

T1S-33



UNIVERSIDADE EDUARDO MONDLANE
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA



Trabalho de licenciatura

FUNÇÕES ORTOGONAIS EMPÍRICAS E
MODELAÇÃO DO CLIMA REGIONAL

Arlindo Oliva Meque

Supervisores:

Prof. Dr. Manuel Pondeca

dr. Genito Maúre

Janeiro de 2005

Declaração de honra

Declaro por minha honra que este trabalho foi realizado com base no material a que se faz referência ao longo do mesmo e que não foi submetido para nenhum outro grau que não seja o indicado, licenciatura em Meteorologia, na Universidade Eduardo Mondlane.

O Autor

...Arlindo Oliva Meque

(Arlindo Oliva Meque)

Dedicatória

Aos meus pais.

Agradecimentos

- Ao meu supervisor, Dr. **Manuel Pondeca**, que, atenciosamente dedicou muito do seu pouco tempo disponível ao desenvolvimento deste trabalho, depositando em mim sua confiança e me despertando interesse na área de modelação numérica.
- Ao meu co-supervisor, dr. **Genito Maúre**, pelos ensinamentos e principalmente pela compreensão, paciência e aceitação das minhas limitações como estudante e como ser humano.
- A todos os professores do Departamento de Física pela sua contribuição na minha formação.
- A todos os trabalhadores do Departamento de Física pela gentileza e colaboração ao longo de todos estes anos, em especial, ao sr. **Frenque Como** pelos seus conselhos.
- Aos amigos revisores, dr. **Orlando de Carvalho**, dr. **Jequissene Alcolete** e **Jaime Gado**, autênticos cúmplices deste tabalho, pela leitura atenta e pelas importantes sugestões apresentadas nas diversas versões deste trabalho.
- Aos colegas **Célio Matuele**, dr. **Luis Magaia**, dr. **André Moiane** e **Luck Injage** e ao dr. **Hipólito Cardoso** pelo apoio e companheirismo em todos os momentos e pela colaboração na realização deste trabalho.
- Aos colegas **Nostado**, **Silinto**, **Saraiva**, dr. **Matimbe**, **Quissico**, dr. **Anselmo**, dr. **Halo**, **Luis João**, **Marrenjo**, **Clousa**, **Maparagem**, dra. **Claire**, **Macucha**, dr. **Nhacuongue**, dr. **Sacramento**, dr. **Sitoe**, dr. **Hagi**, dr. **Ridwan**, **Hermínia** dra. **Ivódia** e a todos os demais pelos momentos que passamos juntos.

-
- Ao Cândido João, Jaime Gado, Altino Luis, Samito Assamo, meus colegas de quarto e a todos os moradores do "SELF" pelas alegrias do nosso convívio que serviram sempre de alento e força nos momentos adversos.
 - A **Laura, Maria de Lurdes e Célia**, minhas irmãs, pela louça lavada, pela comida pronta e principalmente pela nossa convivência.
 - Aos meus sobrinhos Abmério, Chana, Aurélio, Gervásio, Bachane, Tiriza por me lembrarem sempre a beleza e a simplicidade .
 - Ao Garido Lourenço, Isafas Muelegue, Agostinho Trinta, Gregório Atibo, Cassamo Atibo, Gregório Fernando, Félix de Magalhães e Horácio Bive pela grande amizade demonstrada ao longo de todos esses anos.
 - Aos meus pais, **Oliva e Marieta**, pelo seu esforço e dedicação e principalmente por terem sabido compreender o valor da educação.
 - A todos que me ajudaram de alguma forma na realização deste trabalho e que penosamente não estão sendo citados nominalmente.
 - A todos os obstáculos enfrentados ao longo deste trabalho por me terem ajudado a crescer.

Lista de abreviaturas

- ARPS - Atmospheric Regional Prediction System
- BUFR - Binary Universal Format Representation
- CD-ROM - Compact Disc -Read Only Memory
- CPTEC - Centro de Previsão de Tempo e Estudos Climáticos
- ECMWF - European Center for Medium Weather Forecasting
- EOFs - Empirical Orthogonal Functions (Funções Ortogonais Empíricas)
- ERL - Environmental Research Laboratories
- GrADS - Grids Analysis and Display System
- GRIB - Gridded Binary Data
- IPCC - Intergovernmental Panel on Climate Changes
- JMA - Japan Meteorological Agency
- LAPACK - Linear Algebra PACKage
- MGC - Modelos Globais de Circulação
- MM5 - Fifth-Generation Penn State/NCAR Mesoscale Model
- NCAR - National Center for Atmospheric Research
- NCDC - National Climate Data Center
- NCEP - National Centers for Environmental Predictions
- NetCDF - Network Common Data Format
- NOAA - National Oceanic and Atmospheric Administration
- PC - Personal Computers
- RAMS - Regional Atmospheric Modeling System

Resumo

Este trabalho cria a base matemática e computacional para a redução estatística do clima global para o clima regional e local. O método estatístico amplamente utilizado em Meteorologia para tal redução é o de Funções Ortogonais Empíricas (EOFs). As funções ortogonais empíricas têm como principal objectivo a redução da dimensionalidade de um grande conjunto de dados, para tal, as variáveis originais são substituídas por um outro conjunto de variáveis de menor dimensão que o conjunto original. Constatou-se neste trabalho que essa redução de dimensionalidade pode ser atingida considerando-se apenas as primeiras 5 EOFs.

Conteúdo

1	Introdução	2
1.1	Objectivos	5
1.1.1	Gerais	5
1.1.2	Específicos	5
2	Revisão bibliográfica	6
2.1	Tempo, Clima e Variabilidade Climática	6
2.2	Processo de redução de escala	8
2.2.1	O método de redução estatístico-dinâmica	10
2.3	Funções Ortogonais Empíricas (EOFs)	12
2.3.1	Interpretação física das EOFs	17
3	Material e métodos	19
3.1	Material	19
3.2	O sistema operativo Linux	19
3.3	Grids Analysis and Display System (GrADS)	20
3.4	Os dados	21
3.4.1	Projecto Reanalysis	21
3.5	Métodos	23
3.5.1	Cálculo de EOFs e Componentes Principais	24
4	Resultados	26
4.1	Pré-tratamento dos dados	26

5	Conclusões e recomendações	37
5.1	Conclusões	37
5.2	Recomendações	39
A	Programas	i
A.1	Programa para a geração das 15 primeiras EOFs	i
A.2	Programa para a reconstrução dos campos	vi

Lista de Figuras

2.1 Exemplo de uma possível configuração dos dados de vectores f_n e os vectores ortogonais empíricos e_m (Peixoto, 1992).	18
4.1 Primeira EOF	27
4.2 Segunda EOF	28
4.3 Terceira EOF	29
4.4 Valores próprios como função do número de ordem das EOFs	30
4.5 Componentes principais correspondentes à EOF1 e EOF2 para 1948 à 1952 . . .	31
4.6 Componentes principais correspondentes à EOF4 e EOF5 para 1948 à 1952 . . .	32
4.7 Campo original e campo reconstituído na base de 5 EOFs	34
4.8 Campo original e campo reconstituído na base de 10 EOFs	35
4.9 Campo original e campo reconstituído na base de 49 EOFs	36

Capítulo 1

Introdução

As mudanças climáticas são apontadas como sendo uma das principais ameaças ao desenvolvimento sustentável mundial. Estudos realizados com modelos climáticos de circulação geral para diferentes cenários de concentração de gases de efeito estufa cujo aumento se deve à actividade humana apontam para alterações globais significativas da temperatura, precipitação e humidade dos solos (IPCC, 1995).

Entre as consequências possíveis dessas mudanças climáticas, espera-se o aumento da temperatura média da superfície da terra, mudanças nos padrões de chuvas e ventos, derretimento das calotas polares, elevação do nível médio do mar e aumento dos processos de desertificação. Assim, tomando como referência os valores de 1990, projecta-se até ao ano 2100, um aquecimento médio global de 1.4 a 5.8 graus Celsius e um crescimento correspondente do nível médio do mar de 0.09 a 0.88 mm (IPCC, 2001).

Os resultados desses modelos também permitiram identificar mudanças significativas à nível regional e local onde as consequências acima mencionadas poderão ser mais pronunciadas esperando-se um aquecimento mais acentuado do que o da média global em algumas áreas de terra firme. Assim, para a África Austral espera-se (i) um aumento da frequência de ocorrência das ondas de calor e da intensidade dos ventos e chuvas dos ciclones tropicais, (ii) uma maior frequência de ocorrência de eventos de seca e cheias, (iii) Perturbação do balanço hidrológico

local devido às mudanças na variabilidade interanual da quantidade da água das chuvas com consequências na produção de energia hidroeléctrica e no fornecimento de água doce, (iv) proliferação de doenças tropicais, (v) inundações permanentes e perda de terra para a habitação e agricultura devido ao aumento do nível médio do mar e erosão, (vi) perda de biodiversidade nos ecossistemas, deterioração da vegetação, mudanças na distribuição de espécies e esquemas de migração, (vii) um impacto negativo no turismo e (viii) uma possibilidade de grandes movimentos migratórios da população devido aos eventos climáticos extremos como as cheias, ventos fortes e secas (IPCC, 2001).

Um estudo aprofundado dos efeitos das mudanças climáticas ao nível local, incluindo uma quantificação das incertezas dos resultados, é de grande importância para o desenvolvimento sustentável do país e da região, pois, permite o delineamento de políticas apropriadas e a tomada de decisões por parte do governo em sectores como a agricultura e a segurança alimentar, gestão de recursos hídricos e energéticos, desenvolvimento rural e urbano, conservação de recursos florestais e marinhos entre outros.

Esse estudo é efectuado eficientemente através da modelação numérica de processos, o que pressupõe a existência de campos meteorológicos de boa resolução à escala regional e local (Tucci, 2002).

Porém, a resolução horizontal dos modelos globais de circulação, tradicionalmente usados para a modelação das variações climáticas é ainda muito baixa para apreender, em áreas de fisiografia complexa informação adequada para estudos de avaliação de impactos a nível local ou regional. Esses modelos têm uma resolução demasiado baixa para incorporar convenientemente a topografia do terreno, massas de água e linhas costeiras regionais e permitir assim, a extracção de uma informação útil para estudos das variações climáticas sobre seres humanos e sistemas ecológicos (Moriondo, 2002-2003).

Deste modo, impõe-se o desenvolvimento de técnicas de transferência de informação meteorológica para as escalas menores, isto é, técnicas de "downscaling" ou de redução de escala.

São dois os métodos principais utilizados para reduzir o clima e cenários de variações climáticas da escala global para a escala regional e local: a redução estatística e a redução dinâmica (Busch e Heimann 2001, Meier et al., 1999). Na redução estatística, procuram-se, com base em observações e análises meteorológicas, relações estatísticas entre parâmetros do clima local e um certo número de predictores de larga escala. A redução dinâmica por seu turno, denota a prática da introdução de um modelo numérico de mesoescala de alta resolução dentro de um modelo de circulação geral. O modelo de mesoescala é integrado com condições de fronteira (e por vezes também condições iniciais) fornecidas pelo modelo global de circulação.

O método de redução dinâmica pura dos cenários do clima fornecidos pelos Modelos Globais de circulação (MGC) é computacionalmente proibitivo para os nossos recursos pois, este método exige o uso de tecnologia computacional de ponta. Uma técnica que se ajusta melhor às nossas condições computacionais é a combinação cuidadosa das duas técnicas de base acima mencionadas (redução estatística e redução dinâmica) resultando na "redução estatístico-dinâmica".

Por esta razão, o objectivo deste trabalho é lançar a base para o uso do método de redução estatístico-dinâmica do clima global para o clima regional da África Austral. Concretamente, o trabalho está concentrado na criação da base matemática e computacional para a redução estatística do clima global pois, é a parte em falta para a aplicação do método de redução estatístico-dinâmica do clima no Departamento de Física.

A técnica de modelação do clima regional e local foi introduzida nos últimos 15 anos encontrando-se numa fase emergente tanto em Moçambique como em toda a África.

1.1 Objectivos

1.1.1 Gerais

- Iniciar a modelação do clima regional da África Austral pelo método de redução estatístico-dinâmica do clima dos modelos globais;
- Calcular Funções Ortogonais Empíricas como ferramenta para a classificação do clima;

1.1.2 Específicos

- Desenvolver um programa em Fortran para o cálculo de EOFs;
- Usar a temperatura de superfície dos dados de reanálise do NCEP/NCAR para ilustrar o cálculo de EOFs.

Capítulo 2

Revisão bibliográfica

2.1 Tempo, Clima e Variabilidade Climática

O tempo pode ser definido como o estado da atmosfera em determinado instante e lugar. O clima por seu turno, tem sido frequentemente definido como um tempo médio, ou seja, um conjunto de condições normais que dominam uma região, obtidas das médias das observações durante um certo intervalo de tempo. Portanto, o clima é o conjunto de toda a informação estatística sobre o tempo em determinado local (Quadro, 2004).

O clima é essencial para a manutenção da vida na terra e exerce um efeito profundo sobre a segurança alimentar, a salvaguarda de vidas e bens, os recursos hídricos, as actividades de lazer e o desenvolvimento sustentável (Obasi, 2003).

O clima de um local ou região varia, em geral, ao longo do ano como consequência do movimento de translação da terra. Muitos elementos do clima (temperatura e humidade do ar, por exemplo), apresentam também uma marcada variação diurna, associada ao movimento de rotação da terra. Para além da variabilidade de tipo cíclico associada a movimentos astronómicos, muito aproximadamente periódicos, o clima apresenta uma variabilidade natural interna não periódica muito complexa, o que faz com que o clima num dado ano seja diferente do clima de anos anteriores e dos anos seguintes (Peixoto e Oort, 1992).

Sabe-se que ocorreram alterações do clima no passado. No entanto, elas foram devidas à causas naturais. Essa variação do clima em função dos condicionantes naturais do globo terrestre designa-se Variabilidade climática (Tucci, 2002).

As alterações recentes com o aumento de 0.6 graus Celsius da temperatura média global desde o início dos registos efectuados com instrumentos, são em grande parte atribuíveis à actividade humana (IPCC, 1995). Essas alterações do clima devidas à actividade humana dão origem ao termo "Modificação climática" (Tucci, 2002).

Os estudos do clima são feitos eficientemente com base em modelos numéricos que incorporam em vários graus de complexidade, as descrições matemáticas da atmosfera, dos oceanos, da terra, da biosfera e da criosfera (IPCC, 1995).

Cada modelo tem a sua área de cobertura, que se designa por domínio ou campo do modelo. Assim sendo, os modelos numéricos de previsão do tempo e do clima classificam-se em modelos Globais e modelos Regionais (também conhecidos por modelos de área limitada) (Tomé, 2004).

Modelos Globais são modelos que analisam todo o globo terrestre sendo as equações que governam o fluido geofísico resolvidas sobre todo o globo terrestre. Vários centros de previsão de tempo e do clima no Mundo executam rotineiramente modelos globais, dentre eles pode-se citar o NCEP, o ECMWF, a JMA e o CPTEC (Haas, 2002).

Os modelos regionais por seu turno, resolvem as equações que governam o fluido geofísico para uma área localizada sendo por isso necessárias as condições laterais de fronteira. A utilização de modelos regionais é prática rotineira em centros operacionais e de pesquisa em todo o mundo, pois, possibilita previsões destinadas a áreas menores e usuários finais específicos e a análise de fenómenos meteorológicos de escala menor. Existem disponíveis vários destes modelos como o RAMS, o MM5, o modelo Regional ETA, o modelo Brasileiro de alta resolução

(MBAR) e o ARPS (Haas, 2002).

Note-se que alguns destes modelos de previsão numérica, como é o caso do MM5 e ETA, são também usados na modelação do clima regional pelo método de redução dinâmica (De Pondeva, 2001).

Os modelos atmosféricos e climáticos acima citados diferem basicamente quanto à diversidade de opções de parametrizações físicas, métodos de resolução das equações primitivas e métodos de assimilação de dados.

2.2 Processo de redução de escala

Para avaliar os impactos das alterações climáticas usam-se frequentemente modelos globais de circulação. No entanto, algumas das variáveis mais relevantes à escala regional ou local, como por exemplo, as características fisiográficas regionais encontram-se mal representadas ou mesmo ausentes nos MGC (Moriondo, 2002-2003).

As características fisiográficas regionais e as circulações de menor escala a elas associadas podem causar significativa influência no clima regional ou local. São exemplos as brisas marítimas e as circulações induzidas por vales e montanhas (Tucci et.al., 2000).

As brisas marítimas e terrestres são principalmente originadas pela diferença no calor específico entre a terra e a água. Durante o dia, quando a terra e a água são expostas a uma mesma intensidade de radiação solar, a superfície da terra aquece mais que a superfície da água. A terra, relativamente quente, aquece o ar nas suas vizinhanças, diminuindo a sua densidade, conseqüentemente, um gradiente de pressão horizontal se desenvolve entre a terra e a água, com a pressão mais alta sobre a superfície de água. Em resposta ao gradiente, o ar frio se desloca para a terra formando a chamada brisa marítima e, por continuidade, surge um deslocamento de ar da terra para a água em níveis superiores, com ar subindo sobre a terra e descendo sobre

a água (Tucci et al., 2000).

As brisas marítimas são sistemas geralmente confinados nas primeiras camadas da troposfera. Tipicamente, a brisa começa próxima a linha da costa, várias horas depois do nascer do sol e se expande gradualmente tanto sobre a terra como sobre a água atingindo sua máxima extensão ao meio da tarde. Sobre a terra, a extensão da brisa varia de poucas centenas de metros à algumas centenas de quilômetros, podendo alterar significativamente o regime pluviométrico e a intensidade do vento, o que tem impacto na evaporação e na evapotranspiração (Tucci et al., 2000).

Em geral, os processos de circulações atmosféricas acima descritos ocorrem em escalas espaciais que não são resolvidas pela malha dos modelos globais de previsão atmosférica ou climática, de igual maneira os efeitos da topografia ficam bastante suavizados nos modelos globais. Deste modo, impõe-se o desenvolvimento de técnicas de transferência da informação gerada pelos MGC em larga escala para as escalas menores. Essas técnicas são chamadas por técnicas de "downscaling" ou de redução de escala ou de desagregação espacial (Tucci et al., 2000).

As técnicas de redução podem ser enquadradas em duas abordagens distintas conceitualmente:

Redução dinâmica : aqui, os modelos dinâmicos regionais, que são modelos numéricos semelhantes aos MGC mas com resolução mais fina e área limitada, são integrados no tempo usando como dados iniciais e condições de fronteira o resultado da simulação dos MGC (Giorgi, 1990); e

Redução estatística : aqui, procura-se, com base em observações e análises meteorológicas, relações estatísticas entre parâmetros do clima local e um certo número de preditores de larga escala. Essas relações são depois aplicadas aos resultados dos modelos de circulação global com o fim de se obter o clima regional e local correspondente ao clima global e cenários de variação climática globais (Wilby et al., 1998).

Os dois métodos de redução aqui descritos podem ser combinados para formar o método

estatístico-dinâmico de redução do clima global para o clima regional e local (Frey-Buness et al., 1995)

2.2.1 O método de redução estatístico-dinâmica

Giorgi (1990), Tanajura (1996), entre outros, demonstraram que os modelos regionais aninhados à modelos de circulação geral da atmosfera podem efectivamente melhorar a solução gerada por tais modelos globais devido ao facto de os primeiros (modelos regionais), apresentarem alta resolução e em muitos casos possuírem parametrizações ajustadas especificamente para regiões de interesse. A alta resolução beneficia a imposição de forçantes orográficas mais realistas e a uma melhor definição do contraste continente-oceano além de permitir que os tipos de solo e vegetação fiquem melhor caracterizados. No entanto, o método de redução dinâmica pura dos cenários do clima provenientes dos modelos globais é computacionalmente caro e fora do alcance das capacidades de muitos países em desenvolvimento como é o caso de Moçambique.

Um método computacionalmente viável de redução estatístico-dinâmica foi proposto por Frey-Buness et.al.(1995), neste método:

1. Uma série de tempo de muitos anos (vinte ou mais), feita de análises de larga-escala ou da previsão de um modelo de circulação global é partilhada em tempos de larga escala característicos do local ou região.
2. Calcula-se a frequência de ocorrência de cada tipo de tempo de larga escala como a soma do número de ocorrências similares a dividir pelo total de ocorrências de todos os tipos de tempo de larga escala.
3. Faz-se uma integração de curto alcance (24 a 48 horas) de um modelo regional de grande resolução espacial, sendo as condições iniciais e de fronteira características para esse tipo de tempo de larga escala.
4. Calculam-se os parâmetros regionais e locais como médias ponderadas das integrações do modelo regional, usando como pesos da média, as frequências calculadas no ponto 2.

Com este trabalho pretende-se que o Departamento de física inicie a implementação do método de Frey-Buness para a modelação do clima regional e local da África Austral.

Vários métodos estatísticos tais como a regressão múltipla e análise de funções ortogonais empíricas têm sido utilizados para a componente de redução de escala. Recentemente, métodos não lineares como as Redes Neurais têm sido empregues com bons resultados (Muller, 2000).

A análise de regressão é uma metodologia que tem como objectivo ajustar um modelo probabilístico a um conjunto de dados, cada variável é isolada no momento em que as variáveis independentes variam sistematicamente sendo observados os seus efeitos sobre a variável resposta (Wilks, 1995).

As redes neurais por seu turno, podem ser entendidas como conjuntos bem estruturados de unidades de processamento interligados por canais de comunicação, cada qual tendo um determinado peso correspondente a um valor numérico. Elas consistem de várias unidades de processamento (neurónios artificiais) interconetados entre si, formando uma determinada disposição estrutural de camadas. Os modelos neurais procuram aproximar o processamento ao cérebro humano. Têm sido utilizadas principalmente em problemas onde a relação entre as variáveis não é totalmente conhecida e também em problemas onde uma grande quantidade de dados esteja disponível para a simulação. Elas possibilitam a modelação de um fenómeno físico sem o conhecimento da teoria intrínseca ao problema, por essa característica são também chamados por "free models" (Muller, 2000).

Uma outra técnica estatística amplamente usada em Meteorologia e Climatologia é o cálculo de Funções Ortogonais Empíricas (EOFs). A técnica de EOFs permite reduzir as dimensões de grandes séries de dados com o objectivo de avaliar melhor o volume de informação contida nas mesmas (Wilks, 1995).

2.3 Funções Ortogonais Empíricas (EOFs)

A análise de funções ortogonais empíricas é uma das técnicas estatísticas multivariadas mais utilizadas em meteorologia. O método é utilizado para a obtenção dos principais padrões espaciais de variabilidade, a sua variação no tempo e a quantificação da importância relativa desses padrões. Este método baseia-se na resolução do problema de autovalores da matriz de covariância do campo de anomalias de uma variável meteorológica distribuída no espaço e no tempo (por exemplo, temperatura).

Os autovectores obtidos a partir da matriz de dados de anomalias são ortogonais entre si e representam padrões espaciais estacionários ou mapas de variabilidade (EOFs). A porção da variância total explicada por cada EOF é dada pelo seu respectivo autovalor. A evolução temporal de uma EOF é dada pelos coeficientes de expansão calculados projectando-se cada autovector sobre o campo original. Os autovectores (EOFs) e os coeficientes de expansão constituem o modo de variabilidade.

Um dos principais propósitos da análise de EOFs é reduzir um conjunto de dados contendo um grande número de variáveis para um conjunto contendo um número bem menor de novas variáveis. As novas variáveis, denominadas EOFs, são formadas de forma a que :

- Elas não sejam correlacionadas ;
- A primeira EOF explique o máximo possível da variância nos dados;
- Cada EOF sucessiva explique o máximo possível da variância restante;

O tratamento matemático completo que conduz ao método de cálculo de EOFs pode ser encontrado na literatura (por exemplo, Peixoto e Oort, 1992; Venegas, 2001; Hannachi, 2004 e Bjornsson e Venegas, 1997). Neste trabalho descrevemos apenas a "receita" para o seu cálculo .

Consideremos uma série de N mapas nos tempos $t=1\dots N$, onde cada mapa contém medições do campo ψ nas posições $m=1\dots M$, isto é, temos M séries de tempo $\psi_m(t)$, tendo cada uma um

comprimento N . O primeiro e o último tempo ($t=1$ e $t=N$) devem ser os mesmos para todas as M séries.

Primeiramente removemos em cada uma das séries de tempo os seus valores médios, μ_m , e em seguida normalizamos as séries de tempo dividindo cada uma delas pelo seu desvio padrão σ_m .

As séries normalizadas $F_m(t)$ são escritas sob a seguinte forma :

$$F_m(t) = \frac{\psi_m(t) - \mu_m}{\sigma_m} \quad (2.1)$$

onde : μ_m é o valor médio:

$$\mu_m = \frac{1}{N} \sum_{t=1}^N \psi_m(t) \quad (2.2)$$

e σ_m é o desvio padrão :

$$\sigma_m = \sqrt{\frac{1}{N-1} \sum_{t=1}^N \psi_m^2(t)} \quad (2.3)$$

Note-se que, como pré-tratamento dos dados podemos começar por remover das séries de tempo os "trends" conhecidos e de menor importância. Por exemplo, podemos em alguns casos remover a variação sazonal por ser demais conhecida, sendo a sua manifestação nos dados irrelevante para o problema específico que se pretende resolver. Isto pode ser feito calculando-se o ciclo climatológico anual e subtraindo-o do campo $\psi_m(t)$, ficando-se apenas com as anomalias. Portanto, em geral, $\psi_m(t)$ pode tanto representar uma série de tempo dum grandeza "completa" como da sua anomalia obtida após a remoção de um ou mais "trends".

Constrói-se em seguida uma matriz $F = M \times N$, quer dizer, uma matriz de M linhas (posições m) e N colunas (tempo t) a partir dos dados das anomalias (2.1) :

$$F = \begin{bmatrix} F_1(1) & F_1(2) & \cdots & F_1(N) \\ F_2(1) & F_2(2) & \cdots & F_2(N) \\ \cdots & \cdots & \cdots & \cdots \\ F_M(1) & F_M(2) & \cdots & F_M(N) \end{bmatrix} \quad (2.4)$$

A partir da matriz F pode-se derivar a matriz de covariância espacial R_{FF} do campo $F_m(t)$

como resultado da multiplicação da matriz F pela sua transposta F^T :

$$R_{FF} = F \cdot F^T \quad (2.5)$$

Expandindo o produto das matrizes da equação (2.5), resulta:

$$R_{FF} = \begin{bmatrix} \langle F_1 F_1 \rangle & \langle F_1 F_2 \rangle & \dots & \langle F_1 F_M \rangle \\ \langle F_2 F_1 \rangle & \langle F_2 F_2 \rangle & \dots & \langle F_2 F_M \rangle \\ \dots & \dots & \dots & \dots \\ \langle F_M F_1 \rangle & \langle F_M F_2 \rangle & \dots & \langle F_M F_M \rangle \end{bmatrix} \quad (2.6)$$

Onde $\langle F_i F_j \rangle$ é a covariância entre as séries de tempo F_i e F_j (F nas posições i e j) definida como :

$$\langle F_i F_j \rangle = \langle F_j F_i \rangle = \frac{1}{N-1} \sum_{t=1}^N F_i(t) F_j(t) \quad (2.7)$$

Para $i, j = 1 \dots M$. Note-se que, para o caso particular $i = j$, $\langle F_i F_j \rangle = \langle F_i F_i \rangle$ representa a variância na posição i .

A matriz R_{FF} , de dimensão $M \times M$ é quadrática e simétrica mesmo se a matriz F não é quadrada.

Depois de calculada a matriz de covariância, é preciso resolver o problema de valores próprios:

$$R_{FF} \cdot E = E \cdot \Lambda \quad (2.8)$$

O que significa que a matriz R_{FF} foi decomposta em duas matrizes "Λ" e "E". A matriz "Λ" é uma matriz diagonal $M \times M$, contendo os valores próprios λ_k da matriz R_{FF} :

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_M \end{bmatrix} \quad (2.9)$$

Note-se que, dado que R_{FF} é uma matriz simétrica, os seus autovalores são todos reais, isto é, não existe um único autovalor que seja complexo. Este é um facto bem conhecido em álgebra linear. Ademais, R_{FF} é uma matriz semi-positiva definida, pelo que não contém autovalores

negativos. A (possível) existência de autovalores nulos significa que a matriz original (2.4) contém informação redundante, isto é, que as séries de tempo (2.1) não são linearmente independentes. A técnica de EOFs permite eliminar essa informação redundante do conjunto de dados.

Geralmente, os valores próprios em "Λ" são escolhidos em ordem decrescente, assim, $\lambda_1 > \lambda_2 > \dots > \lambda_M$.

A matriz quadrada "E" tem dimensão $M \times M$, as suas colunas estão constituídas pelos vectores E^k que são os vectores próprios da matriz R_{FF} correspondendo aos valores próprios λ_k :

$$E = \begin{bmatrix} E_1^1 & E_1^2 & \dots & E_1^M \\ E_2^1 & E_2^2 & \dots & E_2^M \\ \dots & \dots & \dots & \dots \\ E_M^1 & E_M^2 & \dots & E_M^M \end{bmatrix} \quad (2.10)$$

Cada valor próprio não nulo na matriz "Λ" está associado à coluna do vector próprio E^k na matriz "E". Assim, a real dimensão da matriz "E" é $M \times K$, onde M representa as posições e K representa o número de valores próprios não nulos. Temos pois K formas de decomposição ou EOFs.

A matriz "E" possui a seguinte propriedade : $E \times E^T = E^T \times E = I$, onde I é a matriz identidade. Isto significa que os vectores próprios são mutuamente não correlacionais no espaço, ou seja, eles são mutuamente ortogonais. Os vectores próprios (autovectores) representam padrões espaciais estacionários ou mapas de variabilidade (EOFs).

A evolução do tempo da EOF número k (isto é, como o padrão E^k se desenvolve com o tempo) é dada pela série de tempo $A^k(t)$, que é obtida pela projecção das séries originais dos dados $F_m(t)$ nos vectores E^k :

$$A^k(t) = \sum_{m=1}^M E_m^k F_m(t) \quad (2.11)$$

Onde $m = 1 \dots M$ quantifica as posições, $t = 1 \dots N$ quantifica as etapas e $k = 1 \dots K$ quantifica o modo EOF. $A^k(t)$ representa a componente principal "k". Na notação matricial, a

matriz "A" é obtida multiplicando as matrizes E^T e "F":

$$A = E^T \cdot F \quad (2.12)$$

Onde E^T (a transposta de E) é uma matriz $K \times M$, F é $M \times N$ e, conseqüentemente, A é uma matriz $K \times N$. Os padrões espaciais E^k são ortogonais no espaço e as componentes principais A^k são ortogonais no tempo.

Cada autovalor (valor próprio) é proporcional à percentagem da variância do campo F que é quantificada pelo modo k. Esta percentagem é expressa por :

$$Var(\%) = \frac{\lambda_k}{\sum_{i=1}^k \lambda_i} \times 100 . \quad (2.13)$$

O campo original F pode ser totalmente reconstruído multiplicando cada padrão da função ortogonal empírica E^k pela sua respectiva componente principal A^k e adicionando todos os produtos dos modos k :

$$F_m(t) = \sum_{k=1}^K E_m^k A^k(t) \quad (2.14)$$

Em notação matricial:

$$F = E \cdot A \quad (2.15)$$

Contudo, o objectivo fulcral da decomposição em EOFs é a reconstrução do campo \hat{F} que é uma versão do campo original F, mas, mais comprimido e menos ruidoso. Este procedimento é efectuado truncando a decomposição na equação (2.14), isto é, reconstruindo \hat{F} usando somente os primeiros modos H ($H < K$):

$$\hat{F}_m(t) = \sum_{k=1}^H E_m^k A^k(t) \quad (2.16)$$

Há uma certa arbitrariedade na escolha do número de modos H, mas o princípio mais correcto é o de escolher H de maneira a que \hat{F} quantifique uma larga fracção da variância do campo original F. Em alguns problemas, é suficiente reter os primeiros dois a cinco EOFs para explicar mais de 90% da variância de F.

2.3.1 Interpretação física das EOFs

Na interpretação de EOFs devemos ter em conta que apesar delas serem uma forma estatística mais eficiente de comprimir dados de um dado campo, elas não correspondem necessariamente ao modo do comportamento físico. Contudo, em muitos casos as propriedades de normalidade multivariada e linearidade inerentes ao método de EOFs podem ajudar na captação da "física essencial" contida num conjunto enorme de dados.

De acordo com Peixoto e Oort(1992), uma maneira de entender a ideia base sobre EOFs é imaginarmos que podemos mostrar cada um dos N mapas como um vector f_n no espaço M-dimensional:

$$f_n = \{f_{n1}, f_{n2}, \dots, f_{nM}\} \quad (2.17)$$

no tempo $t = n$.

Cada vector f_n inclui os valores do campo f em todas as posições $m = 1, \dots, M$ para um dado momento do tempo n (isto é, um mapa de f no tempo n). Os N vectores dos dados são direccionados da origem para um ponto no espaço M (fig.2.1).

Se existir alguma correlação entre os N vectores (mapas), então, suas extremidades serão organizadas em grupos ou ao longo de algumas direcções preferidas.

Para se resolver o problema a partir da decomposição de EOFs é preciso encontrar-se uma base ortogonal $\{e_1, e_2, \dots, e_M\}$ no espaço M-dimensional, em vez da base original, de modo que o vector e_1 melhor represente o maior agrupamento dos vectores dos dados originais. O vector e_2 será ortogonal ao primeiro e representará o segundo maior agrupamento dos vectores dos dados originais. O terceiro vector e_3 será ortogonal ao primeiro e ao segundo e representará o terceiro maior agrupamento dos vectores dos dados originais, e assim sucessivamente. Em outras palavras, e_1 quantifica a máxima porção da variância dos dados originais que pode ser explicada com apenas um vector, e_2 quantifica a segunda máxima variância numa direcção perpendicular a e_1 , e_3 quantifica a terceira máxima variância numa direcção perpendicular a e_1 e e_2 , e assim sucessivamente. Este procedimento equivale a encontrar uma série de M vectores

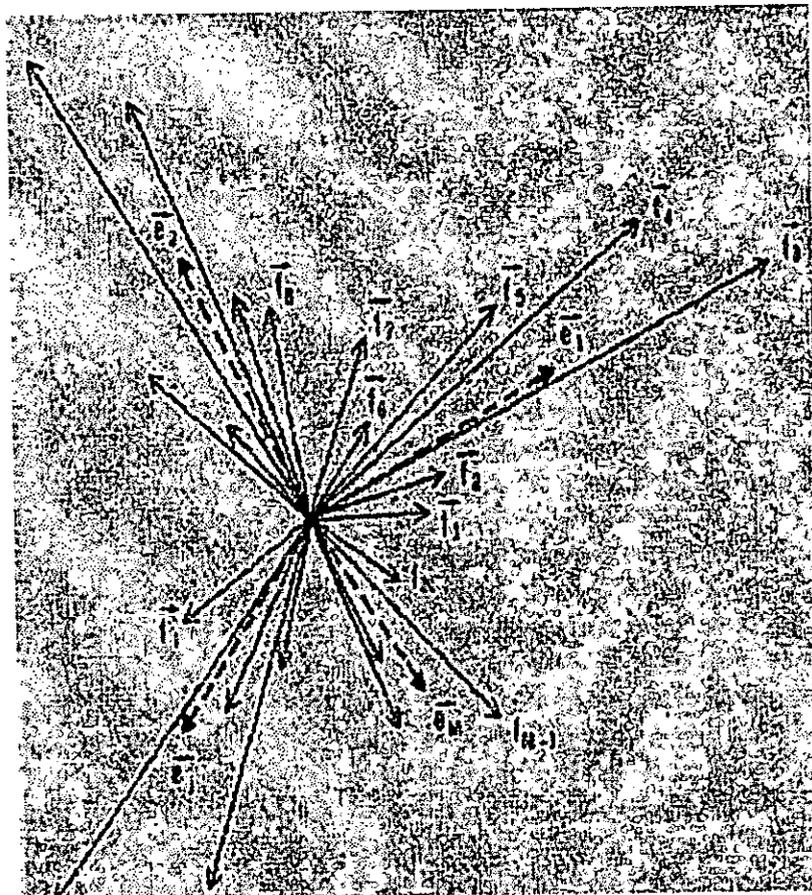


Figura 2.1: Exemplo de uma possível configuração dos dados de vetores f_n e os vetores ortogonais empíricos e_m (Peixoto, 1992).

e_m cuja a orientação é tal que a soma dos quadrados das projeções de todos os vetores f_n em cada e_m é maximizada sequencialmente. Os vetores $e_m, m = 1, \dots, M$ são mutuamente ortogonais, e são as chamadas EOFs.

Capítulo 3

Material e métodos

3.1 Material

Na realização do presente trabalho foi usado um PC Pentium IV com o sistema operativo Linux e compilador de Fortran do "Grupo de Portland", bem como o pacote de visualização, análise e manipulação de dados ambientais denominado GrADS.

3.2 O sistema operativo Linux

O Linux é um sistema operativo de distribuição livre para PCs e "Work stations". Trata-se de um "clone" do sistema Unix, com a característica de ser o primeiro sistema completo de livre distribuição e desenvolvido por um grande grupo de programadores voluntários em todo o Mundo. A nível de aplicativos para Linux existe neste momento um número grande e vasto em todas as áreas tais como a internet, investigação, multimédia e entretenimento.

O sistema Linux pode ser obtido em forma já compilada e com um grande número de programas formando uma distribuição. Existem várias distribuições de Linux (SuSE, Slackware, Debian, RedHat, etc), tendo cada uma as suas vantagens e desvantagens. Para o presente trabalho foi utilizado o Redhat 7.1.

3.3 Grids Analysis and Display System (GrADS)

O GrADS é uma ferramenta interactiva utilizada para a visualização, análise e manipulação de dados ambientais (Simões, 2000).

O GrADS manipula os dados em quatro dimensões distintas, nomeadamente, latitude, longitude, nível e tempo. Suporta dois tipos básicos de dados, que são dados reticulados ou em malhas ("gridded data") e dados de estações ("station data"). Quanto ao formato dos dados existem várias opções: ou estão gravados no formato binário que é o formato próprio do GrADS ou estão compactados nos formatos GRIB ou NetCDF. Os dados são acedidos através de um arquivo descritor, no caso dos formatos GRIB e NetCDF esse arquivo faz também referência a um arquivo de índice que contém informação para a descodificação dos dados.

Os dados e metadados (informação sobre os dados) são guardados em dois arquivos separados. O arquivo de metadados (extensão ctl), também chamado arquivo descritor ou de controle, contém uma descrição completa dos dados no formato ASCII e o nome do arquivo que contém esses dados. Este último, é um arquivo puramente de dados (binário, GRIB ou NetCDF). O arquivo descritor, para além da especificação de um conjunto binário de dados e uma possível referência a um arquivo de índice, contém uma identificação dos dados, um valor que indica a ausência destes, as variáveis utilizadas e a definição do domínio dentro do ambiente dimensional: XDEF (longitude), YDEF (latitude), ZDEF (nível) e TDEF (tempo).

A visualização dos campos de dados pode ser feita de várias formas: por gráficos de linhas, histogramas, mapas de contorno e de sombreamento, vectores, linhas de corrente, etc., e sua gravação é permitida num formato intrínseco ao GrADS e no formato GIF. A conversão para formatos comuns como o postscript e pdf pode ser feita a partir do formato intrínseco ao GrADS (Doty, 1995).

Uma das vantagens ao comparar o GrADS com outros softwares de visualização é que enquanto estes necessitam do uso de programas de conversões fornecidos pelo usuário para

manipular arquivos com o formato GRIB e NetCDF, o GrADS possui um suporte directo ("built-in support") para dados nestes formatos.

3.4 Os dados

Os dados usados neste trabalho são médias mensais da temperatura de superfície para os anos de 1948 a 1998 provenientes do Projecto "Reanalysis" do NCEP e NCAR.

3.4.1 Projecto Reanalysis

O Reanalysis é um projecto realizado pelo NCEP em colaboração com o NCAR com o objectivo de produzir análises históricas globais e melhoradas dos vários parâmetros meteorológicos abrangendo mais de 50 anos. O objectivo é a criação de um banco de dados fiável para o avanço da investigação em meteorologia e climatologia (Kalnay et al.,1996).

As observações usadas no projecto foram sujeitas a testes de controle de qualidade e usadas num sistema variacional tri-dimensional de assimilação de dados. O "output" desse sistema são os campos de análise (convenientemente denominados "reanálise" para distinguí-los das análises prévias, feitas com base no sistema de assimilação "optimal interpolation") dos vários parâmetros meteorológicos.

As equações de balanço usadas no sistema variacional tri-dimensional de análise permitem extrapolar informação de zonas ricas em informação, isto é, onde as observações abundam para zonas pobres em informação de forma a que, mesmo em áreas com ausência relativa de observações o programa possa estimar o estado da atmosfera com uma precisão melhorada.

Foram criados diferentes tipos de arquivos de output para diferentes necessidades do utilizador que são o BUFR (Binary Universal Format Representation), Main synoptic, Reduced "Time Series" e quick look em CD-ROM que contém os dados referentes aos campos atmosféricos mais utilizados bem como campos de superfície e do topo da atmosfera.

A recolha de observações desempenha um papel bastante importante, esta tarefa foi basicamente desempenhada pela NCAR. As observações foram obtidas das seguintes fontes:

- Global Rawinsonde Data;
- COADS (Comprehensive Ocean Atmosphere Data Set ¹) Surface Marine Data;
- Aircraft Data;
- Surface Land Synoptic Data;
- Satellite Sounder Data;
- SSM/I surface wind speeds;
- Satellite Cloud Drift Winds;

O Sistema reanalysis possui três módulos de processamento de dados:

- Pré-processador com controlo de qualidade de dados (QC);
- Módulo de análise com sistema automático de Monitorização;
- Módulo de Output e arquivo;

O objectivo do pré-processamento do reanalysis é reformatar as observações provenientes de diferentes fontes num formato uniforme BUFR. O pré-processador inclui também a preparação das condições de fronteira da superfície (por exemplo, a temperatura da superfície do mar, glaciares, etc).

No módulo de análise é usado o SSI (Spectral Statistical Interpolation) que é uma implementação do esquema variacional de análise tridimensional no espaço espectral. A sua implementação em 1991 pelo NCEP substituindo a análise de interpolação óptima conduziu a uma grande melhoria na análise e previsão globais. Uma vantagem importante do SSI é que o

¹Projecto conjunto da NOAA/ERL, NCAR e NCDC com o auxílio de muitas organizações e países

balanço imposto na análise é válido para todo o globo tornando desnecessária a inicialização de modos normais não lineares. Tal processo de inicialização é usado para a eliminação de desbalanços nas condições iniciais do modelo que resultam em ondas gravitacionais não realistas nas primeiras horas de integração do modelo.

No sistema de assimilação é usado o modelo global espectral do NCEP que inclui a parametrização de todos os processos físicos mais importantes, como por exemplo, a convecção, a precipitação e a radiação. Este modelo apresenta largas vantagens em relação aos anteriores pois:

- melhora a previsão da precipitação;
- melhora o diagnóstico da distribuição das nuvens;
- os padrões de precipitação sobre os trópicos são mais realísticos, com uma distribuição suave e pouca concentração sobre as características orográficas tropicais; e
- foram executadas algumas afinações da nebulosidade e das propriedades ópticas das nuvens de forma a corrigir erros sistemáticos da temperatura e da nebulosidade;

A saída dos dados do reanalysis foi uma componente muito importante no desenvolvimento do projecto. Existem várias possíveis aplicações do output do reanalysis, e alguns deles (por exemplo, o transporte dos gases de estufa, que precisam de todos os transportes turbulentos entre duas quaisquer camadas) têm uma exigência de armazenamento que excede as capacidades do projecto. Por esta razão, foi decidido que cada unidade do output do reanalysis (um mês) incluiria ficheiros de reinício (restart files) que permitem a recriação dos transportes de turbulência e outros dados.

3.5 Métodos

Neste trabalho foi desenvolvido um programa em Fortran para o cálculo de funções ortogonais empíricas e testado o seu funcionamento calculando-se as EOFs de temperatura de superfície para uma série de tempo de 51 anos de dados em malha do projecto "Reanalysis" do

NCEP/NCAR. O programa para o cálculo de EOFs representa uma ferramenta poderosa para, num futuro próximo, completar-se o ponto 1. do método de Frey-Buness.

O Departamento de Física tem estado desde 2001 a correr, com sucesso, o modelo de mesoescala MM5 para previsão do tempo, este mesmo modelo é usado para a componente de redução dinâmica do método de Frey-Buness (ponto 3). Depreende-se pois, que a parte em falta para a aplicação do método de Frey-Buness no Departamento de Física é a redução estatística do clima global. Este trabalho representa o primeiro passo para a redução estatística necessária para a modelação do clima regional da África Austral.

3.5.1 Cálculo de EOFs e Componentes Principais

O programa desenvolvido faz uso da técnica de Lanczos para o cálculo de funções próprias e autovalores de uma matriz real e simétrica. O método de Lanczos é especialmente apropriado para problemas de dimensões muito grandes, isto é, dimensões para as quais o computador não dispõe de memória suficiente para armazenar os elementos da matriz. No método de Lanczos, não é necessário o armazenamento dos elementos da matriz cujos vectores próprios e valores próprios se pretende calcular, bastando apenas fornecer uma subrotina que produz o vector resultante da multiplicação de um dado vector (arbitrário) por essa matriz.

O método de Lanczos é iterativo e permite calcular um conjunto de vectores próprios e autovalores para os modos com os maiores autovalores. O número exacto de vectores próprios e autovalores calculados depende do número de iterações definidas pelo usuário. O método utiliza uma técnica de tri-diagonalização de matrizes o que permite obter os primeiros vectores próprios e autovalores muito antes do fim dessa tri-diagonalização.

O método de Lanczos (ou uma variação dele) é usado no ECMWF para o cálculo dos vectores singulares usados na previsão numérica de "ensemble". A mesma técnica é usada no cálculo dos vectores singulares do MM5 (De Pondecá e Zou, 2001).

Note-se que para problemas de dimensão relativamente pequena, a solução do problema de valores próprios pode ser convenientemente feita com base em pacotes de álgebra linear como o LAPACK.

Capítulo 4

Resultados

No presente trabalho, fizemos o uso das médias mensais da temperatura de superfície desde 1948 à 1998. Os dados são globais e encontram-se numa malha de resolução de 2.5×2.5 , isto é, 2.5 graus de latitude e 2.5 graus de longitude.

4.1 Pré-tratamento dos dados

No cálculo de EOFs é comum a remoção da média temporal de cada série de tempo, por isso, neste trabalho, começamos por remover as médias temporais.

As EOFs foram calculadas para a região compreendida entre as latitudes -50 graus e +10 graus e longitudes -40 graus e +120 graus, usando uma resolução de 2.5 graus de latitude e 2.5 graus de longitude. Isto corresponde a 65 pontos na direcção Este-Oeste e 25 pontos na direcção Norte-Sul.

O programa foi corrido com 100 (cem) iterações de Lanczos o que permitiu a obtenção de 49 valores próprios e 49 EOFs.

Como ilustração, as figuras 4.1, 4.2 e 4.3 mostram as primeiras três EOFs.

Note-se aqui que as EOFs são adimensionais, pois elas foram calculadas para campos renormalizados, isto é, (campo original-média temporal)/desvio padrão.

A figura 4.1 representa a primeira EOF. A primeira EOF está associada com a maior percentagem de variância presente no conjunto dos dados. Geralmente, os contornos tracejados indicam correlações negativas com os dados originais enquanto que os contornos com linhas cheias indicam uma correlação positiva com os dados originais.

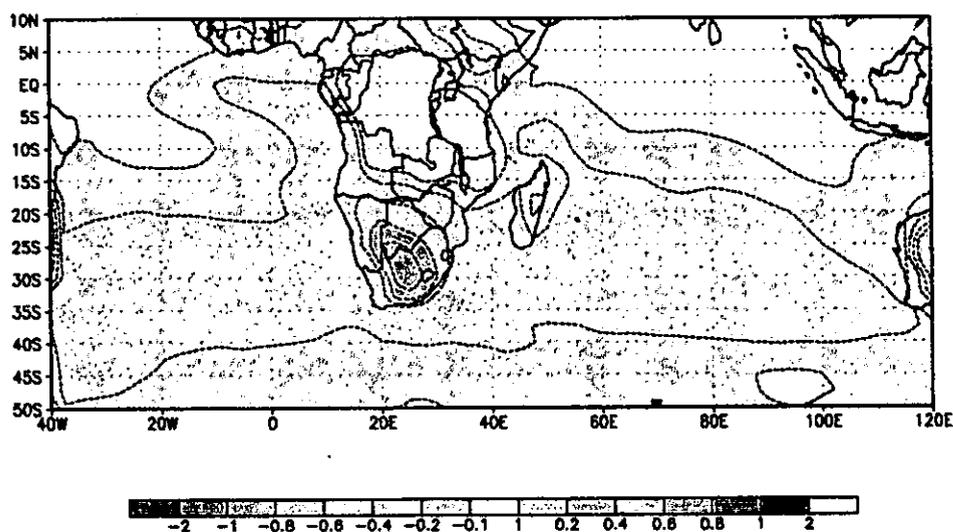


Figura 4.1: Primeira EOF

A figura 4.2 mostra a segunda EOF (EOF2) obtida das anomalias de temperaturas dos dados de Reanalysis do NCEP/NCAR. A EOF2 está associada com a segunda maior variância dos dados, sujeita à condição de ser não correlacionada com a EOF1.

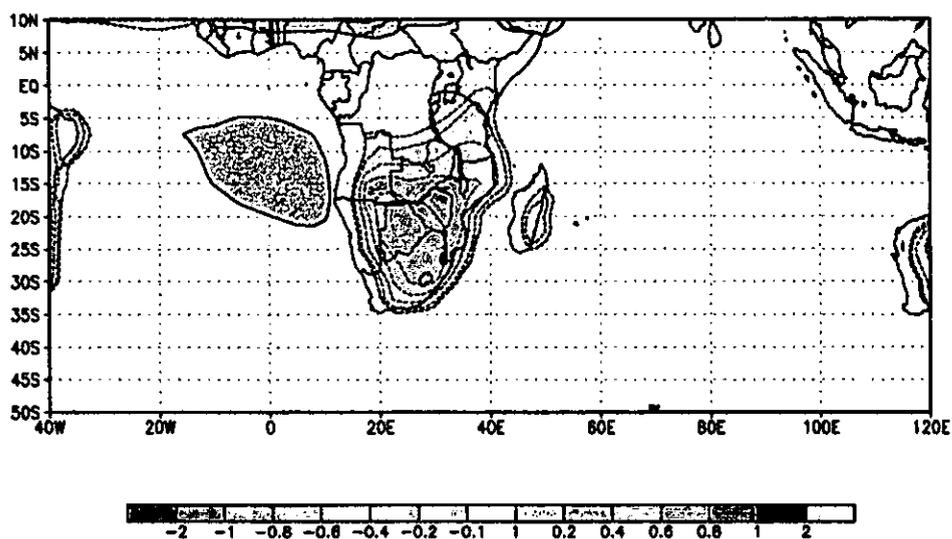


Figura 4.2: Segunda EOF

A figura 4.3 ilustra a terceira EOF (EOF3). A EOF3 está associada com a terceira maior variância dos dados, ela está sujeita à condição de ser não correlacionada nem com a EOF1 e nem Com a EOF2.

Pode-se definir a variância total existente num conjunto de dados multivariados pela soma das variâncias de cada uma das variáveis. Numa matriz de covariâncias, essas variâncias individuais cosnstituem os elementos da diagonal principal. Basta somá-los, portanto, encontrando o traço da matriz para se obter a variabilidade total e , em seguida a contribuição de cada variável.

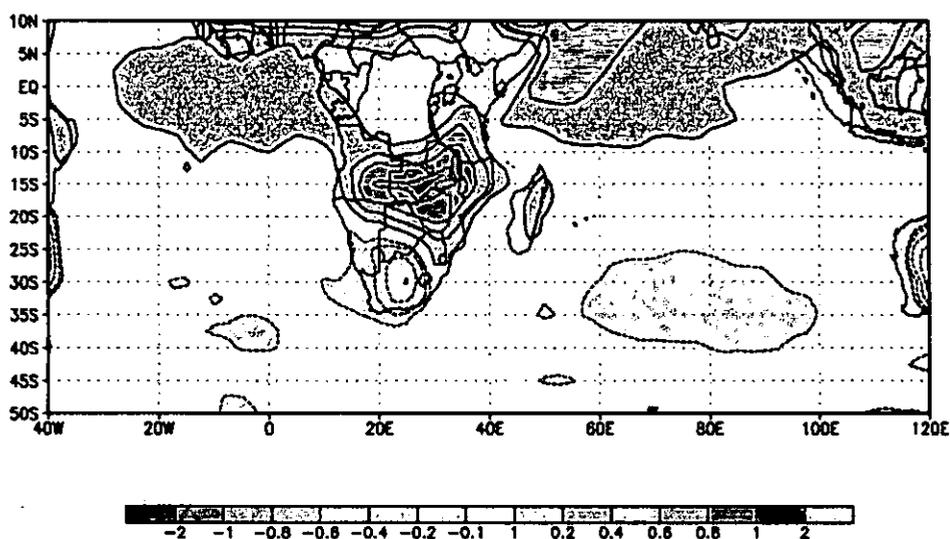


Figura 4.3: Terceira EOF

Existe sempre a questão importante do número de EOFs a ser retido. Não existe uma resposta definitiva para essa questão. Os aspectos a serem considerados incluem a quantidade da variância explicada, o tamanho relativo dos autovalores e a interpretação subjectiva das EOFs. A figura 4.4 tenta ilustrar a quantidade de EOFs a ser retida. Na figura é claramente notável a formação de um "cotovelo" entre o valor próprio 100 e o valor próprio 200. Isto significa que as EOFs que possuem autovalores acima deste intervalo, possuem aproximadamente a mesma magnitude e são relativamente pequenos, por isso, a sua contribuição é desprezível.

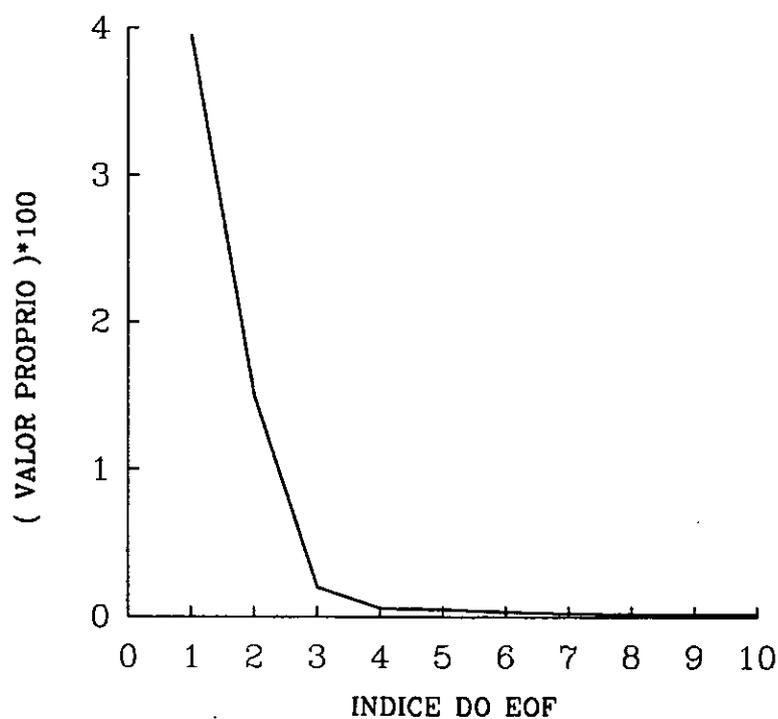


Figura 4.4: Valores próprios como função do número de ordem das EOFs

O método de EOFs é um método de análise que toma toda a variabilidade de um campo meteorológico que evolui com o tempo e a quebra em pequenas oscilações estacionárias (EOFs), cada uma destas oscilações geralmente designada por modo de variabilidade é acompanhada por uma série de tempo e o coeficiente de expansão do modo (componente principal) mostra como esse modo oscila no tempo. Como ilustração, a fig.4.5 mostra as componentes principais para as primeiras duas EOFs durante os primeiros 60 meses, enquanto que a fig.4.6 mostra as componentes principais para a quarta e quinta EOFs durante o mesmo período de tempo.

Na fig.4.5 constata-se que o sinal das componentes principais (coeficientes de expansão) para as EOFs 1 e 2 altera-se, sendo positivo e negativo dentro de um período de 12 meses.

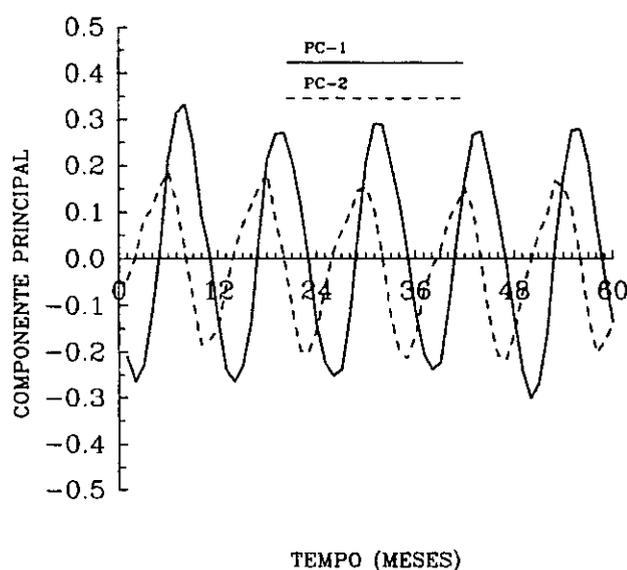


Figura 4.5: Componentes principais correspondentes à EOF1 e EOF2 para 1948 à 1952

Por outro lado, a figura que representa as componentes principais para as EOFs 4 e 5, (fig.4.6), não é de interpretação óbvia. Por exemplo, não há qualquer periodicidade aparente nas séries de tempo das componentes principais. Note-se contudo, que seria possível aplicar-se técnicas de séries de tempo para descobrir sinais (frequências) de maior energia nas componentes principais e tentar relacioná-los com a física e dinâmica conhecidas da atmosfera.

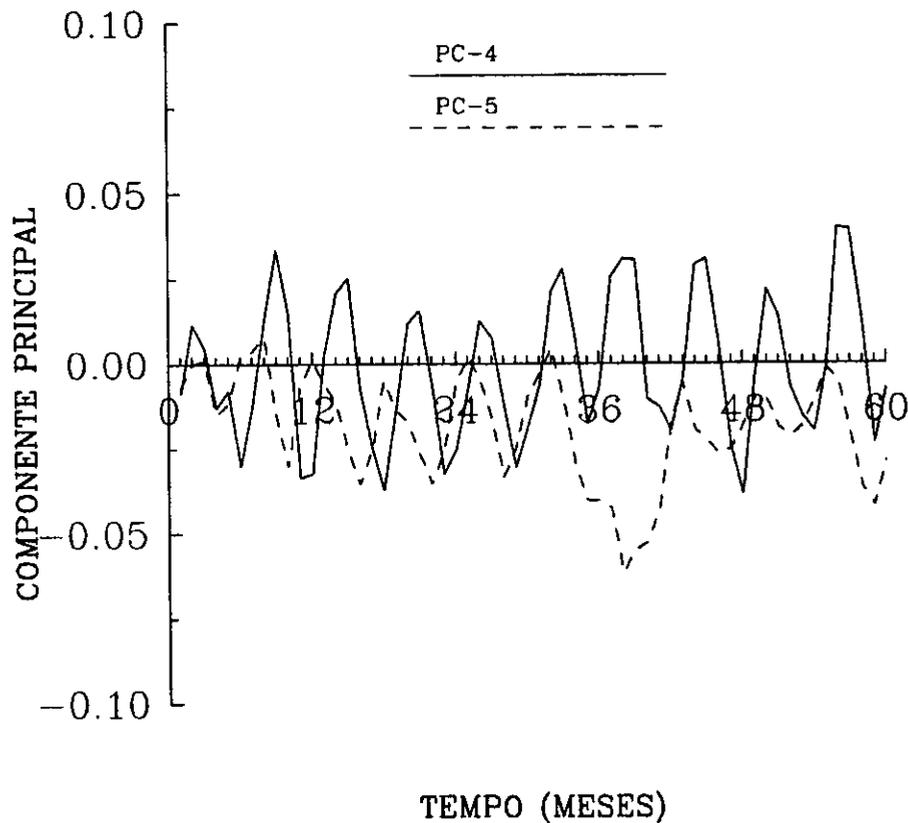


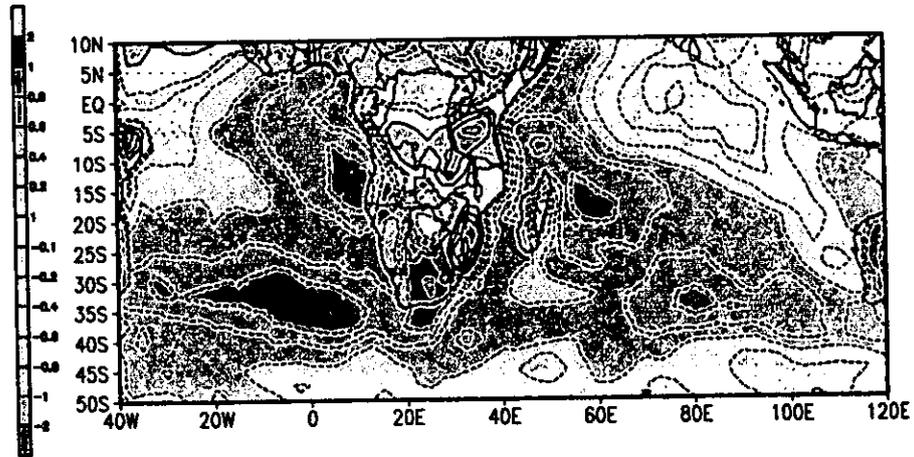
Figura 4.6: Componentes principais correspondentes à EOF4 e EOF5 para 1948 à 1952

Uma das aplicações da decomposição em EOFs é a reconstrução de um campo que é uma versão do campo original, mas, mais comprimido e menos ruidoso. Como ilustração, fez-se neste trabalho, a reconstrução do campo original usando um número limitado de EOFs (figuras 4.7, 4.8 e 4.9). Nessas figuras, cada painel superior mostra o campo de anomalias normalizadas enquanto que os painéis inferiores mostram as reconstruções feitas com base em algumas EOFs.

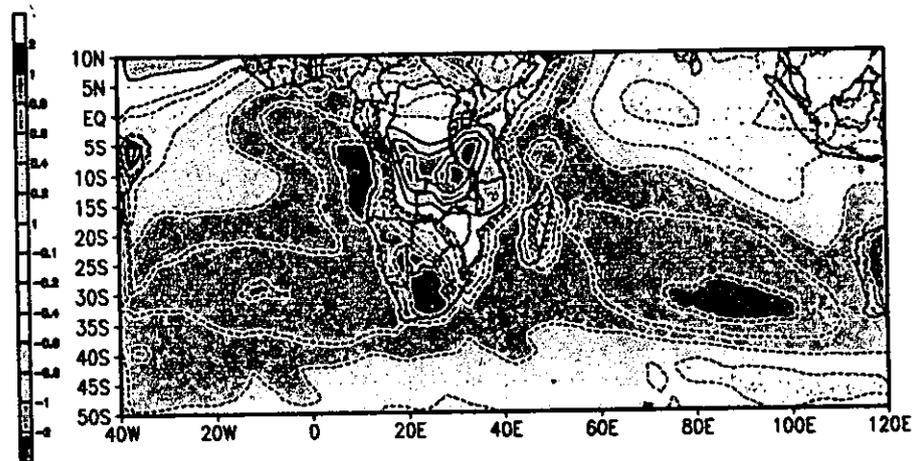
As figuras mostram claramente que o uso de 5 EOFs é suficiente para captar a maior parte da variância do campo original de anomalias. Este resultado está em concordância com a figura 4.4, pois, a figura 4.4 mostra que os valores próprios associados às EOFs caem rapidamente entre a primeira e a quinta EOFs (aproximadamente) e tomam um comportamento assintótico para EOFs com valores crescentes de índice (EOF 6, 7, 8, etc).

ANOMALIAS DE TEMP. MENSAS DE SUPERFICIE
SEPTEMBRO DE 1948

CAMPO ORIGINAL



CAMPO RECONSTITUÍDO COM 5 EOFs

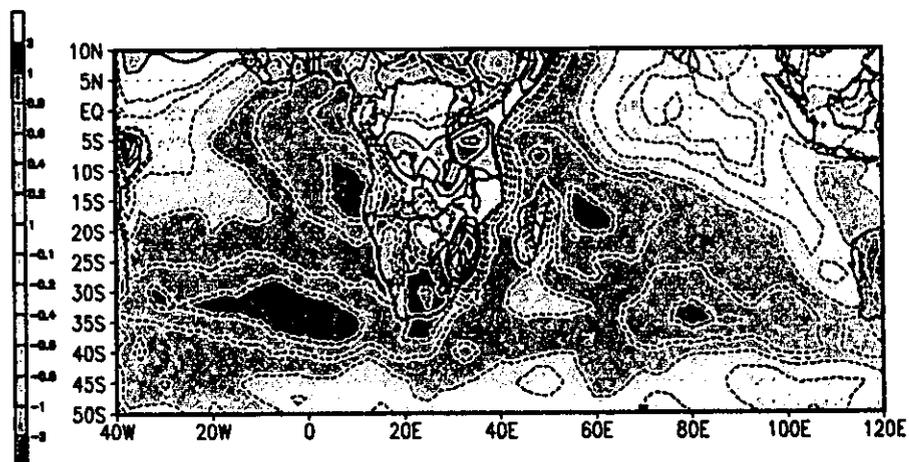


GrADS: COLA/IGES

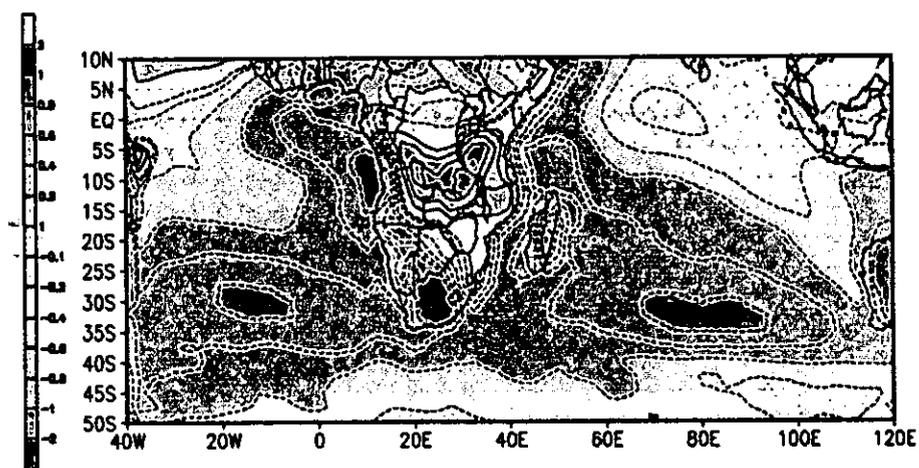
Figura 4.7: Campo original e campo reconstituído na base de 5 EOFs

ANOMALIAS DE TEMP. MENSAS DE SUPERFICIE
SEPTEMBRO DE 1948

CAMPO ORIGINAL



CAMPO RECONSTITUIDO COM 10 EOFs

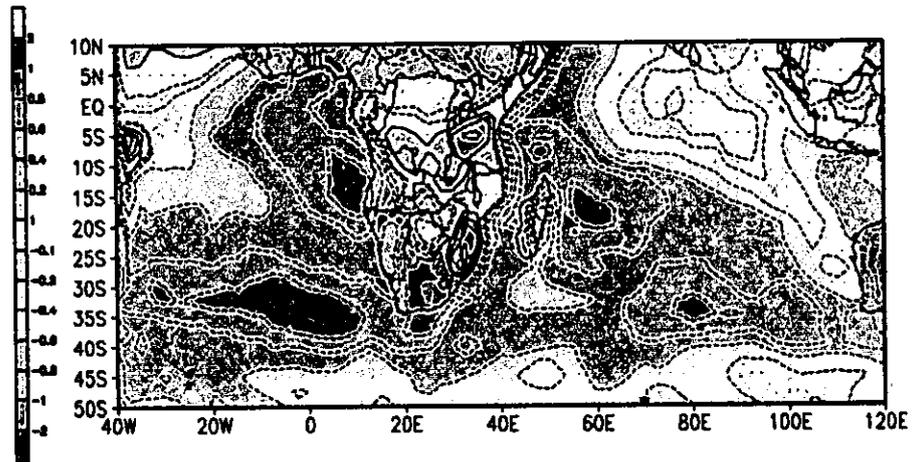


GRADS: COLA/IGES

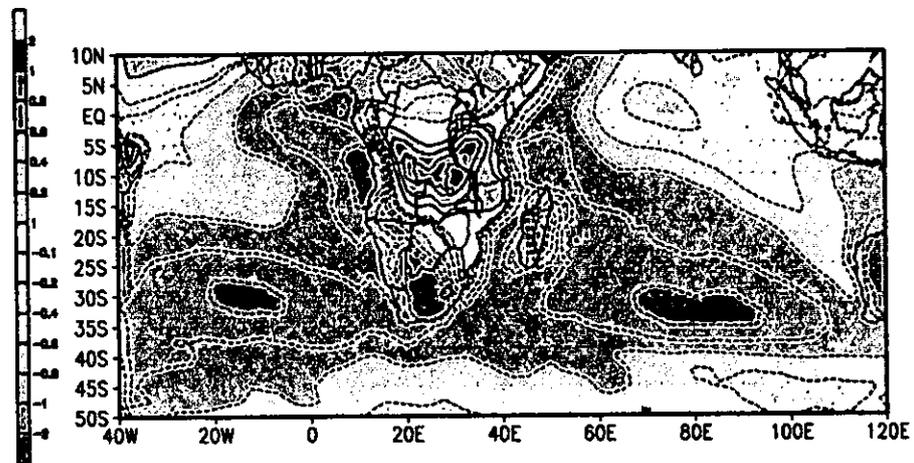
Figura 4.8: Campo original e campo reconstituído na base de 10 EOFs

ANOMALIAS DE TEMP. MENSAIS DE SUPERFICIE
SEPTEMBRO DE 1948

CAMPO ORIGINAL



CAMPO RECONSTITUIDO COM 49 EOFs



GrADS: COLA/IGES

Figura 4.9: Campo original e campo reconstituído na base de 49 EOFs

Capítulo 5

Conclusões e recomendações

5.1 Conclusões

A redução estatística do clima global para o clima regional e local é feita através de relações estabelecidas entre a circulação atmosférica de larga escala e a resposta climatológica local (por exemplo, temperatura ou precipitação local), por isso, um modelo estatístico de redução do clima geralmente pressupõe a classificação da circulação atmosférica de larga escala em padrões de circulação.

São várias as classificações usadas para os padrões de circulação, a título de exemplo, mencione-se a classificação de Lamb (1950, 1972) que é muitas vezes usada para a Europa. A vantagem deste tipo de método consiste no facto de possuir uma boa base física, contudo, um método deste tipo, desenvolvido para uma dada região, não é geralmente transferível para uma outra região. A desvantagem mais significativa é ainda o facto deste tipo de método exigir, amiúde, o uso de enormes recursos humanos, pois os mapas de circulação geral representam muitas vezes um número enorme de séries de tempo.

O nosso objectivo final é o uso da técnica estatístico-dinâmica para a redução do clima global para o clima da nossa região. Precisa-se, para isso, a classificação da circulação atmosférica em padrões de circulação. Para cada padrão, faz-se uma redução dinâmica do clima com a ajuda

de um modelo numérico de mesoescala. A combinação da redução dinâmica com a frequência de ocorrência dos padrões de circulação resultará em informação sobre o clima regional e local.

Este trabalho representa uma pequena contribuição para o objectivo acima mencionado. O trabalho foi basicamente concentrado no desenvolvimento de um programa em Fortran que faz o cálculo de EOFs. Com base no programa foram calculados 49 valores próprios (autovalores) e 49 autovectores, por isso, dizemos de viva voz que o objectivo fundamental do presente trabalho foi satisfatoriamente alcançado. O programa gerado é uma ferramenta poderosa para a repartição de uma série de tempo em diferentes tipos de tempo de larga escala (primeiro passo do método de Frey-Buness).

Conforme a descrição feita na secção 2.3, o objectivo principal das EOFs é a redução da dimensionalidade dos dados, tal redução está bem patente nos exemplos aquí ilustrados, pois, com apenas 5 EOFs foi possível fazer a reconstrução do campo e captar a maior variância do campo original.

Outra conclusão óbvia é que o método de EOFs permite, de facto, a remoção de informação redundante e uma explicação da variância com poucos modos, outra vantagem é a sua adequabilidade no tratamento de largas quantidades de dados, contudo, o método é essencialmente estatístico não assegurando, por isso, uma interpretação directa da "física" associada aos dados.

Os resultados mostrados neste trabalho são de difícil interpretação, pois, a informação contida nos mapas de temperatura de superfície é enorme e a sua sintetização pode não ser óbvia. Note-se que o domínio utilizado neste exemplo comporta 1625 pontos de malha estando cada ponto associado a uma série de tempo com 612 pontos. Os dados originais são pois caracterizados por uma grande variabilidade temporal.

5.2 Recomendações

Em Moçambique e em África em geral, a modelação do clima local e regional encontra-se numa fase emergente. A criação da capacidade de trabalho nesta área deve constituir, sem dúvidas, uma componente importante dos esforços actuais para o desenvolvimento sustentável. As acções concretas a realizar devem ter como objectivo o conhecimento científico do clima da região do passado e do presente e a simulação de cenários credíveis de variações climáticas futuras da região.

Propõe-se ao departamento de física, a criação e validação de um sistema de modelação do clima local com base no método de redução estatístico-dinâmica do clima global. O sistema será de grande utilidade para várias instituições nacionais, como por exemplo, os ministérios da agricultura e desenvolvimento rural, dos recursos minerais e energias e da coordenação da acção ambiental pois, com a realização futura dos estudos de processos, essas instituições terão um valor acrescentado na sua capacidade de adopção de estratégias de desenvolvimento com uma base científica sólida.

Apela-se também, a criação de um banco de dados no Departamento de Física para a investigação em climatologia e mudanças climáticas locais e regionais.

Bibliografia

- [1] Bjornsson, H. e S.A. Venegas, 1997: *A Manual for EOF and SVD Analysis of Climate Data*. Department of Atmospheric and Oceanic Sciences and Centre for Climate and Global Change Research, McGill University.
- [2] Busch U. e D. Heimann, 2001: *Statistical-dynamical extrapolation of nested regional climate simulations*, Clim. Res, 19.
- [3] Da Silva, I.C.S., 2000 : *Sistema de visualização de dados para o centro de pesquisas meteorológicas da UFPel*, dissertação de bacharelato, Universidade Federal de Pelotas, Pelotas.
- [4] De Pondeca, M.S.F.V, 2001a: *The "Energy" Metric Singular Vectors of MM5 Model. Source Code Documentation*, FSU Technical Note. FSU/TN-3-ADJ. The Florida State University, Dept. of Meteorology, Tallahassee.
- [5] De Pondeca, M.S.F.V 2001d: *A description of the operator for the GPS total zenith delay used with the MM5 mesoscale model*, FSU Technical Note. FSU/TN-2-ADJ. The Florida State University, Dept. of Meteorology, Tallahassee.
- [6] Doty, B., 1995: *The Grid Analysis And Display System*. Center for Ocean-Land-Atmosphere Interactions, Department of Meteorology, University of Maryland.
- [7] Frey-Buness A., D. Heimann, e R. Sausen, 1995: *A statistical-dynamical downscaling procedure for global climate simulations*. Theor. Appl. climatol., 50.
- [8] Giorgi, F., 1990: *Simulation of regional climate using a limited area model in a general circulation model*, J. Clim, Vol.3.

- [9] Haas, J., 2002: *Utilização da previsão meteorológica quantitativa de chuva na previsão de afluências*. CIÊNCIAONLINE-INSS n° 1676-5893.
- [10] Hannachi, A., 2004: *A Primer for EOF Analysis of Climate Data*. Department of Meteorology, University of Reading, U.K.
- [11] IPCC GTI, 1995: *Climate Change 1995 - The Science of Climate Change: Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Changes*, J.T.Houghton, L.G.Meira Filho, B.A. Callander, N. Harris, A.Kattenberg e K.Maskell(eds), Cambridge University Press, Cambridge, UK.
- [12] IPCC, 2001: *Climate Change 2001 - Impacts, Adaptation and Vulnerability*. Q. K. Ahmad, Oleg Anisimov, Nigel Amell, Sandra Brown, Ian Burton, Max Campos,..., Tom Wilbanks, Alistair Woodward, David Wratt. Cambridge University Press, Cambridge, UK.
- [13] Kalnay, E.; Deaven, D., Gandi, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J.; Zhu, Y.; Leetmaa, A.; Reynolds, R.; Chelliah, M.; Higgins, W; Wang, J.; Janowiak, J.; Mo, C.K, Ropelewski, C.; Wang, J.; Jenne, R.; Joseph, D., 1996: The NCEP/NCAR 40-year Reanalysis Project, Bulletin of the American Meteorological Society.
- [14] Lamb, H.H., 1950: *Types and Spells of Weather around the year in the British Isles*. Quart.J.R.Met.Soc., 76, 393-438.
- [15] Lamb, H.H., 1972: *British Isles Weather types and a register of daily Sequence of circulation patterns, 1861-1971*. Geophysical Memoir 116, HMSO, London, 85pp.
- [16] Meier, M., Doscher, R., Coward, A.C., Nycander, J., Doos, K., 1999: *Rosby Centre regional Ocean climate model: model description (version 1.0) and first results from the hindcast period 1992-1993*. Reports Oceanography No. 26, SMHI, SE-60176, Norrköping, Sweden.
- [17] Moriondo, M., 2002-2003: *Dowsaling delle temperature minime e massime a partire da un GCM: il caso di Firenze*, tese de mestrado, università degli Studi di Firenze.

- [18] Muller, M., 2000: *Investigação da aplicabilidade do uso de modelos hidrológicos e redes neurais na propagação de vazões*, dissertação de mestrado, UFPR CEHPAR, Curitiba.
- [19] Obasi, G.O.P, 2003: *O nosso clima no futuro*. Comunicação pública do secretário geral da OMM alusiva ao dia meteorológico mundial.
- [20] Peixoto, J.P. e Oort, A.H., 1992 : *Physics of Climate*, American Institute of Physics, New York.
- [21] Quadro, M., 2004: *Curso técnico de meteorologia*. Centro Federal de Educação Tecnológica de Santa Catarina, Brasil. Disponível em: <http://www.cefetsc.edu.br/meteorologia/modulo1/pagina.htm>. Visitado em: 10.07.04.
- [22] Tomé, R.F, 2004: *Previsão do tempo com modelos de mesoescala: Casos de estudo com o modelo MM5 nos Açores*. Universidade de Lisboa, Lisboa.
- [23] Simões, A.C., 2000: *Sistematização e visualização de dados meteorológicos 4D para Portugal continental. Análise objectiva e FDDA*. Universidade de Lisboa, Lisboa.
- [24] Tanajura, C.A.S., 1996: *Modeling and analysis of the South American Summer climate*, Dissertação(PhD), University of Maryland, Maryland.
- [25] Tucci, C.E.M; Clarke, R.T.; Dias, P.L and Collischonn, 2000: *Previsão de vazões com base na previsão climática*, projecto BRA 00/029.
- [26] Tucci, C.E., 2002: *Impactos da variabilidade climática e do uso dos solos nos recursos hídricos*, Agência Nacional de águas (ANA), Brazil.
- [27] Wilby, R.L.; Wigley T.M.L; Conway D.; Jones, P.D; Hewitson B.C; Main J. e Wilks D.S 1998: *Statistical downscaling of general circulation model output: a comparison methods*, Water Resources Research, vol.34.
- [28] Wilks, D.S., 1995: *Statistical Methods in the Atmospheric Sciences*, academic press, New York.

- [29] Venegas, S.A, 2001 : *Statistical Methods for Signal Detection in Climate*, DCESS Report # 2, University of Copenhagen, Denmark.

Apêndice A

Programas

A.1 Programa para a geração das 15 primeiras EOFs

```
*****  
*****  
*****  
*****  
'reset'  
'set display color white'  
*****  *****  *****  
*      Variable Definition  
  
neofs1=16  
  
'open eof.des_0001'  
'open eof.des_0002'  
'open eof.des_0003'  
'open eof.des_0004'  
'open eof.des_0005'
```

```
'open eof.des_0006'  
'open eof.des_0007'  
'open eof.des_0008'  
'open eof.des_0009'  
'open eof.des_0010'  
'open eof.des_0011'  
'open eof.des_0012'  
'open eof.des_0013'  
'open eof.des_0014'  
'open eof.des_0015'
```

```
*****
```

```
xinitial=-40.
```

```
xfinal=120.
```

```
yinitial=-50.
```

```
yfinal=10.
```

```
'set lon 'xinitial ' 'xfinal
```

```
'set lat 'yinitial ' 'yfinal
```

```
lcount=1
```

```
while(lcount<neofs1)
```

```
    'set mpdset mres'
```

```
    if (lcount=1)
```

```
        'set vpage 0.0 11.0 0.0 8.5'
```

```
    endif
```

```
'set parea 1.5 9.5 2.0 7.0'

'suru_rgbset.gs'
'set rgb 16 255 255 255'

'set clevs -2. -1. -.8 -.6 -.4 -.2 -.1
0 1. .2 .4 .6 .8 1. 2.'
'set ccols 59 58 56 48 46 44
42 16 16 22 24 27 29 88 89 '

* 'set clevs -12 -11 -10 -9 -8 -7 -6 -5 -4 -3
-2 -1 -0.5 0 0.5 1 2 3 4 5 6 7 8 9 10 11 12'
* 'set ccols 59 59 58 57 56 49 48 47 46 45
44 43 42 16 16 22 23 24 25 26 27 28 29 87 88 89 89 '

* 'set clevs -12 -11 -10 -9 -8 -7 -6 -5 -4 -3
-2 -1 0.5 0 0.5 1 2 3 4 5 6 7 8 9 10 11 12'
* 'set ccols 59 59 58 57 56 49 48 47 46 45 44
43 42 41 21 22 23 24 25 26 27 28 29 87 88 89 89 '

* 'set clevs -12 -11 -10 -9 -8 -7 -6 -5 -4
-3 -2 -1 0.5 -0.25 0 0.25 0.5 1 2 3 4 5 6 7 8 9 10 11 12'
* 'set ccols 59 59 58 57 56 49 48 47 46 45
44 43 42 41 16 16 21 22 23 24 25 26 27 28 29 87 88 89 89 '

'set gxout shaded'
'd eof.'lcount
'run my_cbarn_first 0.75 0 5.5 1.7'
'set gxout contour'
```

```
'set ccolor 1'
'set cthick 4'
'set black 0 0'
'set clab off'
'set clevs -2. -1. -.8 -.6 -.4 -.2 -.1
  1. .2 .4 .6 .8 1. 2.'
'd eof.'lcount

'enable print gg_'lcount
'print'
'disable print'

if (lcount<(neofs1-1))
  'clear'
endif

lcount=lcount+1
endwhile
'!gxps -c -i gg_1 -o gg_1.ps'
'!gxps -c -i gg_2 -o gg_2.ps'
'!gxps -c -i gg_3 -o gg_3.ps'
'!gxps -c -i gg_4 -o gg_4.ps'
'!gxps -c -i gg_5 -o gg_5.ps'
'!gxps -c -i gg_6 -o gg_6.ps'
'!gxps -c -i gg_7 -o gg_7.ps'
'!gxps -c -i gg_8 -o gg_8.ps'
'!gxps -c -i gg_9 -o gg_9.ps'
'!gxps -c -i gg_10 -o gg_10.ps'
```

```
'!gxps -c -i gg_11 -o gg_11.ps'
'!gxps -c -i gg_12 -o gg_12.ps'
'!gxps -c -i gg_13 -o gg_13.ps'
'!gxps -c -i gg_14 -o gg_14.ps'
'!gxps -c -i gg_14 -o gg_15.ps'

'!cat gg_1.ps gg_2.ps gg_3.ps gg_4.ps gg_5.ps
gg_6.ps gg_7.ps gg_8.ps gg_9.ps gg_10.ps gg_11.ps
gg_12.ps gg_13.ps gg_14.ps gg_15.ps > allplots.ps'

* 'print'
* 'set strsiz 0.12 0.12'
* 'draw string 1.5 10.6 2m TEMP DIFF (K) (EXP6 - NEW OPR)'
* 'draw string 1.5 10.4 INITIAL TIME IS 12Z 13 MAY 2004 '
* 'draw string 4.7 10.4 OOH-FCT '

* 'set strsiz 0.15 0.15'
* 'draw string 1.5 10.6 24h ACCUM. PRECIP (mm) VALID 12Z 14 May 04'
* 'draw string 5.5 10.4 24H-FCT '

* 'draw string 1.5 10.0 24H-FCT '
* 'draw string 1.5 5.25 36H-FCT '

*'enable print gg'
*'print'
*'disable print'
*'!gxps -c -i gg -o gg.ps'
'quit'
```

```
*****
```

?

A.2 Programa para a reconstrução dos campos

```
PROGRAM recons_field
include '../dev/param_dims.incl'
PARAMETER (NEOFS=49, LTIME=9)

C

DIMENSION A(NX,NY),B(NX,NY)
DIMENSION EV(KT2),BNDEV(KT2),AUX(NT)

dimension F(nx,ny),D(nx*ny)

OPEN (13, FILE='results.out',
+      FORM='FORMATTED',STATUS='OLD')
OPEN (14, FILE='results2.out',
+      FORM='FORMATTED',STATUS='OLD')
REWIND(13)
REWIND(14)

READ (13,*) KT
READ (13,*) (EV(J),J=1,KT)
READ (13,*) (BNDEV(J),J=1,KT)
PRINT*, 'in recons_field, KT-1=',KT

READ (14,*) KT
READ (14,*) (EV(J),J=1,KT)
READ (14,*) (BNDEV(J),J=1,KT)
PRINT*, 'in recons_field, KT-2=',KT
```

```
DO J=1,NY
DO I=1,NX
  B(I,J)=0.
ENDDO
ENDDO

DO NN=1,NEOFS
  READ (13,*) N1
  READ (13,*) A
  READ (14,*) N2
  READ (14,*) AUX
  PRINT*, 'in recons_field, NN,N1,N2=', NN,N1,N2
  DO J=1,NY
  DO I=1,NX
    B(I,J)=B(I,J)+AUX(LTIME)*A(I,J)
  ENDDO
  ENDDO
ENDDO
CALL outgrad_rec(B,NX,NY,LTIME,
+   RLONMAPO,DLONMAP,RLATMAPO,DLATMAP)

END
```

C

C=====

```
subroutine outgrad_rec(field,nx,ny,kselect,
+   rlonmap0,dlonmap,rlatmap0,dlatmap)
```

```
dimension field(nx,ny)
character*50 dsname,dsctl,title
character*4 clun
data title/'EOF ANALYSIS'/
character*80 datdes(1000)
character*1 blank
data blank/' '/
data undef/-9.99e33/

np=1
kout=1
startp=1.
pinc=1.

write (clun(1:4),'(i4.4)') kselect
dsname='rec.dat_'//clun
dsctl='rec.des_'//clun

open(70,file=dsctl,form='formatted')
open(80,file=dsname,form='unformatted')

do i=1,1000
  write(datdes(i),10)(blank,k=1,80)
10  format(80a1)
end do
write(datdes(1),100)'^'//dsname
100 format('DSET ',a50)
```

```
write(datdes(2),200)
200 format('options little_endian')
write(datdes(3),300)title
300 format('TITLE ',a50)
write(datdes(4),400)undef
400 format('UNDEF ',e11.2)
write(datdes(5),500)nx,r lonmap0,d lonmap
500 format('XDEF ',i5,' LINEAR ',f7.2,f7.2)
write(datdes(6),600)ny,r latmap0,d latmap
600 format('YDEF ',i5,' LINEAR ',f7.2,f7.2)
next=7
write(datdes(next),700)np,startp,pinc
700 format('ZDEF ',i5,' LINEAR ',f7.2,f7.2)
next=next+1
write(datdes(next),800)kout
800 format('TDEF ',i5,' LINEAR 0Z23may1992 24hr')
next=next+1
write(datdes(next),810)
810 format('VARS 1')
next=next+1
write(datdes(next),910)np
910 format('rec ',i5,' 99 rec ')
next=next+1
write(datdes(next),980)
980 format('ENDVARS')
last=next
write(70,2000)(datdes(i),i=1,last)
2000 format(a80)
```

```
write (35,*) field
```

```
write (80) field
```

```
end
```