmosférico mais simples, de complexidade muito aquém de um modelo de circulação geral da atmosfera (MCGA). Um exemplo de tal modelo pode ser encontrado em Latif e Flügel (1991). Em Chang et al. (2001) e em Barnett et al. (1993) o HCM provou ser uma ferramenta extremamente valiosa na modelagem nos estudos de interação oceano-atmosfera tropical.

Sarda e Plaut (1996) propuseram dois esquemas de previsão híbridos onde conhecimentos de previsão estatísticos e dinâmicos são combinados. Coelho et al, (2006) estudou a combinação de modelos de previsão dinâmicos com modelos de previsão estatísticos com o objetivo de analisar se há ou não melhoras na habilidade da previsão na África do Sul.

Neste estudo, a simulação usando o modelo híbrido representa uma combinação da TSM com a saída do modelo numérico como preditores, usando a precipitação observada como preditando para produzir previsões baseadas em ambas as técnicas e, com isso tirar proveito dos aspectos positivos de ambos os modelos na tentativa de superar algumas das limitações das duas abordagens individualmente.

3 MATERIAL E MÉTODOS

3.1 DADOS OBSERVADOS

3.1.1 TSM

Foram utilizados neste trabalho os dados referentes à Temperatura da Superfície do Mar (TSM) dos Oceanos Atlântico e Pacífico, numa grade de $1^\circ \times 1^\circ$, obtidos do *National Center for Environmental Prediction* (NCEP/NOAA) que utiliza interpolação ótima para combinar informações do *Comprehensive Ocean-Atmosphere Data Set* (COADS), da rede de bóias sobre os oceanos e de satélites no período de 1981 a 2008, para os meses de Outubro, Novembro, Dezembro, Janeiro e Fevereiro.

A série de dados de TSM para uso no CPT está disponível através da biblioteca de dados do IRI podendo ser baixados a partir do site: <http://iridl.ldeo.columbia.edu>. Mais informações sobre essa série de dados podem ser encontrados em Reynolds (1988), Reynolds e Smith (1994) e Reynolds et al. (2002).

Podem-se destacar em particular algumas regiões do Oceano Pacífico tropical que foram representativas na descrição dos domínios ótimos de TSM como preditor para a precipitação no Estado do Ceará. São elas:

*Niño 1+2 (0° - 10° S; 80° - 90° W): região que tipicamente aquece primeiro quando o El Niño se desenvolve, localizada próxima à costa oeste da América do Sul;

*Niño 3 (5° S- 5° N; 150° W- 90° W): região do Pacífico Tropical que tem maior variabilidade na temperatura da superfície do mar sobre a escala temporal do El Niño;

*Niño 3.4 (5°S-5°N; 170°W-120°W): região de grande variabilidade nas escalas de tempo do El Niño e mais próxima (em relação à região do Niño 3) da área onde as mudanças na TSM locais são importantes para deslocar a região de precipitação localizada no Pacífico Oeste.

*Niño 4 (5°S-5°N; 160°E-150°W): região onde as mudanças na TSM conduzem a valores de aproximadamente 27,5°C, o que é visto como um importante limiar para a produção de convecção profunda.

Tais regiões são destacadas na Figura 3.1.

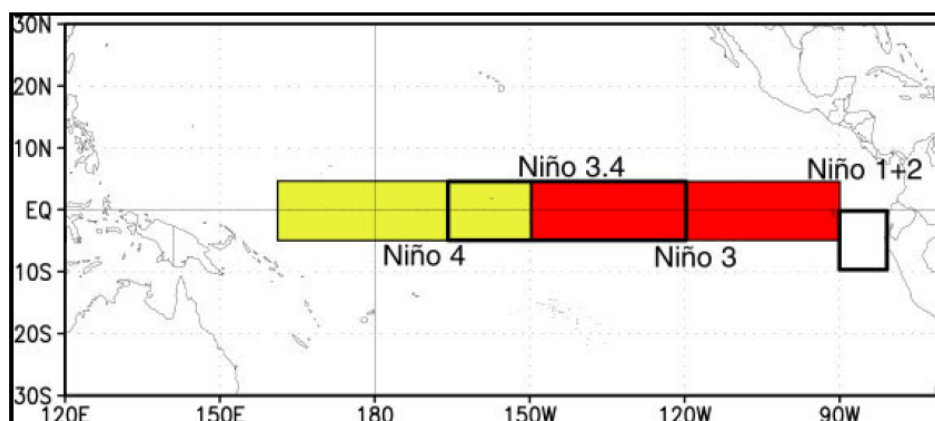


Figura 3.1 Regiões do Oceano Pacífico onde é feito o monitoramento da TSM.

Fonte: Golden Gate Weather Services (2008).

O ENOS e o Dipolo do Atlântico Tropical interferem diretamente no posicionamento da Zona de Convergência Intertropical (ZCIT) que por sua vez altera a distribuição pluviométrica sobre a bacia do Atlântico e norte da América do Sul, em especial a Região Nordeste do Brasil (HASTENRATH, 1984; NOBRE e SHUKLA, 1996).

O padrão de Dipolo gera anomalias de TSM com sinais opostos ao norte e ao sul do equador sobre o Atlântico Tropical (HASTENRATH e HELLER, 1977; MOURA e SHUKLA, 1981). A ocorrência de gradientes meridionais de anomalias de TSM por sua vez influencia a posição latitudinal da ZCIT, induzindo mudanças na distribuição sazonal da precipitação pluviométrica sobre o Atlântico Equatorial, da parte norte do Nordeste do Brasil, até a parte central da Amazônia (NOBRE e SHUKLA, 1996).



Figura 3.2 – Regiões de influencia da TSM na precipitação no Norte e Nordeste do Brasil.
Fonte: Silva et al (2007)

3.1.2 PRECIPITAÇÃO

O segundo conjunto de dados é referente aos totais de chuvas dos trimestres Novembro-Dezembro-Janeiro (NDJ), Dezembro-Janeiro-Fevereiro (DJF), Janeiro-Fevereiro-Março (JFM), Fevereiro-Março-Abril (FMA), Março-Abril-Maio (MAM) e Abril-Maio-Junho (AMJ), registrados em 36 anos de observação (1970-2005), para 8 Regiões Pluviometricamente Homogêneas e 11 Unidades de Planejamento (Bacias) localizadas no Estado do Ceará, cedidos pela Fundação Cearense de Meteorologia e Recursos Hídricos (FUNCEME). As médias sobre as áreas foram obtidas a partir das observações pontuais em pluviômetros usando polígonos de Thiessen.

A Figura 3.3 mostra as áreas das 8 Regiões Pluviometricamente Homogêneas e das 11 Unidades de Planejamento (Bacias), respectivamente, localizadas no Estado do Ceará.

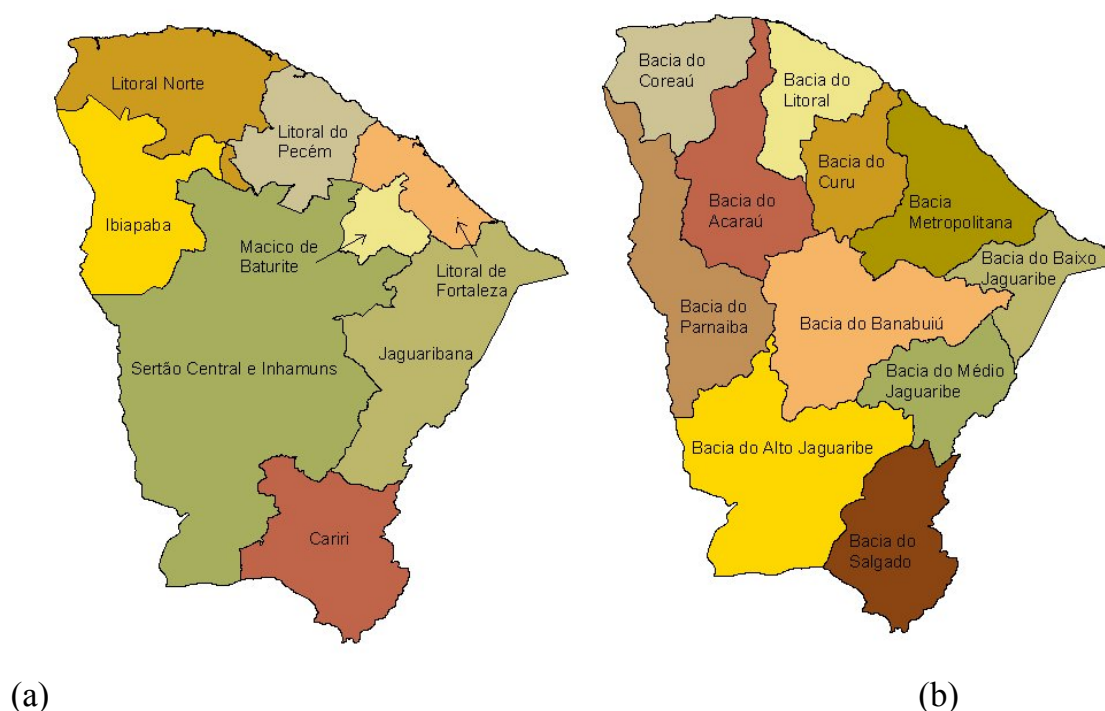


Figura 3.3 (a) Regiões Pluviometricamente Homogêneas e (b) Unidades de Planejamento (bacias). Fonte: FUNCEME

Um segundo conjunto de dados de precipitação também foi utilizado como base para aferição dos resultados da previsão do modelo dinâmico, bem como para a produção da previsão híbrida, pois o primeiro conjunto de dados de precipitação fornecido pela FUNCEME tem apenas registros referentes ao Estado do Ceará. Estes dados são provenientes da *university of EAST Anglia climate research unit* (CRU), referentes à climatologia de precipitação de 1971-2000, interpolada para uma grade com resolução de $0,5^\circ$ conforme descrito por New et al (1999) e New et al (2001).

3.2 DOWNSCALING DINÂMICO DA FUNCEME

No sistema de previsão da FUNCEME, o RSM 97 realiza o *downscaling* do modelo global ECHAM 4.5 para todo o ano e o RAMS 4.4 (*Regional Atmospheric Modeling System*) para o primeiro semestre. 20 membros do ECHAM 4.5 (10 com cada cenário de TSM, isto é, prevista e persistida),

oriundos do IRI, são utilizados para alimentar os dois modelos regionais. A Figura 3.4 mostra a versão do Sistema de Previsão Climática Regional da FUNCEME, implantado em 2005.

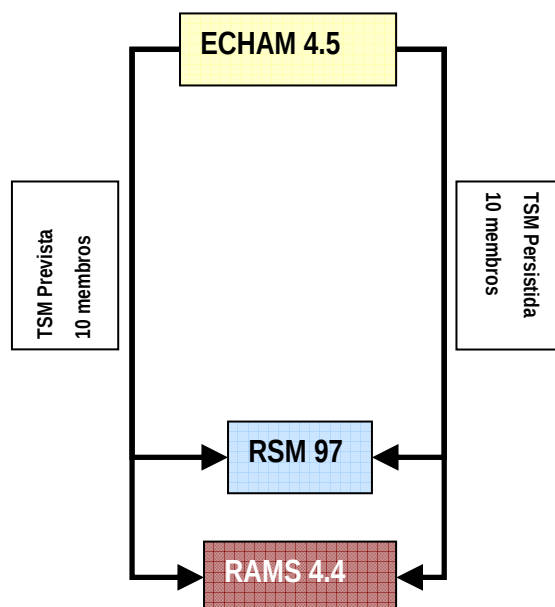


Figura 3.4 – Diagrama esquemático do Sistema de Previsão Climática Regional da FUNCEME, implantado em 2005.

3.2.1 RSM

O modelo espectral regional (RSM) desenvolvido no Centro de Modelagem de Meio Ambiente do *NCEP (National Center for Environmental Prediction)* por Juang e Kanamitsu (1994), resolve equações primitivas para o fluido atmosférico usando uma aproximação hidrostática (assume-se que há equilíbrio entre a força de gravidade e a componente vertical do gradiente de pressão) para a equação de momentum vertical, com as perturbações (desvios em relação ao escoamento de grande escala) como variáveis prognósticas. São empregadas as coordenadas do tipo *sigma-p* para representar o escoamento sobre os padrões topográficos. Parametrizações de convecção (versão simplificada de Arakawa e Schubert, 1974), radiação (de onda longa e onda curta com variação diurna), solo (duas camadas, com a espécie de solo especificada pelo usuário) e vegetação (fração de cobertura vegetal especificada pelo usuário) são utilizadas.

Todos os detalhes acerca da configuração do modelo RSM para o Nordeste brasileiro são fornecidos por Sun et al. (2005).

É um modelo bastante simplificado, com fortes limitações no que se refere à representação de determinados processos físicos, no entanto possui, pelo menos em grades menos refinadas, um ótimo desempenho computacional.

Na FUNCEME, o modelo RSM é utilizado para previsão climática com um *ensemble* de 20 membros, alimentados pelos dados do modelo ECHAM 4.5, sendo 10 forçados com Temperatura da Superfície do Mar (TSM) prevista (modelo dinâmico na Bacia do Pacífico e empírico na Bacia do Atlântico) e 10 com anomalia de TSM persistida. A grade do modelo inclui o Nordeste Brasileiro e parcela significativa do Atlântico Intertropical, incluindo a região da ZCIT (Figura 3.5).



Figura 3.5 – Grade do RSM-97/FUNCEME

3.2.2. RAMS

Os modelos atmosféricos de mesoescala utilizam técnicas de simulação numérica, com a finalidade de estudar o comportamento da atmosfera, e permitem também a análise combinatória de modificação de poucos parâmetros para, a partir de variáveis meteorológicas geradas pelo modelo numérico, identificar a contribuição separada de cada fator ou a participação conjunta de vários fatores ao mesmo tempo, numa mesma situação (PIELKE, 1984).

O RAMS (*Regional Atmospheric Modelling System*) é uma ferramenta de simulação numérica de multiescala e multipropósitos para estudos atmosféricos. Historicamente, diferentes rotinas numéricas, desde a escala de nuvens até a mesoescala, vêm sendo usadas basicamente para fins de pesquisa em ciências atmosféricas e para variadas aplicações. A versão básica e inicial do modelo encontra-se descrita em Pielke et al. (1992). Nele podem ser encontradas a filosofia básica, a estrutura deste modelo e as suas aplicações mais frequentes, incluindo as referentes à formação de nuvens, chuva e granizo de todos os tipos de sistemas atmosféricos, os efeitos de topografia sobre o movimento atmosférico e nebulosidade. Previsões regionais de tempo têm no RAMS uma ferramenta poderosa, essencial para os avisos de mau tempo, inclusive tempestades severas e enchentes.

Na FUNCEME, o modelo RAMS é utilizado para previsão climática com um *ensemble* de 20 membros, alimentados pelos dados do modelo ECHAM 4.5, do IRI (*International Research Institute For Climate and Society*) forçado por TSM prevista (10 membros) e anomalia de TSM persistida (10 membros). Foi adotada uma grade polar estereográfica de 80 pontos nas duas direções horizontais (latitudinal e longitudinal), com espaçamento de 40 km. A mesma foi centrada em 10° S, 40° W, possibilitando incluir todo o Nordeste Brasileiro (Figura 3.6).

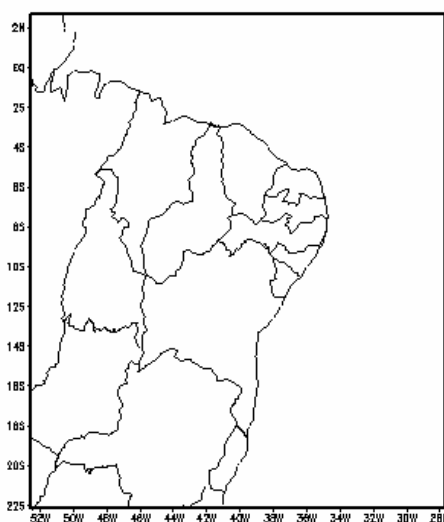


Figura 3.6 – Grade horizontal utilizada no modelo RAMS

Na vertical, para que fosse proporcionada uma boa representação do transporte convectivo, bem como dos processos de troca na camada limite, usou-se uma grade vertical de 41 níveis, com espaçamento variável entre 80m (próximo à superfície) e 1km (em altos níveis).

O modelo foi inicializado a partir de dados do ECHAM 4.5, tanto para as simulações de climatologia (anos anteriores), quanto para a previsão (com a diferença que, neste último caso, o modelo é forçado tanto com TSM prevista, quanto com anomalia de TSM persistida).

No que diz respeito às parametrizações físicas, é usado o esquema de Harrington (1997) para radiação em ondas curtas e longas chamado a cada 20 minutos, turbulência resolvida com o esquema de Mellor & Yamada (1982), microfísica de nuvens parametrizada conforme o esquema proposto por Walko et al. (1995) e interação com superfície resolvida pelo submodelo LEAF2, descrito por Walko et al. (2000). Na parametrização de convecção, utiliza-se o esquema de Kuo (1974) modificado.

Neste trabalho, nos limitamos a utilizar, para fins de construção da metodologia de correção dos resultados de modelos dinâmicos e da previsão híbrida, empírica e dinâmica, ao uso dos dados gerados pelo modelo RAMS.

3.3 CPT

O uso de ferramentas estatísticas no prognóstico climático, complementando as previsões fornecidas pelos modelos numéricos, apresenta-se como uma possibilidade bastante propícia, notadamente em regiões geográficas onde os modelos numéricos disponíveis ainda apresentam destreza insuficiente. Hoje, existe a facilidade de se gerar previsões estatísticas pela utilização do programa *Climate Prediction Tool* (CPT), desenvolvido e disponibilizado pelo *International Research Institute for Climate and Society* (IRI, 2006). Trata-se de um aplicativo *Windows* para construção de modelos de previsão climática

sazonal, que possibilita as opções de se realizar tanto a análise por correlação canônica (ACC) quanto à regressão por componentes principais (RCP).

Uma maneira de analisar a relação entre dois conjuntos de dados é a regressão múltipla. Na regressão múltipla, uma variável é explicada por uma combinação linear de outras variáveis. A variável a ser explicada é chamada de preditando e as variáveis explicativas são chamadas preditoras. Na técnica de ACC, por outro lado, busca-se correlacionar diretamente dois conjuntos de variáveis. A idéia consiste em se desenvolver uma combinação linear em cada um dos conjuntos de variáveis, de forma que a correlação entre os dois conjuntos seja maximizada. O CPT facilita bastante a realização da ACC entre dois conjuntos de dados associados, cada um deles, a uma grade específica, normalmente cobrindo diferentes regiões geográficas do globo.

3.4 ESTRATÉGIAS DE USO DA ACC

3.4.1 PREVISÃO EMPÍRICA

Totais de precipitação para os trimestres NDJ a AMJ sobre o Ceará, tratados como preditandos (Y), são previstos utilizando a técnica de ACC. Como preditores (X) são usados as médias mensais de TSM em regiões selecionadas do Atlântico e Pacífico com uma defasagem de dois meses. Em outras palavras, o modelo utilizado no presente trabalho teve como base o campo de anomalias da TSM no mês $n - 2$ como preditor, e o campo de anomalias de precipitação para o trimestre de n a $n + 2$ como preditando (por exemplo, a TSM de dezembro foi utilizada na previsão da precipitação de fevereiro a abril).

Ambos os campos foram pré-filtrados com Funções Ortogonais Empíricas (FOE), (HOREL, 1981), para se eliminar os ruídos contidos nos dados originais. Na pré-filtragem das variáveis X e Y calculam-se as FOE dessas variáveis separadamente. O modelo é estabelecido usando um número de

autovetores que retêm em torno de 80% da variância original de cada variável. Essa pré-filtragem faz com que a ACC focalize-se nos modos de variabilidade dominantes dentro de cada conjunto de dados de entrada, e em princípio os ruídos de pequena escala são filtrados.

Posteriormente constrói-se uma matriz de correlação cruzada entre as séries de componentes principais de X e Y. As dimensões dessa matriz são então reduzidas ao número de modos retidos pelo preditor e predictando. Da transposta desta matriz, são obtidos os autovetores e autovalores canônicos para o X e Y.

A função temporal canônica do preditor é encontrada através de combinações lineares entre os autovetores canônicos e a série de componentes principais do preditor (utilizada na construção da matriz) para cada modo. O número de modos utilizados neste caso é o menor número de modos do preditor e predictando, fornecidos pela FOE. A equação de regressão expressando os modos canônicos é derivada em termos de variáveis originais. Assim, define-se um coeficiente de regressão para converter a função temporal canônica do preditor, em uma função temporal canônica do predictando.

Finalmente, uma equação preditiva que relaciona preditor e predictando é construída e a previsão pode ser realizada. A Figura 3.7 ilustra esquematicamente as diversas etapas para previsão da precipitação.

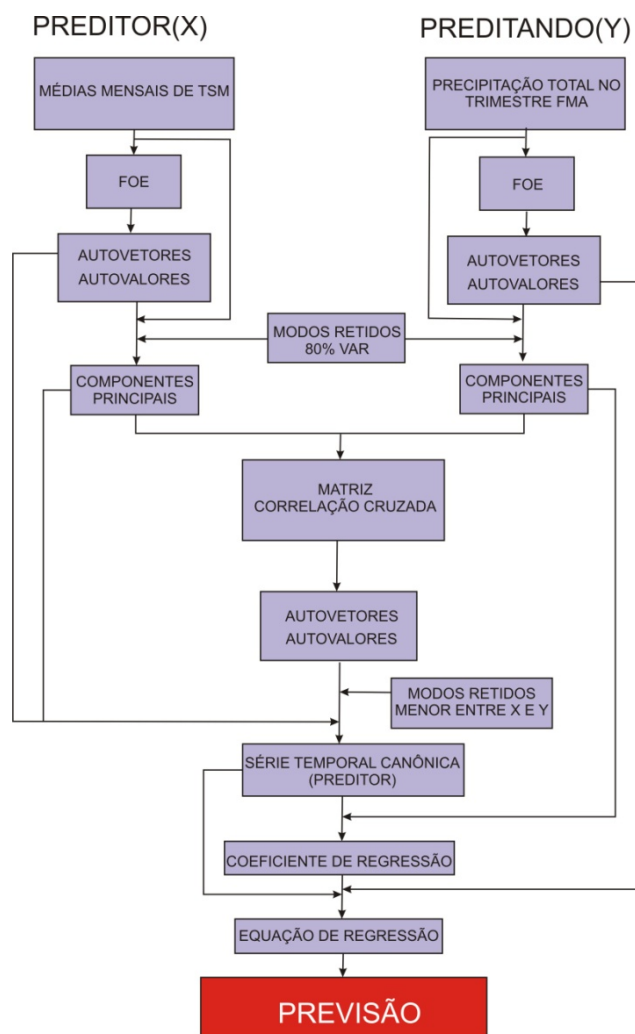


Figura 3.7 - Esquema ilustrativo das etapas utilizadas para a previsão da precipitação a partir da temperatura da superfície do mar, utilizando ACC.

3.4.2 CORREÇÃO ESTATÍSTICA DOS RESULTADOS DE DOWNSCALING DINÂMICO

Na previsão climática por modelos dinâmicos, os modelos atmosféricos globais dependem de uma boa previsão de TSM, pois os erros e as incertezas nessa previsão se propagam para os modelos regionais, por meio do *downscaling* dinâmico.

É através do *downscaling* que as informações do modelo global são detalhadas com bases físicas para uma maior resolução, favorecendo-se da

capacidade dos modelos de mesoescala em representarem a interação do escoamento de grande escala com os forçantes locais (topografia, sistemas de brisa, solo e vegetação, etc.). No entanto, se a previsão do modelo global não for boa, o modelo regional não consegue realizar uma boa previsão já que este é forçado pelas fronteiras do modelo global. Por isso, ao mesmo tempo em que o *downscaling* dinâmico visa a obtenção de um maior detalhamento das previsões climáticas, pode acrescentar novos erros (intrínsecos às limitações dos esquemas numéricos e parametrizações do modelo de mesoescala) aos erros já existentes.

Toda essa discussão leva a uma visão clara da complexidade e até certo ponto, fragilidade do *downscaling* dinâmico, apesar de ele apresentar o que de melhor possa ser feito com a teoria existente. Fontes de imprecisão são, assim, carregadas para dentro dos modelos de mesoescala, denunciando as suas limitações. Portanto, é imensamente temerário creditar total confiança em um dado modelo, aceitando totalmente os seus resultados como valores finais do processo de previsão. Quando muito, a reunião dos resultados dos diversos modelos disponíveis com as demais ferramentas de análise, poderia se associar com resultados em conjunto – “*ensemble*” – permitindo a indicação das incertezas das previsões aos usuários.

Ao mesmo tempo, é possível corrigir os resultados de modelos dinâmicos utilizando ferramentas estatísticas, sejam elas técnicas de regressão, ou técnicas de estatística multivariada. Neste trabalho, propomos o uso da ACC para corrigir a saída dos modelos utilizados no *downscaling* dinâmico da FUNCEME. A Figura 3.8 ilustra esquematicamente as etapas da correção estatística da previsão da precipitação aqui proposta.

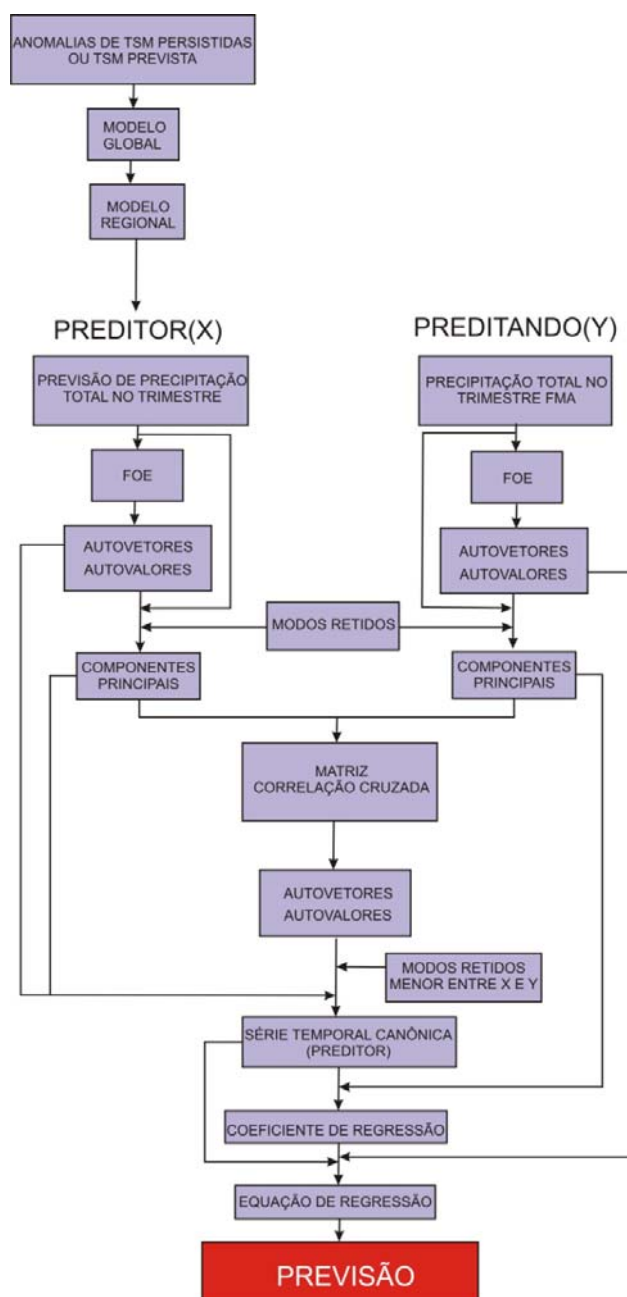


Figura 3.8 - Esquema ilustrativo das etapas utilizadas para a correção estatística dos resultados da previsão dinâmica da precipitação.

3.4.3 PREVISÃO HÍBRIDA

Usando as idéias já comentadas aqui de que a TSM é um bom preditor para a precipitação sobre o NEB e que os modelos dinâmicos produzem previsões com alta destreza sobre a região e se baseando no fato de que modelos empíricos e dinâmicos podem ter habilidades de previsão complementares, foi

desenvolvido neste trabalho um modelo “híbrido” com o objetivo de combinar os conhecimentos de previsão estatísticos e dinâmicos e com isso tirar proveito dos aspectos positivos de ambos os modelos na tentativa de superar algumas das limitações das duas abordagens individualmente.

Neste trabalho, a simulação usando o modelo híbrido representa uma combinação da TSM com a saída do modelo numérico como preditores, usando a precipitação observada como preditando.

A Figura 3.9 ilustra esquematicamente as etapas da hibridização aqui realizada.

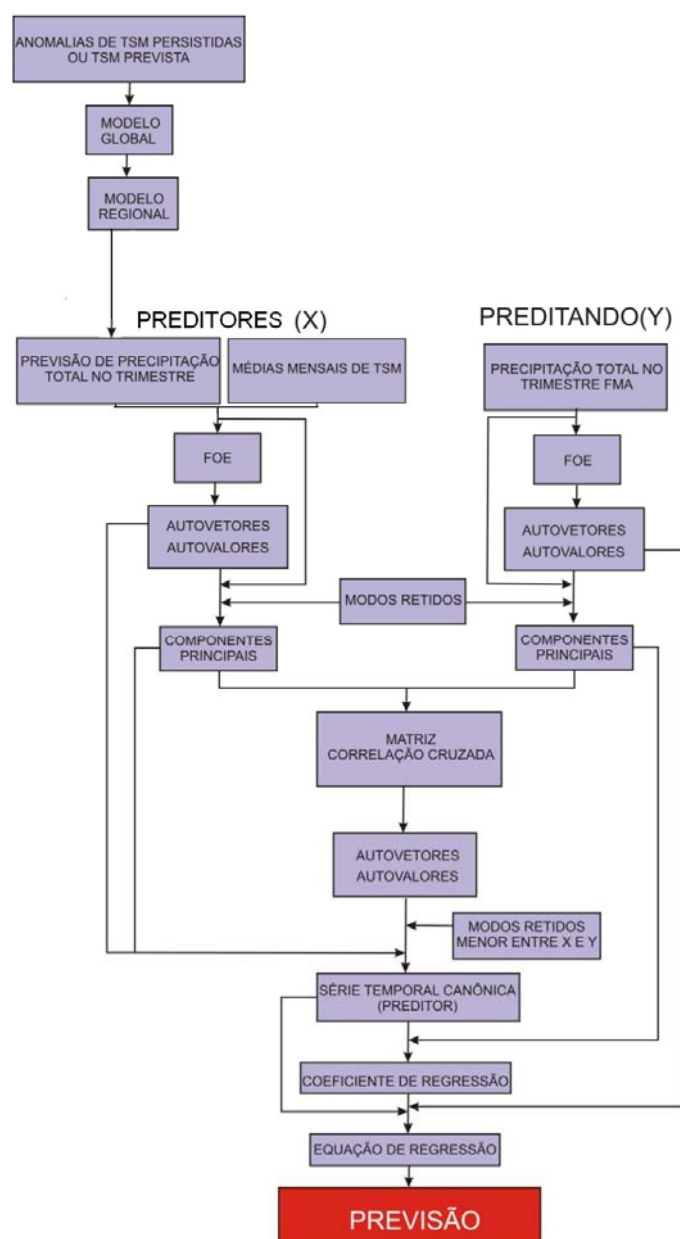


Figura 3.9 - Esquema ilustrativo das etapas utilizadas para a construção do modelo híbrido.

3.5 MÉTRICAS E DESTREZA DAS PREVISÕES

Quando se desenvolvem sistemas de previsão, é importante validar seus resultados de forma a quantificar seu poder discriminativo e identificar um procedimento ou método como bom ou não para determinada análise. Entretanto, devemos levar em conta que a simples quantificação de acertos numa amostra de dados não necessariamente reflete o quão eficiente o sistema é, pois essa quantificação dependerá essencialmente da qualidade e distribuição dos dados nesta amostra.

Para analisar os resultados deste trabalho, foram usadas a análise dos coeficientes de correlação de Pearson e Spearman e a construção de curvas ROC (*Receiver Operating Characteristic Curve*).

3.5.1 COEFICIENTE DE CORRELAÇÃO DE PEARSON

É um método comumente conhecido para medir a correlação entre duas variáveis. Este método de correlação foi estudado por Francis Galton e seu aluno Karl Pearson, em 1897.

O coeficiente de correlação r é definido como:

$$r_{xy} = r = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y}$$

O Estimador do Método de Coeficiente de Correlação Linear de Pearson é dado por:

$$\hat{\rho} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

A covariância é uma medida que expressa a variação conjunta de duas variáveis, cuja expressão é dada por:

$$\text{cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$$

Para se utilizar deste coeficiente se deve levar em conta algumas considerações:

- 1) A relação entre as duas variáveis deva ser linear.
- 2) As variáveis envolvidas sejam aleatórias.
- 3) As duas variáveis tenham uma distribuição normal bivariada conjunta, o que equivale a dizer que para cada X dado, a variável Y é normalmente distribuída. Esta consideração é necessária para fazer inferências estatísticas (teste de hipótese e intervalo de confiança), sendo desnecessário quando se tratar de estudos amostrais.

O coeficiente de correlação de Pearson é interpretado como a medida do grau de relacionamento linear entre duas variáveis aleatórias, por isso ele tem ênfase na predição do grau de dependência entre duas variáveis aleatórias.

O coeficiente de correlação de Pearson tem sua magnitude compreendida entre os valores -1 e 1 . Essa magnitude, em módulo, mede a “força” da relação entre as duas variáveis. Assim, um coeficiente igual a $0,6$ indica maior grau de dependência linear que um de $0,3$. Um coeficiente de valor zero indica a total ausência de relacionamento linear entre as variáveis e coeficientes de valor 1 e -1 indicam uma perfeita dependência linear entre elas.

O sinal do coeficiente de correlação $(+/-)$ define a direção da relação entre as duas variáveis (positiva ou negativa).

Uma correlação positiva sugere que, enquanto os valores das componentes de uma variável aumentam, os valores das componentes da outra variável aumentam e vice-versa. Para uma correlação negativa, acontece o oposto: enquanto uns aumentam, os outros diminuem e vice-versa.

Entretanto, na prática ocorrem diferentes valores do coeficiente de correlação (r). A interpretação do valor do coeficiente de correlação (r) depende muito dos objetivos de sua utilização e as razões pelas quais este é calculado. Segundo Callegari-Jacques (2006), o coeficiente de correlação pode ser avaliado qualitativamente da seguinte forma:

Tabela 3.1 – Análise qualitativa do grau de correlação entre duas variáveis.

Fonte: Callegari-Jacques (2006)

r	Magnitude da correlação
0,00	Nula
$0,00 \leq r < 0,30$	Fraca
$0,30 \leq r < 0,60$	Moderada
$0,60 \leq r < 0,90$	Forte
$0,90 \leq r < 1,00$	Muito Forte
1,00	Plena ou Perfeita

3.5.2 COEFICIENTE DE CORRELAÇÃO DE SPEARMAN

É um método estatístico baseado em postos e foi introduzida por Spearman em 1904 (SIEGEL, 1975). O coeficiente ρ de Spearman mede a intensidade da relação entre variáveis ordinais. Usa, em vez do valor observado, apenas a ordem das observações. Basicamente equivale ao coeficiente de correlação de Pearson aplicado à dados ordenados. Por este motivo o coeficiente de correlação de Spearman não é sensível a assimetrias na distribuição, nem à presença de *outliers*, não exigindo, portanto que os dados decorram de duas amostras normais.

O coeficiente de correlação de Spearman é aplicado em variáveis intervalares como alternativa ao r de Pearson, quando neste último se viola a normalidade. Nos casos em que os dados não formam uma nuvem “bem comportada”, com alguns pontos muito afastados dos restantes, ou em que parece existir uma relação crescente ou decrescente em formato de curva, o coeficiente ρ de Spearman é mais adequado.

Contudo, o coeficiente de Pearson é uma opção mais atraente para calcular a correlação linear entre duas variáveis quantitativas, pois esse coeficiente usa os valores observados para X e Y enquanto que o coeficiente de Spearman usa suas ordens. Ao substituir os valores das variáveis pelos seus

postos perde-se a magnitude de cada observação. Quando houverem postos empatados, o procedimento usual é atribuir a média dos postos que lhes caberiam se não tivesse havido empate, e novamente perde-se informação. Para amostras grandes esse problema é ainda pior, pois o número de postos empatados tende a aumentar.

Uma fórmula simples para calcular o coeficiente ρ de Spearman é dada por:

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}$$

em que n é o número de pares (X_i, Y_i) e $d_i = (\text{postos de } X_i \text{ dentre os valores de } X) - (\text{postos de } Y_i \text{ dentre os valores de } Y)$.

Se os postos de X são exatamente iguais aos postos de Y , então todos os d_i serão zero e ρ será 1.

Do mesmo modo que o coeficiente de correlação de Pearson (r), a magnitude do coeficiente ρ de Spearman varia entre -1 e 1. Quanto mais próximo estiver destes extremos, maior será a associação entre as variáveis. O sinal negativo da correlação significa que as variáveis variam em sentido contrário, isto é, as categorias mais elevadas de uma variável estão associadas a categorias mais baixas da outra variável.

Conforme Bunchaft e Kellner (2001), as correlações ordinais não podem ser interpretadas da mesma maneira que as correlações de Pearson. Inicialmente, não mostram necessariamente tendência linear, mas podem ser consideradas como índices de monotonicidade, ou seja, para aumentos positivos da correlação, aumentos no valor de X correspondem a aumentos no valor de Y , e para coeficientes negativos ocorre o oposto.

3.5.3 ÁREAS ROC

A análise ROC (*Receiver Operating Characteristic*) é um método gráfico para avaliação, organização e seleção de sistemas de diagnóstico e/ou

predição utilizando classificadores baseado em seus desempenhos. Para realizar estas análises, utilizam-se os gráficos ROC para mostrar o limiar entre taxas de acertos e alarmes falsos (taxas de erros) de classificadores. Neste trabalho foram utilizados na previsão climática, para se avaliar a qualidade das predições de eventos de precipitação.

Foram restritos nessa discussão os problemas de classificação binária, (onde uma instância pode assumir valores no conjunto P, N, positivo e negativo respectivamente).

Dado um padrão (exemplo) de teste e um classificador, a classificação se enquadra em quatro possíveis categorias:

Verdadeiro Positivo - VP: Ocorre quando o classificador prediz uma resposta do tipo 'positivo' e essa resposta está de acordo com os dados reais (acerto).

Verdadeiro Negativo - VN: Ocorre quando o classificador prediz uma resposta do tipo 'negativo' e essa resposta está de acordo com os dados reais (acerto).

Falso Positivo - FP: Ocorre quando o classificador prediz uma resposta do tipo 'positivo'. Entretanto, na realidade a resposta deveria ser 'negativo' (erro).

Falso Negativo - FN: Ocorre quando o classificador prediz uma resposta do tipo 'negativo'. Entretanto, na realidade a resposta deveria ser 'positivo' (erro).

Uma maneira natural de apresentar as estatísticas para a avaliação de um modelo de classificação é por meio de uma tabulação cruzada entre a classe prevista pelo modelo e a classe real dos exemplos. Essa tabulação é conhecida como tabela de contingência (também chamada de matriz de confusão).

Esta matriz serve como base para muitas métricas que podem ser aplicadas a classificação. A Tabela 3.2 mostra a matriz de confusão.

		Valor Verdadeiro (confirmado por análise)	
		positivos	negativos
Valor Previsto (predito pelo teste)	positivos	VP Verdadeiro Positivo	FP Falso Positivo
	negativos	FN Falso Negativo	VN Verdadeiro Negativo

Tabela 3.2 - Tabela de Contingência (matriz de confusão)

Se dividirmos cada entrada na matriz mostrada na Tabela pelo tamanho da amostra, cada entrada dessa matriz representará uma estimativa da probabilidade conjunta da classe real do exemplo e da predição dada pelo modelo.

Uma opção à avaliação utilizando medidas é o uso de gráficos e/ou diagramas. Gráficos permitem uma melhor visualização da multidimensionalidade do problema de avaliação. O gráfico ROC é baseado na probabilidade de detecção, ou taxa de verdadeiros positivos (taxa de VP), e na probabilidade de falsos alarmes, ou taxa de falsos positivos (taxa de FP).

Para se obter o ponto no espaço ROC correspondente a um modelo de classificação, calcula-se a taxa de verdadeiros e falsos positivos (taxa de VP e taxa de FP) desse modelo a partir da sua matriz de contingência.

$$\text{Taxa de VP} = \frac{VP}{P} \quad \text{e} \quad \text{Taxa de FP} = \frac{FP}{N}$$

onde, P e N são, respectivamente, o número total de amostras positivas e negativas na classe real.

$$P = VP + FN$$

$$N = VN + FP$$

Uma análise mais refinada pode ser feita pela decomposição das probabilidades conjuntas em probabilidades condicionais e marginais.

Probabilidades condicionais podem ser obtidas a partir das leis básicas de probabilidade:

$$P(X,Y) = P(X | Y)P(Y) = P(Y | X)P(X)$$

na qual $P(X|Y)$ é a probabilidade condicional de X ser verdade, dado que Y é verdade. É importante ressaltar que $P(X,Y)=P(Y,X)$, mas $P(X|Y)\neq P(Y|X)$. Dado um conjunto de exemplos, todas essas probabilidades podem ser estimadas como proporções. Mais especificamente

$$P(X|Y) = p(X,Y)/p(Y) \text{ e } P(Y|X) = p(X,Y)/p(X)$$

Essas probabilidades condicionais também podem ser calculadas para as quatro entradas da matriz e podem ser representadas como matrizes.

A probabilidade $P(X|Y)$ é importante para o usuário do modelo, uma vez que ela dá a probabilidade de que a classe seja positiva, dado que a previsão feita pelo modelo é positiva. Essa probabilidade também é conhecida como *confiança*. Entretanto, em termos de avaliação do modelo, $P(Y|X)$ é muito mais útil. Uma das vantagens da fatoração de $P(X,Y)$ em $P(Y|X)$ e $P(X)$ é que $P(Y|X)$ é condicional ao valor de X, isto é, é condicional à proporção de exemplos entre as classes. Essa probabilidade condicional é frequentemente conhecida como *crença* ou *verossimilhança*, uma vez que ela especifica a probabilidade de que uma predição particular é feita dada a ocorrência de uma observação específica. Essa probabilidade indica o quanto um modelo é capaz de discriminar os casos entre as possíveis classes.

Para se construir o gráfico ROC plota-se a taxa de FP no eixo dos ordenadas - eixo x e taxa de VP no eixo das abscissas - eixo y. A Figura 3.10 mostra um exemplo de gráfico ROC.

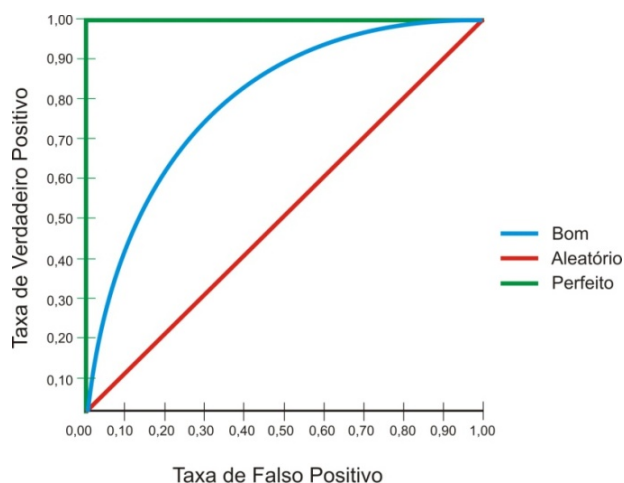


Figura 3.10 - Curvas ROC de diferentes classificadores

Assim, um gráfico ROC representa compromissos entre benefícios (taxa de VP) e custos (taxa de FP). Observando essa Figura, citamos alguns pontos importantes no espaço ROC:

O ponto (0,0) representa a estratégia de nunca classificar um exemplo como positivo. Modelos que correspondem a esse ponto não apresentam nenhum falso positivo, mas também não conseguem classificar nenhum verdadeiro positivo. A estratégia inversa, de sempre classificar um novo exemplo como positivo, é representada pelo ponto (1,1).

O ponto (0,1) representa o modelo perfeito, ou seja, máxima quantidade de amostras positivas classificadas verdadeiramente (VP) e nenhuma amostra negativa classificada como positiva (FP).

O ponto (1,0) representa o modelo que sempre faz previsões erradas.

Modelos próximos ao canto inferior esquerdo são conhecidos como "conservativos": eles fazem uma classificação positiva somente se têm grande segurança na classificação. Como consequência, eles cometem poucos erros falsos positivos, mas frequentemente têm baixas taxas de verdadeiros positivos.

Modelos próximos ao canto superior direito são conhecidos como "liberais": eles predizem a classe positiva com maior frequência, de tal maneira que classificam a maioria dos exemplos positivos corretamente, mas, geralmente, com altas taxas de falsos positivos.

A linha diagonal ascendente (0,0) – (1,1) representa um modelo de comportamento estocástico: cada ponto (p, p) pode ser obtido pela previsão da classe positiva com probabilidade p e da classe negativa com probabilidade 1–p.

Pontos pertencentes ao triângulo superior esquerdo a essa diagonal representam modelos que desempenham melhor que o aleatório e pontos pertencentes ao triângulo inferior direito representam modelos piores que o aleatório. Idealmente, a curva deve ser convexa e sempre crescente.

A maneira mais utilizada para comparar curvas ROC é o cálculo da área sob a curva (*area under curve* – AUC).

A AUC é uma fração da área de um quadrado de lado um e seu valor está sempre entre 0 e 1. Além disso, a AUC também é numericamente igual à

probabilidade de, dados dois exemplos de classes distintas, o exemplo positivo seja ordenado primeiramente que um exemplo negativo.

A área abaixo da curva ROC é um dos índices mais utilizados para sumariar a "qualidade" da curva. Teoricamente, quanto maior a AUC, melhor o sistema.

Analiticamente, a área abaixo da curva ROC pode ser determinada através de:

- métodos de resolução numérica, tipo regra do trapézio;
- métodos estatísticos: relação com a estatística de Wilcoxon-Mann-Witney e
- estimativa de máxima verosimilhança.

4 RESULTADOS E DISCUSSÃO

4.1 PREVISÃO EMPÍRICA PARA AS UNIDADES DE PLANEJAMENTO (BACIAS)

Nesta seção discutiremos os resultados da previsão empírica onde usamos a TSM como preditor e a precipitação observada nos trimestres de DJF a AMJ sobre as Unidades de Planejamento (Bacias) como preditando usando o módulo de ACC do CPT. Foram feitos experimentos similares com as Regiões Homogêneas apresentando resultados próximos aos das Bacias.

Em seguida à descrição da configuração do modelo, são mostradas as métricas de destreza da previsão: os coeficientes de correlação de Pearson e Spearman e as áreas ROC entre as séries observadas de precipitação e as séries previstas pelo CPT.

4.1.1 CONFIGURAÇÃO DO CPT

Diversas posições e extensões dos domínios oceânicos foram testadas neste estudo a partir dos arquivos de saída do CPT (com seus respectivos valores de *goodness index* e correlações de Pearson e Spearman) para determinar os diferentes domínios de TSM que se correlacionam com a precipitação sobre o Estado do Ceará e com isso serem usados na previsão da quadra chuvosa do Estado. Foi observado que uma maior previsibilidade das precipitações no Ceará para cada um dos trimestres pode ser explicada usando os domínios descritos em detalhe na tabela 4.1.

Tabela 4.1 - Posições dos domínios oceânicos que foram usados para a previsão empírica.

TSM DOS OCEANOS	ÁREA OCEÂNICA		PERÍODOS		MÉDIAS DAS MÉTRICAS		
	LATITUDE	LONGITUDE	PREDITOR	PREDITANDO	GOODNESS ÍNDICE	PEARSON	SPEARMAN
Atlântico	7°S-22°S	15°W-5°W	Outubro	DJF	0,597	0,413	0,536
<u>Niño</u> 3.4	8°S-8°N	170°E-140°W	Novembro	JFM	0,522	0,445	0,397
<u>Niño</u> 3.4	7°S-7°N	180°W-130°W	Dezembro	FMA	0,611	0,495	0,495
<u>Niño</u> 3	10°S-EQ	150°W-90°W	Janeiro	MAM	0,632	0,444	0,566
Atlântico + Pacífico	15°S-10°N	80°W-112°W	Fevereiro	AMJ	0,760	0,632	0,587

Foram aplicados os seguintes parâmetros de entrada no CPT: séries de precipitação com 36 anos (1970-2005), período de treinamento com climatologia de 25 anos (1981-2005) para as bacias e de 24 anos (1981-2004) para as regiões pluviometricamente homogêneas, comprimento da janela de validação cruzada de 5 anos e o número máximo de modos igual a 4.

O comprimento da janela de validação cruzada define quantos anos são deixados de fora quando realizamos a validação cruzada ou para se calcular estatisticamente o desempenho. Para mais informação sobre os procedimentos e as propriedades de validação cruzada referir Michaelson (1987) e Livezey (1995).

4.1.2 MODOS DE VARIABILIDADE

Como não há diferenças relevantes entre os resultados para as bacias e para as regiões pluviometricamente homogêneas, optamos por mostrar os resultados apenas referentes às bacias.

Utilizando-se dos melhores domínios de TSM, sugeridos pelo *goodness index* e usando a precipitação total nos trimestres Dezembro-Janeiro-Fevereiro, Janeiro-Fevereiro-Março, Fevereiro-Março-Abril, Março-Abril-Maio e Abril-Maio-Junho, foram obtidos as cargas canônicas da TSM e da precipitação

sobre as 11 bacias além da evolução temporal dos escores para essas duas variáveis como descritas nas Figuras de 4.1 a 4.5.

As cargas canônicas (*canonical loading*) representam as correlações entre as variáveis originais e o índice do grupo em questão, e são obtidas pela matriz de correlação entre as variáveis originais e canônicas. Podem ser utilizadas para interpretar as variáveis na função, sendo que quanto mais alta a carga canônica maior a importância desta variável no par canônico.

A Figura 4.1 indica que chuvas acima do normal em todo o Ceará durante o trimestre DJF estão relacionadas ao aquecimento em uma dada região do Atlântico Sul, cujo campo de TSM deve ser importante na modulação da atividade e/ou no posicionamento dos vórtices ciclônicos e das frentes frias (principais sistemas que influenciam a pré-estação chuvosa sobre o Ceará). Portanto, é possível que o Atlântico Sul exerça algum controle sobre as configurações desses sistemas meteorológicos, conferindo algum grau de previsibilidade para a pré-estação.

Na Figura 4.2, o que se nota é um aquecimento do Pacífico Oeste e um resfriamento do Pacífico Leste favorecendo a chuva no Estado do Ceará no período JFM.

Através das figuras 4.3 e 4.4 pode-se notar nas duas estações seguintes (FMA e MAM) um padrão canônico inverso entre a TSM (sinal negativo) e a precipitação (sinal positivo), indicando o fato já conhecido de que eventos de El Niño, em que há aquecimento da bacia do Pacífico, estão correlacionados com precipitação abaixo da média sobre as diversas regiões do Estado do Ceará (e vice-versa).

Apesar de o campo de TSM no Atlântico reconhecidamente ser responsável direto pelo posicionamento da ZCIT e, portanto, pela modulação da chuva sobre o norte do NEB e o Ceará, o que estes resultados sugerem é que a maior parte da previsibilidade para esse período recai sobre o Pacífico. No período de Novembro a Janeiro, os resultados sugerem que as condições que irão prevalecer sobre a bacia do Atlântico ainda não estão completamente

estabelecidas, podendo mudar de maneira relativamente rápida, ao passo que o comportamento da bacia do Pacífico evolui de maneira mais previsível.

No que diz respeito às precipitações em AMJ, a carga canônica da TSM mostra uma influência distribuída entre as duas Bacias (Figura 4.5), indicando que o estado do Atlântico a partir de Fevereiro é relativamente mais estável.

São mostrados também os padrões de comportamento semelhante entre as curvas dos escores temporais das FOE do preditor (vermelho) e preditando (verde).

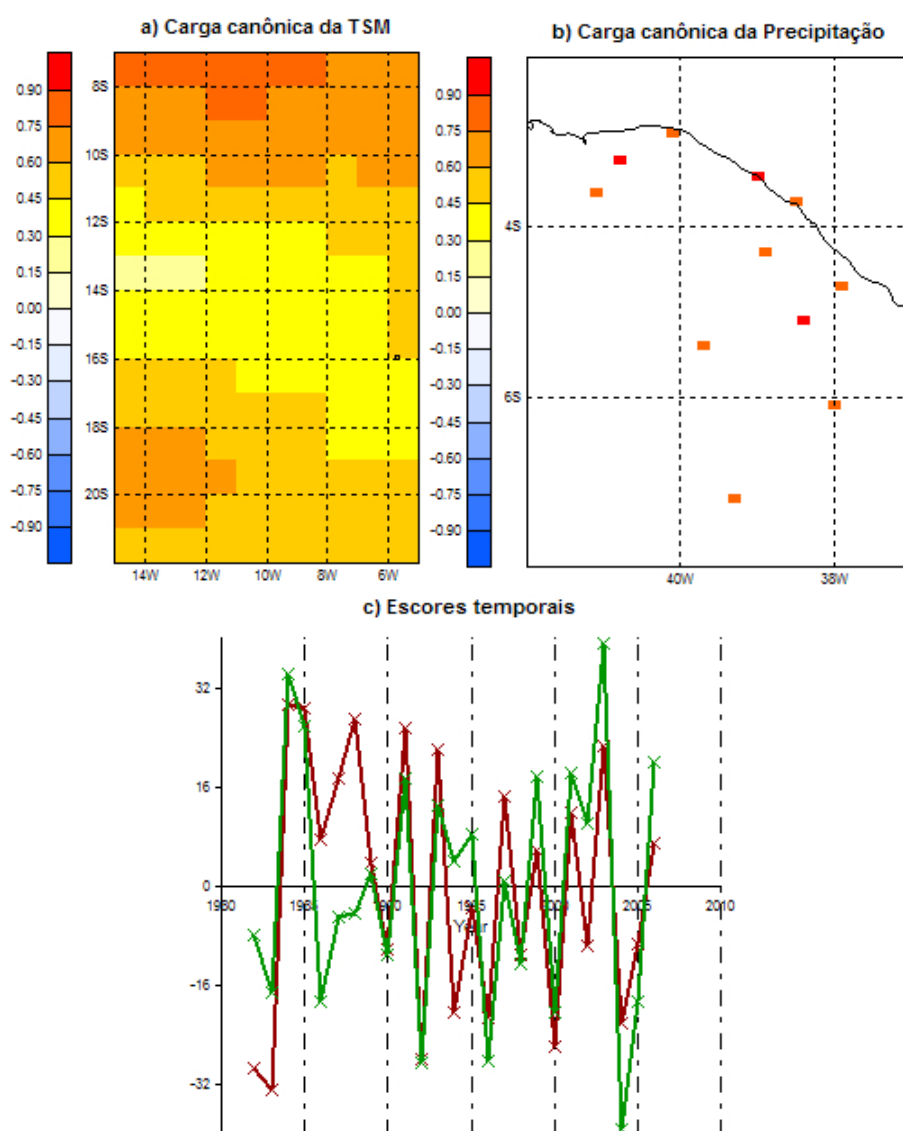


Figura 4.1 - a) Carga canônica da anomalia de TSM em outubro, b) Carga canônica da precipitação em DJF e c) Escores temporais para o primeiro modo dominante de correlações canônicas: TSM em vermelho e precipitação em verde.

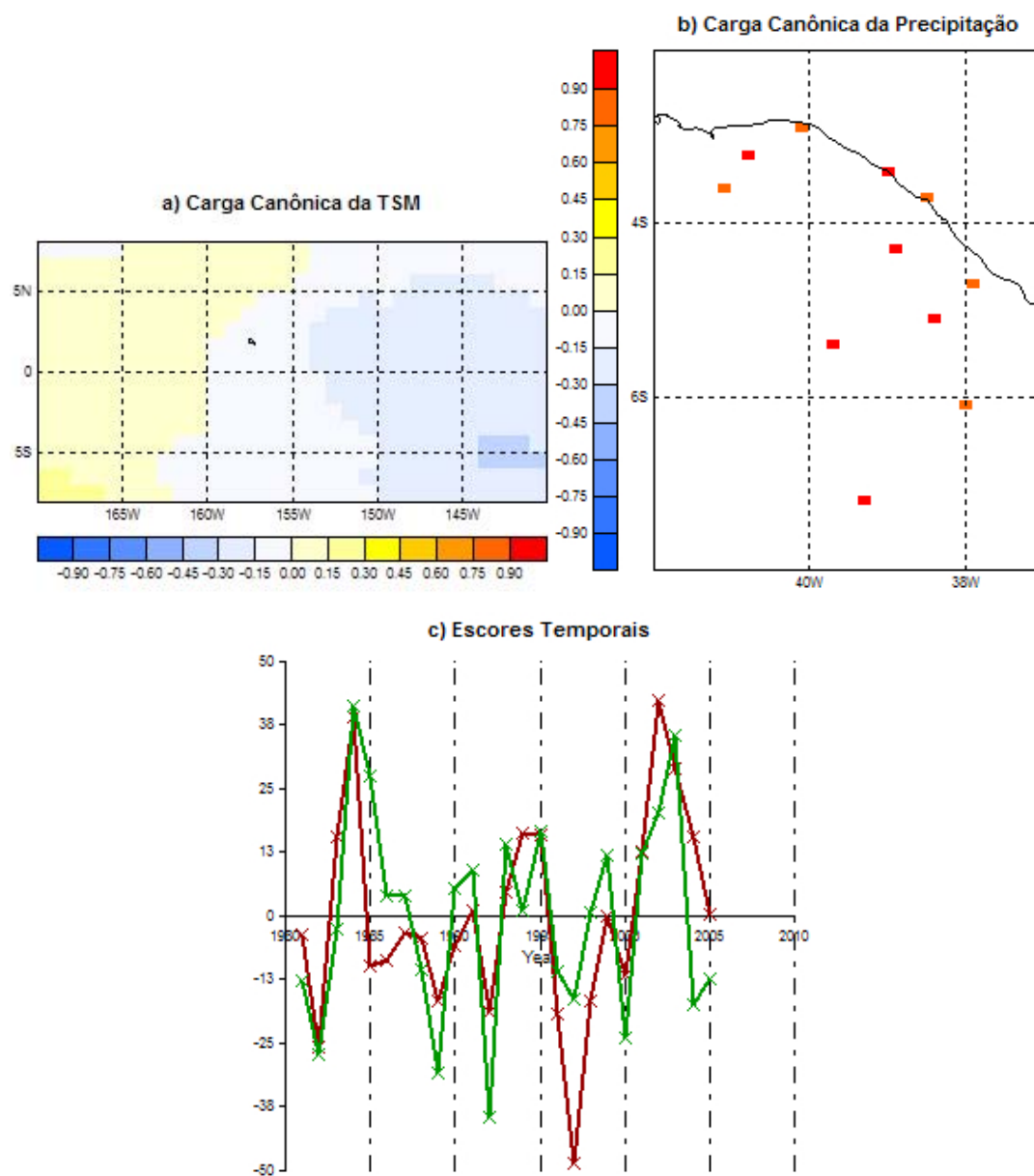


Figura 4.2 – Como na Figura 4.1, mas para a TSM em novembro e a precipitação em JFM.

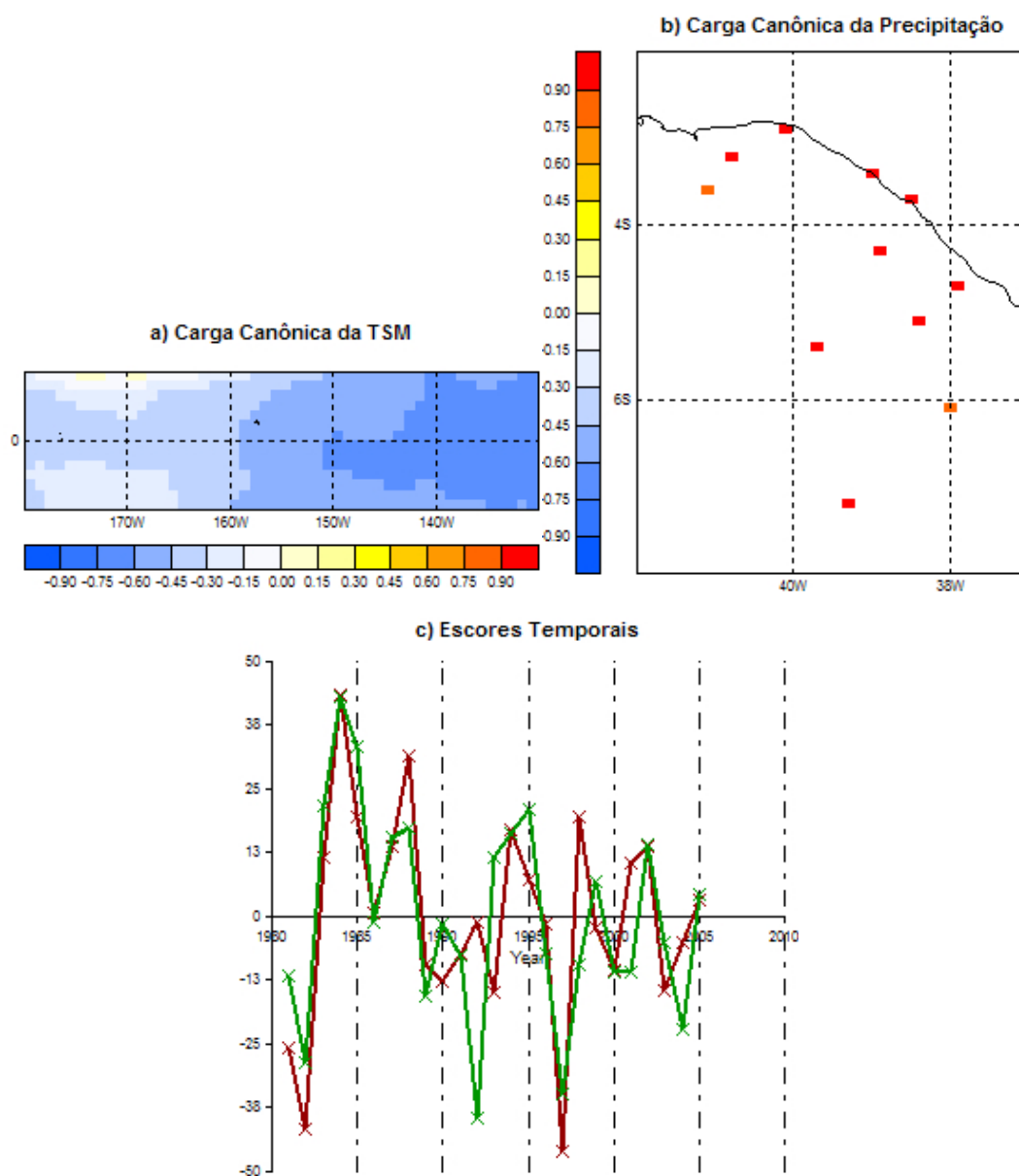


Figura 4.3 – Como na Figura 4.1, mas para a TSM em dezembro e a precipitação em FMA.

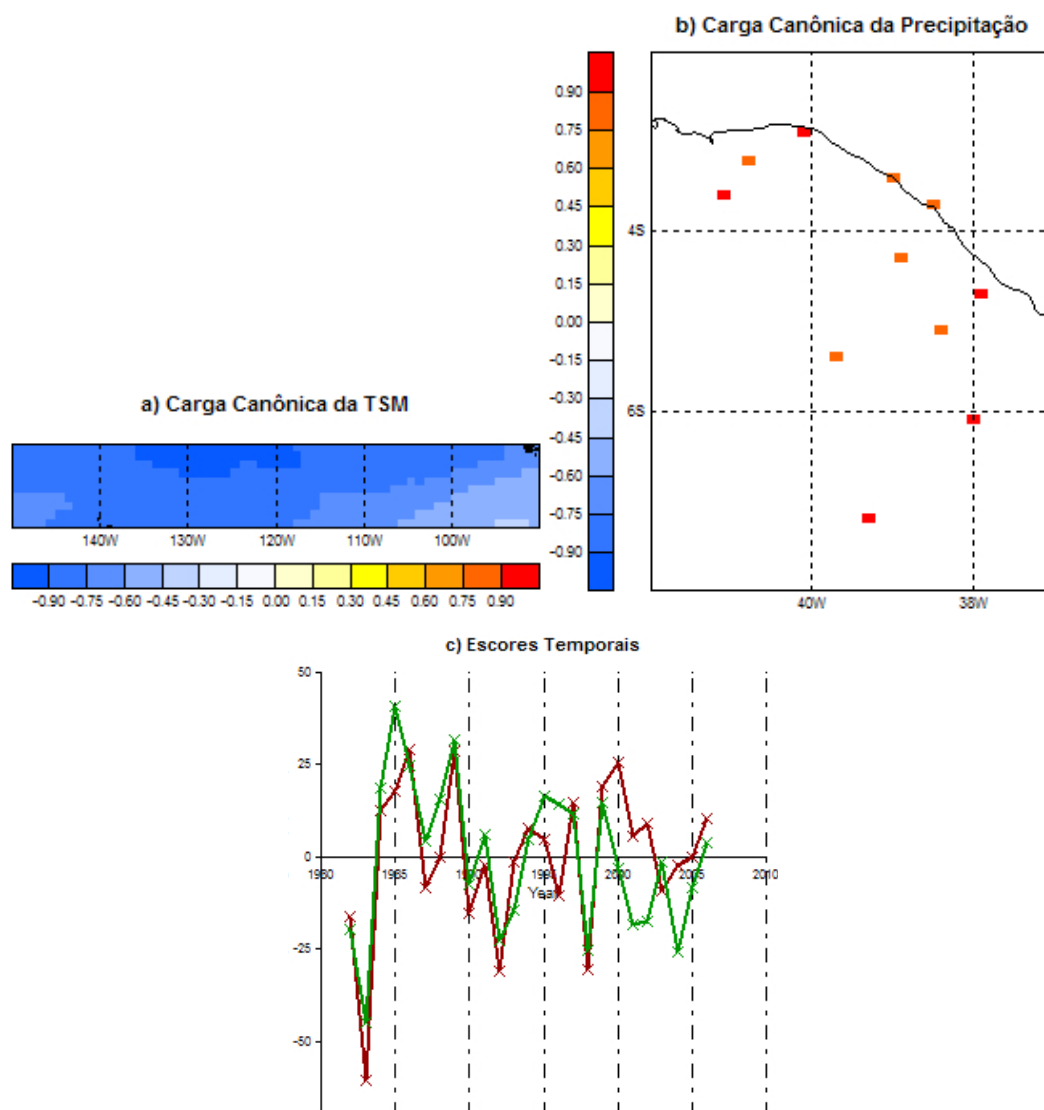


Figura 4.4 - Como na Figura 4.1, mas para a TSM em janeiro e a precipitação em MAM.

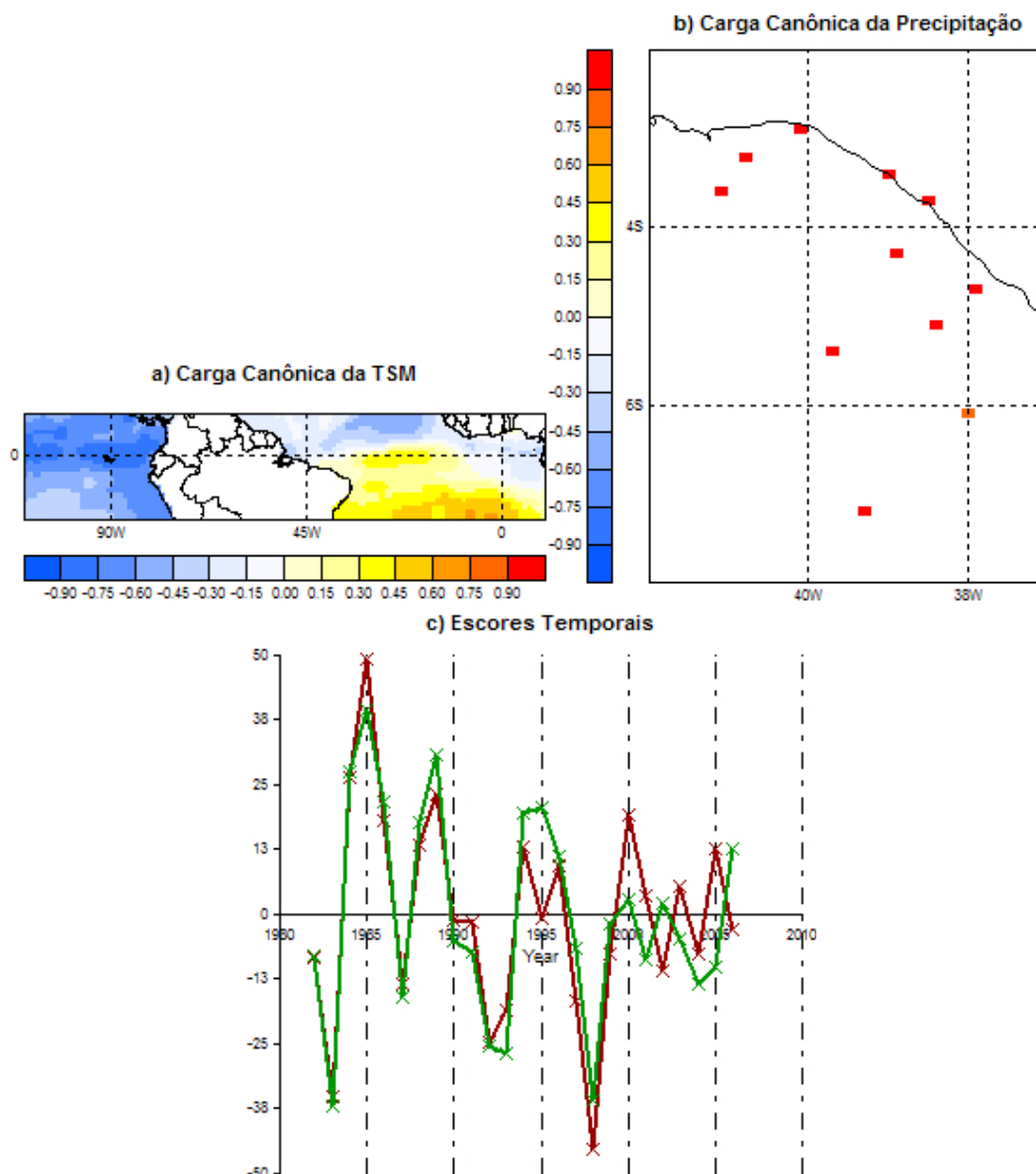


Figura 4.5 – Como na Figura 4.1, mas para a TSM em fevereiro e a precipitação em AMJ.

4.1.3 DESTREZA DO MODELO

Os coeficientes de correlação de Pearson (CP) e Spearman (CS) e as áreas ROC entre as séries observadas de precipitação e as séries previstas pelo CPT no período de 1981 a 2005 são indicados respectivamente nas Figuras 4.6 a 4.9. É importante ressaltar que enquanto CP utiliza os valores absolutos dessas variáveis para medir o grau de associação linear entre elas, CS utiliza variáveis qualitativas ordinais, não enfatizando os extremos, como o faz CP.

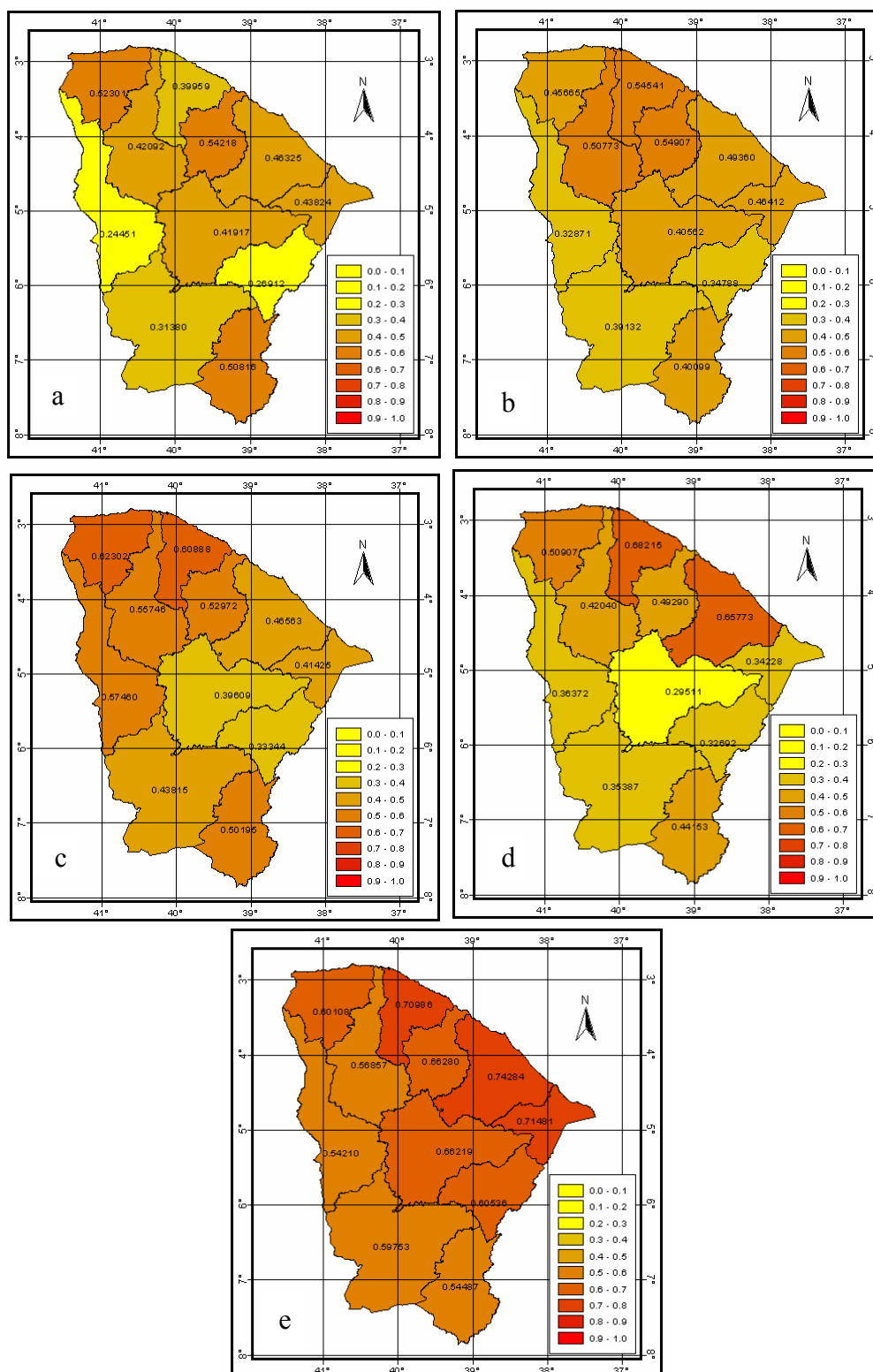


Figura 4.6 – Coeficientes de correlação de Pearson entre as séries observadas de precipitação e as séries previstas pelo CPT no período de 1981 a 2005 para as bacias no Estado do Ceará para os trimestres: (a) DJF, (b) JFM, (c) FMA, (d) MAM e (e) AMJ.

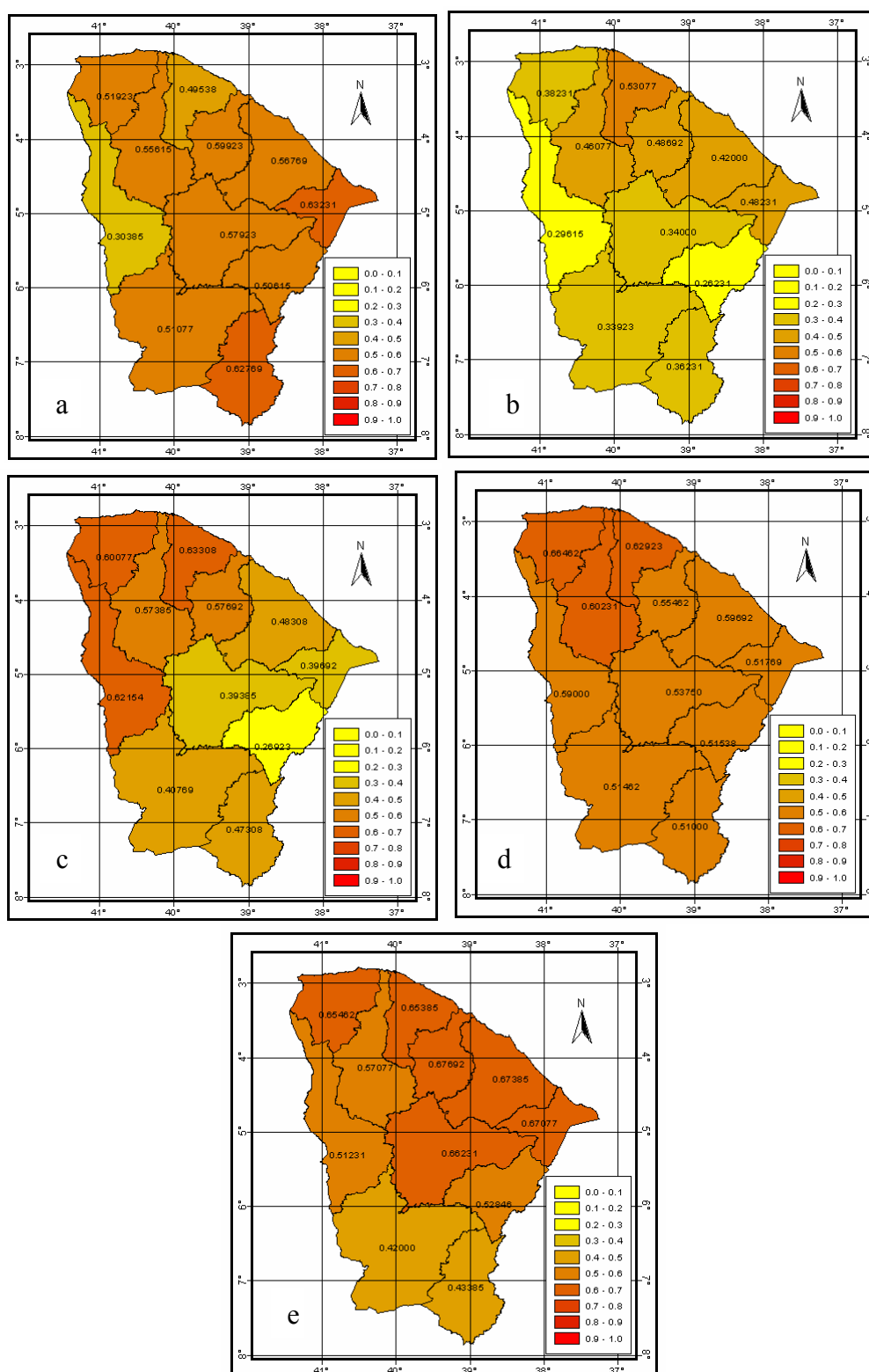


Figura 4.7 – Como na Figura 4.6, mas para o coeficiente de correlação de Spearman.

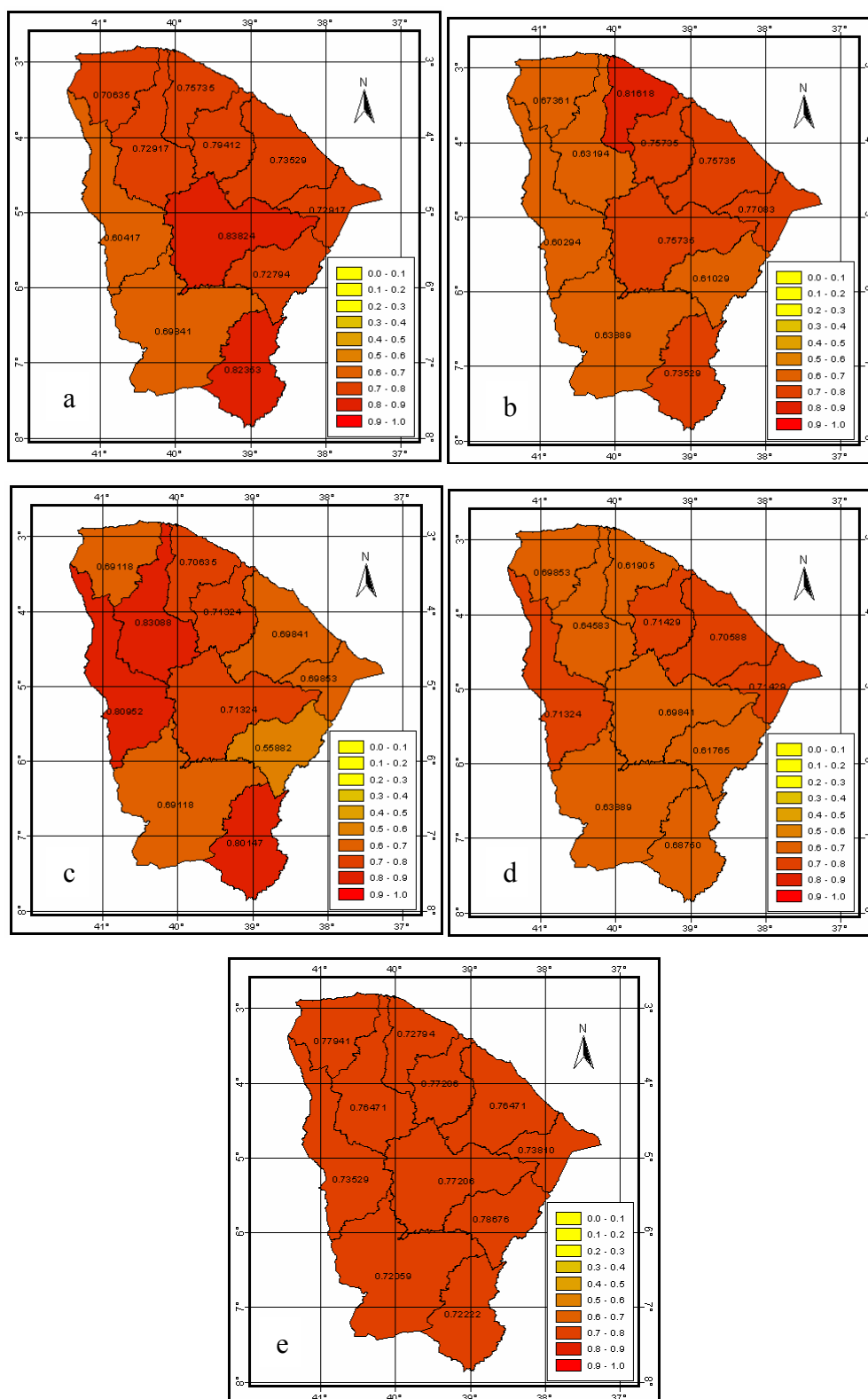


Figura 4.8 – Como na Figura 4.6, mas para as áreas ROC - Above.

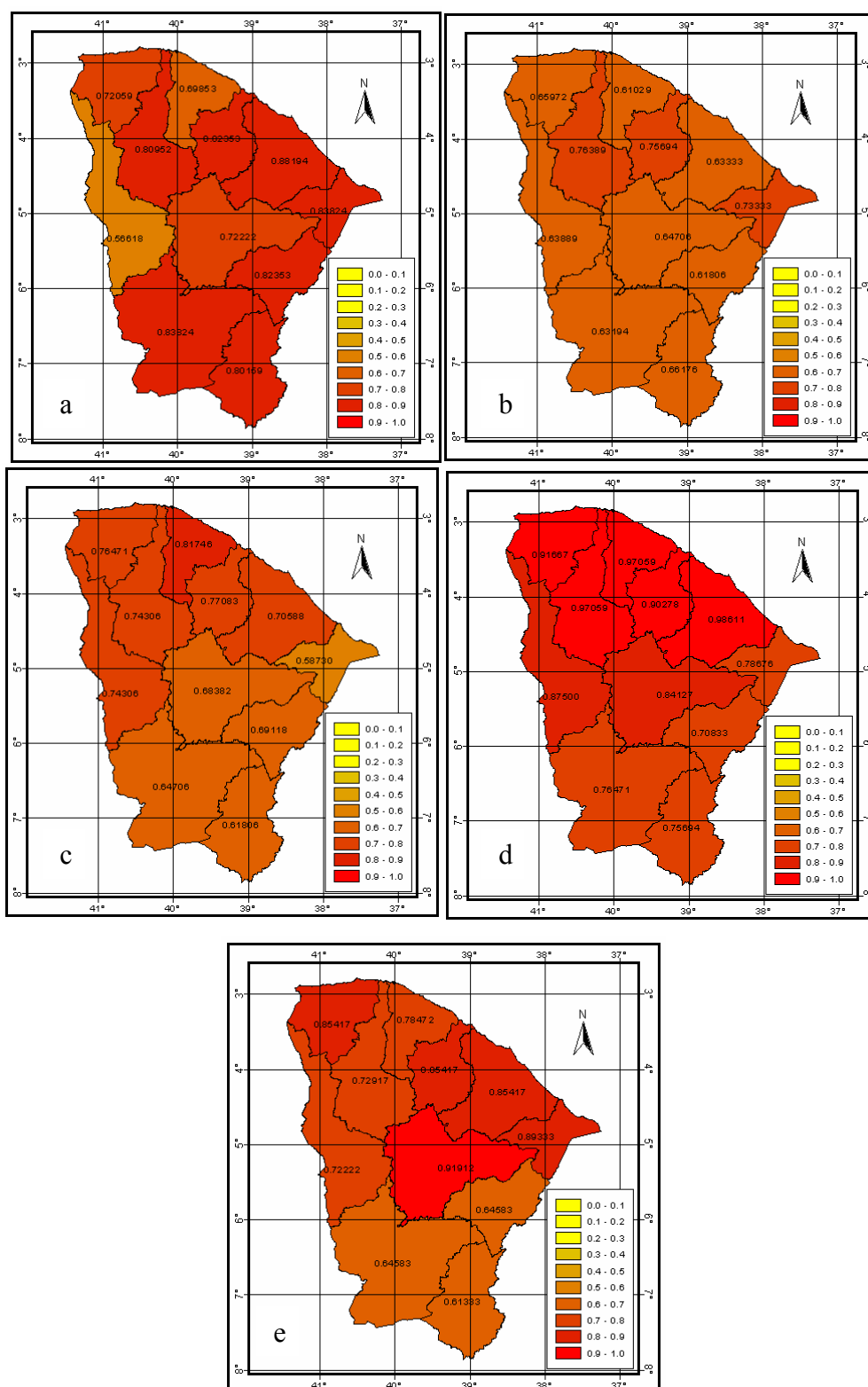


Figura 4.9 – Como na Figura 4.6, mas para as áreas ROC - Below.

Em DJF, os maiores valores de CP aparecem sobre as bacias do Coreaú, do Curu e Salgado, enquanto os mais altos valores de CS ocorrem para

as bacias do Baixo Jaguaribe e Salgado. Os resultados sugerem que a precipitação sobre esta última apresenta maior previsibilidade no período citado.

No trimestre JFM, os maiores valores de CP ocorrem nas bacias do Litoral e do Curu. A bacia do Litoral também apresenta altos valores de CS. Os valores relativamente elevados de CP e CS ao longo da pré-estação e do início da estação chuvosa se contrapõem de certa maneira à noção comum de que a previsibilidade das chuvas durante esse período é baixa.

No início da estação chuvosa (FMA), os valores de CP e CS em geral são semelhantes, com valores acima de 0,50 em todas as bacias do noroeste do Estado, com destaque para as bacias do Coreaú e Litoral (ambas as métricas acima de 0,60). Em contrapartida, valores relativamente pequenos de CP e CS (em torno a abaixo de 0,40) aparecem sobre as bacias do Banabuiú, Médio e Baixo Jaguaribe. No trimestre seguinte (MAM), os valores de CS tendem a ser maiores do que os de CP, com destaque para as bacias do Coreaú, Acaraú e Litoral (maiores valores de CS) e para as bacias do Litoral e Metropolitana (maiores valores de CP).

Ao final da estação chuvosa/início da pós-estação (AMJ), os valores de CP e CS atingem seus máximos na maior parte das bacias (CP chega a exceder 0,70 nas bacias do Litoral, Baixo Jaguaribe e Metropolitana).

No que diz respeito às áreas ROC, Buizza et al. (1999) sugere que a área sobre a curva ROC acima de 0,8 é um indicador de alta destreza por parte do modelo preditor, e uma área de 0,7 é o seu limite útil.

De acordo com este critério, em DJF, todas as Bacias, com exceção da Bacia do Parnaíba e do Alto Jaguaribe, apresentaram bom desempenho para discriminar eventos de maior precipitação (área sob a curva ROC-*above* de 0,70 a 0,83), conforme a Figura 4.8a. Além disso, nesse trimestre, apenas as Bacias do Litoral e do Parnaíba apresentaram área sob a curva ROC-*below* abaixo de 0,70, sendo que as demais apresentam bom desempenho para discriminar eventos de menor precipitação (área sob a curva ROC-*below* de 0,72 a 0,88), conforme a Figura 4.9a.

Para JFM, somente metade das Bacias apresentaram área sob a curva ROC-*above* acima de 0,70 (Figura 4.8b) com destaque para a Bacia do Litoral. Por sua vez, nesse trimestre, a área sob a curva ROC-*below* apresentou baixos valores para a maioria das Bacias, sendo que apenas três (Bacias do Curu, Coreaú e do Alto Jaguaribe) apresentaram valores acima de 0,70 (Figura 4.9b).

Da mesma forma que em JFM, o trimestre FMA apresentou área sob a curva ROC-*above* acima de 0,70 em seis Bacias dentre as quais a Bacia do Acaraú se destaca (0,83), de acordo com a Figura 4.8c. Para esse mesmo trimestre, a área sob a curva ROC-*below* acima de 0,70 aparece para as Bacias mais ao norte do Estado com destaque para a Bacia do Litoral como mostra a Figura 4.9c.

No trimestre MAM apenas quatro Bacias apresentaram valores de área sob a curva ROC-*above* em torno de 0,70 (Figura 4.8d). No entanto, para o mesmo trimestre, os valores de área sob a curva ROC-*below* foram bem maiores com destaque para as Bacias do Curu (0,90), Coreaú (0,91), Acaraú (0,97), Litoral (0,97) e em especial a Bacia Metropolitana (0,98) como indicado na Figura 4.9d.

Já na Figura 4.8e, no último trimestre analisado (AMJ), todas as Bacias indicaram desempenho satisfatório para os valores de área sob a curva ROC-*above* (de 0,72 a 0,78). No que diz respeito à área sob a curva ROC-*below*, para o mesmo trimestre, todas as Bacias, com exceção de apenas três, apresentam valores acima de 0,70 com destaque para a Bacia do Banabuiú (0,91) (Figura 4.9e).

Em suma, o comportamento do índice ROC-*above* e ROC-*below* para a maioria das regiões em todos os trimestres estudados mostra que as previsões do modelo encontram-se na faixa útil (acima de 0,7), em diversos casos excedendo o valor para uma discriminação considerada “boa”, isto é, com valores acima de 0,8.

De um modo geral, os resultados para as correlações de Pearson e Spearman e as áreas ROC demonstram razoável habilidade do modelo empírico

baseado em ACC em simular a variabilidade interanual da precipitação sobre o Estado do Ceará.

4.2 CORREÇÃO DA MODELAGEM NUMÉRICA

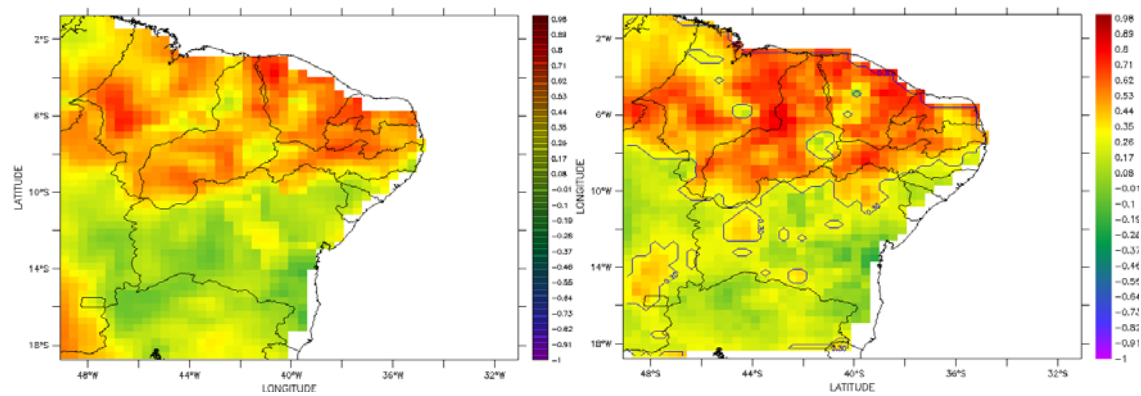
Tem sido documentado que os modelos dinâmicos podem fornecer hábeis previsões sazonais, ou seja, previsões que são melhores do que a climatologia, particularmente nos trópicos (STOCKDALE et al. 1998), mas também tem sido demonstrado que a habilidade de previsão para a precipitação, pode ser melhorado através de técnicas estatísticas para correção dos resultados do modelo bruto. Feddersen et al. (1999) mostraram como a correção estatística usando a análise de decomposição em valores singulares (DVS), poderia melhorar a habilidade de simulações de precipitação sazonal feita com um modelo de circulação geral da atmosfera forçados pela TSM observada.

No presente estudo, para corrigir o modelo numérico, foi usada a análise de correlação canônica para a construção de modelos que compararam os resultados do modelo numérico puro, de 30 anos, com o modelo numérico corrigido para os mesmos 30 anos. Esse modelo numérico corrigido relaciona a precipitação observada (preditando) com a precipitação prevista (preditor) resultado do modelo numérico RAMS nas mesmas regiões.

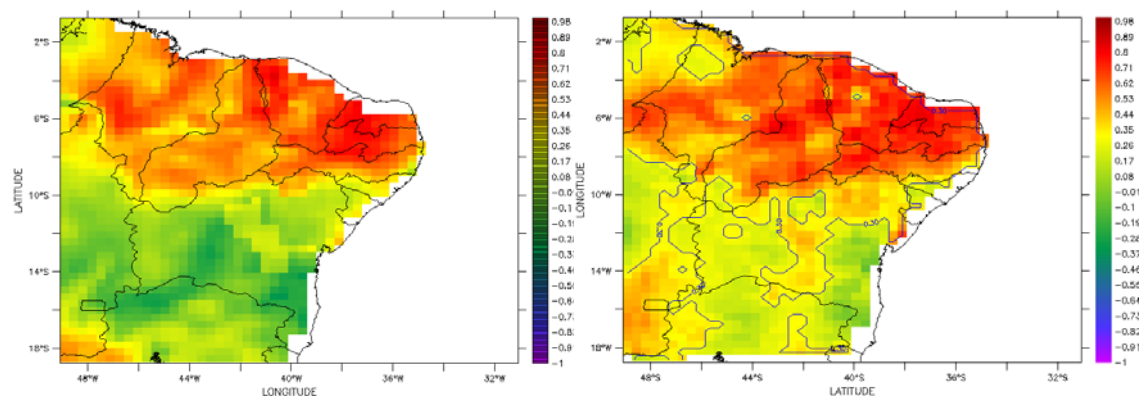
Para se investigar o sentido de concordância entre a precipitação no NEB simulada pelo modelo numérico puro com a do modelo numérico corrigido, aplicaram-se os coeficientes de correlação de Pearson e Spearman e as áreas ROC.

A Figura 4.10 faz uma comparação da correlação de Pearson entre os resultados não corrigidos do modelo regional e a precipitação observada (esquerda) com a correlação de Pearson entre os resultados corrigidos via CPT e a precipitação observada (direita) para os trimestres JFM, FMA, MAM e AMJ.

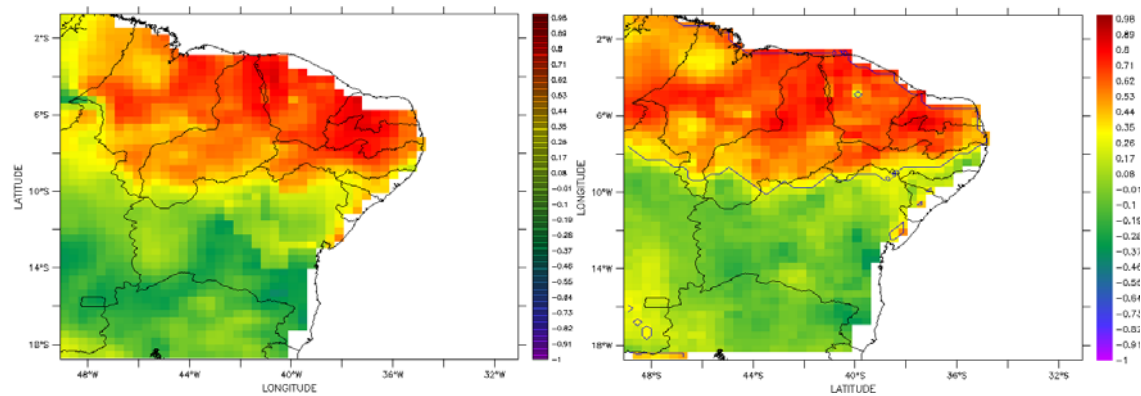
a) trimestre JFM



b) trimestre FMA



c) trimestre MAM



d) trimestre AMJ

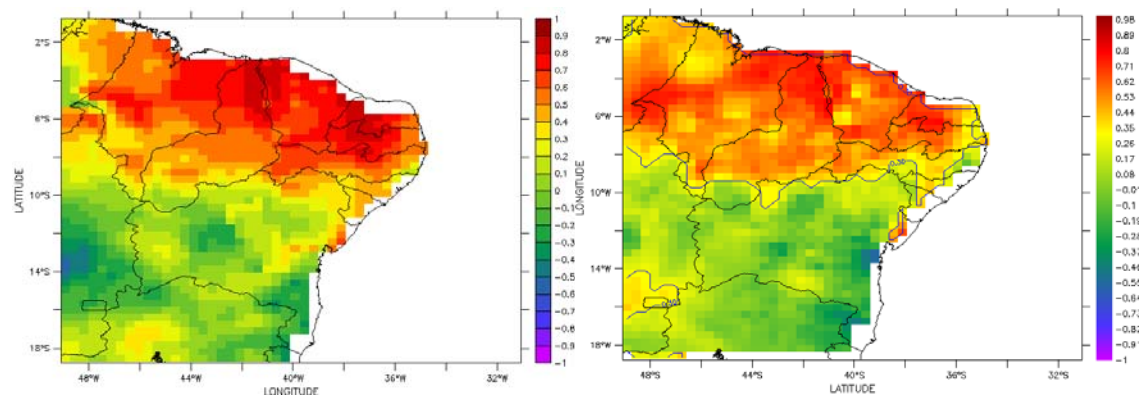
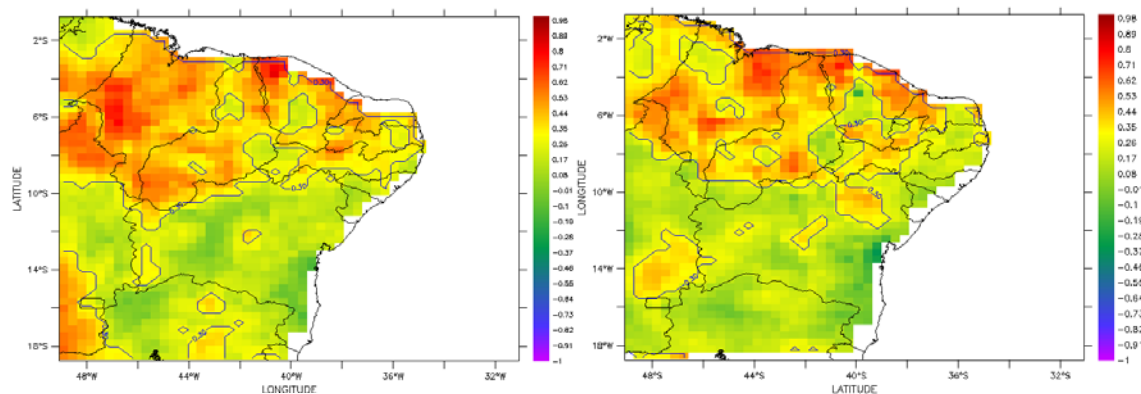


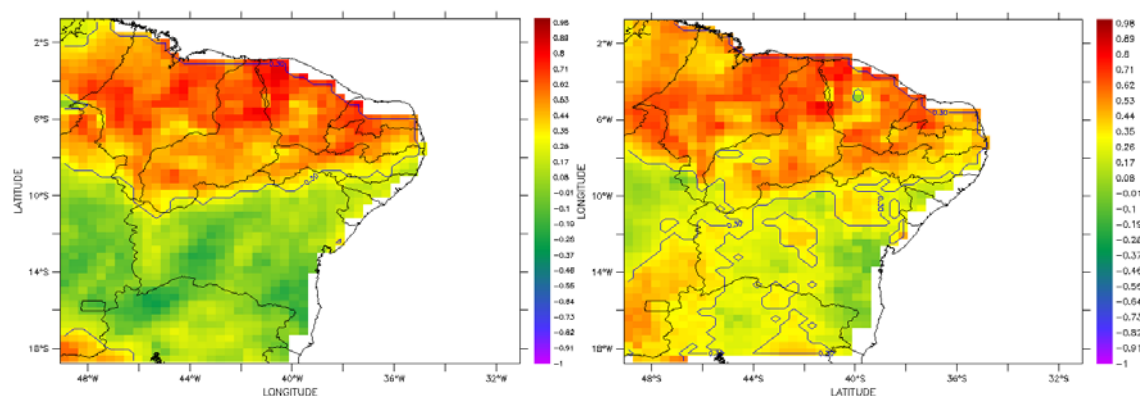
Figura 4.10 – Coeficientes de correlação de Pearson entre os resultados não corrigidos do modelo regional e a precipitação observada (esquerda) e a correlação de Pearson entre os resultados corrigidos através do CPT e a precipitação observada (direita) para os trimestres: (a) DJF, (b) JFM, (c) FMA, (d) MAM e (e) AMJ.

A Figura 4.11 mostra a comparação da correlação de Spearman entre os resultados não corrigidos do modelo regional e a precipitação observada (esquerda) com a correlação de Spearman entre os resultados corrigidos via CPT e a precipitação observada (direita) para os trimestres JFM, FMA, MAM e AMJ.

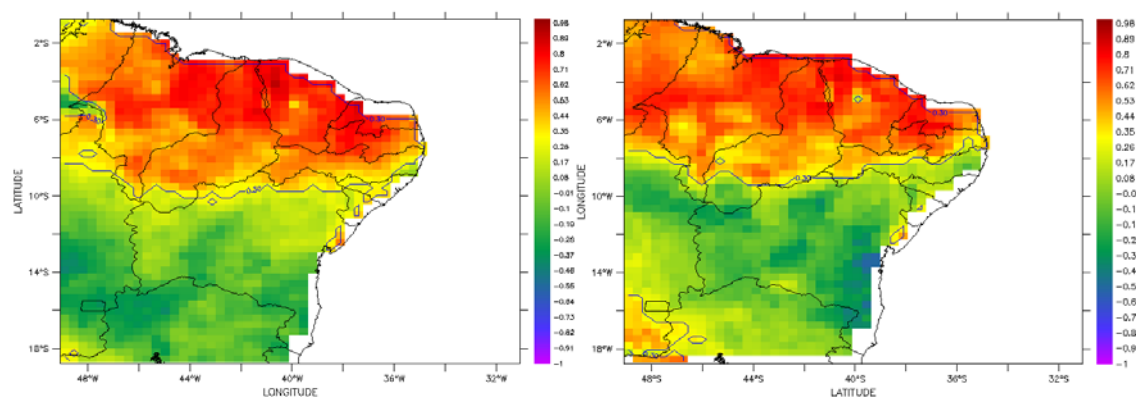
a) trimestre JFM



b) trimestre FMA



c) trimestre MAM



d) trimestre AMJ

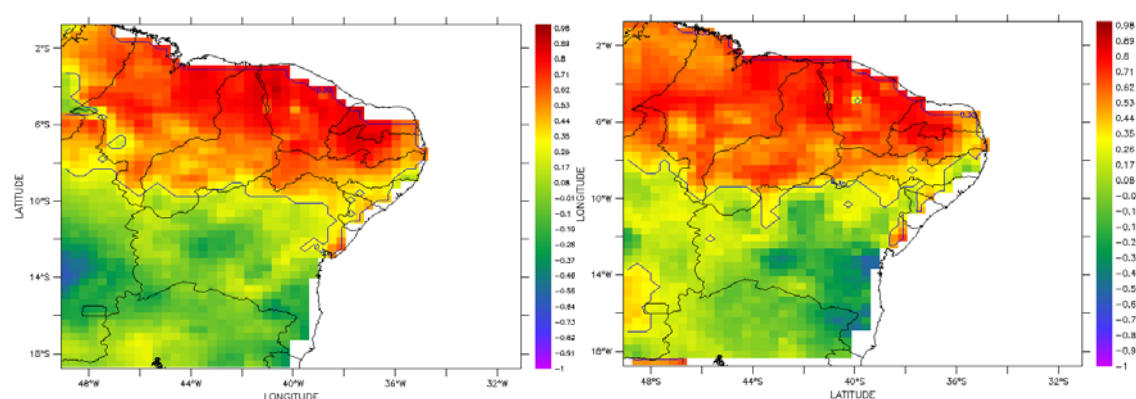
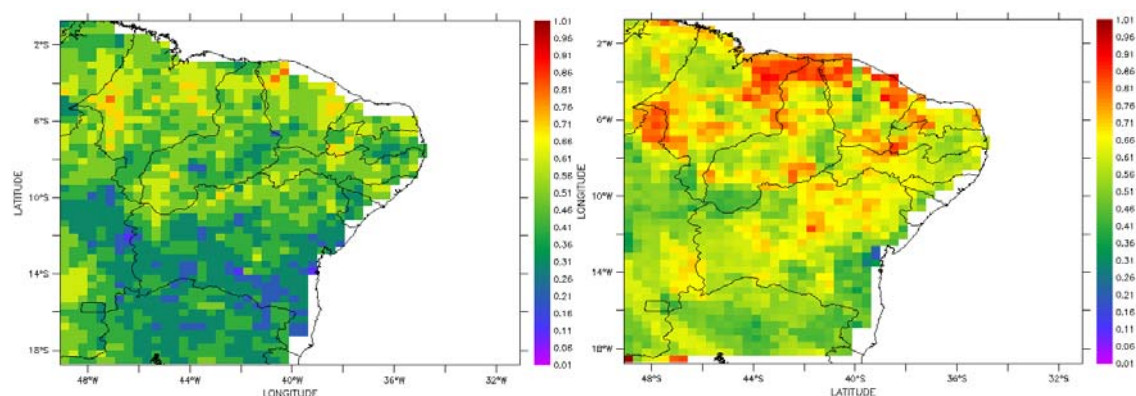


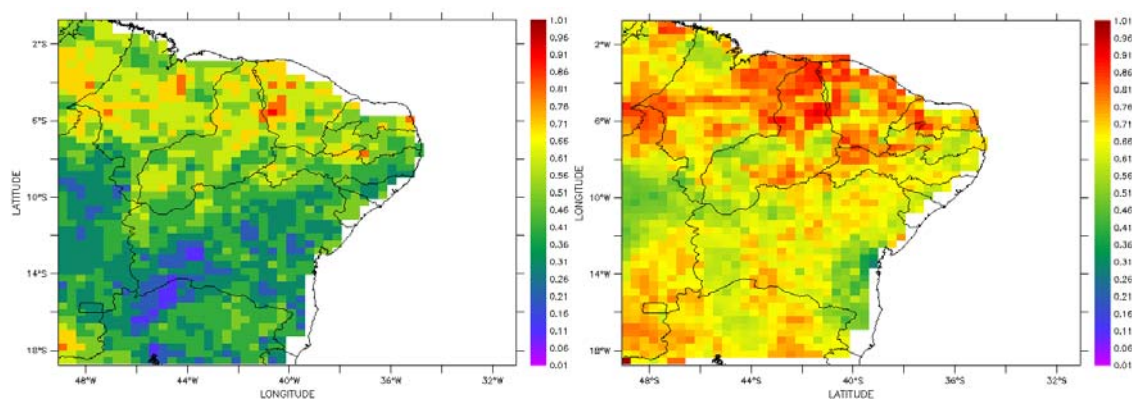
Figura 4.11 – Como na Figura 4.10, mas para o coeficiente de correlação de Spearman.

A Figura 4.12 mostra a comparação das áreas ROC-ABOVE entre os resultados não corrigidos do modelo regional e a precipitação observada (esquerda) com as áreas ROC-ABOVE entre os resultados corrigidos via CPT e a precipitação observada (direita) para os trimestres JFM, FMA, MAM e AMJ.

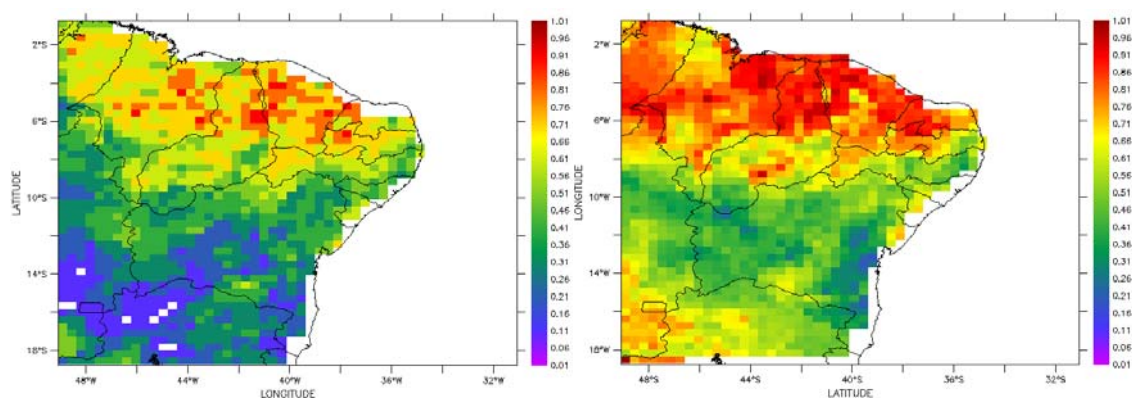
a) trimestre JFM



b) trimestre FMA



c) trimestre MAM



d) trimestre AMJ

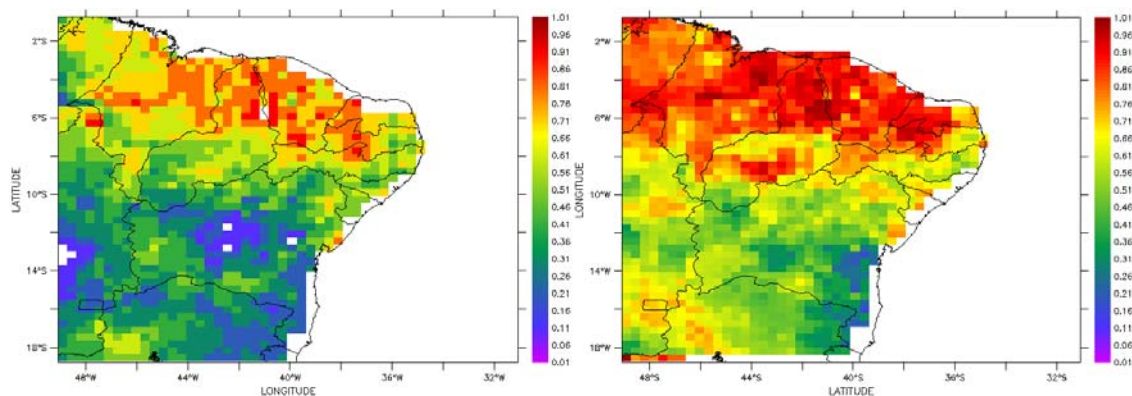
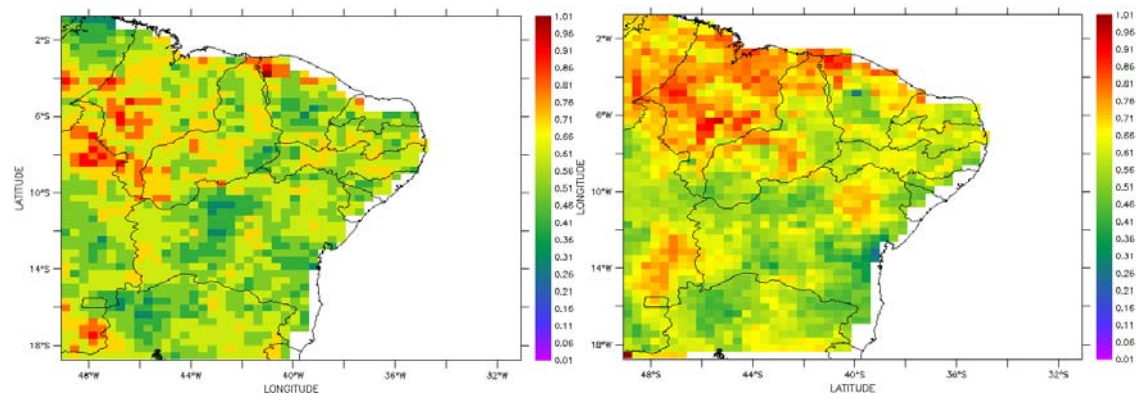


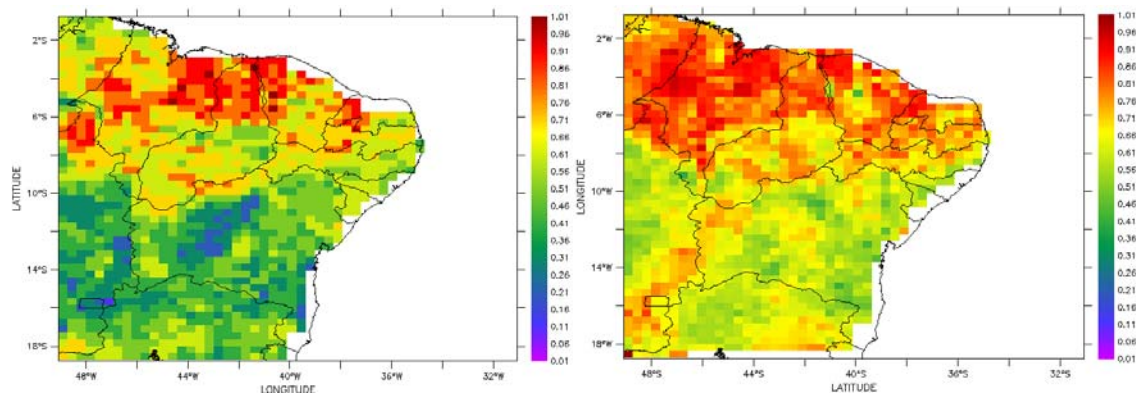
Figura 4.12 – Como na Figura 4.10, mas para as áreas ROC - Above.

Na Figura 4.13 é apresentada a comparação das áreas ROC-BELOW entre os resultados não corrigidos do modelo regional e a precipitação observada (esquerda) e as áreas ROC-BELOW entre os resultados corrigidos via CPT e a precipitação observada (direita) para os trimestres JFM, FMA, MAM e AMJ.

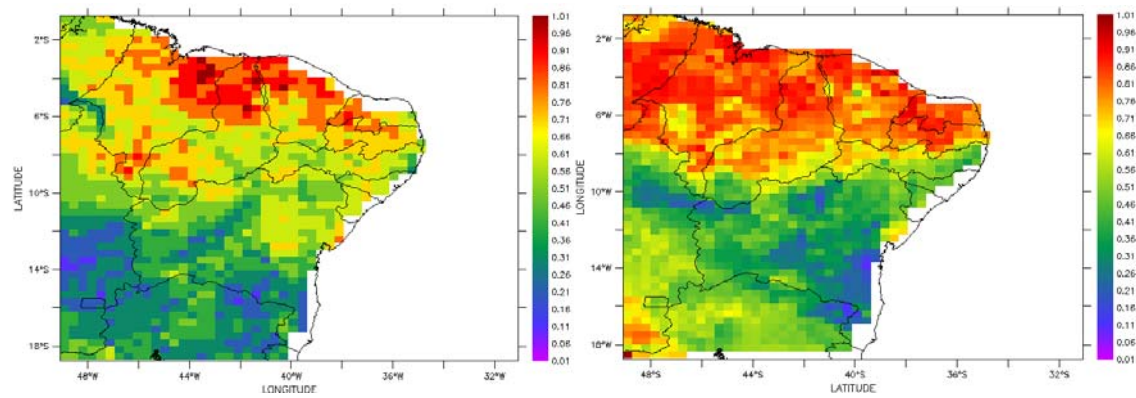
a) trimestre JFM



b) trimestre FMA



c) trimestre MAM



d) trimestre AMJ

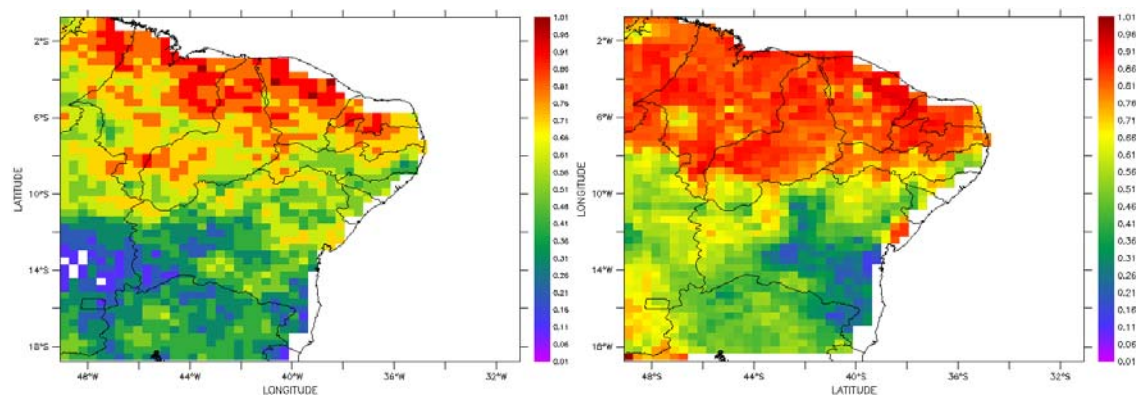


Figura 4.13 – Como na Figura 4.10, mas para as áreas ROC - *Below*.

No que tange à correlação de Pearson, nota-se que em JFM os resultados indicados pelo modelo corrigido são melhorados em todo o Piauí, leste do Maranhão e no Ceará (principalmente Região Jaguaribana e Cariri), enquanto no restante do NE os resultados mantiveram aproximadamente seus valores. Para o trimestre FMA, observam-se melhores resultados praticamente em toda a Bahia, em parte do Piauí e Ceará (especialmente Cariri), além do oeste de Pernambuco. Já para o trimestre MAM não foram percebidas melhorias significativas, de um modo geral. É interessante notar que, em AMJ, período final da estação chuvosa, o modelo puro produziu melhores resultados que o modelo corrigido.

No que se refere à correlação de Spearman, o que se percebe é que em JFM, o modelo corrigido não apresentou resultados satisfatórios para a correção do modelo numérico puro. Em FMA o modelo corrigido só apresentou melhores resultados na Bahia. Para os trimestres MAM e AMJ o modelo corrigido não apresentou bons resultados em contrapartida ao modelo puro.

Analisando as áreas ROC-*Above* no trimestre JFM percebe-se uma significativa melhora do modelo corrigido em todo o NE, principalmente no norte do NE. O que se observa nos trimestres seguintes (FMA, MAM e AMJ) é uma gradativa melhora, trimestre a trimestre, dos índices ROC-*Above* para as mesmas regiões do NE.

Finalmente a análise das áreas ROC-*Below* indica que no trimestre JFM o modelo corrigido apresentou melhores resultados em todo o NE, principalmente no centro do Piauí, norte do Ceará e em todo o Maranhão. No trimestre FMA, os resultados são ainda melhores no NE com destaque para o Maranhão. Analisando o trimestre MAM nota-se que o modelo corrigido não supera o modelo puro na Bahia, mais continua se mantendo melhor no norte do NE. No ultimo trimestre (AMJ) o modelo corrigido permanece indicando bom desempenho para discriminar eventos de menor precipitação.

4.3 PREVISÃO HÍBRIDA

Nesta seção são comparados os resultados do modelo híbrido com os resultados da previsão empírica pura (que usa somente TSM como preditor) e também com os resultados da previsão numérica pura e da previsão numérica corrigida com base de dados de 19 anos. A redução do período se deve à interseção entre os conjuntos de dados utilizados (TSM disponibilizada pela NOAA, de 1981 a 2009 e precipitação simulada pelo *downscaling* da FUNCEME, de 1971 a 2000, sendo que a previsão empírica da precipitação, pelo menos para os primeiros trimestres, requer dados do ano anterior).

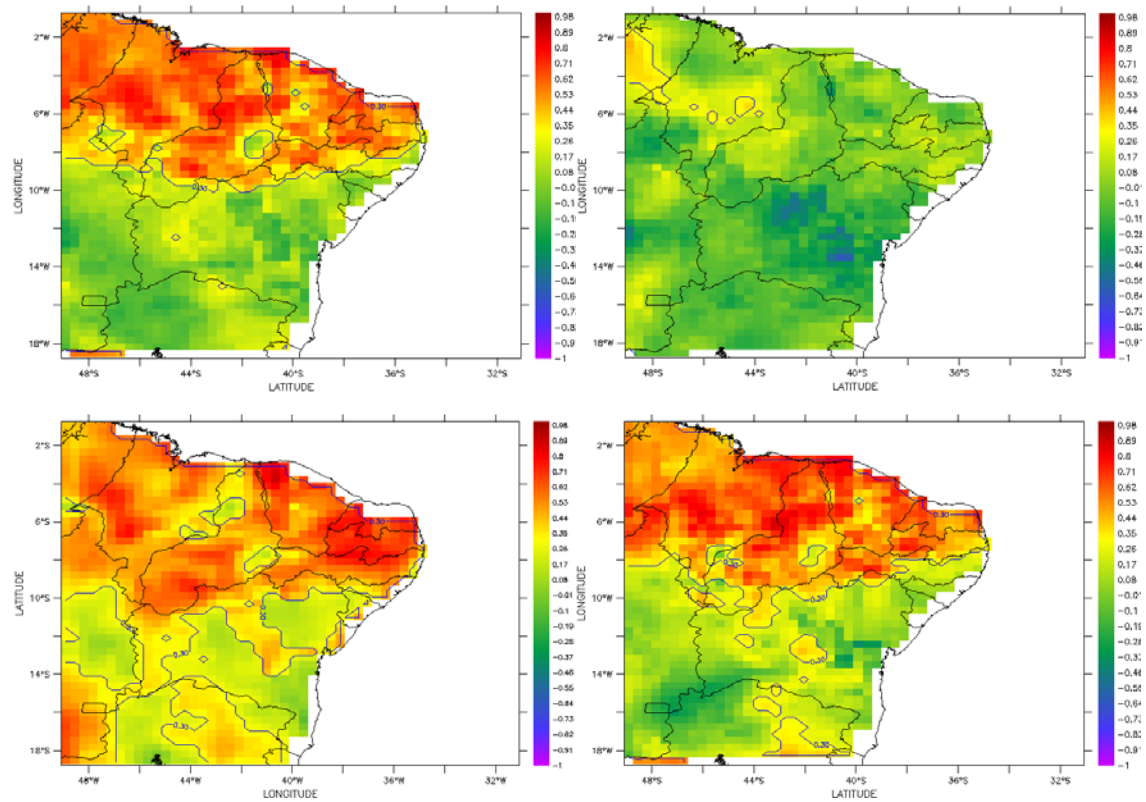
O modelo híbrido faz uso da combinação da TSM com a saída do modelo numérico como preditores, utilizando a precipitação observada (dados da CRU) como preditando.

Para se comparar a qualidade da previsão gerada pelo modelo híbrido com a previsão empírica pura, com a previsão numérica pura e com a previsão numérica corrigida para o NEB, aplicaram-se novamente os coeficientes de correlação de Pearson e Spearman e as áreas ROC.

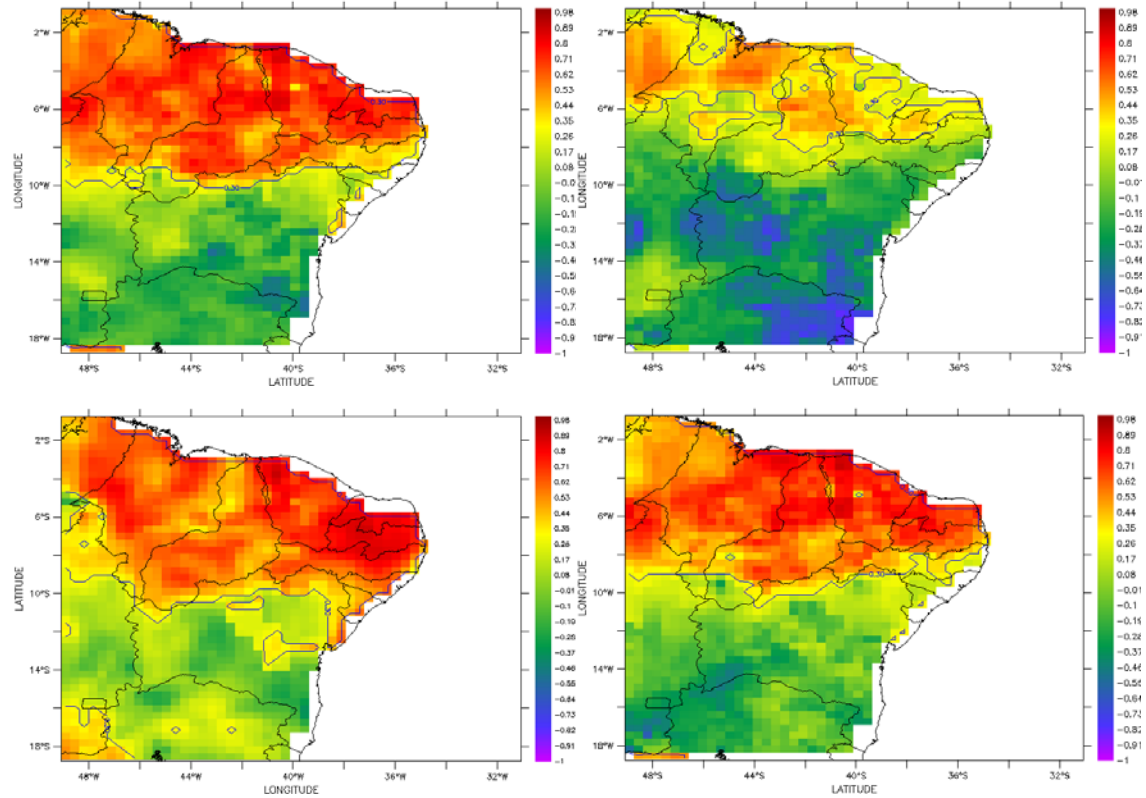
4.3.1 PEARSON

A Figura 4.14 mostra a comparação da correlação de Pearson entre a precipitação observada e o resultado do modelo híbrido (esquerda-acima) com a correlação de Pearson entre a precipitação observada e os resultados, via CPT, da previsão empírica pura (direita-acima), previsão numérica pura (esquerda-abixo) e a previsão numérica corrigida (direita-abaixo) com base de dados de 19 anos para os trimestres JFM a AMJ.

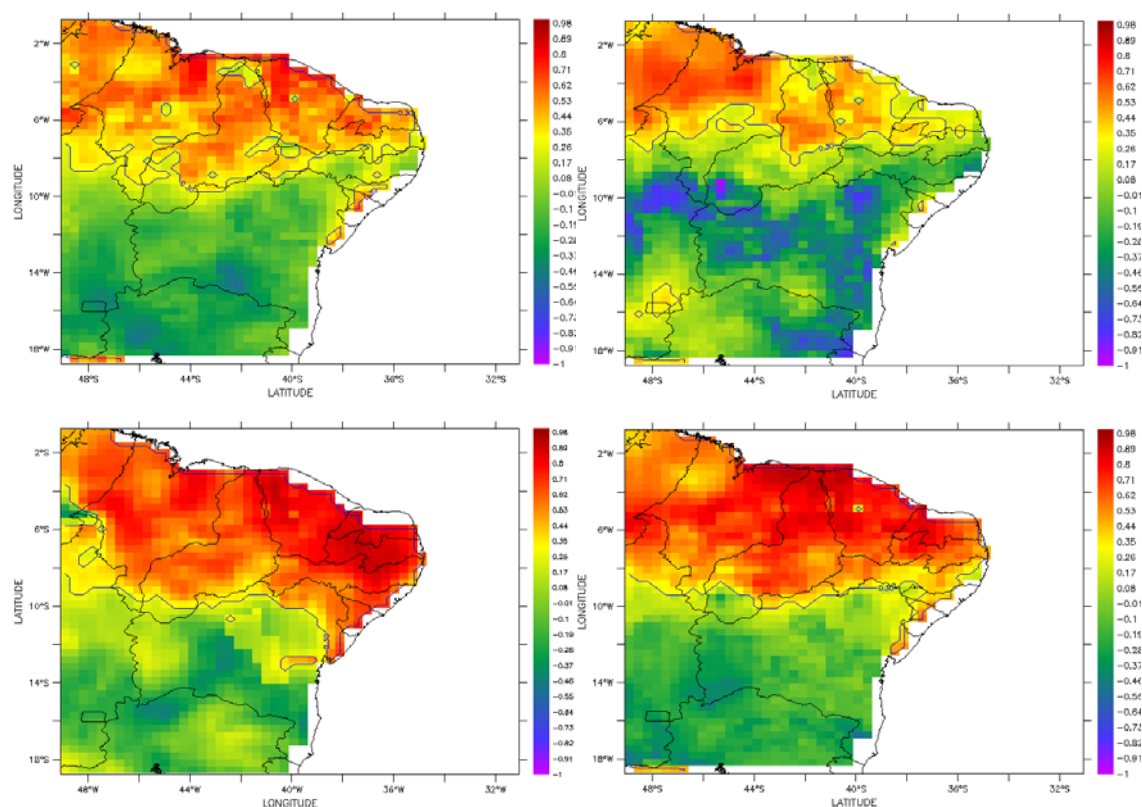
a) trimestre JFM



b) trimestre FMA



c) trimestre MAM



d) trimestre AMJ

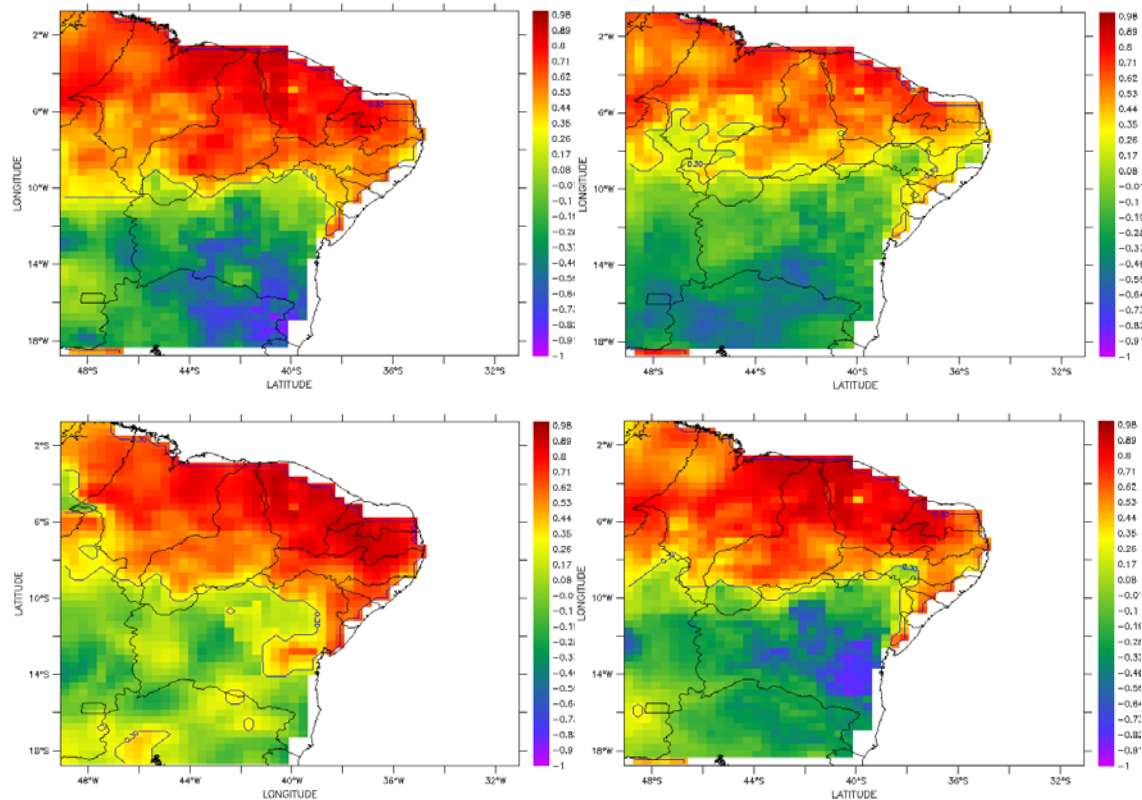


Figura 4.14 – Coeficientes de correlação de Pearson entre a precipitação observada e o resultado do modelo híbrido (esquerda-acima) com a correlação de Pearson entre a precipitação observada

e os resultados, via CPT, da previsão empírica pura (direita-acima), previsão numérica pura (esquerda-abaixo) e a previsão numérica corrigida (direita-abaixo) para os trimestres: (a) JFM, (b) FMA, (c) MAM e d) AMJ.

Os resultados mostram que o modelo híbrido, para o trimestre JFM, superou em muito o modelo empírico puro para todo o NE, mas, quando comparado ao modelo numérico puro apresentou melhores resultados apenas no noroeste do NE, tendo inclusive resultados não muito bons na Bahia. Quando comparado com o modelo numérico corrigido, mostrou-se um pouco melhor apenas no norte do Maranhão e extremo leste do Pará.

Analisando as Figuras comparativas para o trimestre FMA, percebe-se que o modelo híbrido continua superando o modelo empírico puro em todo o NE e, quando comparado com o modelo numérico puro, indica melhores resultados apenas no Maranhão, Piauí e Ceará e piores resultados na Bahia. Novamente se compararmos o modelo híbrido ao modelo numérico corrigido, vemos que ele só apresenta melhores resultados no norte do Maranhão e sul do Piauí e Ceará.

No trimestre MAM, o modelo híbrido só mostra melhores resultados, quando comparado ao modelo empírico puro, percebendo-se, para esse trimestre, a superioridade do modelo numérico puro em relação aos outros.

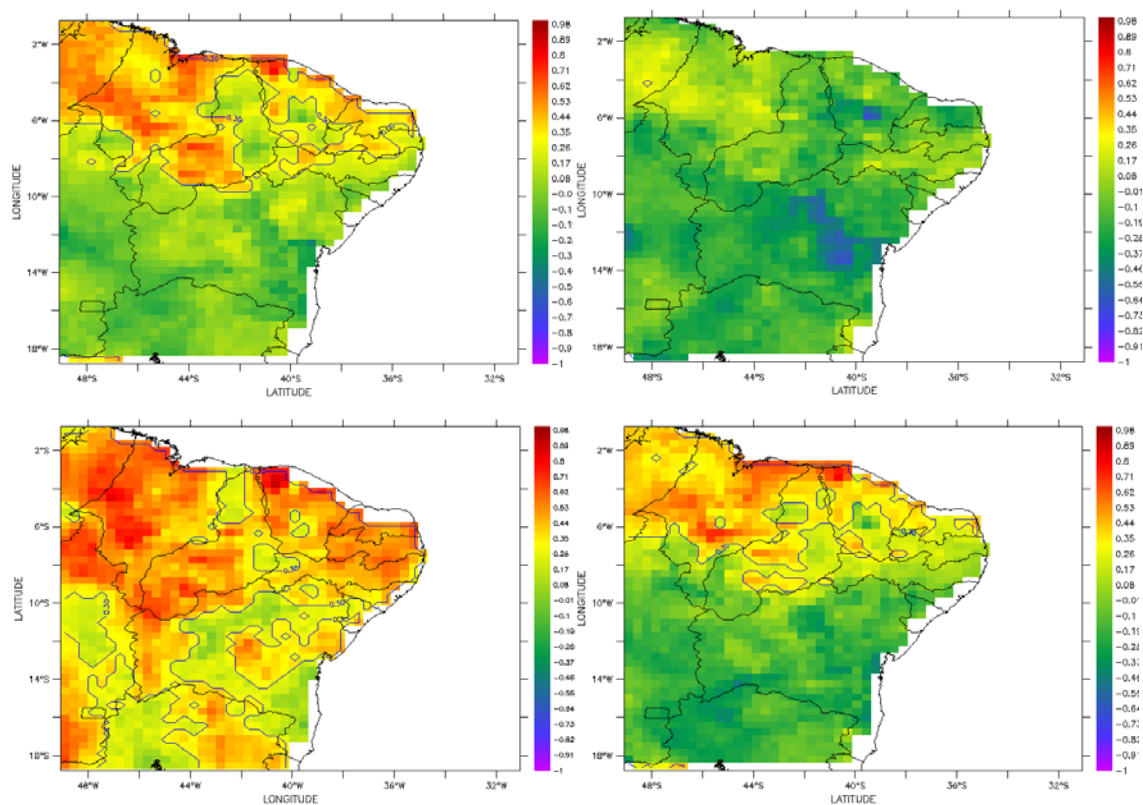
Para o trimestre AMJ, o modelo híbrido mantém sua superioridade em relação ao modelo empírico e continua relativamente melhor que o modelo numérico puro apenas no norte do NE, não sendo tão bom na Bahia. O que se nota em relação ao modelo numérico corrigido é que seus índices são muito semelhantes.

4.3.2 SPEARMAN

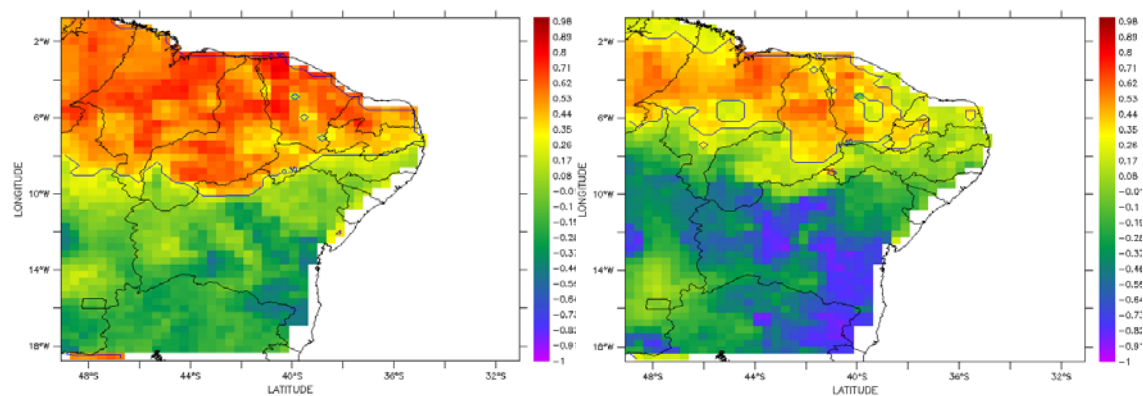
Na Figura 4.15, apresentamos a comparação da correlação de Spearman entre a precipitação observada e o resultado do modelo híbrido (esquerda-acima) com a correlação de Spearman entre a precipitação observada e os resultados, via CPT, da previsão empírica pura (direita-acima), previsão

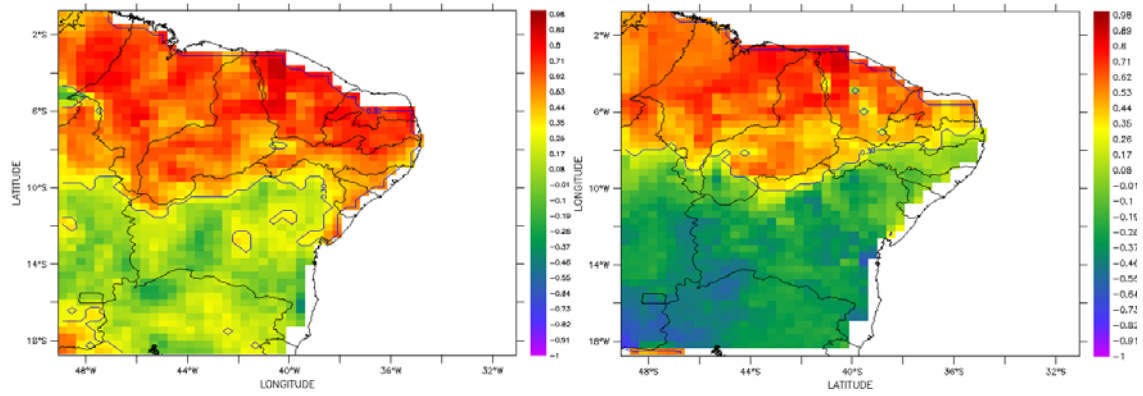
numérica pura (esquerda-abaixo) e a previsão numérica corrigida (direita-abaixo) com base de dados de 19 anos para os quatro trimestres estudados.

a) trimestre JFM

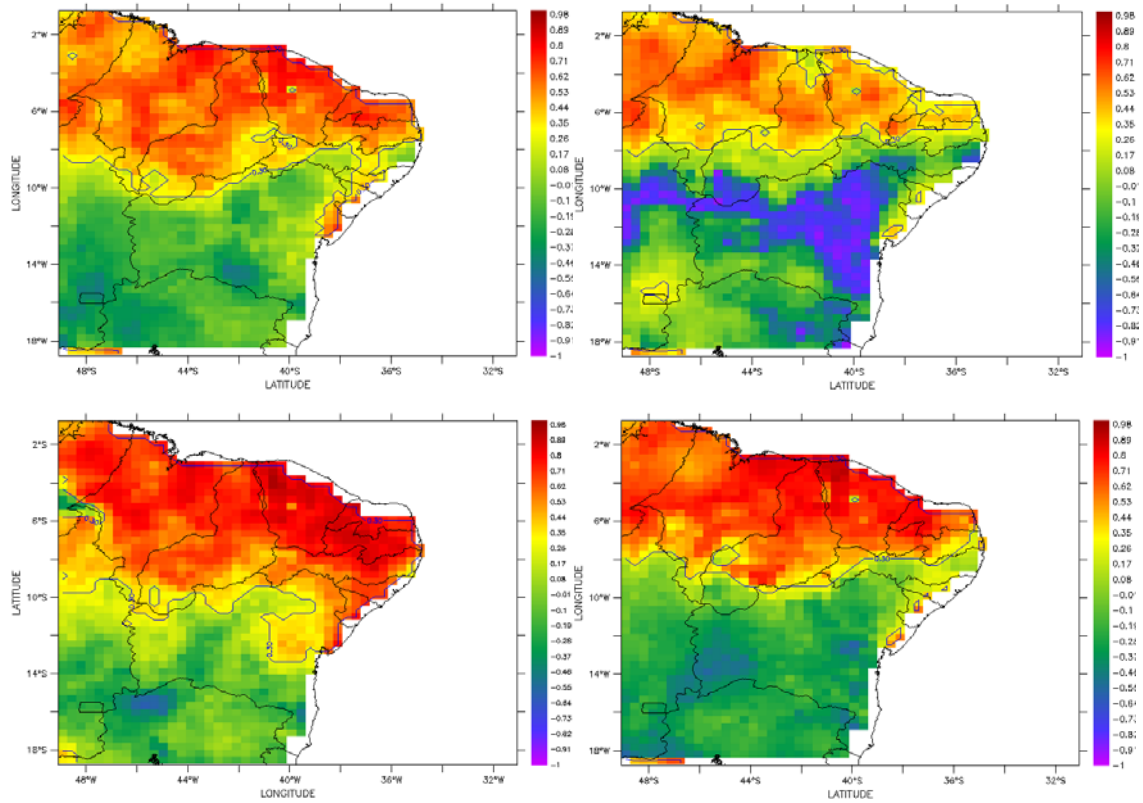


b) trimestre FMA

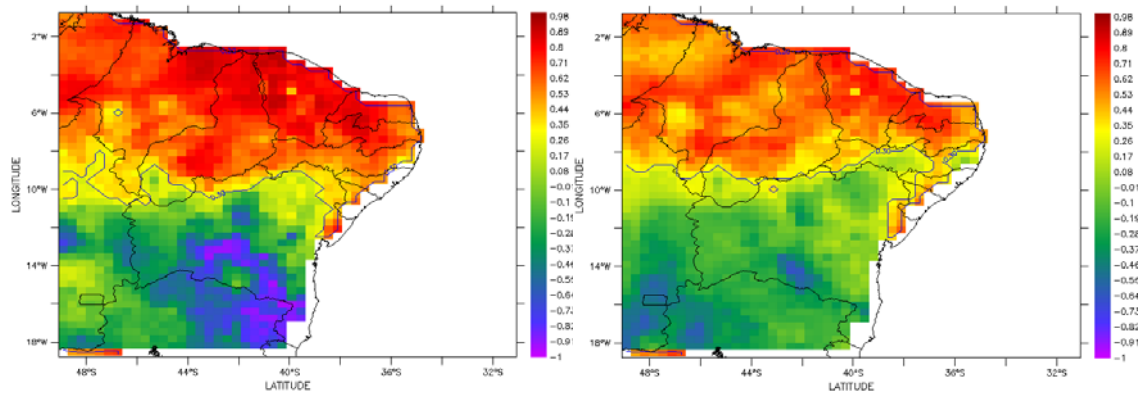




c) trimestre MAM



d) trimestre AMJ



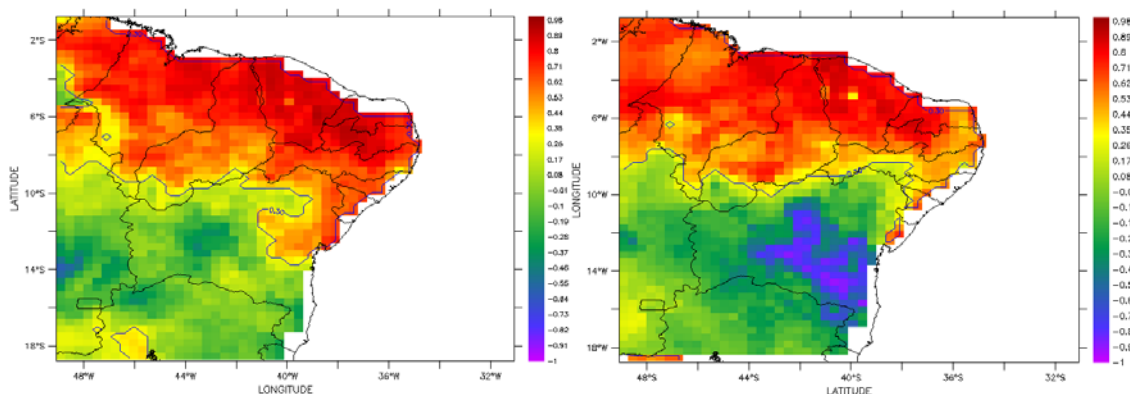


Figura 4.15 – Como na Figura 4.14, mas para o coeficiente de correlação de Spearman.

Por meio desses resultados observamos que o modelo híbrido, para o trimestre JFM, continua superando o modelo empírico puro para todo o NE, no entanto, quando comparado com o modelo numérico puro não apresentou resultados significativamente melhores em nenhuma região do NE. Quando comparado com o modelo numérico corrigido mostrou-se um pouco melhor apenas no Maranhão e centro do Piauí.

Investigando as Figuras para o trimestre FMA, percebe-se que o modelo híbrido ainda supera o modelo empírico puro em todo o NE e, quando comparado com o modelo numérico puro, apresenta semelhanças no Maranhão, Piauí e Ceará, mas com piores resultados para o restante do NE. Ao compararmos o modelo híbrido com o modelo numérico corrigido, notamos grande semelhança nos resultados e pouca melhora apenas na Bahia.

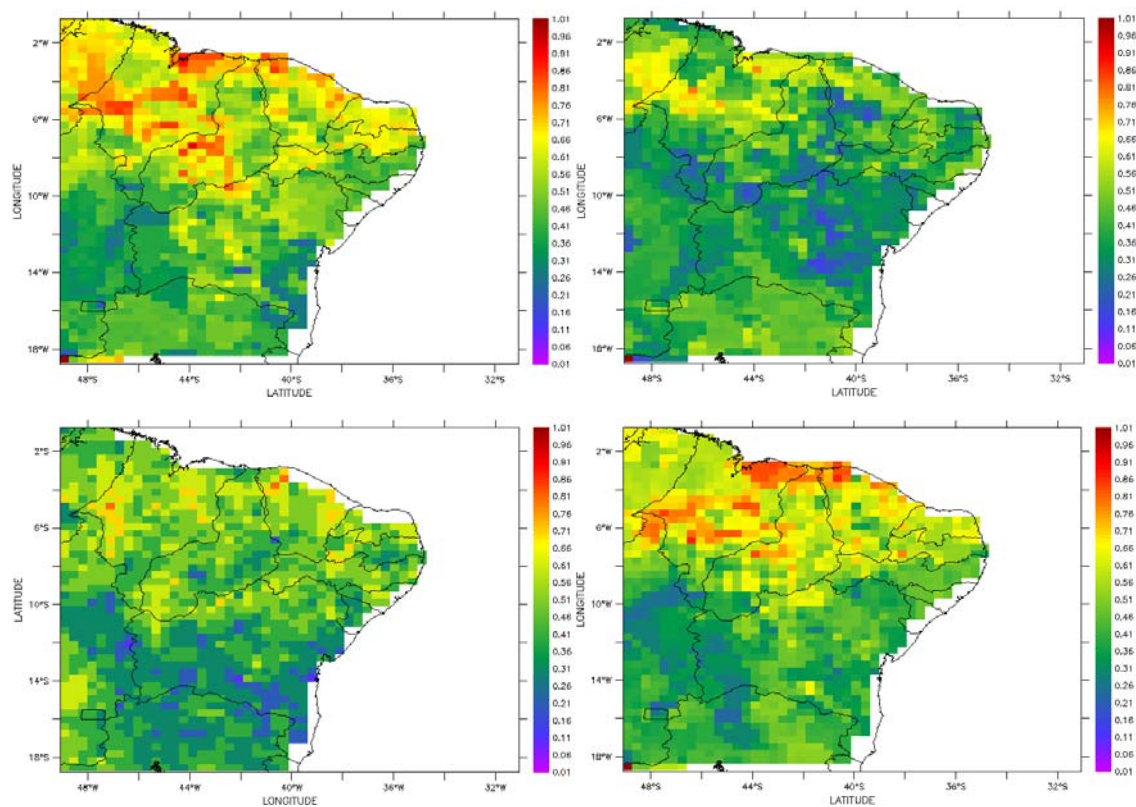
No trimestre MAM, assim como aconteceu com a correlação de Pearson, o modelo híbrido só mostra melhores resultados quando comparado ao modelo empírico puro, percebendo-se novamente que, para esse trimestre, existe uma superioridade do modelo numérico puro em relação aos outros.

Para o trimestre AMJ, o modelo híbrido mantém sua superioridade em relação ao modelo empírico e continua relativamente melhor que o modelo numérico puro apenas no norte do NE, não sendo tão bom na Bahia. O que se nota em relação ao modelo numérico corrigido e ao modelo híbrido é que seus índices são bastante semelhantes.

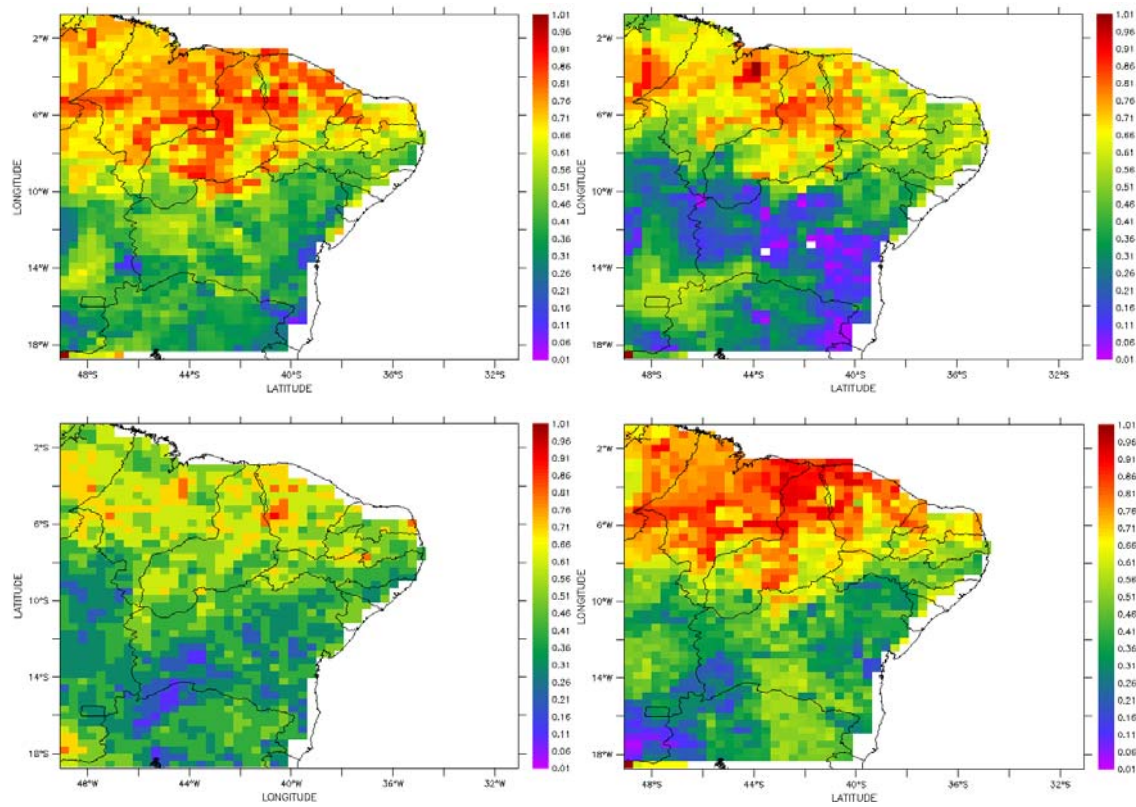
4.3.3 ÁREAS ROC - ABOVE

Através da Figura 4.16, é feita a comparação das Áreas ROC - Above entre a precipitação observada e o resultado do modelo híbrido (esquerda-acima) com as Áreas ROC - Above entre a precipitação observada e os resultados, via CPT, da previsão empírica pura (direita-acima), previsão numérica pura (esquerda-abixo) e a previsão numérica corrigida (direita-abaixo) com base de dados de 19 anos para os trimestres JFM a AMJ.

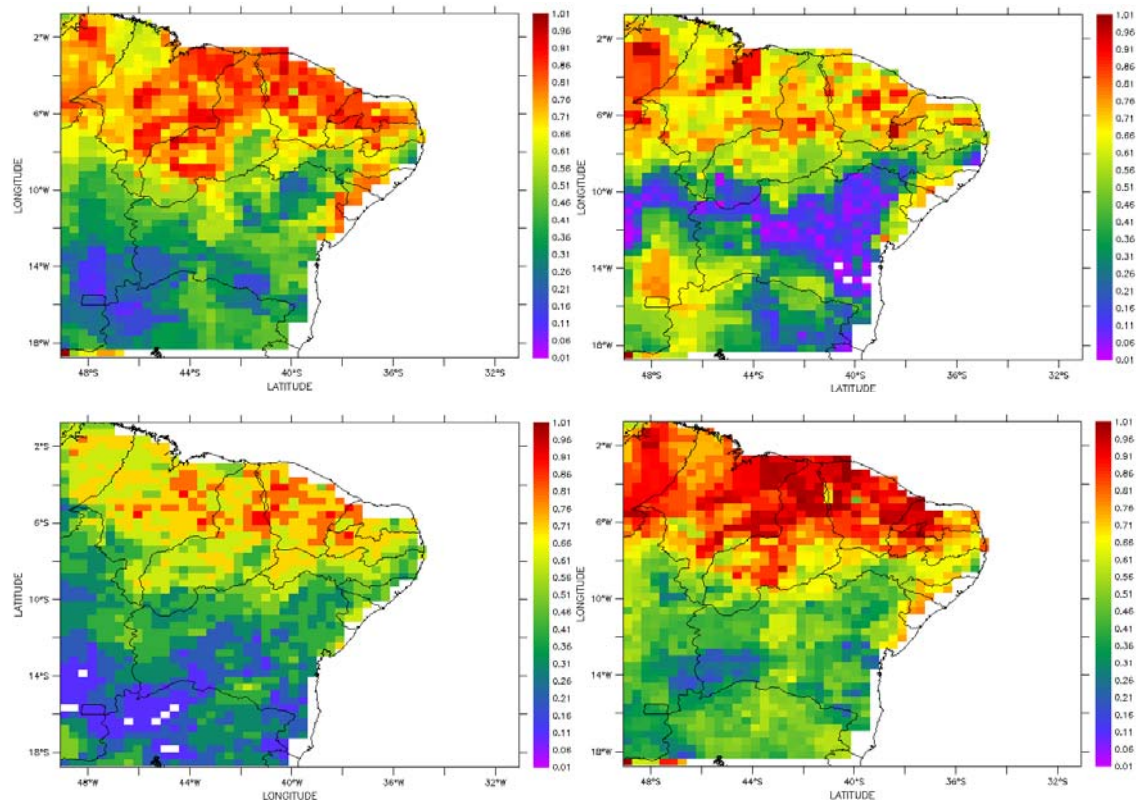
a) trimestre JFM



b) trimestre FMA



c) trimestre MAM



d) trimestre AMJ

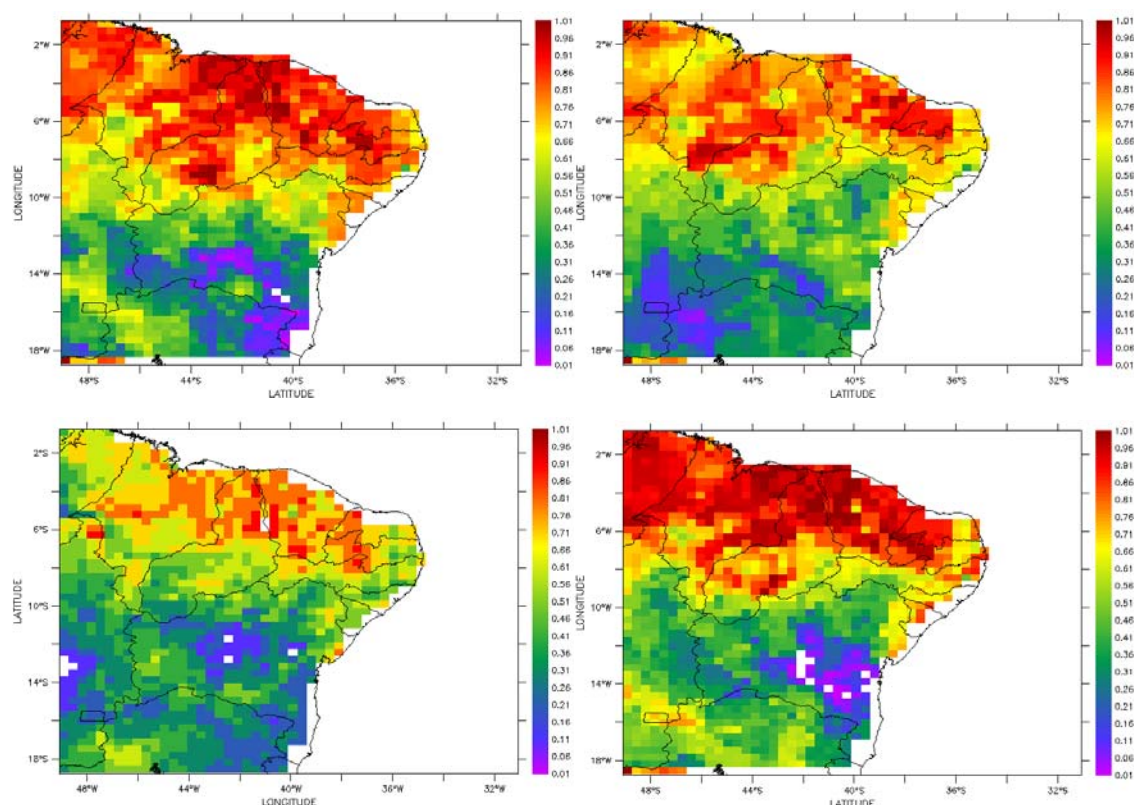


Figura 4.16 – Como na Figura 4.14, mas para as áreas ROC - Above.

Os resultados observados, para o trimestre JFM, indicam que o modelo híbrido apresenta valores mais altos para o índice ROC-Above, superando o modelo empírico e o modelo numérico puro para praticamente todo o NE, no entanto, quando comparado com o modelo numérico corrigido, o ganho não é tão bom assim.

Nas Figuras que representam o trimestre FMA, percebe-se que o modelo híbrido ainda continua superando o modelo empírico e o modelo numérico puro em todo o NE. Ao compararmos o modelo híbrido com o modelo numérico corrigido, notamos uma semelhança nos resultados no norte do NE e uma pequena perda na Bahia.

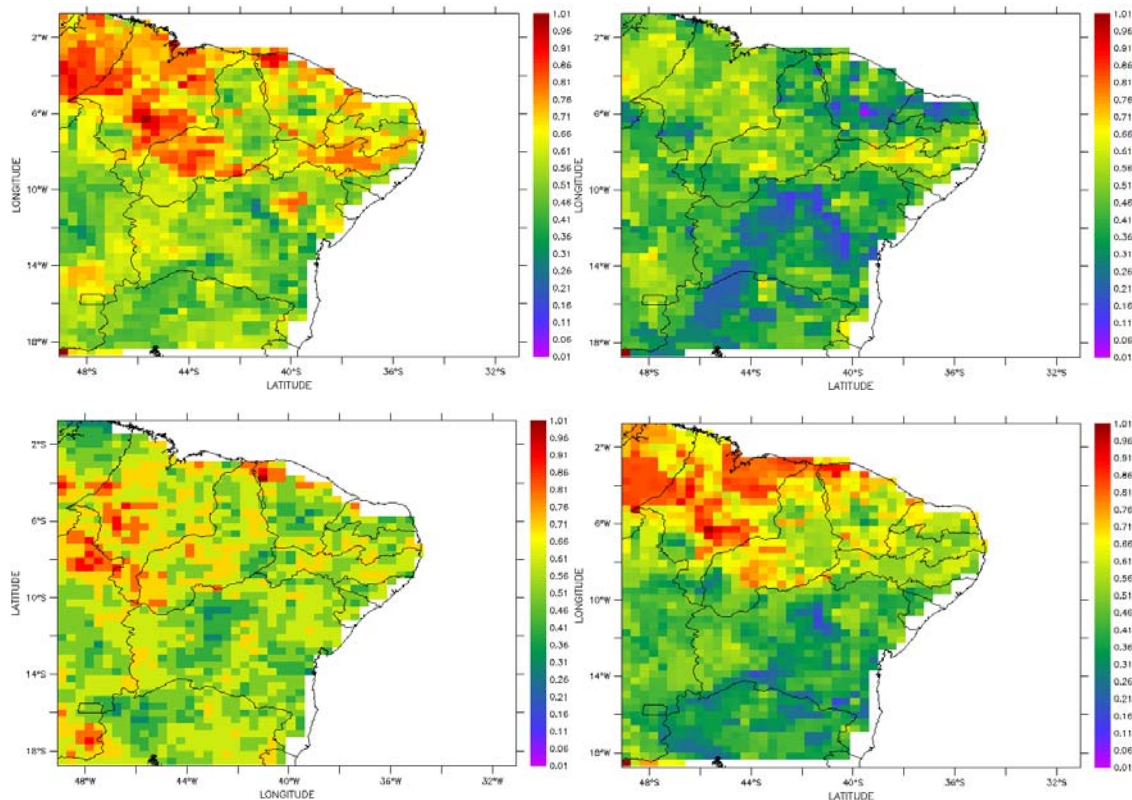
No trimestre MAM, assim como aconteceu com o trimestre anterior, o modelo híbrido mostra melhores resultados quando comparado ao modelo empírico e o modelo numérico puro, no entanto, o modelo numérico corrigido indica, de um modo geral, melhores resultados principalmente no Maranhão, Piauí e Ceará.

Para o trimestre AMJ o modelo híbrido mantém sua superioridade em relação ao modelo empírico e ao modelo numérico puro em todo NE, porém, em relação ao modelo numérico corrigido, ele volta a apresentar seus índices ROC-*Above* inferiores em todo o NE.

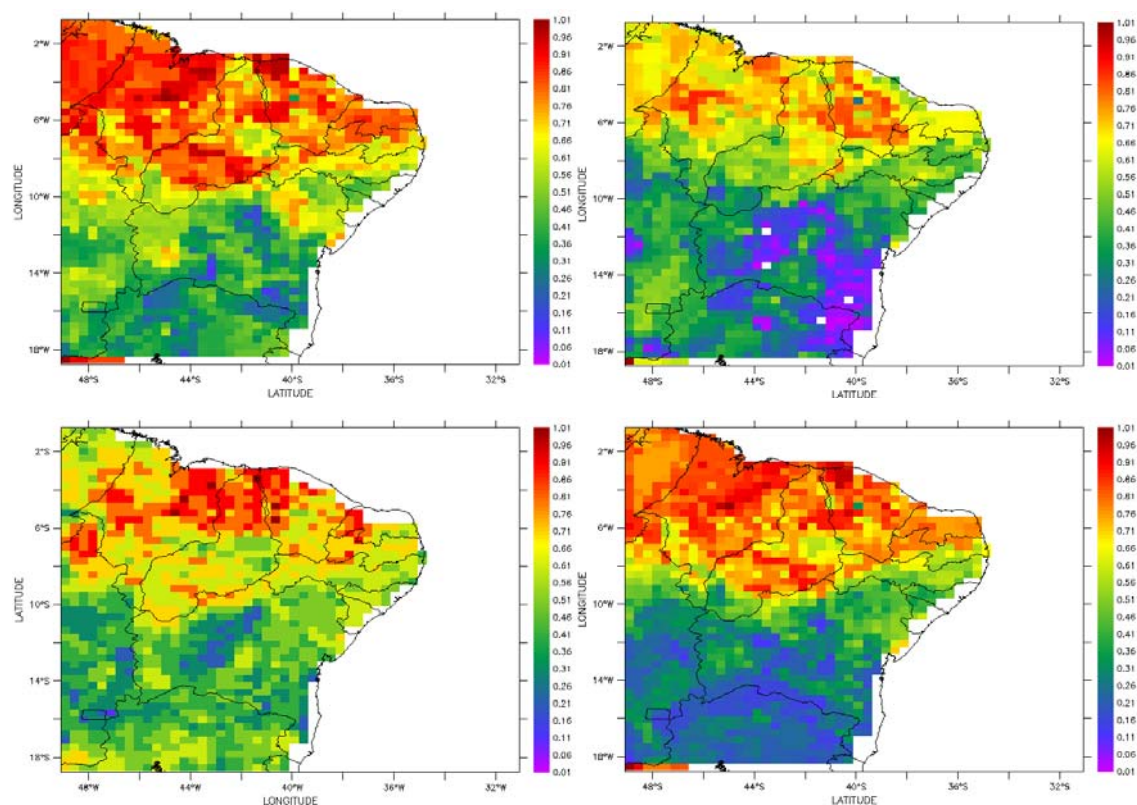
4.3.3 ÁREAS ROC - *BELOW*

Finalmente, a Figura 4.17 mostra a comparação das Áreas ROC - *Below* entre a precipitação observada e o resultado do modelo híbrido (esquerda-acima) com as Áreas ROC - *Below* entre a precipitação observada e os resultados, via CPT, da previsão empírica pura (direita-acima), previsão numérica pura (esquerda-abaixo) e a previsão numérica corrigida (direita-abaixo) com base de dados de 19 anos para os quatro trimestres estudados.

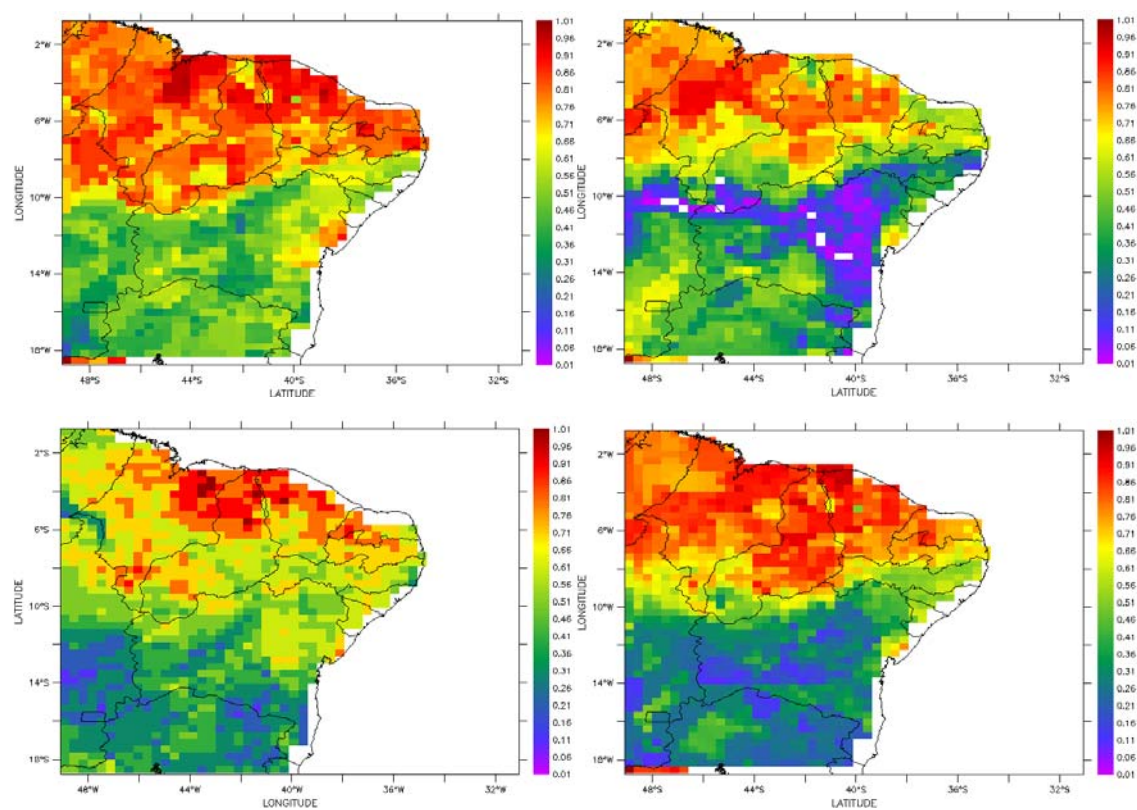
a) trimestre JFM



b) trimestre FMA



c) trimestre MAM



d) trimestre AMJ

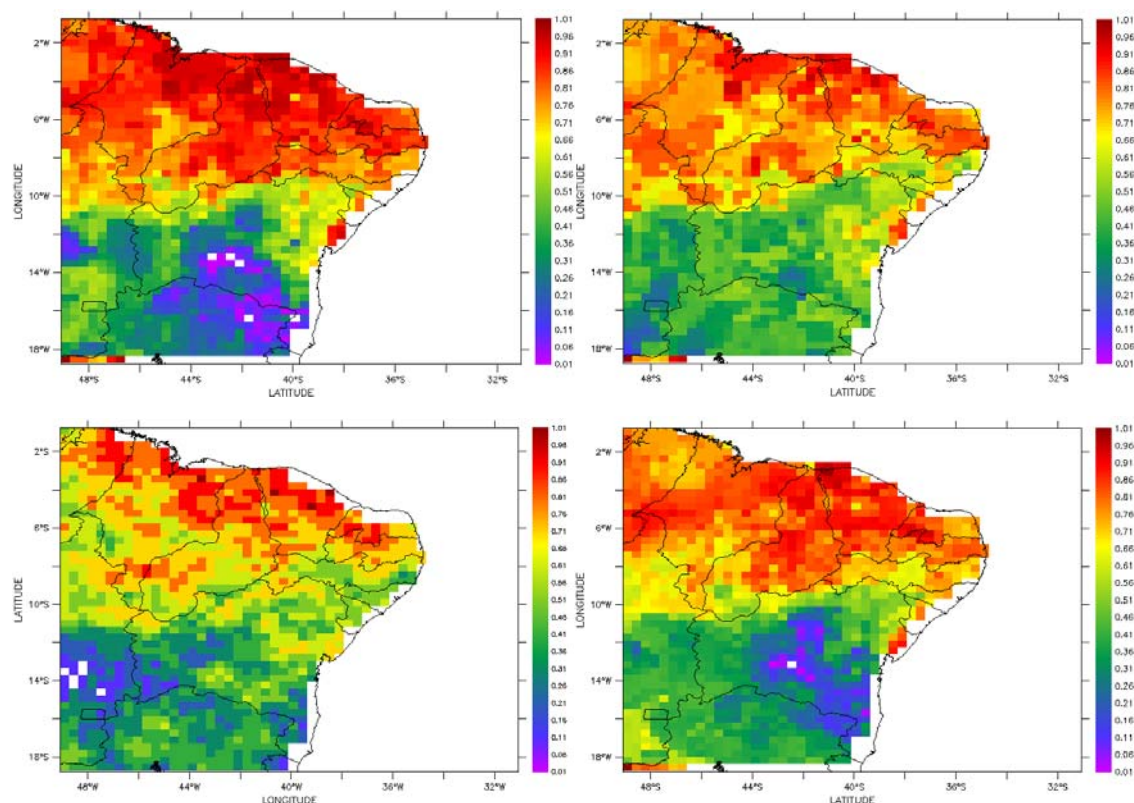


Figura 4.17 – Como na Figura 4.14, mas para as áreas ROC - *Below*.

Os resultados mostram que, para o trimestre JFM, o modelo híbrido superou o modelo empírico para todo o NE. Quando comparado ao modelo numérico puro apresentou melhores resultados apenas no Maranhão, Piauí e Pernambuco, tendo inclusive resultados não muito bons na Bahia. Quando comparado com o modelo numérico corrigido, mostrou-se um pouco melhor apenas no Maranhão, extremo leste do Pará, centro do Piauí e todo Pernambuco.

Avaliando as Figuras para o trimestre FMA, percebe-se que o modelo híbrido continua superando o modelo empírico puro em todo o NE e, quando comparado com o modelo numérico puro, indica melhores resultados no Pará, Maranhão, Piauí, Rio Grande do Norte e Ceará e resultados não muito bons na Bahia. Novamente se compararmos o modelo híbrido ao modelo numérico corrigido, vemos que ele apresenta semelhanças em todo o NE.

No trimestre MAM, o modelo híbrido mostra melhores resultados quando comparado a todos os outros modelos, percebendo-se, para esse

trimestre, a mesma superioridade em todo o NE com respeito às áreas ROC-*Below*.

Finalmente para o trimestre AMJ, o modelo híbrido mantém sua superioridade em relação a todos os outros modelos, mais nesse caso apenas para o norte do NE, pois para a Bahia ele só apresentou melhores resultados quando comparado ao modelo numérico corrigido.

5 CONCLUSÕES E PERSPECTIVAS

Neste trabalho, foram apresentadas diversas possibilidades de uso da ACC aplicada à previsão climática sazonal sobre o Nordeste Brasileiro. Inicialmente, foi construído um modelo de previsão empírica onde os totais de precipitação para os trimestres DJF a AMJ sobre o Ceará, tratados como predictandos, foram previstos utilizando a técnica de ACC. Como preditores, foram usados as médias mensais de TSM em regiões selecionadas do Atlântico e Pacífico com uma defasagem de dois meses. O modelo empírico apresentou, em geral, um desempenho satisfatório, inclusive nos períodos de pré e pós estação chuvosa (geralmente considerados de mais baixa previsibilidade), podendo, portanto, ser usado operacionalmente na previsão sazonal do Estado do Ceará.

Em seguida, propomos o uso da ACC para corrigir a saída dos modelos utilizados no *downscaling* dinâmico da FUNCEME (no caso, esta técnica foi testada com os resultados do RAMS, um dos modelos utilizados na previsão). Foi usada a análise de correlação canônica para a construção de um modelo de correção que utiliza, como preditor, os resultados do modelo numérico puro, durante 30 anos (1971-2000), relacionando, com estes dados, a precipitação observada (predictando). Foram aplicados os coeficientes de correlação de Pearson e Spearman e as áreas ROC para investigar a concordância entre as séries de precipitação observada com o modelo numérico corrigido para a grade deste modelo (que cobre todo o NEB). Os resultados, em geral, mostram que a ACC contribui para elevar os índices de destreza do modelo numérico (em algumas regiões, esta melhoria é mais significativa; em outras, é apenas marginal).

Neste trabalho, a técnica de ACC também foi utilizada para a construção de um modelo híbrido, representado por uma combinação da TSM com a saída do modelo numérico como preditores, tendo a precipitação observada como predictando. O modelo híbrido apresentou resultados

sistematicamente superiores ao modelo empírico, mas seu desempenho não chega a superar o do modelo dinâmico, sem correção, ou com correção. No entanto, embora seja evidente que os resultados aqui apresentados do modelo híbrido não foram nitidamente superiores ao modelo corrigido, deve-se alertar que o modelo dinâmico (e, portanto, o dinâmico corrigido) está com destreza superestimada, pois foi alimentado por um modelo global forçado por TSMs observadas. Essa destreza só seria obtida, portanto, com uma previsão de TSM perfeita. É provável, assim, que, sendo os resultados do modelo dinâmico não tão superiores aos do modelo empírico puro, como mostrado neste trabalho, um modelo híbrido possa efetivamente contribuir para a previsão, ao mesclar informações de modelos com destrezas similares (modelo empírico com base na TSM e modelo dinâmico). Os modestos resultados com a previsão híbrida aqui apresentados são, portanto, ainda inconclusivos.

Será posteriormente estudada a implementação operacional tanto da previsão empírica (como já foi testada na FUNCEME no XI Workshop Internacional de Avaliação Climática para o Semi-árido Nordeste em 2009 e será feita novamente em 2010) quanto da correção e da previsão híbrida.

Enfim, como foi mostrado na modelagem híbrida, o uso de ACC pode ser usada para mesclar preditores, compondo assim uma previsão de “super-ensemble”, juntando as previsões de vários modelos como, por exemplo, os diversos modelos do “super-ensemble” da FUNCEME em fase de implantação (RAMS e RSM forçados por diferentes modelos globais).

REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, M. A.; VETTER, D. A análise de relações entre conjuntos de variáveis na matriz geográfica: correlação canônica. In: FAISSOL, S. Tendências atuais na geografia urbano/ regional: teorização e quantificação. Rio de Janeiro: IBGE, 1978. p. 133-144.

ALVES, J. M. B., et al. A study inter-comparative of statistical-dynamical seasonal forecast precipitation in nordeste's Brazil. *Rev. bras. meteorol.* [online]. 2007, vol.22, n.3 [cited 2009-12-01], pp. 354-372.

AMBRIZZI, T.; COELHO, C. A. S.; UVO, C. B. Exploring the impacts of the tropical Pacific SST on the precipitation patterns over South America during ENSO periods, *Theoretical and Applied Climatology*, 71, 185-197, 2002.

APTE, P. G. -1990-*Textbook of Econometrics*. McGraw- Hill Company, 1st Ed., 570 pp.

ARAKAWA, A.; SCHUBERT, W. H. Interaction of a cumulus ensemble with the largescale environment, Part I. *Journal of the Atmospheric Sciences*, Boston, v. 31, n. 3, p. 674-704, 1974.

BARNETT, T. P.; PREISENDORFER, R. – 1987 - Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.*, 115:1825-1849.

BARNETT, T. P.; GRAHAM, N.; LATIF, M.; PAZAN, S. and WHITE, W. 1993: ENSO and ENSO-related predictability. Part I: Prediction of equatorial

Pacific sea surface temperature with a hybrid coupled ocean–atmosphere model. *J. Climate*, 6, 1545–1566.

BARNSTON, A. G. - 1994 - Linear Statistical short term climate predictive skill in the Northern Hemisphere. *Jornal of Climate*, 7: 1513-1564.

BARNSTON, A. G.; ROPELEWSKI, C. F. Prediction of ENSO episodes using canonical correlation analysis. *Journal of Climate*, v. 5, n. 11, p. 1316-1345, Nov. 1992.

BARNSTON, A. G.; SMITH, T. M. Specification and prediction of global surface temperature and precipitation from global SST using ACC. *Journal of Climate*, v.9, n. 11, p. 2660- 2697, 1996.

BARROSO, L. P. Análise Multivariada. 48^a Reunião da RBRAS e 10^o SEGRO – 7 a 11 de julho de 2003 – Lavras MG. Departamento de Ciências e Exatas. Universidade Federal de Lavras.

BRANKOVIC, C.; PALMER, T. N. 1997: Atmospheric seasonal predictability and estimates of ensemble size. *Monthly Weather Review*, vol. 125, 859-874.

BUIZZA, R.; HOLLINGSWORTH, A.; LALAURETTE, F.; GHELLI, A. Probabilistic predictions of precipitation using the ECMWF ensemble prediction system. *Weather and Forecasting*, v. 14, n. 2, p. 169-189, 1999.

BUNCHAFT, G.; KELLNER, S. R. O. *Estatística sem mistérios*. 2. ed. Petrópolis, RJ: Vozes, 2001. 230 p.

CALLEGARI-JACQUES S. M. *Bioestatística: princípios e aplicações*. Porto Alegre: Artmed, 2006.

CARDOSO, A. O.; SILVA DIAS, P. L. Identificação de trimestres extremos no regime pluviométrico do Sul e Sudeste do Brasil e relação com a TSM. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 2002, Foz de Iguaçu-PR. Anais: SBMet, 2002. 1 CD-ROM.

CARDOSO, A. O.; SILVA DIAS, P. L. Variabilidade da TSM do Atlântico e Pacífico e temperatura na cidade de São Paulo. Revista Brasileira de Meteorologia, Rio de Janeiro, v.19, n.3, p. 307-324, 2004.

CHANG, P.; GIESE, B. S.; JI, L.; SEIDEL, H. F. Decadal Change in the South Tropical Pacific in a Global Assimilation Analysis, Geophys. Res. Lett., 28, 3461–3464, 2001.

COELHO, C. A. S.; STEPHENSON, D. B.; BALMASEDA, M.; DOBLAS-REYES, F. J.; VAN OLDENBORGH, G. J., 2006: Towards an integrated seasonal forecasting system for South America. J. Climate. Vol. 19, No. 15, 3704-3721.

COSTA, A. A. Aplicação de um modelo de correlações canônicas para previsão de precipitação sobre o Estado do Ceará. Fundação Cearense de Meteorologia e Recursos Hídricos. Fortaleza-Ce. 5p. 2007.

DIAZ, A. F.; STUDZINSKI, C. D.; MECHOSO, C. R. Relationships between precipitation anomalies in Uruguay and Southern Brazil and sea surface temperature in the Pacific and Atlantic oceans. Journal of Climate, Boston, v. 11, n. 2, p. 251-271, 1998.

DICKINSON, R. E.; ERRICO, R. M. et al. A regional climate model for the western U.S. Clim. Change, Berlin, v.15, 383-422. 1989.

DINIZ, G. B. *Determinação de preditores informativos visando um modelo de previsão climática de temperatura máxima e mínima para regiões homogêneas do Rio Grande do Sul*. 2002. 167 f. Tese (Doutorado - Agrometeorologia) – Programa de Pós-Graduação em Fitotecnia, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.

FEDDERSEN, H.; NAVARRA, A.; WARD, M. N. Reduction of Model Systematic Error by Statistical Correction for Dynamical Seasonal Predictions. *Journal of Climate*, v. 12, p.1974-1989,1999.

GLAHN H. R. Canonical correlation analysis and its relationship to discriminant analysis and multiple regression. *Journal of Atmosphere Science*, v. 25, n. 1, p. 23-31, Jan.1968.

GLAHN, H. R.; LOWRY, D. A. The use of model output statistics (mos) in objective weather forecasting. *J. Appl. Meteor.*, v. 11, p. 1203-1211, Dec. 1972.

GRAHAM, N. E.; MICHAELSEN, J.; BARNETT, T. P. - 1987a - An Investigation of El Niño-Southern Oscillation cycle with statistical models. Part 1: Predictor field characteristics. *J. Geophys. Res.* 92:14251-14270.

GRAHAM, N. E.; MICHAELSEN, J.; BARNETT, T. P. – 1987b - An Investigation of El Niño-Southern Oscillation cycle with statistical models. Part 2: Model results. *J. Geophys. Res.* 92:14271-14289.

GRIMM, A. M.; FERRAZ, S. E. T.; GOMES, J. Precipitation anomalies in Southern Brazil associated with El Niño and La Niña events. *Journal of Climate*, Boston, v.11,n.11, p. 2863-2880, 1998.

HARRINGTON, J. Y. *The effects of radiative and microphysical processes on simulated warm and transition season arctic stratus*. 1997. 239 p. Tese (PhD in Meteorology), Colorado State University, Fort Collins.

HASTENRATH, S.; HELLER, L. (1977): Dynamics of climatic hazards in north-east Brazil. *Quart. J. R. Meteor. Soc.*, 110, 411-425.

HASTENRATH, S.; LAMB, P., 1977. *Climate Atlas of the Tropical Atlantic and Eastern Pacific Oceans*. University of Wisconsin Press, Madison, WI.

HASTENRATH, S. (1984): Interannual variability and annual cycle: mechanisms of circulation and climate in the tropical Atlantic. *Mon. Wea. Rev.*, 112, 1097-1107.

HOTELLING, H., 1936: Relationships between two, sets of variates. *Biometrika*, 28, 321-377

HOREL, H. D., 1981: A rotated principal component analysis of the interannual variability of the Northern Hemisphere 500 mb height field. *Monthly Weather Review*. Vol. 109 (10), 2080-2092.

JACKSON, J. E, 1991. *A user Guide to Principal Components*, John Wiley and Sons Jolliffe IT, 1986. *Principal Component Analysis*. Nova York, Springer-Verlag.

Johnson, Richard A.; Wichern, Dean W. *Applied Multivariate Statistical Analysis*, 3ª edição, 1992

JUANG, H. M. H.; KANAMITSU, M. The NMC nested regional spectral model. *Monthly Weather Review*, Boston, v.122, n. 1, p. 3-26, 1994.

KHAN, V. M.; KIM, I. S.; SARAIVA, J. M. A relação entre as anomalias de TSM no Hemisfério Sul com as anomalias da precipitação no Sul do Brasil. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 10.; CONGRESSO DA FLISMET, 8., 1998, Brasília. Anais... Rio de Janeiro: SBMet, 1998. 1 CD-ROM.

KUMAR, A.; HOERLING, M. P.; JI, M.; SARDESHMUKH, P., 1996, 'Assessing a GCM's suitability for making seasonal predictions', J. Climate, 9, 115–129.

KUO, H. L. Further studies of the parameterization of the influence of cumulus convection on large-scale flow. J. Atmos. Sci., v. 31, n. 5, p.1232-1240, 1974.

LATIF, M.; FLÜGEL, M., 1991. 'An investigation of short-range climate predictability in the tropical Pacific', J. Geophys. Res., 96, 2661–2673.

LEE, G. W.; XU, M.; WILSON, J. W.: Improved precipitation forecast by correcting phase and intensity errors of a meso-scale numerical model. National Center for Atmospheric Research, Boulder, Colorado, 2007.

LEE, G. W.; ZAWADZKI, I. 2006: Correction of phase and amplitude errors in numerical precipitation forecasts. European conference on radar meteorology and hydrology, Barcelona, Spain.

LEUNG, L. R.; HAMLET, A. F.; LETTENMAIER, D. P.; KUMAR, A. "Simulations of the ENSO Hydroclimate Signals in the Pacific Northwest Columbia River Basin", Bull. Am. Soc., 80(11), 2313-2329, 1999

LIVEZEY, R. E., 1995: The evaluation of forecasts. Analysis of Climate Variability: Applications of Statistical Techniques, H. von Storch and A. Navarra, Eds., Springer-Verlag, 177–195.

LIVEZEY, R. E.; SMITH, T. M. 1999a: Considerations for use of the Barnett and Preisendorfer (1987) algorithm for canonical correlation analysis of climate variations. *J. Climate*, 12, 303–305.

MARKHAN, C. G.; MCLAIN, D. R., 1977. ‘Sea surface temperature related to rain in Ceara, northeast Brazil’, *Nature*, 265, 320-323.

MARQUES, J. R. Variabilidade espacial e temporal de precipitação pluvial no Rio Grande do Sul e sua relação com indicadores oceânicos. 2005. 210f. Tese (Doutorado-Agrometeorologia) – Programa de Pós-graduação em Fitotecnia, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.

MASON, S.; GRAHAM, N. Conditional probabilities, relative operating characteristics and relative operating levels. *Weather and Forecasting*, v. 14, n. 5, p. 713-725, 1999.

MASS, C. F, 2003: IFPS and the future of the National Weather Service. *Wea. Forecasting*, 18, 75-79.

MELLOR, G. L.;YAMADA, T. Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysical Physics and Space Physics*, v. 20, p. 851.875, 1982.

MENEZES, R. H. N. Tendência prognóstica das chuvas no nordeste do Brasil em relação às anomalias de TSM do Atlântico e Pacífico tropicais. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 10.; CONGRESSO DA FLISMET, 8., 1998, Brasília. Anais... Rio de Janeiro: SBMet, 1998. CL-98141. 1 CD-ROM.

MICHAELSON, J., 1987: Cross-validation in statistical climate forecast models. *J. Climate Appl. Meteor.*, 26, 1589–1600.

MINUZZI, R. B.; SEDIYAMA, G. C. Relação do dipolo do Atlântico tropical com o período chuvoso da região Sudeste do Brasil. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 13, 2004, Fortaleza. Anais... Rio de Janeiro: SBMet, 2002. PL-0057. 1 CD-ROM.

MO, R.; STRAUS, D. M., 2002: Statistical-dynamical seasonal prediction based on principal component regression of GCM ensemble integrations. *Mon. Wea. Rev.*, 130, 2167-2187.

MONTECINOS, A.; DÍAZ, A.; ACEITUNO, P. Seasonal diagnostic and predictability of rainfall in Subtropical South America based on Tropical Pacific SST. *Journal of Climate*, v 13, n. 4, p. 746-758, 2000.

MOURA, A. D.; SHUKLA, J. On the dynamics of the droughts in Northeast Brazil: Observations, theory and numerical experiments with a general circulations model. *Journal of the Atmospheric Science*, v.38, n.12, p. 2653-2673, 1981.

NEELIN, J. D., 1989: Interannual Oscillations in an ocean GCMsimple atmospheric model, *Philos. Trans. Roy. Soc. London Ser A*, 329, 189–205.

NEELIN, J. D., 1990: A hybrid coupled general circulation model for El Niño studies, *J. Atmos. Sci.*, 47, 674–693.

NEW, M.; HULME, M.; JONES, P. D.: Representing twentieth century space-time climate variability. Part 1: development of a 1961-90 mean monthly terrestrial climatology. *Journal of Climate* 12, 829-856, 1999.

NEW, M.; LISTER, D.; HULME, M.; MAKIN, I.: A high-resolution data set of surface climate over global land areas. *Climate Research* 21, 1-25, 2001.

NOBRE, P.; SHUKLA, J. 1996. Variations of sea surface temperature, wind stress, and rainfall over the tropical Atlantic and South America. *Journal of Climate*. v. 9, p.2464–2479.

PALMER, T. N.; ANDERSON, D. L. T., 1994. ‘The prospects for seasonal forecasting’, *Q. J. Roy. Meteorol. Soc.*, 120, 755–793.

PEZZI, L. P. - 1998 – Previsões Regionais Sazonais e Comparações com Observações para o Sul do Brasil durante o Episódio El Niño Oscilação Sul 97/98. Workshop sobre o El Niño 1997/98: Previsões, Impactos e Lições para o Futuro. Anais do X CBMet e VII FLISMET. Brasília – DF - EN98014.

PEZZI, L. P.; UBARANA, V.; REPELLI, C. Desempenho e previsão de um modelo regional estatístico para a Região Sul do Brasil. *Revista Brasileira de Geofísica*, v. 18, n.2, p. 129-146, 2000.

PIELKE, R. A.; *Mesoscale Meteorological Modeling*. Academic Press, EUA, 1984, 612 p.

_____ et al.; A Comprehensive Meteorological Modeling System – RAMS. *Met. Atmos. Phys.*, 49, p. 69-91. 1992.

PISCIATTANO, G.; DÍAZ, A.; CAZES, G.; MECHOSO, C. El Niño-Oscillation impact on rainfall in Uruguay. *Journal of Climate*, Boston, v.7, n.8, p. 1286 1301, 1994.

REPELLI, C. A.; ALVES, J. M. B., 1996. Uso de Análise de Correlações Canônicas para Prognosticar a Variabilidade Espacial da Precipitação Sazonal

Sobre o Nordeste do Brasil. Revista Brasileira de Meteorologia, v. 11, 1/2, 67-75.

REPELLI, C. A.; NOBRE, P. 2004: Statistical prediction of sea surface temperature over the tropical Atlantic. International Journal Climatology, 24, p.45–55.

REYNOLDS, R. W., 1988: A real-time global sea surface temperature analysis. J. Climate, 1, 75-86.

REYNOLDS, R. W.; SMITH, T. M. 1994: Improved global sea surface temperature analyses. J. Climate, 7, 929-948.

REYNOLDS, R. W., RAYNER, N. A.; SMITH, T. M.; STOKES, D. C.; WANG, W., 2002: An improved in situ and satellite SST analysis for climate. J. Climate, 15, 1609-1625.

SARDA, J.; PLAUT, G.; PIRES, C.; VAUTARD, R. 1996: Statistical and dynamical long-range atmospheric forecasts: Experimental comparison and hybridization. Tellus, 48A, 518–537.

SIEGEL, SIDNEY. *Estatística não-paramétrica: para as ciências do comportamento*. São Paulo: McGraw-Hill do Brasil, 1975. 350 p.

SILVA, E. M. et al . Uma aplicação de conjuntos difusos na otimização do prognóstico de consenso sazonal de chuva no Nordeste do Brasil. Rev. bras. meteorol., São Paulo, v. 22, n. 1, Apr. 2007

SILVA, V. P. R. *Comportamento estatístico de séries temporais de pluviometria no Nordeste do Brasil*. Campina Grande: DCA/ CCT/UFPB, 1992. 81p. Dissertação Mestrado.

SILVA, V. P. R.; SOUSA, F. A. S.; CAVALCANTI, E. P.; SOUZA, E. P.; SILVA, B. B. Teleconnections between sea-surface temperature anomalies and air temperature in northeast Brazil. *Journal of Atmospheric and Solar-Terrestrial Physics*, v. 68, p. 781-792, 2006.

SMITH, R. E.; LIVEZEY, R. E., Covariability of aspects of North American Climate with Global Sea Surface Temperatures on Interannual to Interdecadal Timescales. *J. of Climate*, 12, 289-302, 1999.

STEED, R. C.; MASS, C. F. 2004: Bias removal on a mesoscale forecast grid. [Available online from <http://www.mmm.ucar.edu/mm5/workshop/ws04/Session2/Steed.Rick.pdf>]

STEPHENSON, D. B., 2000: Use of the odds ratio for diagnosing forecast skill. *Wea. Forecasting*, 15(2), 221-232.

STERN, W.; MIYAKODA, K., 1995. 'Feasibility of seasonal forecasts inferred from multiple GCM simulations', *J. Climate*, 8, 1071–1085.

STEYN, D. G.; MCKENDRY, I. G. Quantitative and Qualitative Evaluation of a Three-Dimensional Mesoscale Numerical Model Simulation of a Sea Breeze in Complex Terrain. *Monthly Weather Review*, v. 116, n. 10, p. 1914-1926, out. 1988.

STOCKDALE, T. N.; ANDERSON, D. L. T.; ALVES, J. O. S.; BALMASEDA, M. A., 1998b. 'Global seasonal rainfall forecasts using a coupled ocean–atmosphere model', *Nature*, 392, 370–373.

STUDZINSKI, C. D. – 1995 – *Um estudo da precipitação na Região Sul do Brasil e sua Relação com os Oceanos Pacífico e Atlântico Tropical Sul*. São José dos Campos. INPE. Dissertação de Mestrado em Meteorologia, 87pp

SUN, L.; MONCUNILL, D. F.; LI, H.; MOURA, A. D.; SOUZA FILHO, F. A. “Climate Downscaling over Nordeste, Brazil, Using the NCEP RSM97”, *Journal of Climate*, 18, 551-567, 2005.

UVO, C. B.; REPELLI, C. A.; ZEBIAK, S. E.; KUSHNIR, Y. The relationships between Tropical Pacific and Atlantic SST and Northeast Brazil monthly precipitation. *Journal of Climate*, v.11, n. 10, p. 551-562, 1998.

WALKO, R. L.; BAND, L. E.; BARON, J.; KITTEL, T. G. F.; LAMMERS, R.; LEE, T. J.; OJIMA, D.; PIELKE, R. A.; TAYLOR, C.; TAGUE, C.; TREMBACK, C. J.; VIDALE, P. J., 2000. Coupled atmosphere-biophysicshydrology models for environmental modeling. *Journal of Applied Meteorology* 39: 931–944.

WALKO, R. L.; TREMBACK, C. J.; PIELKE, R. A.; COTTON, W. R.. An interactive nesting algorithm for stretched grids and variable nesting ratios. *J. Appl. Meteor.*, 34, 1995, 994-999.

WANG, H.; ZHOU, G.; ZHAO, Y., 2000: An effective method for correcting the seasonal-interannual prediction of summer climate anomaly. *Adv. Atmos. Sci.*, 17(2), 234-240.

WANG, Y.; LEUNG, L. R.; MCGREGOR, J. L.; LEE, D. K.; WANG, W. C.; DING, Y.; KIMURA, F. “Regional Climate Modeling: Progress, Challenges, and Prospects”, *Journal of Meteorological Society of Japan*, 82 (6), 1599-1628, 2004

WARD, N. M.; FOLLAND, C. K., 1991. 'Prediction of seasonal rainfall in the north Nordeste of Brazil using eigenvectors of sea-surface temperatures', *Int. J. Climatol.*, 11, 711–743.

WARD, M. N.; NAVARRA, A. 1997: Pattern analysis of SST-forced variability in ensemble GCM simulations: Examples over Europe and the tropical Pacific. *J. Climate*, 10, 2210–2220.

WOOD, A. W.; MAURER, E. P.; KUMAR, A.; LETTENMAIER, D. P. Long-range experimental hydrologic forecasting for the eastern United States. *Journal of Geophysical Research*. V.107, n.20, p.6(1-15). 2002.

WILKS, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

XAVIER, T. M. B. S.; XAVIER, A. F. S.; SILVA DIAS, P. L.; SILVA DIAS, M. A. F. (2000e). Modelos Estocásticos de Previsão Regional da Chuva no Ceará: Avaliação para as Quadras Chuvosas de 1997 a 2000. *Anais do XI CBMET-Congresso Brasileiro de Meteorologia* (em CD-ROM) Rio de Janeiro, 16-20 de Outubro.

ZENG, Q. C.; YU, R.; PENG, G. K.; CHAI, F. X. (1994), Research on “Ya- An-Tian-Lou.” part III: Physical structure and possible mechanism (in Chinese), *Chin. J. Atmos. Sci.*, 18, 649– 659.

CHEN, M.; DICKINSON, R. E.; ZENG, X.; HAHMANN, A. N. (1996). “Comparison of Precipitation Observed Over the Continental United States to that Simulated by Climate Model”, *Journal of Climate*, 9, 2233-2249

ANEXO

São mostradas as diversas posições e extensões dos domínios oceânicos testadas neste estudo a partir dos arquivos de saída do CPT para determinar os diferentes domínios de TSM que se correlacionam com a precipitação sobre o Estado do Ceará e com isso serem usados na previsão da quadra chuvosa do Estado.

Posições dos domínios oceânicos que foram testados (usando a precipitação total no trimestre FMA nas regiões homogêneas).

	LONGITUDE	LATITUDE	FOE (TSM)	FOE (PRECIP)	ACC	GOODNESS INDEX
GLOBAL	(ZERO,360)	(90S,90N)	4	4	2	-0,045
		(40S,40N)	4	3	2	-0,049
		(30S,30N)	4	3	2	-0,001
		(20S,20N)	1	1	1	0,039
ATLÂNTICO	(70W,20E)	(40S,40N)	2	4	2	-0,285
		(30S,30N)	3	4	1	-0,245
		(20S,20N)	2	2	1	-0,205
		(EQ,30N)	1	4	1	-0,185
		(30S,EQ)	1	1	1	-0,394
PACÍFICO	(70W,130E)	(40S,40N)	2	1	1	-0,013
		(30S,30N)	2	1	1	0,117
		(20S,20N)	2	2	2	0,121
		(10S,10N)	2	2	2	0,088
ATLÂNTICO + PACÍFICO	(130E,20E)	(40S,40N)	4	3	2	-0,011
		(30S,30N)	3	3	2	-0,003
		(20S,20N)	4	4	2	0,103
		(10S,10N)	4	2	1	0,106
NIÑO 1	(90W,80W)	(10S,5S)	1	1	1	-0,126
NIÑO 2	(90W,80W)	(5S,EQ)	1	1	1	0,03
NIÑO 1+2	(90W,80W)	(10S,EQ)	1	1	1	-0,023
NIÑO 3	(150W,90W)	(5S,5N)	1	1	1	0,146
NIÑO 4	(150W,160E)	(5S,5N)	3	1	1	-0,057
NIÑO 3.4	(170W,120W)	(5S,5N)	2	3	1	0,252

Posições dos domínios oceânicos que foram testados (usando a precipitação total no trimestre FMA nas bacias).

	LONGITUDE	LATITUDE	FOE (TSM)	FOE (PRECIP)	ACC	GOODNESS INDEX
GLOBAL	(ZERO,360)	(90S,90N)	4	4	2	0,002
		(40S,40N)	4	4	2	-0,064
		(30S,30N)	4	3	2	-0,037
		(20S,20N)	1	1	1	0,061
ATLÂNTICO	(70W,20E)	(40S,40N)	4	4	1	-0,35
		(30S,30N)	4	4	1	-0,216
		(20S,20N)	4	4	2	-0,253
		(EQ,30N)	4	3	3	-0,252
		(30S,EQ)	1	1	1	-0,385
PACÍFICO	(70W,130E)	(40S,40N)	2	2	1	0,161
		(30S,30N)	2	1	1	0,153
		(20S,20N)	2	1	1	0,11
		(10S,10N)	2	2	2	0,098
ATLÂNTICO + PACÍFICO	(130E,20E)	(40S,40N)	4	1	1	-0,047
		(30S,30N)	3	4	2	-0,025
		(20S,20N)	4	1	1	0,052
		(10S,10N)	4	1	1	0,083
NIÑO 1	(90W,80W)	(10S,5S)	1	1	1	-0,135
NIÑO 2	(90W,80W)	(5S,EQ)	1	1	1	0,049
NIÑO 1+2	(90W,80W)	(10S,EQ)	1	1	1	-0,016
NIÑO 3	(150W,90W)	(5S,5N)	1	1	1	0,173
NIÑO 4	(150W,160E)	(5S,5N)	1	1	1	-0,044
NIÑO 3.4	(170W,120W)	(5S,5N)	2	1	1	0,223

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)