

Classificação Automática de Sistemas Precipitantes com
Radar Meteorológico

Relatório Final de Atividades de Pesquisa de Iniciação
Científica

Aluno I.C.: Leandro Meneguelli Biondo

N^oUSP: 5090095

leandromet@gmail.com

Orientador: Prof. Dr. Augusto José Pereira Filho

DCA - IAG - USP

Período: Setembro 2009 - Agosto 2010

Projeto realizado com auxílio de bolsa de Iniciação Científica PIBIC/CNPq

18 de Agosto de 2010

Sumário

1	Introdução	4
1.1	O Radar Meteorológico	5
1.2	Mapas de precipitação	7
1.3	O Radar meteorológico de São Paulo	8
2	Metodologia	9
2.1	Classificação de sistemas precipitantes com radar meteorológico	9
2.1.1	Convecção Isolada	9
2.1.2	Brisa Marítma	10
2.1.3	Linhas de Instabilidade	10
2.1.4	Bandas Dispersas de Precipitação	11
2.1.5	Frentes Frias	11
2.2	Características de textura de imagens digitais	12
2.3	Programação JAVA	13
2.3.1	Eclipse	14
2.3.2	ImageJ	14
3	Resultados	14
3.1	Características Visuais das Imagens do RMSP	14
3.1.1	Convecção Isolada	15

3.1.2	Brisa Marítma	15
3.1.3	Linhas de Instabilidade	15
3.1.4	Banda Dispersa	16
3.1.5	Frentes Frias	16
3.2	Análise Estatística de Imagens	16
4	Conclusão	22

Resumo

Este relatório descreve a tentativa de classificar automaticamente sistemas precipitantes que ocorrem na região coberta pelo Radar Meteorológico de São Paulo. Para esta classificação utilizou-se um método de análise estatística de textura nas imagens de radar em que haviam eventos precipitantes significativos.

Eventos ocorridos entre 2003 e 2007 foram classificados visualmente por intensidade de precipitação (Convectivos e Extratiformes) e por 5 tipos principais de sistemas (Convecção Isolada, Brisa Marítima, Linhas de Instabilidade, Bandas Dispersas e Frentes Frias). Os eventos classificados foram então submetidos a análise de textura com o objetivo de encontrar variáveis estatísticas com valores que identificassem os sistemas.

Os resultados obtidos indicam algumas grandezas que diferenciam os sistemas em situações extremas de intensidade, mas apenas de forma a eliminar um ou dois sistemas dos possíveis para classificação, restando ao menos 3 possibilidades de classificação. Foi verificado, ainda, que a ocorrência das situações em que é possível a eliminação de algum dos sistemas prováveis é baixa em relação ao total de eventos analisados.

1 Introdução

Foi proposto a análise estatística de imagens provenientes do Radar Meteorológico de São Paulo (RMSP), na tentativa de identificar a relação entre os sistemas meteorológicos presentes na imagem e a estatística da precipitação estimada pelo RMSP.

Foram utilizados dados do radar armazenados entre os anos de 2003 e 2007, com os quais foram realizada a classificação visual subjetiva dos sistemas presentes nas imagens, com o objetivo de se familiarizar com os resultados apresentados pelo radar. Entre as informações importante estão o tempo de duração de cada sistema, a intensidade e área coberta por precipitação, e o deslocamento médio aparente do sistema na imagem do RMSP.

Para os dados utilizados, foi considerada a mesma classificação apresentada por José Felipe Farias em 2009 para previsão de curtíssimo prazo, além de utilizar sua classificação dos sistemas como base para a identificação dos arquivos processados.

1.1 O Radar Meteorológico

Radar é um sistema eletromagnético para detecção e localização de objetos. O princípio de funcionamento do radar meteorológico consiste em emitir através de uma antena um pulso de energia eletromagnética de duração da ordem de μs e de comprimento de onda λ (cm). Geram pulsos de onda eletromagnética altamente energética para alcance a longas distâncias. Estes sinais viajam a velocidade da luz.

Quando esta energia é interceptada por um alvo pequeno (por exemplo, uma gota de chuva) se dispersa em todas as direções (**Fig. 1**), de forma que uma fração retorna na direção do pulso emitido pelo radar e captado por um refletor, normalmente localizado na mesma antena. A distancia deste obstáculo é determinada do tempo transcorrido entre a emissão e recepção do pulso.

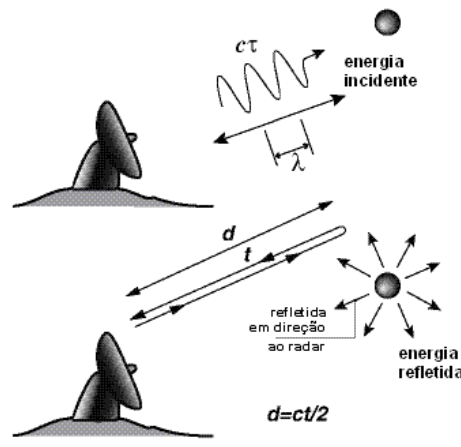


Figura 1: Esquema de transmissão e recepção de energia por uma gota.

O pulso é transmitido à velocidade da luz, c . Decorrido um tempo t entre a emissão e recepção do sinal, a distancia até a gota é expressa em $d = ct/2$. O radar mede a energia retro-espalhada em sua direção pelas gotas de água situadas no interior de certo volume V , determinada pela metade da duração do pulso (**Fig. 2**).

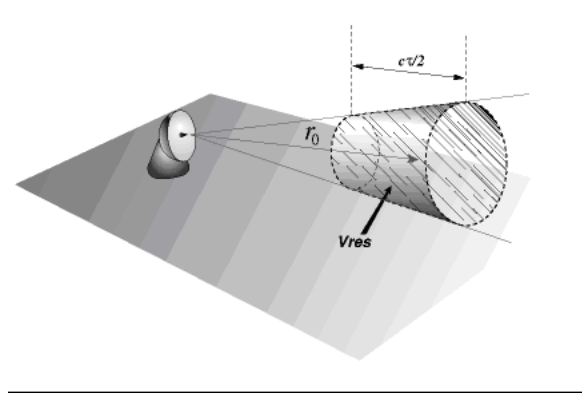


Figura 2: Esquema do volume iluminado pelo feixe do radar com volume V .

A energia medida pode ser expressa como:

$$\bar{P}(r_0) = \frac{C}{L^2(r_0)r_0^2}Z(r_0) \quad (1)$$

Onde, P é a potência média retro-espalhada pelas gotas situadas no interior do volume V a uma distancia r_0 do radar. A potência média resulta de n pulsos, tal que o sinal recebido seja estável; C é denominada constante do radar que agrupa uma série de características do radar, tais como o comprimento de onda, forma e largura do feixe, comprimento do pulso, potência transmitida, ganho da antena, entre outros; Z é a refletividade das gotas no volume V iluminado pelo radar.

A refletividade Z é definida como a integral dos diâmetros das gotas no interior de um volume elevadas a sexta potencia:

$$Z = \int_0^\infty N(D)D^6 dD \quad (2)$$

onde, $N(D)$ é o número de gotas de diâmetro D .

A intensidade de chuva (R) pode ser expressa como o fluxo de água através de uma superfície, e pode ser obtida como:

$$R = \frac{\pi}{6} \int_0^\infty V(D)N(D)D^3 dD \quad (3)$$

Das equações (2) e (3) pode ser obtida uma relação de transformação da taxa de precipitação (em $mm h^{-1}$) em refletividade ($mm^6 m^{-3}$):

$$Z = aR^b \quad (4)$$

Em geral, os parâmetros a e b variam significativamente. Há um grande número de relações ZR experimentais. Estas relações variam de acordo com o tipo de chuva e localização geográfica.

1.2 Mapas de precipitação

Mapas horizontais de precipitação são obtidos da refletividade do radar ou da taxa de chuva estimada a partir desta em um determinado elemento de área ou pixel. O radar meteorológico executa um movimento de rotação e elevação de modo a varrer a troposfera em azimute e elevação, respectivamente (**Fig. 3**).

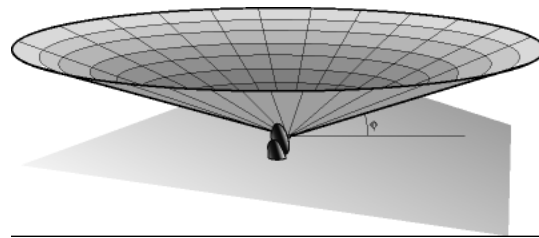


Figura 3: Esquema de varredura do radar meteorológico para um dado ângulo de elevação (ϕ).

Estas medições de refletividade em coordenadas esféricas é processada para gerar mapas horizontais de chuva: a) PPI, Plan Position Indicator, correspondente a refletividade numa da elevação projetada num plano horizontal (**Fig. 4a**) e; b) CAPPI, Constant Altitude Plan Position Indicator, a refletividade num plano a uma altitude constante é obtida por meio de segmentos de diversas elevações mais próximas da altitude de interpolação (**Fig. 4b**).

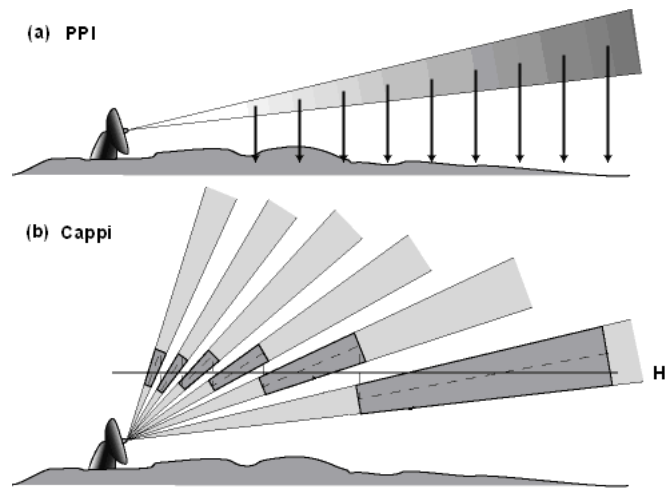


Figura 4: Ilustração de um PPI (a) e CAPPI (b) do radar meteorológico.

1.3 O Radar meteorológico de São Paulo

O Radar Meteorológico de São Paulo (RSP) opera com comprimento de onda de 10 cm, o que o classifica como um radar banda S e fornece entre outros produtos mapas de taxa de precipitação a cada 5 minutos, com resolução espacial de $2 \text{ km} \times 2 \text{ km}$. O radar meteorológico de São Paulo cobre uma área circular de 240 km e abrange boa parte do litoral paulista, sul de Minas Gerais, Vale do Paraíba e sul do Rio de Janeiro (Fig 5). O RSP permite monitorar a evolução espaço-temporal da chuva sobre as bacias hidrográficas da região leste do Estado de São Paulo. Trata-se de uma importante ferramenta destinada a previsão de chuva a curtíssimo prazo para a previsão hidrológica até 3 horas a frente. [Silva Dias, 2005]

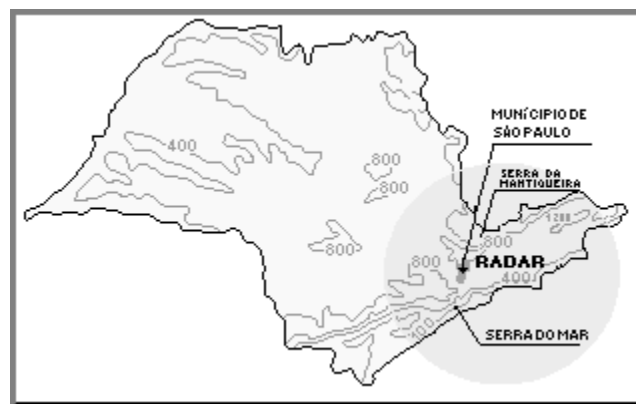


Figura 5: Área de cobertura do radar meteorológico de São Paulo

2 Metodologia

2.1 Classificação de sistemas precipitantes com radar meteorológico

Foram utilizados neste estudo campos de precipitação estimados na altitude de 3 km, com resolução espacial de 2-km x 2-km e resolução temporal de 5 minutos dos anos de 2004 a 2007. Estes campos foram usados para classificar a precipitação em convectiva e estratiforme, como também identificar cinco tipos de sistemas precipitantes que atuam durante todo o ano no leste do Estado de São Paulo, dentro da área de cobertura do RSP (Pereira Filho e Silva Dias, 2005). Os dados binários do RSP foram tratados e convertidos em campos de precipitação por meio de programas em linguagem fortran, shell scripts e gerados campos de precipitação com o software GrADS.

Os eventos foram classificados de acordo com as intensidades de precipitação; em Convectivos (C) e Estratiformes (E). Posteriormente, os sistemas serão classificados quanto à sua morfologia, em cinco tipos; Convecção Isolada (CI), precipitação associada a Brisa Marítima (BM), Linhas de Instabilidade (LI), Bandas Dispersas de precipitação (BD) e Frentes Frias (FF). Todos os mapas de chuva foram utilizados para a classificação dos sistemas, conforme acima. A primeira etapa consistiu da classificação de sistemas em convectivos e estratiformes. Na segunda etapa foram analisados o grau aparente de organização dos ecos de chuva, sua distribuição horizontal, deslocamento e duração, classificados em cinco tipos básicos de sistemas meteorológicos: precipitações associadas à convecção isolada (CI) e a brisa marítima (BM), linhas de instabilidade (LI), bandas dispersas de precipitação (BD) e frentes frias (FF).

2.1.1 Convecção Isolada

A CI geralmente se forma à tarde sem orientação aparente e pode estar, ou não, associada a algum sistema de grande escala; surge e se dissipa, geralmente, próximo ao local de formação. São células convectivas ordinárias sem deslocamento aparente ou de deslocamento muito lento, e sem região preferencial para formação, sobre o continente. [J. Farias, 2009]

Eventos de CI ocorrem preferencialmente entre 13:30 HL e 18:30 HL, de acordo com o ciclo do aquecimento diurno. Estas células ocorrem em sua grande maioria no interior do continente, seu deslocamento, quando observado, é de oeste para leste na maior parte das vezes. Podem dar origem, na fase de decaimento, a áreas de precipitação estratiformes mais extensas do que as células na fase madura de desenvolvimento. [J. Farias, 2009]

2.1.2 Brisa Marítma

A BM pode dar origem a sistemas precipitantes convectivos e estratiformes. Os eventos convectivos induzidos pela BM se deslocam para oeste na forma de CI, com características morfológicas semelhantes a CI. Em geral, desenvolvem-se entre 14:30 HL e 19:00 HL e tendem a formar linhas de células convectivas ao longo da costa, as quais são favorecidas pela Serra do Mar. Este tipo de sistema alcança a RMSP, onde normalmente produz precipitação convectiva intensa. [J. Farias, 2009]

Em alguns casos, a BM dá origem a pequenas bandas ou áreas de precipitação estratiforme, com pequeno tempo de vida. Na Figura 4, é possível observar o deslocamento de núcleos convectivos alinhados em direção ao interior do continente, associada à frente de BM. [J. Farias, 2009]

2.1.3 Linhas de Instabilidade

As LI apresentam em sua dianteira células convectivas intensas com área de precipitação estratiforme na sua retaguarda, deslocando-se rapidamente, em torno de 10 a 12 $m s^{-1}$. Surgem geralmente da fusão de ecos convectivos isolados ou podem estar associados a outros sistemas, em geral à dianteira ou retaguarda de FF. As LI foram observadas em todos os horários, com predominância entre 15:00 HL e 17 HL, com ciclo de vida entre 4 a 12 horas. Foram observados deslocamentos de LI para todas as direções, mas preferencialmente de noroeste para sudeste, perpendiculares ou paralelas ao movimento de sistemas frontais. [J. Farias, 2009]

2.1.4 Bandas Dispersas de Precipitação

Deslocamento de Nordeste para Sudeste, sem organização espacial, com extensa área de chuva estratiforme com presença de áreas menores com chuva convectiva. As BD não apresentam organização espacial e, geralmente, estão associadas a um sistema de grande escala, tais como frentes estacionárias. Em geral, elas se deslocam de noroeste para sudeste e com chuva estratiforme e convecção embebida nestas. O período de vida destes sistemas pode chegar a até 5 dias, de acordo com os dados de radar. As BD convectivas de maior longevidade estão, em geral, associadas aos episódios de ZCAS.[J. Farias, 2009]

Episódios de ZCAS têm variabilidade associada a entrada de frentes frias (em geral), a propagações para leste da Oscilação de Madden-Julia (no período de 30-60 dias) e a propagações de ondas de rossby (2-30 dias)[Cavalcanti,2009]. Estes eventos que influenciam na variabilidade da ZCAS estão, por sua vez, associados uns aos outros.

2.1.5 Frentes Frias

Uma frente fria é formada pelo deslocamento de uma massa de ar mais frio que se move em uma região inicialmente mais quente. No Brasil as frentes frias são formadas principalmente pelo deslocamento de massas polares secas que se movem de sudoeste para nordeste sobre a região continental (Sul e Sudeste, podendo chegar até o Nordeste) e a parte adjacente do Oceano Atlântico Sul.[Cavalcanti,2009]

A precipitação gerada pelas frentes frias na região do RMSP utiliza umidade existente na atmosfera quente existente na localidade antes do deslocamento do sistema.As frentes frias atuam na região do RMSP durante todo o ano, sendo mais convectivas durante o verão e estratiformes no inverno, quando a velocidade de passagem da frente é maior. O tempo de deslocamento médio destas frentes na região é de 18-20 horas[Santos, 2006]. Um critério utilizado para a identificação objetiva de frentes frias no Brasil é identificar uma redução de temperatura em 925hPa, de pelo menos $2^{\circ}C$, e/ou a existência de ventos de sul maiores ou iguais a $2\frac{m}{s}$. [Cavalcanti,2009]

2.2 Características de textura de imagens digitais

Há diferentes métodos e definições de textura de imagens. Pode-se definir textura de imagem como sendo a organização espacial das células da imagem relacionadas às suas tonalidades. A organização espacial pode ser aleatória, ter uma dependência entre pares de células numa dada vizinhança ou até de n-células. Essa dependência pode ser estrutural, probabilística ou funcional. A metodologia aqui utilizada deriva dos métodos de análise estatística de textura utilizados por Oliveira e Pereira (2007).

Na metodologia estatística, a textura é descrita por um conjunto de medidas estatísticas (e.g., contraste, homogeneidade) da distribuição espacial das tonalidades das células. Dado um ângulo Θ e uma distância d , essas medidas são calculadas sobre a distribuição de probabilidades $P_{d,\Theta}(m)$, sendo m a diferença das tonalidades entre duas células a uma distância d na direção Θ . A estimativa de $P_{d,\Theta}(m)$ pode ser feita por meio da frequência de ocorrência das tonalidades. Abaixo apresenta-se algumas medidas utilizadas [Oliveira e Pereira, 2007]:

- Média: $\mu_{d,\Theta} = \sum_m m P_{d,\Theta}(m)$
- Desvio Padrão: $\sigma_{d,\Theta} = \left[\sum_m (m - \mu_{d,\Theta})^2 P_{d,\Theta}(m) \right]^{1/2}$
- Contraste: $C_{d,\Theta} = \sum_m m^2 P_{d,\Theta}(m)$
- Segundo Momento Angular: $M_{d,\Theta} = \sum_m [P_{d,\Theta}(m)]^2$
- Entropia de Shannon: $S_{d,\Theta} = - \sum_m P_{d,\Theta}(m) \log [P_{d,\Theta}(m)]$
- Homogeneidade Local: $H_{d,\Theta} = \sum_m \frac{P_{d,\Theta}(m)}{[1+m^2]}$
- Sombra de Aglomerados: $SA_{d,\Theta} = \frac{[\sum_m (m - \mu_{d,\Theta})^3 P_{d,\Theta}(m)]}{\sigma_{d,\Theta}^3}$
- Proeminência de Aglomerados: $PA_{d,\Theta} = \frac{[\sum_m (m - \mu_{d,\Theta})^4 P_{d,\Theta}(m)]}{\sigma_{d,\Theta}^4 - 3}$

A metodologia espectral é baseada nas propriedades do espectro de Fourier (e/ou ondaletas) e geralmente é utilizada para detectar periodicidades globais na imagem. Como informações retiradas do espectro de Fourier destacam-se a direção principal e os períodos espaciais dos padrões. A direção principal dos padrões de textura pode ser estimada por meio dos picos no espectro, enquanto que suas localizações (dos picos) no plano de frequências fornecem os períodos espaciais dos padrões. Além

disso, eliminando-se as componentes periódicas da imagem, por meio de filtros, sobram apenas os elementos não-periódicos que podem ser descritos por técnicas estatísticas.

Em paralelo ao estudo formal da análise de textura, aplicar-se-á os algoritmos desenvolvidos para a criação de padrões artificiais que serão utilizados como exemplos. Por fim, os algoritmos estudados serão aplicados aos dados de precipitação obtidos pelo radar meteorológico de São Paulo (Fig. 5) entre os anos de 1998 a 2007. Os dados consistem de padrões compostos por 240 x 240 pixels numa área de 240 km de raio (Fig. 5) que serão inicialmente classificados visualmente, os eventos entre 2004 e 2007. Os eventos de 1998 a 2003 foram classificados por Silva (2006) e os de 2003 a 2007 por Farias (2009): bandas de chuva; convecção isolada; frentes de brisas marítimas; frentes frias e linhas de instabilidade. A metodologia desenvolvida por Oliveira e Pereira Filho (2007) será aplicada para se obter os padrões de precipitação e comparadas com as obtidas visualmente.

2.3 Programação JAVA

Java é uma linguagem de código aberto, fortemente baseada no paradigma de orientação a objetos. As linguagens de programação com este paradigma tendem a oferecer uma alta reutilização de software, desenvolvimento sempre evolutivo, construções complexas a partir de construções simples e a manutenção e atualização dos softwares tende a ser simplificada.

Um programa Java é executado em uma máquina virtual (JVM) que funciona como um intermediário de código entre o código e o equipamento utilizado, dessa forma é possível executar um programa, escrito na linguagem, em qualquer equipamento que tenha a esta JVM instalada (independente de tecnologia de hardware e versão de sistema operacional). A linguagem oferece, também, diversos recursos e facilidades de interação com outras tecnologias e linguagens, possibilitando a execução de parte dos métodos em outra linguagem caso esta ofereça vantagens de desempenho ou facilidade de desenvolvimento (caso seja necessário utilizar scripts previamente existentes em C ou FORTRAN, por exemplo).

Para o armazenamento e organização dos dados foi escolhido um banco de dados SQL, com o uso de um servidor MySQL, que é acessível ao programa Java com o uso do Connector/J, que oferece conectividade entre ambos. O banco de dados SQL foi utilizado para armazenar os resultados e o local

em que se encontravam os arquivos, sendo que a vantagem do uso de um banco de dados deste tipo é poder fazer classificações, consultas, alterações e inclusão de dados de modo rápido e mantendo a integridade dos dados.

Foi utilizada a interface de desenvolvimento Eclipse, também de código aberto, que auxilia no desenvolvimento do código ao permitir inclusão de classes e testes de execução de forma prática e rápida. Parte dos dados e resultados foram verificados com o software ImageJ, desenvolvido em Java e utilizado em diversas áreas (medicina, astronomia, etc.) para a análise e processamento de imagens.

2.3.1 Eclipse

O Eclipse é um ambiente de programação multi-plataforma, multi-linguagem e livre. Foi utilizado por oferecer revisão de código (debug), listar variáveis e locais no código em que elas são utilizadas, impedir que sejam realizados erros comuns de programação que podem levar a uma queda de produtividade de desenvolvimento (problemas com parênteses, variáveis de tipos diferentes e etc.) e, além disso, ser largamente utilizado por programadores e instrutores Java.

2.3.2 ImageJ

Um poderoso programa de processamento de imagens, desenvolvido em Java e permite que sejam feitos programas e scripts também em Java, de modo a executar métodos matemáticos diretamente na imagem e tornar possível acompanhar visualmente os resultados e testá-los. O ImageJ oferece embutidas classes que realizam análises estatísticas e ajustes de propriedades de imagem, além do suporte a diversos formatos (JPEG, GIF, PNG, etc.) e modos de imagem (8bits, RGB 8 e 16bits, etc.). Neste estudo o ImageJ foi utilizado apenas para verificar resultados encontrados, mas é possível integrá-lo a um programa Java e aproveitar seus recursos.

3 Resultados

3.1 Características Visuais das Imagens do RMSP

Foram utilizadas imagens selecionadas por José Felipe da Silva Farias (2009), que foram classificadas conforme os sistemas presentes na área do RMSP, esta classificação foi feita visualmente de acordo com o tamanho, localização, intensidade e movimentação aparente apresentados em sequência cronológica.

A resolução das imagens é de 240x240 pixels, sendo que a intensidade de precipitação registrada pode ser de 0, 1, 5, 8, 10, 15, 20, 25, 30, 40, 50, 75 ou 100 milímetros de chuva. Esta intensidade pode ser visualizada em tons de cinza (8 bits, 256 tons) ou realçada da forma que for conveniente.

A seguir serão descritas características visuais observadas nos eventos aqui analisados.

3.1.1 Convecção Isolada

Ocorrem células de tamanho reduzido (2 a 20 pixels de raio), múltiplas, independentes em sua posição mas relacionadas por seu desenvolvimento com o tempo. Estes sistemas são detectados pelo RMSP apenas sobre o continente, na maior parte das vezes nos quadrantes NE e NW da imagem. Apresentam direção de deslocamento predominante de Oeste para Leste (W-E) e velocidade média de $18 \frac{km}{h}$ [J. Farias, 2009].

3.1.2 Brisa Marítma

Os eventos de brisa marítma são organizados com relação a sua localização e seu desenvolvimento com o tempo, apresentam uma faixa de precipitação com intensidade variável. Área grande, sendo que sua dimensão na diagonal principal é consideravelmente maior (até 10 vezes) do que sua largura, acompanhando as proporções da serra que induz a precipitação.

Ocorre obrigatoriamente sobre e após a Serra do Mar, que na imagens analisadas fica sobre a diagonal principal. Sendo assim, a precipitação provocada por Brisa Marítma ocorre principalmente nos quadrantes NW e NE da imagem. Sua direção de deslocamento predominante é de Sudeste para Noroeste (SE-NW), com velocidade média de $40 \frac{km}{h}$ [J. Farias, 2009].

3.1.3 Linhas de Instabilidade

Quando presente, a precipitação devido a linhas de instabilidade apresenta pouca variação de posição com o tempo, tendo intensidade variável, a área ocupada pelo sistema é grande, da ordem de 80 pixels

de raio. Bem organizado quanto ao deslocamento e forma. As Linhas de Instabilidade podem ocorrer em qualquer região da imagem, sendo importante o seu deslocamento, que tem direção predominante de Noroeste para Sudeste (NW-SE) e velocidade média de $37 \frac{km}{h}$ [J. Farias, 2009].

3.1.4 Banda Dispersa

Apresenta uma forma desorganizada, mas seu deslocamento tem uma direção comum, em alguns casos apresentam pequenos núcleos de máximos de precipitação. Direção de deslocamento predominante de Norte para Sudeste (N-SE) e velocidade média de $45 \frac{km}{h}$. [J. Farias, 2009]

3.1.5 Frentes Frias

Bem organizadas tanto quanto a forma (por vezes parecida com a das Brisas, mas na diagonal secundária, e um pouco mais larga) e o deslocamento. A passagem do sistema é rápida, podendo produzir precipitação em qualquer região da imagem, apresenta deslocamento predominante de Sudoeste para Nordeste (SW-NE). [J. Farias, 2009]

As imagens de frentes frias apresentam variações de intensidade maiores na direção da diagonal principal, que tende a ser perpendicular ao seu deslocamento. A precipitação pode ocupar áreas grandes (~ 100 pixels de raio) quando o sistema é intenso, mas pode também apresentar diversas regiões menores de precipitação (~ 15 pixels de raio) quando o sistema é menos intenso. Em ambos os casos é mantida uma coerência de deslocamento e de área total sob influência do sistema.

3.2 Análise Estatística de Imagens

Fez-se necessário, antes de iniciar a análise, uma homogeneização dos arquivos de imagem a serem utilizados, pois parte deles se encontrava armazenada em formato binário proveniente do RMSP (compactado para facilidade de transmissão) e parte em formato binário “.gra”, gerados para visualização no programa GRADS, como os arquivos do RMSP poderiam ser descompactados para o formato de leitura do GRADS (com perda de informações de cabeçalho) mas o inverso não poderia ser feito, foi escolhido este formato de arquivos para o desenvolvimento de programas.

Desenvolvido um programa na linguagem JAVA que realiza a leitura de arquivos “.gra”, posteriormente calcula as variáveis estatísticas de primeira ordem da textura da imagem (estatística de toda a imagem). Foi utilizado o programa ImageJ para verificação destes cálculos estatísticos em uma primeira fase, quando foram identificados problemas com relação ao resultado esperado e ao calculado. Posteriormente foi identificado que haviam erros de cálculo devido à concepção do histograma da imagem realizado, que não era compatível com os cálculos estatísticos.

Após reestruturada a estatística do programa passaram a ser obtidos resultados compatíveis com as verificações no ImageJ, passando então para o estabelecimento da melhor forma de armazenar e organizar os resultados. Nesta etapa foi feita uma tentativa de criar arquivos de texto independentes para cada imagem, mas este processo se mostrou dispendioso quando havia necessidade de buscar determinados tipos de sistema ou determinadas características estatísticas das imagens.

Assim, numa tentativa de acelerar a consulta aos resultados, foi desenvolvido um banco de dados MySQL (com o auxílio do driver JDBC Connector/J) em que são armazenados um ponteiro do arquivo de imagem, o tipo de sistema classificado visualmente e os resultados estatísticos encontrados.

A comunicação entre o banco de dados e o programa desenvolvido é simples, mas houveram problemas de implantação devido a precisão de ponto flutuante utilizada por cada ambiente nos tipos de variáveis disponíveis, o que inicialmente gerou erros na exportação dos resultados para o banco. A maior parte dos problemas puderam ser contornados, sendo que em um teste realizado com ~20.000 arquivos não classificados foi registrado um índice de arquivos não processados ou armazenados menor do que 5% (800 arquivos). A velocidade de processamento com o cálculo e armazenamento dos arquivos é de cerca de 1.200/min.

Este banco de dados possibilita agregar, no futuro, outras análises de textura (de outra ordem) ou novas análises de outros dados, sem necessidade de alteração dos dados já armazenados. Permite, também, desenvolver um software de análise dos dados armazenados para realizar novos testes estatísticos de acordo com as propriedades já obtidas. Para a análise inicial dos resultados, os dados do banco foram exportados para o formato de uma planilha com a análise estatística de uma primeira parcela de dados, totalizando 5259 imagens, sendo elas:

- 1.756 imagens de Bandas Dispersas (BD);

- 489 imagens de Brisa Marítima (BM);
- 784 imagens de Convecção Isolada (CI);
- 1.276 imagens de Frentes Frias (FF);
- 954 imagens de Linhas de Instabilidade (LI).

Destas imagens foram calculadas as propriedades estatísticas do item 2.2, sendo que os cálculos foram realizados apenas com os pixels em que havia indicação de chuva (pixel não nulo). Foi introduzida, também, uma variável “Média Total” que é a média de precipitação medida incluindo todos os pixels da imagem. Desta forma, as variáveis examinadas foram Tipo de Sistema (**tipo_sistema**, que pode ser 1 (CI), 2(BM), 3(LI), 4(BD) ou 5(FF)), Média (**Vm**), Desvio Padrão (**Dv**), Média Total (**totalmean**), Contraste (**Co**), Segundo Momento Angular (**Sma**), Entropia de Shannon (**Es**), Homogeneidade Local (**HI**), Sombra de Aglomerados (**Sa**) e Proeminência de Aglomerados (**Pa**).

Os resultados obtidos indicam que os sistemas têm diferenças sutis em algumas variáveis, e diferenças maiores em outras. Apesar de ser possível identificar diferenças em relação aos conjuntos de variáveis quando comparadas entre sistemas, essas diferenças tendem a ocorrer apenas em uma parcela reduzida dos dados analisados, sendo que a maior parte dos dados se concentra em bacias grandes que são compartilhadas pelos sistemas.

Aparentemente, apesar de não ser possível criar uma função cronológica, por exemplo, que identifique o tipo de sistema presente na imagem, ou ainda identificar qualquer sistema presente em qualquer imagem do RMSP apenas com estas variáveis, é possível eliminar alguns sistemas baseado em algumas informações estatísticas.

Foi feita uma análise no programa Statistica para identificar a possibilidade das características de textura aqui utilizadas poderem descrever o sistema meteorológico presente na imagem. Em uma primeira passagem foram selecionadas as variáveis que tinham distância euclidiana menor nas relações identificadas na Análise de Agrupamento (**Fig.6**).

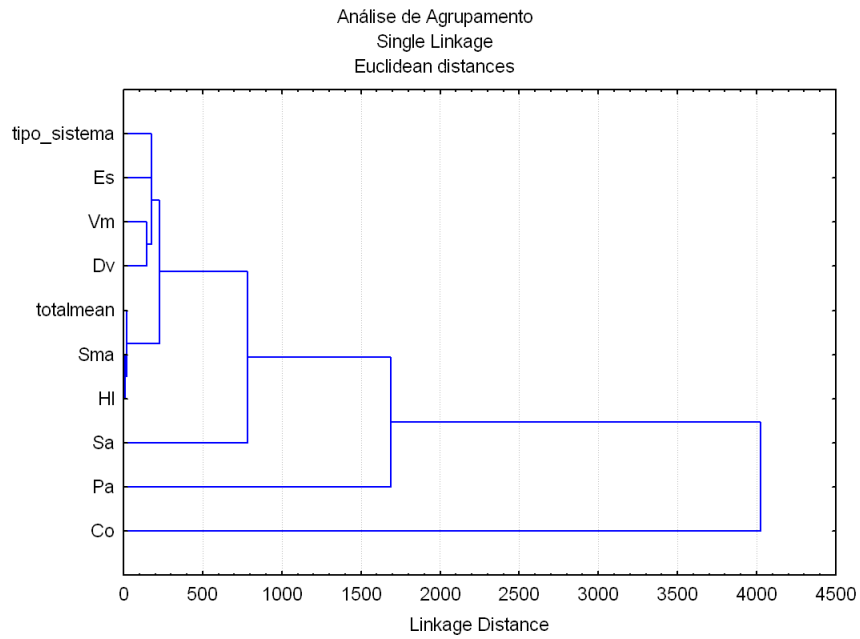


Figura 6: Agrupamento (clustering) das variáveis independentes encontradas, indicam o grau de relação existente entre elas

Este resultado indica haver uma dependência maior entre as variáveis *tiposistema*, *Es*, *Vm* e *Dv*, enquanto as variáveis *Pa* e *Co* são as que tem menor possibilidade de descrever o tipo de sistema presente.

Posteriormente foi realizado um levantamento de correlações entre estas 10 propriedades que estaria relacionadas entre si, para identificar se haveria uma forma de descrever o tipo de sistema com o uso de todas ou parte delas (**Fig.7**).

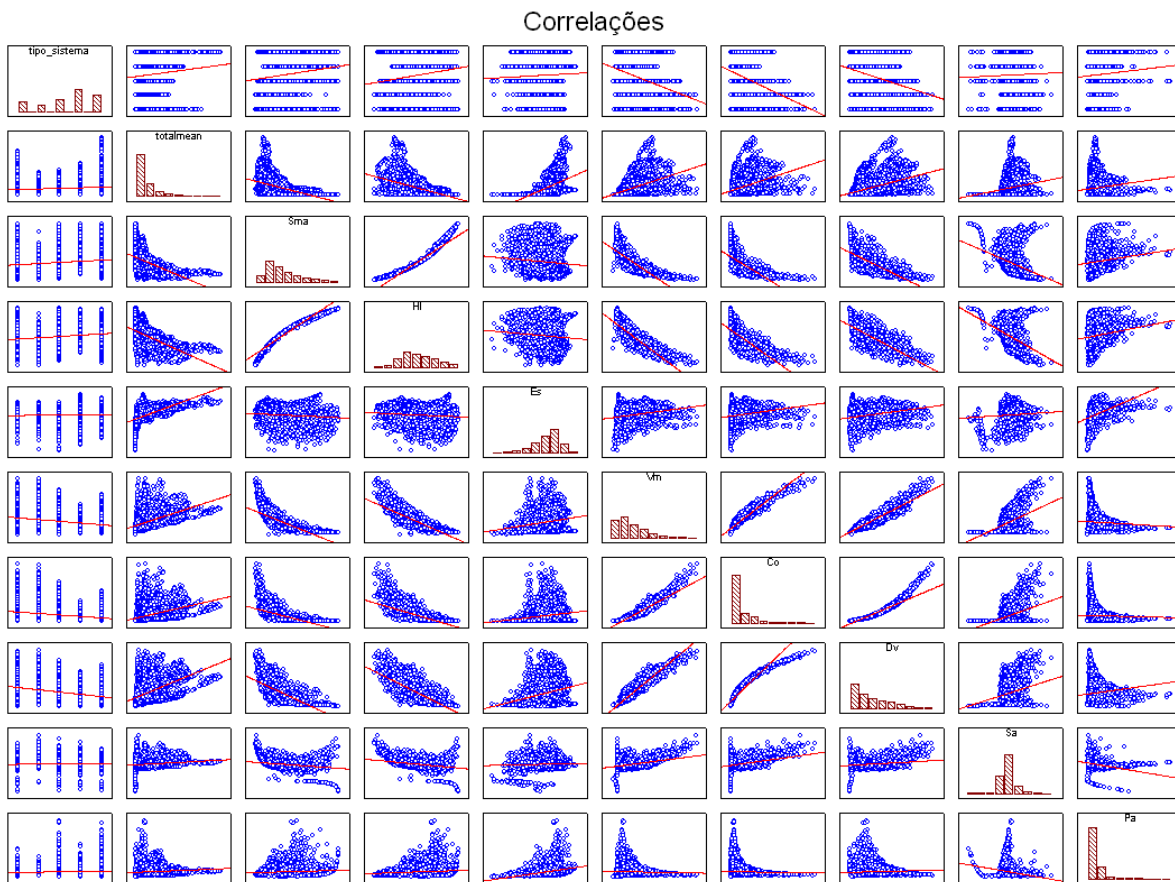


Figura 7: conjunto de correlações entre os pares de variáveis, indicam que apesar de as variáveis se relacionarem, nenhuma das correlações se destaca o suficiente para explicar os sistemas 1 (CI), 2(BM), 3(LI), 4(BD) e 5(FF).

Com esta análise verificou-se que, apesar de algumas das variáveis serem bem descritas uma pela outra, nenhuma delas descreve bem a variável *tiposistema*, sendo que a que mais se aproxima disto é a *Es*. A tentativa seguinte foi examinar a relação de *Es* com *tiposistema* em um gráfico que apresentasse a variação de ambos, este gráfico é apresentado na **Fig. 8**.

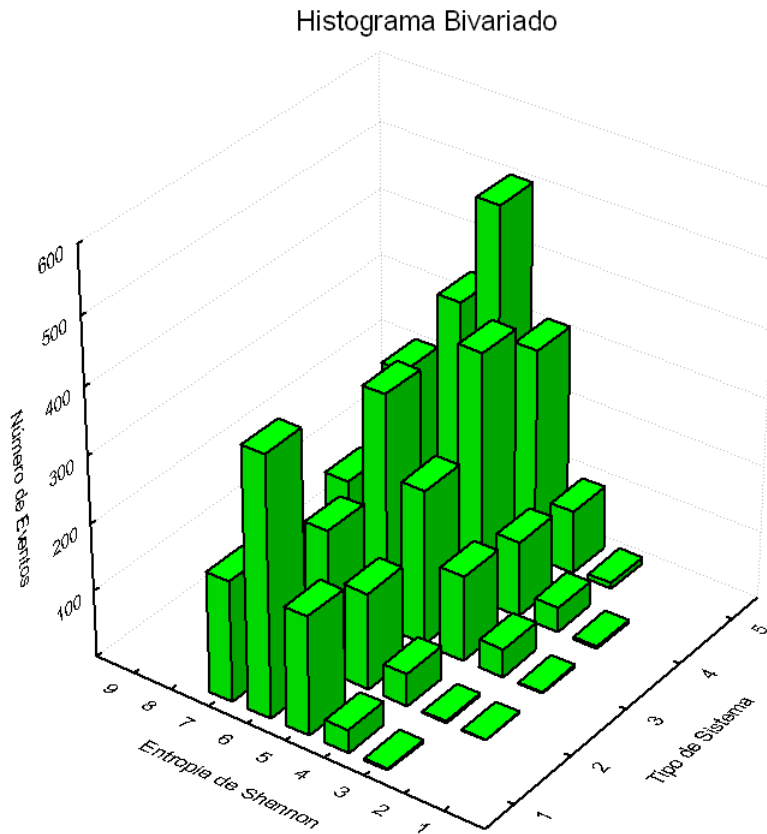


Figura 8: Exemplo do problema encontrado com as variáveis e sua correlação com o tipo de sistema, apesar de alguns casos se destacarem nos picos (neste caso as frentes frias ou sistema 5 têm valores extremos maiores que todos os outros), não há como diferenciar o tipo de sistema da maior parte das imagens apenas pela entropia de Shannon, o mesmo ocorre para o restante das variáveis estudadas e para o conjunto delas.

Pela **Fig.8** é possível identificar diferenças entre os sistemas para Es em torno de 2, por exemplo, em que apenas sistemas dos tipos 2, 3 e 4 aparecem, mas não seria possível identificar o tipo de sistema, apenas eliminar a possibilidade de serem os sistemas 1 ou 5. Mesmo esta possibilidade de eliminar não é importante, uma vez que uma parcela muito pequena (cerca de 5 imagens para o evento 4) das imagens tem Es em torno de 2.

Outro exemplo é a análise de Contraste e Valor Médio (variáveis dependentes), que indica não haver imagens com sistemas classificados como Bandas Dispersas com Co maior que 150 ou Vm maior que 5. Caso esta informação seja confirmada, seria possível afirmar que em uma imagem com Co 200 e/ou Vm 6 não contém Bandas Dispersas.

Da mesma forma, uma imagem com uma quantidade de sítios não nulos maior que 50.000 não seria de BM, um contraste maior que 250 não seria de FF, ou uma imagem com média menor do que 3 e precipitação máxima 100 não teria uma BD. Novamente a quantidade de imagens que apresentam estes valores não é representativa do total de imagens analisada de cada evento.

Foram, ainda, encontrados resultados como BM não apresentarem média total maior do que 0,35 (ou 35% da imagem preenchida com pelo menos 1mm/h de chuva) ou Sombra de Aglomerado maior que 100, FF não apresentarem média maior do que 10, ou ainda que uma quantidade total de chuva maior do que 40.000mm indica a presença de uma FF.

Como todos estes casos são de extremos das variáveis, fica limitado o alcance de possibilidade de identificar um sistema qualquer na região do RMSP utilizando as propriedades de textura utilizadas. Uma alternativa não muito complicada de implementar já com o mesmo programa e banco de dados, seria dividir a região do RMSP em 4 quadrantes, calcular as mesmas variáveis, e verificar se há maiores diferenças entre sistemas quando analisados desta forma.

A divisão em quadrantes poderia, por exemplo, identificar que a precipitação não é de BM por haver precipitação concentrada no quadrante SE, em que não ocorre este tipo de sistema. Seria ainda um primeiro passo antes de se calcular a estatística de ordem superior, comparando vizinho a vizinho.

4 Conclusão

Quanto ao objetivo principal do estudo, não foi possível detectar o tipo de sistema presente na imagem de radar apenas com as variáveis aqui apresentadas, da forma que o problema foi abordado. Há possibilidade de uma abordagem diferente, como dividir as imagens em setores ou aplicar outras variáveis, oferecer um resultado que permita elaborar um modelo de detecção automática do tipo de sistema presente na imagem.

A metodologia de análise de textura desenvolvida por Oliveira e Pereira (2007) se mostrou ineficiente para os objetivos pretendidos com este projeto, indicando que as variáveis estatísticas utilizadas, da forma aqui aplicadas, não são capazes de distinguir diferentes tipos de sistemas em imagens de radar. Isto ocorre devido a semelhança de distribuição de precipitação nos diferentes tipos de sistema estudados,

que mesmo visualmente por vezes necessitam de mais de uma imagem de uma série temporal para que os sistema seja identificado (com ajuda do movimento médio do sistema, sua origem, direção de propagação e velocidade aparente, por exemplo).

Com relação as atividade de Iniciação Científica realizadas, foi importante o contato do aluno com as imagens de radar, assim o aprendizado necessário para interpretá-las e realizar a organização e processamento das mesmas. O projeto, apesar de simples em seu conceito, envolveu uma quantidade grande de dados (cerca de 40GB entre arquivos binários e imagens) e informações a serem extraídas.

As ferramentas utilizadas também propiciaram um conhecimento fundamental para uma maior qualidade e facilidade no processamento e análise de dados em geral, que serão de grande importância em atividades e projetos futuros.

Referências

- [Cavalcanti,2009] Cavalcanti, I. F. A.; Ferreira, N.J.; Silva Dias,M. A. F.; Justi da Silva,M. G. A., 2009: Tempo e Clima no Brasil. Oficina de Textos, 1a Ed. 464p.
- [J. Farias, 2009] Farias, J. F. S., 2009: Previsão de chuva a curtíssimo prazo na área de abrangência do radar meteorológico de São Paulo. Dissertação de Mestrado IAG USP. 97p.
- [Santos, 2006] Santos Silva, F. D., 2006: Análise Objetiva Estatística da Precipitação Estimada com Radar e medida por uma Rede Telemétrica. Dissertação de Mestrado IAG USP. 115p.
- [Oliveira e Pereira, 2007] OLIVEIRA, E. E PEREIRA FILHO, A.J., 2007: Revisiting the statistical texture approach on digital image analysis. Submitted to QJRMS.
- [Silva Dias, 2005] PEREIRA FILHO, A. J.; SILVA, F. D S.: The morphology of tropical rainfall systems and their hydrological significance. In: 32nd conference on radar meteorology, 2005, Albuquerque. Boston: American Meteorological Society, 2005.
- [1] Documentação Java EE 1.6 32bits - <http://java.sun.com/reference/index.jsp#documentation>
- [2] Documentação Eclipse 3.5.1 - <http://www.eclipse.org/documentation/>

[3] Documentação Connector/J 5.1 - dev.mysql.com/doc/refman/5.0/en/connector-j.html

[4] Documentação MySQL Server 5.1- dev.mysql.com/doc/

[5] Documentação ImageJ 1.42q - <http://rsbweb.nih.gov/ij/docs/index.html>