

Aplicação em Banco de Dados Espaciais Para o Rastreamento de Células Convectivas

Marcos Lima Rodrigues¹, Stephan Stephany², Lúbia Vinhas³, Karine Reis Ferreira³, Gilberto Ribeiro de Queiroz³, João Victor Cal Garcia⁴, Carlos Frederico de Angelis⁴

¹Programa de Pós-graduação Computação Aplicada (CAP)

Instituto Nacional de Pesquisas Espaciais (INPE)

Caixa Postal 515 - 12.245-97 – São José dos Campos – SP – Brasil

²Laboratório Associado de Computação e Matemática Aplicada (LAC)

Instituto Nacional de Pesquisas Espaciais (INPE)

Caixa Postal 515 - 12.245-97 – São José dos Campos – SP – Brasil

³Divisão de Processamento de Imagens (DPI)

Instituto Nacional de Pesquisas Espaciais (INPE)

Caixa Postal 515 - 12.245-97 – São José dos Campos – SP – Brasil

⁴Centro Nacional de Monitoramento e Alertas de Desastres Naturais (CEMADEN)

Rodovia Presidente Dutra, Km 40 SP-RJ - 12630-000 – Cachoeira Paulista – SP – Brasil

{marcos.rodriques,stephan.stephany, lubia.vinhas}@inpe.br

{karine.ferreira, gilberto.queiroz}@inpe.br

{joao.garcia, carlos.angelis}@cemaden.gov.br

Abstract. *In this work, we propose the development of a system for thunderstorms cell tracking, based on a spatial database. We also introduce a that performs spatio-temporal clustering operations of lightning data using the DBSCAN algorithm. The spatial database query language were used to define the time window, in order to characterize cells associated with storms. In the analyzed period it was identified and tracked 13 precipitation systems in the radar coverage area located in Jaraguari-MS. Identified cells exhibited good consistency with the intensity core of the thunderstorms.*

Resumo. *Neste trabalho, propomos o desenvolvimento de um sistema para o rastreamento de células de tempestades, com o suporte de banco de dados espaciais (SBDE). O método empregado realiza operações de agrupamento espaço-temporal de descargas atmosféricas, através do algoritmo DBSCAN. O SBDE foi usado para dar suporte a consultas com a definição das janelas de tempo do monitoramento. A fim de caracterizar células associadas a tempestades, no período analisado foram identificados e rastreados 13 sistemas precipitantes na área de alcance do radar de Jaraguari-MS. As células identificadas exibiram boa coerência com a parte mais intensa das tempestades.*

1. Introdução

O Brasil, assim como outros países têm experimentado, principalmente na última década um incremento no número de ocorrências de desastres naturais, vide Figura 1. Estudos

apontam que existe uma conexão com os fenômenos das mudanças climáticas, bem como uma relação com a ocupação desordenada das cidades [Xavier et al. 2014]. Como forma de mitigar os problemas causados por essas ocorrências, é muito importante que se estude a construção de aplicações que processem dados coletados em tempo real com o intuito de auxiliar no monitoramento e alerta de eventos meteorológicos extremos. As tempestades são eventos extremos que podem causar desastres naturais com perdas de vidas, como deslizamentos de terra e inundações.

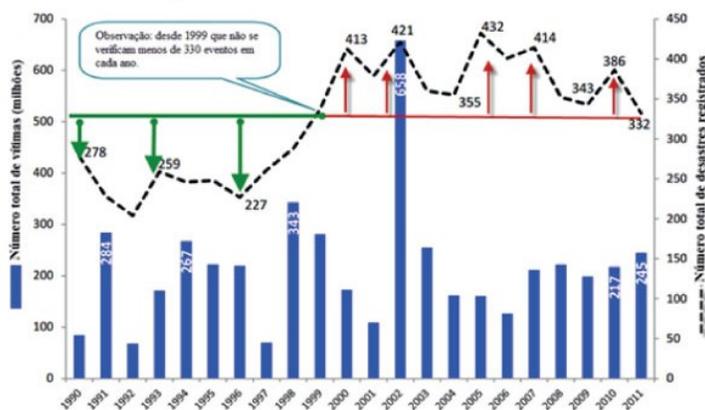


Figura 1. Incremento no número de desastres naturais no Brasil. Adaptada de [MATA-LIMA et al. 2013]

Existem diversos métodos para a realização do rastreamento (*tracking*) de células de tempestades, normalmente definidos para a identificação de áreas com precipitação intensa. É comum a utilização de métodos computacionais e estatísticos para determinação do centro de gravidade de áreas definidas a partir de um limiar relacionado à característica da tempestade. Outro método utilizado é a extrapolação de um campo que pode ser de refletividade ou de precipitação, registrado por um satélite ou radar meteorológico. Pode-se ainda definir células de tempestades a partir da definição de agrupamentos de descargas atmosféricas baseados em função da densidade [Steinacker et al. 2000].

Outros trabalhos recomendam a utilização, de forma complementar, da combinação dos dados de descargas atmosféricas, radares meteorológicos e imagens obtidas por satélites meteorológicos, a fim de aumentar a eficiência do processo de monitoramento e previsão de curto alcance de tempestades ([Steinacker et al. 2000] e [Bonelli and Marcacci 2008]). Isso permite analisar com maior acurácia todo o ciclo de vida das tempestades, pois o processo de eletrificação está relacionado ao início do desenvolvimento do processo de convecção e dura até depois da maturação do sistema em média entre 10 e 20 minutos após a primeira descarga ocorre então a precipitação [Steinacker et al. 2000].

Segundo Betz et al. [2008], o registro de descargas elétricas totais, ou seja, de ambos os tipos, intra-nuvem (*IC*) e nuvem-solo (*CG*), servem como um indicativo e precursor de condições de tempestades extremas, especialmente quando a taxa de *IC* aumenta de forma abrupta num curto espaço de tempo. O trabalho apresentado por [Liu and Heckman 2011] utiliza ambos os tipos de descarga atmosférica para a emissão de alertas de tempestades, através do rastreamento de células de precipitação convectiva. O

uso de funções de densidade e limiares sobre elas permite a delimitação das células de interesse.

Este trabalho tem como objetivo caracterizar e rastrear tempestades a partir de dados de descargas atmosféricas, explorando as capacidades de um Sistema Gerenciador de Banco de Dados (SGBD) com extensão espacial. A abordagem aqui apresentada faz uso de operações espaço-temporais, providas pelo SBDE, para identificar as células de tempestades a partir de agrupamento de descargas atmosféricas. Os sensores possibilitam o registro de descargas atmosféricas, até mesmo em pequenas quantidades, de forma localizada ao longo do tempo, tornando possível a identificação de células associadas a tempestades mais fracas em diferentes instantes de tempo, favorecendo assim o seu agrupamento e rastreamento [Betz et al. 2008].

Uma contribuição importante do trabalho é aumentar o conjunto de ferramentas disponíveis em Meteorologia, através da integração de ferramentas de Geoinformática, na solução de problemas desse domínio. Para isso, descrevemos como foram utilizadas ferramentas baseadas no PostgreSQL e sua extensão espacial PostGIS. Esta solução se mostra uma alternativa robusta e sem custos, com potencial de atender a demandas operacionais de um centro para o monitoramento e alerta de desastres.

2. Região do Estudo e Dados

Nesse trabalho foram utilizados registros de descargas atmosféricas da rede Earth Networks em parceria com a BrasilDAT¹, e as imagens do radar meteorológico instalado no município de Jaraguari-MS, ambos fornecidos pelo Centro Nacional de Monitoramento e Alerta de Desastres Naturais (Cemaden). O Cemaden foi criado pelo governo federal em resposta às catástrofes ocorridas em 2011 em todo o país, especialmente na região serrana do Rio de Janeiro. Sua missão é “realizar o monitoramento das ameaças naturais em áreas de riscos em municípios brasileiros suscetíveis à ocorrência de desastres naturais, além de realizar pesquisas e inovações tecnológicas que possam contribuir para a melhoria de seu sistema de alerta antecipado, com o objetivo final de reduzir o número de vítimas fatais e prejuízos materiais em todo o país” [CEMADEN 2016].

A rede BrasilDAT possui 56 sensores instalados no Brasil, e emprega tecnologia que possibilita detecção, localização e identificação do tipo de descarga ocorrida. A Figura 2 apresenta os locais de instalação dos sensores. Segundo [Naccarato et al. 2016], a rede apresenta eficiência de detecção entre 70 e 85%, precisão na localização entre 400 e 700 metros e capacidade de caracterização do tipo de descarga entre 60 e 80%.

O radar instalado no município de Jaraguari-MS é do tipo banda S e fornece uma varredura volumétrica a cada 10 minutos com 13 elevações de *Plan Position Indicator* (PPI). Os PPIs são então reprojatados para formar uma imagem de refletividade em altitude constante denominada *Constant Altitude Plan Position Indicator* (CAPPI), como demonstra a Figura 3. Esse tipo de imagem é comumente utilizada em Meteorologia para observação e estimativa de precipitação. Neste trabalho, servirá para a validação dos resultados alcançados, pela aplicação desenvolvida.

¹Rede BrasilDAT - <https://www.earthnetworks.com/networks/brazil/>



Figura 2. Localização dos sensores da rede BrasilDat. Adaptada de [Naccarato et al. 2016]

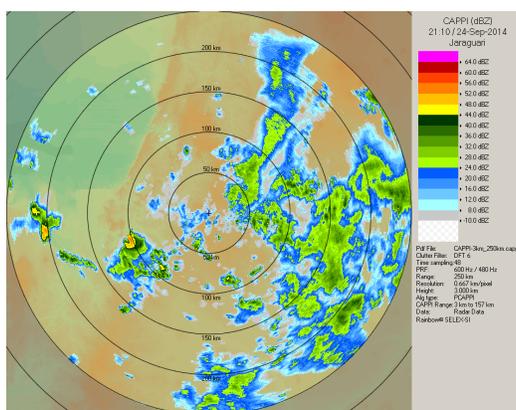


Figura 3. Imagem do CAPPI para o radar de Jaraguari do dia 24/09/2014 às 21:10 UTC. Fonte: Screenshot do software Rainbow 5.

3. Metodologia

Foi criado um banco de dados com os atributos e localização espacial dos registros de descargas atmosféricas de ambos os tipos *IC* e *CG*. Por se tratar de um estudo exploratório, que visa comprovar a possibilidade de se usar um banco de dados geográficos para o acompanhamento de tempestades, usou-se apenas um subconjunto dos dados de descargas elétricas disponíveis. Os dados analisados são referentes ao dia 24/09/2014, quando foram registrados eventos meteorológicos extremos em toda região Centro-Oeste do Brasil, especialmente na região de alcance do radar. Os dados são disponibilizados no formato *Universal ASCII Lightning Format* (UALF); estes são arquivos de texto formatados em colunas de atributos, onde cada linha é referente a uma detecção registrada com os seus atributos como: tipo, data e horário, localização, pico de corrente e altitude (no caso de descargas *IC*), conforme Tabela 1.

Foi desenvolvido um *Shell Script* em linguagem *Bourn Again Shell* (BASH), que

Tabela 1. Extrato das informações contidas no arquivo de dados de descargas atmosféricas.

tipo	datahora	latitude	longitude	pico_corrente	altura_ic
IC	2014-09-24 00:00:00.807	-26.8998734	-57.8639393	-8801	16175
IC	2014-09-24 00:00:01.362	-28.4522589	-54.8191756	7511	13838
CG	2014-09-24 00:00:01.592	-7.0376185	-54.8472766	-15698	0
IC	2014-09-24 00:00:01.644	-28.3011662	-54.7069739	-12911	10561

serviu de interface para a inserção dos registros no banco. Essa rotina teve como finalidade facilitar o carregamento dos dados, uma vez que em um dia foram registrados mais de 120 mil ocorrências de descargas para toda a América do Sul. Como resultado desse pré-processamento, foi gerado um arquivo com comandos SQL para inserção no banco de dados.

O uso de funções de cálculo de densidade permite transformar as observações discretas das descargas em uma superfície de densidade; e a aplicação de limiares sobre elas permitem a delimitação das células. No presente trabalho, a densidade dos pontos de descargas são determinadas com o auxílio do algoritmo de agrupamento (*clustering*) denominado *Density Based Spatial Clustering of Application with Noise* (DBSCAN) [Ester et al. 1996], que atualmente faz parte do core da extensão espacial PostGIS. O algoritmo DBSCAN é baseado na conectividade entre os pontos através da densidade de pontos de vizinhança, a abordagem é caracterizada pela definição de pontos de centro e de borda do agrupamento a partir da definição de 2 parâmetros: densidade de vizinhança (**Eps**), ou seja, distância máxima entre os pontos e a quantidade mínima de pontos (**MinPts**) de vizinhança para que ele pertença ao agrupamento encontrado. Esses parâmetros foram determinados empiricamente a partir de vários testes e da análise dos resultados obtidos. Ele apresenta como principais vantagens a descoberta automática do número de agrupamentos, o que facilita o processo para agrupamentos não conhecidos. Além da capacidade de identificar agrupamentos com formatos arbitrários e eliminação de *outliers* [Cassiano 2014].

Além da tabela para o armazenamento dos dados descargas atmosféricas, o banco possui as seguintes tabelas: *cells_table* e *cells_table_tmp*. Essa última serve para armazenar as células identificadas na última janela de tempo processada e possibilitar a determinação de continuidade das tempestades, através da intersecção entre as células de janelas de tempo distintas; todo o registro nessa tabela é mapeado para a tabela *cells_table* através de uma função gatilho ou *trigger*. A detecção e o rastreamento das células de tempestades são baseados na sobreposição espacial e temporal dos dados de descargas atmosféricas [Li et al. 2008], para isso foi criada uma função em PL/pgSQL, a qual recebe como parâmetros a data e hora inicial e final para o processamento, além dos valores de *Eps* e *MinPts* para o algoritmo de *clustering*. A partir desse intervalo, são definidas janelas de tempo de 10 minutos para o agrupamento e os seguintes parâmetros para o algoritmo de *clustering*: $Eps \approx 8km$ e $MinPts = 2$. Para cada agrupamento foi então definido um círculo envolvente para delimitar as células de descargas associadas a tempestades, essa configuração gerou agrupamentos bem ajustados aos eventos de interesse como demonstra Figura 4. A partir da sobreposição espaço-temporal dessas células, tornou-se possível a identificação, e rastreamento das tempestades pela conexão entre as cen-

troides das células “pais” e suas descendentes células “filhas” ([Liu and Heckman 2011] e [Meyer et al. 2013]); essa é uma abordagem simples que possibilita a investigação e extração de informações sobre o ciclo de vida dos sistemas atuantes [Betz et al. 2008].

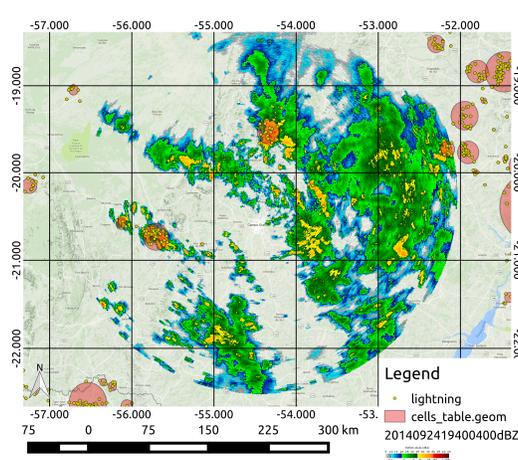


Figura 4. Mapa que evidência a criação do círculo envolvente dos agrupamentos determinados a partir das descargas atmosféricas na janela de 19:40 UTC à 19:50 UTC.

A Figura 5 exibe um fluxograma que ilustra a aplicação da metodologia descrita. Identificando a entrada de dados, processos envolvidos e os resultados alcançados, com a identificação e registros das células de tempestades na tabela *cells_table* na estrutura do banco de dados, bem como a geração de mapas para a visualização através de um Sistema Informações Geográficas (SIG).

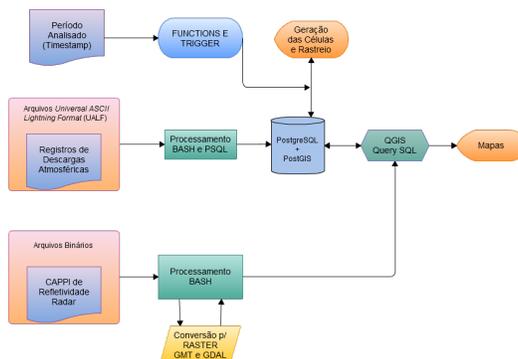


Figura 5. Fluxograma dos principais passos descritos na metodologia, para a identificação e rastreamento das células de tempestades a partir do processo de clustering dos dados de descargas atmosféricas.

4. Resultados

A data escolhida para esse estudo apresentou um grande número de descargas em todo o país, especialmente concentradas nas regiões Centro-Oeste e Sul, devido ao avanço de uma frente fria, ver Figura 6.

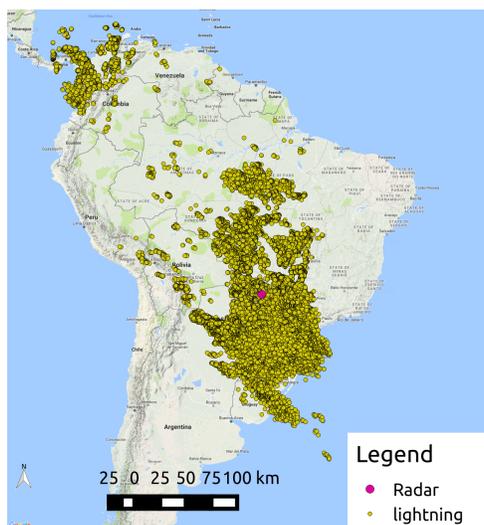


Figura 6. Mapa com a distribuição total de descargas atmosférica para o dia 24/09/2014, foram registrados 128663 eventos entre IC e CG.

A partir do total de registros de ocorrência de descargas atmosféricas, considerou-se o seguinte intervalo de tempo: das 19:30 UTC às 22:50 UTC para análise, e a resolução temporal da janela definida. Foram geradas 1114 células, das quais 116 foram registradas dentro da área de cobertura do radar, conforme demonstra a Figura 7.

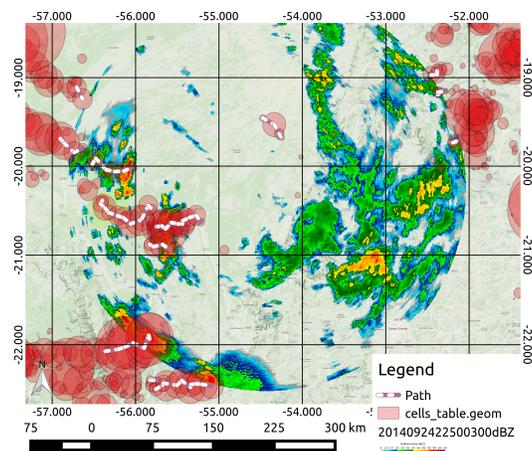


Figura 7. Mapa com o total de agrupamentos criados próximos a área de cobertura do radar, cada célula está associada a um núcleo de tempestade no tempo.

A partir da intersecção do caminho seguido por essas células com a área de do alcance do radar, foram identificados ao longo do período o acompanhamento de 13 sistemas convectivos com deslocamentos territoriais importantes, conforme destaca a Figura

8. Cada cluster que foi gerado foi associado a um sistema convectivo identificado pelo radar por meio de comparação visual.

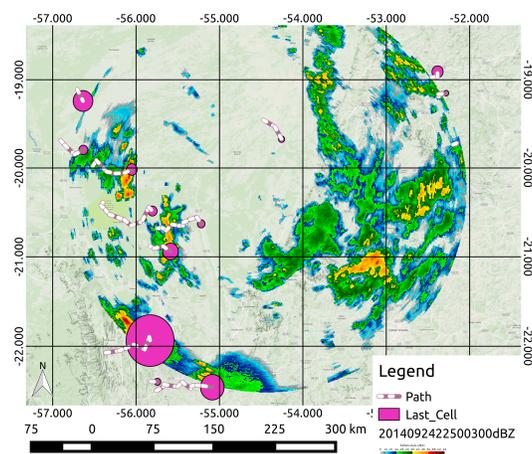


Figura 8. Rastreamento dos sistemas precipitantes que que interagiram com o radar de Jaraguari durante o período de análise.

Além de identificar e rastrear a células associadas a tempestades, o sistema permite a coleta de informações sobre o ciclo de vida das mesmas. Essa informação pode ser determinante para órgãos como a defesa civil, fornecendo informações que ajudem a definir a severidade de um evento, favorecendo assim a tomada de decisão por parte de seus agentes. A Figura 9 apresenta, em detalhes, o rastreamento e a posição da última célula que demonstra o fim do ciclo de vida desses sistemas.

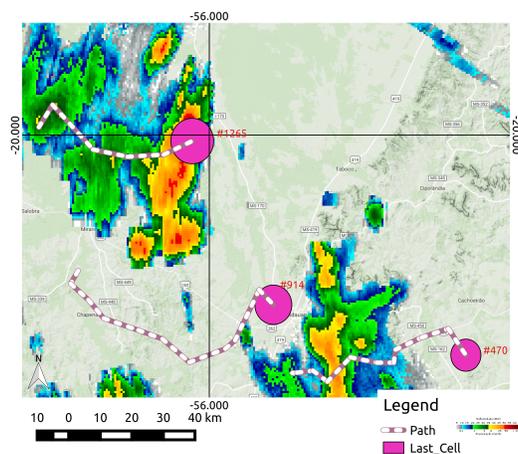


Figura 9. Destaque de sistemas rastreados na área de alcance do radar entre o intervalo de 19:30 UTC à 22:50 UTC.

A partir da informação de identificação das células rastreadas, foram então exportadas do banco de dados algumas das características desses sistemas. A Figura 10 apresenta uma visão geral dessas informações, para tempestades registradas na data de 24 de setembro de 2014. Essas informações permitem acompanhar a iniciação, desenvolvimento e extinção dos sistemas convectivos atuantes. O processo ilustrado pela Figura 5

permite perceber que é possível a operacionalização desse método, e a exibição de forma automática das células de tempestades identificadas ao longo do tempo, bem como o caminho feito por elas. Esses resultados podem ser exibidos através de uma interface web baseada em um servidor de mapas, para que possa auxiliar o trabalho de equipes responsáveis pelo monitoramento e alerta de desastres naturais ligados a eventos extremos de precipitação.

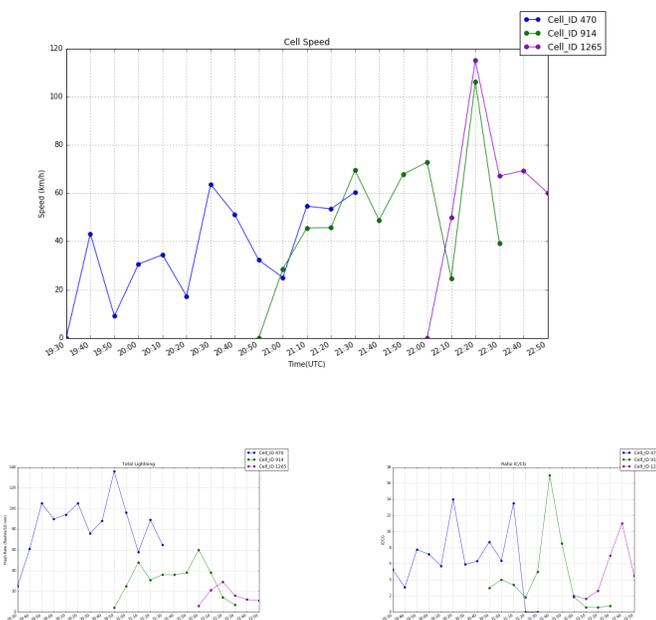


Figura 10. Informações de velocidade, total de descargas e relação entre as descargas do tipo IC/CG para as células de tempestades com ID 470, 914, 1265.

5. Conclusão

O processo de identificação e rastreamento de tempestades é uma etapa importante para o desenvolvimento de previsões de curto alcance (*nowcasting*), geralmente realizadas a partir da extrapolação da informação em instantes anteriores [Betz et al. 2008]. O método aqui desenvolvido, com o uso de operações espaciais e do algoritmo de *clustering* DBSCAN, apresentou boa coerência com os núcleos de tempestades visualizados na imagem do radar. Como trabalho futuro recomenda-se a adoção de um algoritmo de *clustering* que não dependa da ordem dos dados e que seja flexível na definição da densidade local dos elementos agrupados, como por exemplo o apresentado por [Birant and Kut 2007], favorecendo assim a delimitação de *clusters* adjacentes. As operações providas pelo SBDE facilitaram o processo de geração da informação da continuidade dos sistemas precipitantes devido à sobreposição espacial das células de tempestades.

Por se tratar de um trabalho em andamento, a estrutura desenvolvida para esse estudo foi baseada num aporte inicial de dados e imagens para validação dos resultados, porém para sua utilização de forma operacional, bastaria algumas pequenas adaptações como:

1. Modificar o script para adquirir os dados de descarga de forma automática e em tempo quase real a partir do *webservice* do Cemaden;
2. Utilizar um agendador de tarefas como o *crontab*², para a execução da função PL/pgSQL responsável pelo processo de identificação e rastreamento de tempestades;
3. Integrar a consulta de resultados do processamento registrado no banco, a um sistema servidor de mapas como o *GeoServer*³, para exibição das células de tempestades e da sua informação de rastreamento.

A partir da integração desse trabalho com um sistema servidor de mapas, é possível exibir em *layers* as células de tempestades identificadas e sua informação de rastreamento, constituindo assim um ferramental importante para os órgãos que necessitam de informações sobre tempestades no menor tempo possível. A escalabilidade da abordagem proposta será analisada num trabalho futuro, mas foram gerados agrupamentos para uma região extensa, que corresponde a quase toda a América do Sul, apenas a validação desses agrupamentos ficou restrita ao radar de Jaguari.

Referências

- [Betz et al. 2008] Betz, H. D., Schmidt, K., Oettinger, W. P., and Montag, B. (2008). Cell-tracking with lightning data from LINET. *Advances in Geosciences*, 17:55–61.
- [Birant and Kut 2007] Birant, D. and Kut, A. (2007). ST-DBSCAN: An algorithm for clustering spatial-temporal data. *Data and Knowledge Engineering*, 60(1):208–221.
- [Bonelli and Marcacci 2008] Bonelli, P. and Marcacci, P. (2008). Thunderstorm nowcasting by means of lightning and radar data: Algorithms and applications in northern Italy. *Natural Hazards and Earth System Sciences*, 8(5):1187–1198.
- [Cassiano 2014] Cassiano, K. M. (2014). Análise de Séries Temporais Usando Análise Espectral Singular (SSA) e Clusterização de Suas Componentes Baseada em Densidade. *din.uem.br*, page 172.
- [CEMADEN 2016] CEMADEN (2016). Missão.
- [Ester et al. 1996] Ester, M., Kriegel, H.-p., S, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, pages 226–231.
- [Li et al. 2008] Li, X., Ramachandran, R., Movva, S., Graves, S., Plale, B., and Vijayakumar, N. (2008). Storm Clustering for Data driven Weather Forecasting. In *24th Conference on IIPS*, pages 1–5.
- [Liu and Heckman 2011] Liu, C. and Heckman, S. (2011). The application of total lightning detection and cell tracking for severe weather prediction. *WMO Technical Conference on Instruments and . . .*
- [MATA-LIMA et al. 2013] MATA-LIMA, H., ALVINO-BORBA, A., PINHEIRO, A., MATA-LIMA, A., and ALMEIDA, J. A. (2013). Impactos dos desastres naturais nos sistemas ambiental e socioeconômico: o que faz a diferença. *Ambient. soc., São Paulo*, 16(3).

²Manual do *crontab* agendador de tarefas do Linux, disponível em: <http://man7.org/linux/man-pages/man5/crontab.5.html>, consultado em set. 2016.

³*GeoServer* - servidor de mapas de código aberto, disponível em: <http://geoserver.org/>, consultado em set. 2016.

- [Meyer et al. 2013] Meyer, V. K., Höller, H., and Betz, H. D. (2013). Automated thunderstorm tracking: Utilization of three-dimensional lightning and radar data. *Atmospheric Chemistry and Physics*, 13(10):5137–5150.
- [Naccarato et al. 2016] Naccarato, K. P., Santos, W. A., Carretero, M. A., Moura, C., and Tikami, A. (2016). Total Lightning Flash Detection from Space A CubeSat Approach. *Proceedings of the ILDC*.
- [Steinacker et al. 2000] Steinacker, R., Dorninger, M., Wolfelmaier, F., and Krennert, T. (2000). Automatic Tracking of Convective Cells and Cell Complexes from Lightning and Radar Data. *Meteorology and Atmospheric Physics*, 72:101–110.
- [Xavier et al. 2014] Xavier, D. R., Barcellos, C., Barros, H. d. S., Magalhães, M. d. A. F. M., de Matos, V. P., and Pedroso, M. d. M. (2014). Organização, disponibilização e possibilidades de análise de dados sobre desastres de origem climática e seus impactos sobre a saúde no Brasil. *Ciência & Saúde Coletiva*, 19(9):3657–3668.