



**AVALIAÇÃO DO MODELO ETA
EM SITUAÇÕES DE CHUVAS INTENSAS ASSOCIADAS
COM DESLIZAMENTOS**

**RELATÓRIO FINAL DE PROJETO DE INICIAÇÃO CIENTÍFICA
(PIBIC/CNPq/INPE)**

**Jessica Motta Guimarães (UFRJ, Bolsista PIBIC/CNPq)
E-mail: jessica.meteoro@gmail.com**

**Dra. Chou Sin Chan (DMD/CPTEC/INPE, Orientadora)
E-mail: chou@cptec.inpe.br**

COLABORADORA

Dra. Claudine Pereira Dereczynski (IGEO/UFRJ)

Julho de 2008

O presente relatório enumera as atividades realizadas no decorrer da Bolsa de Iniciação Científica PIBIC/ CNPq no período de julho de 2007 até julho de 2008.

SUMÁRIO

1. INTRODUÇÃO.....	2
2. METODOLOGIA.....	3
3. RESULTADOS PRELIMINARES E TRABALHOS FUTUROS	6
4. REFERÊNCIAS BIBLIOGRÁFICAS	9
ANEXO I Trabalho Submetido e aceito no Congresso Brasileiro de Meteorologia..	150
ANEXO II Programa desenvolvido FORTRAN	15.

1. INTRODUÇÃO

As previsões numéricas de tempo produzidas atualmente ainda não permitem determinar com grande precisão e confiabilidade a localização e intensidade exatas de um evento chuvoso intenso. Casos de precipitação extrema estão comumente associados com uma série de graves problemas, tais como ocorrência de deslizamentos e enchentes, além de provocarem um grande número de mortes, acidentes e prejuízos materiais.

Indicadores do potencial de ocorrência de um evento crítico são importantes para que um alerta seja emitido com suficiente antecedência à defesa civil e população em geral, para tomadas de providências que reduzam os possíveis prejuízos.

Diversas técnicas vêm sendo utilizadas com a finalidade de indicar de forma eficiente a ocorrência de eventos considerados extremos, associados com deslizamentos e demais problemas por eles provocados. Já foram realizados estudos utilizando índices de instabilidade atmosférica com a finalidade de prever eventos que poderiam provocar chuvas intensas e estar associados com deslizamentos e outros problemas. Índices tais como o ILEV, K, ITT, CILEV, CK, CITT (os três últimos índices correspondem à combinação dos três primeiros com o efeito da convergência de umidade em 850 hPa) foram gerados para diferentes resoluções espaciais do modelo Eta (15km, 10km, 5km) para situações em que ocorreram chuvas intensas localizadas na região da Serra do Mar.

De forma geral conclui-se que os índices conseguiram localizar áreas onde há maior instabilidade próximas ao local do evento. Cada índice apresentou um desempenho característico, representando tanto localmente quanto temporalmente os eventos de formas distintas. Os índices localizaram melhor os máximos de instabilidade quando associados com a convergência de umidade. Para os estudos específicos feitos, os índices K e CILEV apresentaram melhores resultados de forma geral.

Na atual fase de estudo, outra forma de implementar a previsibilidade de eventos extremos de precipitação vem sendo estudada. Assim sendo, no presente trabalho será apresentada a metodologia do trabalho desenvolvida e que consiste em utilizar técnicas de agrupamento com a finalidade de distinguir com antecedência eventos meteorológicos propensos a provocar desastres como deslizamentos e enchentes dos demais eventos. Um artigo (ver Anexo I) contendo os resultados preliminares desta fase do trabalho foi aceito para publicação nos Anais do XV Congresso Brasileiro de Meteorologia. Tal evento será realizado em São Paulo (SP) no período de 24 a 29 de agosto de 2008.

2. METODOLOGIA

Inicialmente foram coletados todos os casos de ZCAS ocorridos entre os anos de 1998 e 2007, baseados na análise mensal da Revista Climanálise e elaborada a Tabela 1. Posteriormente foram identificados aqueles casos associados com a ocorrência de deslizamentos, segundo os registros de acidentes do Instituto de Pesquisas Tecnológicas (IPT), referentes a deslizamentos ocorridos no Brasil, assim como notícias publicadas na mídia.

Tabela 1 - Datas de todos os casos catalogados de ZCAS entre 1998 e 2007. A coluna da direita indica os casos associados com ocorrência de deslizamentos.

Período ZCAS	Deslizamentos
14-19/11/1997	
12-16/02/1998	X
20-25/11/1998	
06-18/01/1999	X
24/10/1999 - 03/11/1999	
17-25/11/1999	

08-14/12/1999	
16-20/12/1999	
01-08/01/2000	X
21-24/01/2000	
6-13/02/2000	
01-08/12/2000	
17-22/12/2000	
01-04/01/2001	
01-06/11/2001	
16-21/11/2001	
17-21/12/2001	
24-28/12/2001	
04-07/02/2002	
16-24/02/2002	
10-16/12/2002	
27/12/2002 - 07/01/2003	
13-19/01/2003	X
25/1/2003 - 01/02/2003	X
02-06/01/2004	
10-20/01/2004	
25-29/01/2004	X
07-11/02/2004	
20-24/02/2004	
20-25/11/2004	
09-14/12/2004	
21-25/12/2004	
17-21/01/2005	X
13-22/02/2005	
01-07/03/2005	X
15-20/03/2005	X
10-15/11/2005	
17-21/11/2005	
24-28/11/2005	
11-16/12/2005	X
24-29/12/2005	
01-08/01/2006	X
27/01/2006 - 02/02/2006	X
08-13/02/2006	X
07-16/03/2006	
10-14/11/2006	
26/11/2006 - 02/12/2006	
07-16/12/2006	X
27/12/2006 - 16/01/2007	X
22-27/01/2007	X
30/01/2007 - 09/02/2007	X
12-17/02/2007	
19-23/03/2007	X
Total - 54	Total - 18

A partir dessas datas foram selecionadas rodadas do modelo Eta com resolução espacial de 40 km, com previsão para 48 horas para o primeiro dia de cada evento de ZCAS listado na tabela anterior. Foram utilizadas rodadas com pouca resolução tendo em vista que o evento estudado é de escala sinótica por definição. Foi selecionado um ponto no domínio onde foram observadas grandes taxas de precipitação, próximo à

cidade de Caraguatatuba (lat: -23,6° e lon: -45,4°) para extrair os dados das variáveis estudadas. As variáveis que serão utilizadas no método de agrupamento são:

- a. Geopotencial em 500 hPa
- b. Umidade em 500 hPa
- c. Vento em 850 hPa
- d. Precipitação acumulada em 72 h
- e. Divergência de Umidade em 850 hPa
- f. Espessura entre os níveis 500 e 1000 hPa
- g. Vorticidade em 500 hPa

A partir dessa fase foi elaborado um programa em linguagem FORTRAN (ver Anexo II) que visa separar todos os casos listados em dois grupos distintos:

- a) Grupo A: casos de ZCAS sem ocorrência de deslizamentos.
- b) Grupo B: casos de ZCAS com ocorrência de deslizamentos.

Existem diversas técnicas de análise grupos, porém a que mais se mostrou conveniente para o trabalho aqui proposto foi a K-Means (SILVA, 1989). Essa metodologia é propícia para situações em que se desejam poucos grupos finais, e, portanto, grupos com vários membros.

O método K-Means é um dos métodos de partição mais conhecidos e utilizados. É uma técnica que se baseia na análise de centróides que servem como representação dos membros de cada grupo. Para agrupar os diversos membros, os valores de cada um são comparados com o valor dos centróides, seguindo o esquema a seguir:

Sendo k o número final de grupos desejados (dois grupos finais, no estudo aqui apresentado), e D um conjunto de dados com n objetos.

- (1) Devem ser escolhidos arbitrariamente k objetos do grupo D como os centros iniciais dos grupos;
- (2) Deve-se atribuir cada objeto para o grupo que tenha o centro mais próximo do valor do objeto;
- (3) A partir desse momento são calculados os novos centros dos grupos, que correspondem às médias dos objetos dos grupos;

(4) Os passos (2) e (3) devem ser repetidos até não ocorrerem mais mudanças nos centros antigos e os calculados.

Partindo desse algoritmo foi desenvolvido o programa em anexo e foram obtidos os resultados que serão aqui apresentados.

3. RESULTADOS PRELIMINARES E TRABALHOS FUTUROS

Inicialmente foram separados 23 casos dos 55 listados na tabela 1, pois algumas rodadas do modelo encontravam-se com problemas nas datas iniciais. Dentre esses 23 casos, 11 deles não estavam associados com deslizamentos e 12 estavam associados com esse tipo de desastre. A primeira variável a ser testada foi geopotencial em 500 hPa. A seguir é apresentada uma lista com os 23 casos testados inicialmente.

Tabela 2 – Informações sobre os 23 casos analisados inicialmente.

Número de referência	Período da ZCAS	Data inicial - todos as 12 Z	Deslizamentos
1	25-29/01/2004	23/01/2004	X
2	20-25/11/2004	18/11/2004	
3	09-14/12/2004	7/12/2004	
4	21-25/12/2004	19/12/2004	
5	17-21/01/2005	15/1/2005	X
6	13-22/02/2005	12/2/2005	
7	01-07/03/2005	28/2/2005	X
8	15-20/03/2005	13/3/2005	X
9	10-15/11/2005	8/11/2005	
10	17-21/11/2005	15/11/2005	
11	24-28/11/2005	22/11/2005	
12	11-16/12/2005	9/12/2005	X
13	24-29/12/2005	22/12/2005	
14	01-08/01/2006	30/12/2005	X
15	27/01/2006 - 02/02/2006	25/1/2006	X
16	09-13/02/2006	7/2/2006	X
17	07-16/03/2006	5/3/2006	
18	10-14/11/2006	8/11/2006	
19	26/11/2006 - 02/12/2006	24/11/2006	
20	07-16/12/2006	5/12/2006	X
21	27/12/2006 - 16/01/2007	25/12/2006	X
22	22-27/01/2007	20/1/2007	X
23	30/01/2007 - 09/02/2007	28/1/2007	X

O programa agrupou os casos da seguinte forma:

Grupo A		Grupo B	
4.000000	5791.320	1.000000	5857.070
6.000000	5771.360	2.000000	5831.550
9.000000	5789.830	3.000000	5843.850
16.000000	5782.230	5.000000	5854.360
18.000000	5815.790	7.000000	5868.790
22.000000	5814.210	8.000000	5866.910
		10.000000	5847.630
		11.000000	5860.830
		12.000000	5843.800
		13.000000	5868.510
		14.000000	5878.860
		15.000000	5829.990
		17.000000	5856.200
		19.000000	5866.010
		20.000000	5861.700
		21.000000	5861.180
		23.000000	5837.510

Temos que para cada grupo apresentado a primeira coluna representa os números de referência de cada membro e que a segunda coluna mostra os valores de geopotencial comparados. O grupo A é o referente aos casos em que não ocorreram deslizamentos, já o grupo B está relacionado aos casos em que ocorreram deslizamentos. De forma geral conclui-se que os casos com menores valores de geopotencial foram separados no grupo A, como não associados com deslizamentos e os com valores mais elevados no grupo B, como casos em que ocorreram deslizamentos. Tendo em vista a tabela 1 anteriormente apresentada, podemos fazer a seguinte análise:

Tabela 3 – Análise do resultado obtido com o programa utilizado

Sem deslizamentos	Com deslizamentos	Grupo A (programa)	Grupo B (programa)
2	1	4	1
3	5	6	2
4	7	9	3
6	8	16	5
9	12	18	7
10	14	22	8
11	15		10
13	16		11
17	20		12
18	21		13
19	22		14
	23		15
			17
			19
			20
			21
			23

Assim sendo o programa se mostrou relativamente eficiente, pois dos 6 membros agrupados como não associados com deslizamentos, 4 deles realmente não provocaram esse tipo de desastre (taxa de acerto de 67%). Porém, dos 19 casos separados como associados com deslizamentos, apenas 10 realmente estiveram associados de fato (taxa de acerto de 53%). Índices de acerto e de alarme falso serão calculados para avaliar a metodologia.

Futuramente as outras variáveis serão testadas da mesma forma, assim como o grupo de casos a serem estudados será completado com as datas faltantes. O método estudado pode se mostrar eficiente e é prático, pois requer o uso de um programa simples, não requerendo grande custo computacional e de tempo. Também será testada se o aumento na resolução do modelo Eta implicará em melhores resultados nesse tipo de análise. Já que inicialmente os dados utilizados são provenientes de rodadas de baixa resolução espacial (40km).

4. REFERÊNCIAS BIBLIOGRÁFICAS

SILVA, F. A. S. e. Identificação de regiões pluviometricamente homogêneas segundo um método de classificação automática. Campina Grande, Pb, 1989, 102 p. Dissertação (Mestrado em Engenharia Civil), Universidade Federal da Paraíba, 1989.

Quadro, M. F. L. Estudos de episódios de Zona de Convergência do Atlântico Sul (ZCAS) sobre a América do Sul. São José dos Campos. 94p. (INPE-6341-TDI/593). Dissertação (Mestrado em Meteorologia) – Instituto Nacional de Pesquisas Espaciais (INPE), 1994.

INPE/CPTEC. Climanálise (1998 a Nov 2007) - Boletim de Monitoramento e Análise Climática

ANEXO I

Trabalho submetido e aceito no Congresso Brasileiro de Meteorologia

ANÁLISE DE CASOS DE ZCAS UTILIZANDO TÉCNICAS DE AGRUPAMENTO

Jessica Motta Guimarães¹

Sin Chan Chou²

Claudine Dereczynski¹

¹Instituto de Geociências, Departamento de Meteorologia, Universidade Federal do Rio de Janeiro, RJ-BRA. jessica.meteoro@gmail.com

²Instituto Nacional de Pesquisas Espaciais – Cachoeira Paulista – SP, Brasil.

RESUMO: No presente trabalho a técnica de análise de agrupamento é utilizada a fim de classificar diferentes casos de Zonas de Convergência do Atlântico Sul (ZCAS) como associados ou não a eventos de deslizamentos. Dentre todos os casos observados de ZCAS ocorridos no período de 1998 a 2007 (53 casos), 18 estiveram associados com ocorrência de deslizamentos na Região Sudeste. A Análise foi baseada na técnica de agrupamento conhecida como K-means. Os dados de entrada do programa K-means foram as alturas geopotenciais em 500 hPa previstas pelo modelo Eta (40 km) com 48 h de antecedência. Essa análise compara diversas variáveis a partir de um valor referência inicial e assim estabelece uma separação dos casos considerados mais severos dos menos críticos. O método provou ser eficiente em separar casos severos associados a deslizamentos, daqueles sem deslizamento, mostrando menores valores de geopotencial no primeiro caso.

ABSTRACT: In this work, cluster analysis is applied to Eta Model forecasts to identify cases of South Atlantic Convergence Zone (SACZ) associated with landslides in the Serra do Mar, around Sao Paulo region. Among the 53 SACZ cases which occurred during the period 1998-2007, 18 cases were associated with landslides in the Southeast Region. The analysis was based on the k-means clustering procedure and was applied to distinguish more severe cases (associated with landslides) from the others. The 48-h forecasts of 500-hPa geopotential height was used as input data to the K-means program. The method proved to be efficient in grouping severe cases from less intense ones, showing lower values of 500-hPa geopotential height in the first group.

Palavras-chave: agrupamento, modelo Eta, ZCAS, deslizamentos, , cluster analysis, Eta Model, SACZ, landslides.

1. INTRODUÇÃO

As previsões numéricas de tempo produzidas atualmente ainda não permitem determinar com grande acurácia e confiabilidade a localização e intensidade exatas de um evento chuvoso intenso. Casos de precipitação extrema estão comumente associados com uma série de graves problemas, tais como ocorrência de deslizamentos e enchentes, além de provocarem um grande número de mortes, acidentes e prejuízos materiais.

Indicadores do potencial de ocorrência de um evento crítico são importantes para que um alerta seja emitido com suficiente antecedência à defesa civil e população em geral, para tomadas de providências que reduzam os possíveis prejuízos.

No presente trabalho será apresentada uma metodologia de agrupamento para destacar, com antecedência, eventos meteorológicos propensos a provocar desastres como deslizamentos e enchentes, dos demais eventos.

Os métodos de agrupamento são comumente utilizados na meteorologia para os mais diversos tipos de pesquisa. De acordo com Everitt (1993) várias técnicas de agrupamento hierárquico têm sido propostas, destacando-se, dentre as mais utilizadas: ligação simples (single linkage method); ligação completa (complete linkage method); método centróide (centroid method); método da mediana (median method); método da média dos grupos (group average method) e método da variância mínima (Ward's method). Esses métodos já foram utilizados com sucesso para estudo pluviométrico no Brasil (Silva, 1989), assim como para outros estudos que envolvem classificação de variáveis.

2. METODOLOGIA

Inicialmente foram selecionados os casos de Zonas de Convergência do Atlântico Sul (ZCAS) ocorridos entre os anos de 1998 e 2007 (53 casos), baseados na análise mensal da Revista Climanálise (INPE, 2007). Posteriormente foram identificados aqueles casos associados com a ocorrência de deslizamentos (18 casos), segundo os registros de acidentes do Instituto de Pesquisas Tecnológicas (IPT), referentes a deslizamentos ocorridos no Brasil, assim como notícias publicadas na mídia.

A partir dessas datas foram realizadas integrações com o modelo com resolução espacial de 40 km, com previsão de 48 horas para o primeiro dia de cada evento de ZCAS listado na tabela anterior. A resolução do modelo não foi aumentada, tendo em vista que o evento estudado é de escala sinótica por definição. Foi selecionada a caixa do modelo contendo a cidade de Caraguatatuba (lat: -23,6° e lon: -45,4°) onde foram observadas grandes taxas de precipitação. Desta caixa, foi extraída a variável altura geopotencial em 500 hPa. Os casos listados foram separados em dois grupos distintos:

Grupo A: casos que não estiveram associados com a ocorrência de deslizamentos.

Grupo B: casos que estiveram associados com a ocorrência de deslizamentos.

Existem diversas técnicas de análise de grupos, porém a que mais se mostrou conveniente para o trabalho aqui proposto foi a K-Means (referência do site). Essa metodologia é propícia para situações em que se desejam poucos grupos finais, e, portanto, grupos com vários membros. Inicialmente é selecionado aleatoriamente pelo usuário um caso de cada grupo (A e B) e o valor da variável é considerado como sendo o centróide do grupo, ou seja, um valor

representativo do mesmo. Neste trabalho, por exemplo, o primeiro centróide do grupo A foi o valor da altura geopotencial e 500 hPa para o caso 6 (5771 mgp) e o primeiro centróide do grupo B foi o valor de altura geopotencial em 500 hPa para o caso 19 (5866 mgp).

A seguir o programa compara os valores de geopotencial de todos os casos com os valores dos centróides (5771 mgp e 5866 mgp) e verifica se o valor dessa variável analisada está mais próximo do grupo A ou B. Após os dois grupos estarem formados, calcula-se novamente o valor do centróide da média das alturas geopotenciais nos grupos A e B. O procedimento de comparação dos valores das variáveis com os novos centróides é repetido até que não haja mais mudança dos membros nos grupos formados.

3. RESULTADOS

Inicialmente foram separados 23 casos dos 55 listados na tabela 1, para realização de testes. Dentre esses 23 casos, 11 deles não estavam associados com deslizamentos e 12 estavam associados com esse tipo de desastre. A variável testada foi altura geopotencial em 500 hPa. A seguir é apresentada uma lista com os 23 casos testados inicialmente.

N. Ref	Período da ZCAS	Condição inicial – 12Z	Deslizamentos
1	25-29/01/2004	23/01/2004	X
2	20-25/11/2004	18/11/2004	
3	09-14/12/2004	7/12/2004	
4	21-25/12/2004	19/12/2004	
5	17-21/01/2005	15/1/2005	X
6	13-22/02/2005	12/2/2005	
7	01-07/03/2005	28/2/2005	X
8	15-20/03/2005	13/3/2005	X
9	10-15/11/2005	8/11/2005	
10	17-21/11/2005	15/11/2005	
11	24-28/11/2005	22/11/2005	
12	11-16/12/2005	9/12/2005	X
13	24-29/12/2005	22/12/2005	
14	01-08/01/2006	30/12/2005	X
15	27/01/2006 - 02/02/2006	25/1/2006	X
16	09-13/02/2006	7/2/2006	X
17	07-16/03/2006	5/3/2006	
18	10-14/11/2006	8/11/2006	
19	26/11/2006 - 02/12/2006	24/11/2006	
20	07-16/12/2006	5/12/2006	X
21	27/12/2006 - 16/01/2007	25/12/2006	X
22	22-27/01/2007	20/1/2007	X
23	30/01/2007 - 09/02/2007	28/1/2007	X

Tabela 1 – Lista com os 23 casos estudados, contendo seus números de referência, o período no qual a ZCAS ocorreu, a data inicial do caso estudado e o registro da ocorrência de deslizamentos.

O programa agrupou os casos da seguinte forma:

Grupo A		Grupo B	
Referência	Geopotencial (mgp)	Referência	Geopotencial (mgp)
4	5791,320	1	5857,070
6	5771,360	2	5831,550
9	5789,830	3	5843,850

16	5782,230	5	5854,360
18	5815,790	7	5868,790
22	5814,210	8	5866,910
		10	5847,630
		11	5860,830
		12	5843,800
		13	5868,510
		14	5878,860
		15	5829,990
		17	5856,200
		19	5866,010
		20	5861,700
		21	5861,180
		23	5837,510

Tabela 2 – Resultado obtido com o programa K Means utilizado para fazer o agrupamento. As duas primeiras colunas correspondem ao grupo A e as duas últimas ao grupo B, ambos formados pelo programa.

Para cada grupo apresentado, a primeira coluna representa os números de referência de cada membro e que a segunda mostra os valores de geopotencial comparados. O grupo A é o referente aos casos em que não ocorreram deslizamentos, já o grupo B está relacionado aos casos em que ocorreram deslizamentos. De forma geral conclui-se que os casos com menores valores de geopotencial foram separados no grupo A, como não associados com deslizamentos e os com valores mais elevados no grupo B, como casos em que ocorreram deslizamentos. Tendo em vista a tabela 1 anteriormente apresentada e os resultados do programa K Means, foi possível construir a tabela 3.

Observação Sem deslizamentos	Observação Com deslizamentos	Grupo A (programa)	Grupo B (programa)
2	1	4	1
3	5	6	2
4	7	9	3
6	8	16	5
9	12	18	7
10	14	22	8
11	15		10
13	16		11
17	20		12
18	21		13
19	22		14
	23		15
			17
			19
			20
			21
			23

Tabela 3 – Análise do resultado obtido com o programa de agrupamento comparado com os dados existentes sobre os casos estudados.

Assim sendo o programa se mostrou relativamente eficiente, pois dos 6 membros agrupados como não associados com deslizamentos, 4 deles realmente não provocaram esse tipo de desastre (taxa de acerto de 67%). Porém, dos 19 casos separados como associados com deslizamentos, apenas 10 realmente estiveram associados de fato (taxa de acerto de 53%). Futuramente as outras variáveis serão testadas da mesma forma, assim como o grupo de casos a serem estudados será completado com as datas faltantes. Dentre os casos associados com a ocorrência de deslizamentos, o que apresentou valor mais alto de altura geopotencial foi o caso 14, com 5879 mgp. Esse valor pode ser utilizado como referência para previsões de casos de deslizamentos.

4. CONCLUSÕES

O presente trabalho visa aplicar técnicas de agrupamento para analisar casos de ZCAS ocorridos entre os anos de 1998 e 2007 na região Sudeste. O objetivo do trabalho é separar casos que provocaram deslizamentos dos que não provocaram, através do método de análise de agrupamento conhecido como K-means.

O método estudado pode se mostrar eficiente e é prático, pois requer o uso de um programa simples, não requerendo grande custo computacional e de tempo. Também será testada se o aumento na resolução do modelo Eta implicará em melhores resultados nesse tipo de análise. Já que inicialmente os dados utilizados são provenientes de rodadas com resolução espacial relativamente baixa (40km).

5. REFERÊNCIAS BIBLIOGRÁFICAS

EVERITT, B.S. Cluster analysis. 3 ed. London: Heinemann Educational Books, 1993. 122p.

SILVA, F. A. S. e. Identificação de regiões pluviometricamente homogêneas segundo um método de classificação automática. Campina Grande, Pb, 1989, 102 p. Dissertação (Mestrado em Engenharia Civil), Universidade Federal da Paraíba, 1989.

Climanálise Boletim de Monitoramento e Análise Climática INPE, São José dos Campos. Brasil. (1998 a 2007) -

6. AGRADECIMENTOS

O primeiro autor é bolsista do programa CNPq/PIBIC/INPE no 101532/2006-7. Este trabalho é parcialmente financiado pelo projeto FAPESP No. 04/09649-0, e parcialmente apoiado pelo CNPQ No.308725/2007-7.

ANEXO II

Programa desenvolvido em FORTRAN para separar os casos em dois grupos distintos, casos sem e com deslizamentos (grupos A e B, respectivamente), através do método de análise K-Means.

```
PROGRAM K_MEANS
  parameter (n=55,ica=20,icb=30)
  parameter (e=10)
  REAL Y(n),da(n),db(n),IGA(n),IGB(n)
  INTEGER :: k,la,lb,Amed,Bmed

!1-Ler os arquivos e determinar quantos clusters finais serao formados
!   write(*,*) 'Quantos clusters finais sao desejados?'
!   read(*,*) k
  k=2
!ler o conjunto de dados (pontos)
  do i=1,n
    read(10,*) (Y(i))
  enddo

!2-Definir quais serao os centroides iniciais para o agrupamento
  CA= Y(ica) !recebe a media do grupo de zcas sem deslizamento
  CB= Y(icb) !recebe a media do grupo de zcas com deslizamento

!3-Loop que calcula as distancias relativas entre os membros e os
centroides
  iter=0
  100 iter=iter+1
  print*, ' ITER:', iter
  la=1
  lb=1
  do i=1,n
    da(i)=abs(Y(i)-CA)
    db(i)=abs(Y(i)-CB)
    if (da(i) .LT. db(i))then
      IGA(la)=i
      la=la+1
    else
      IGB(lb)=i
      lb=lb+1
    endif
  print*,i,da(i),db(i),la,lb,IGA(la),igb(lb)
  enddo

  write(*,*) (IGA(i),i=1,la-1)
  write(*,*) (IGB(i),i=1,lb-1)

!4-Calcular as novas medias (centroides) considerando os grupos
formados
  Amed=0. !media de todos os membros do grupo formado
  Bmed=0. !media de todos os membros do grupo formado
  do i=1,la
    amed=amed+Y(IGA(i))
  enddo
```

```

    amed=amed/float(la)
    do i=1,lb
    bmed=bmed+Y(IGB(i))
    enddo
    bmed=bmed/float(lb)

!5-Compara o novo centroide com o anterior
    if (((Amed - CA).lt.E).and.((Bmed - CB).LT.E)) then !valor "e"
pequeno a ser discutido
        write (*,*) 'os membros do grupo A sao'
        do i=1,la-1
        write(*,*) IGA(i),Y(IGA(i))
        enddo
        write(*,*) 'os membros do grupo B sao'
        do i=1,lb-1
        write(*,*) IGB(i),Y(IGB(i))
        enddo
    else
        CA=Amed
        CB=Bmed
        goto 100
    endif

stop
end

```