



Ministério da  
Ciência e Tecnologia



INPE-16610-RPQ/834

## USO DE REDES NEURAIS PARA A PREVISÃO DE DESCARGAS ATMOSFÉRICAS NO SUDESTE DO BRASIL

Gisele dos Santos Zepka

Relatório final da disciplina Princípios e Aplicações de Mineração de Dados  
(CAP-359) do Programa de Pós-Graduação em Computação Aplicada, ministrada  
pelo professor Rafael Santos.

Registro do documento original:

<<http://urlib.net/sid.inpe.br/mtc-m18@80/2009/10.01.19.08>>

INPE  
São José dos Campos  
2009

## **PUBLICADO POR:**

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3945-6911/6923

Fax: (012) 3945-6919

E-mail: [pubtc@sid.inpe.br](mailto:pubtc@sid.inpe.br)

## **CONSELHO DE EDITORAÇÃO:**

### **Presidente:**

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

### **Membros:**

Dr<sup>a</sup> Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr. Haroldo Fraga de Campos Velho - Centro de Tecnologias Especiais (CTE)

Dr<sup>a</sup> Inez Staciarini Batista - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Dr. Ralf Gielow - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dr. Wilson Yamaguti - Coordenação Engenharia e Tecnologia Espacial (ETE)

## **BIBLIOTECA DIGITAL:**

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Jefferson Andrade Ancelmo - Serviço de Informação e Documentação (SID)

Simone A. Del-Ducca Barbedo - Serviço de Informação e Documentação (SID)

## **REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:**

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Marilúcia Santos Melo Cid - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

## **EDITORAÇÃO ELETRÔNICA:**

Viveca Sant´Ana Lemos - Serviço de Informação e Documentação (SID)

## **RESUMO**

Identificar, com razoável antecedência e confiabilidade, condições favoráveis à ocorrência de descargas atmosféricas é essencial para a veiculação adequada de alertas e antecipar a adoção de estratégias que eliminem ou minimizem o impacto negativo deste fenômeno meteorológico. O objetivo deste trabalho é apresentar resultados preliminares de um estudo inicial, a partir de correlações entre dados de descargas atmosféricas e variáveis meteorológicas do modelo WRF, para o desenvolvimento de um previsor de descargas atmosféricas usando redes neurais artificiais. As características e os tipos de relâmpagos e uma breve descrição do modelo WRF serão aqui discutidos. A seguir o banco de dados utilizado neste trabalho e a arquitetura de rede neural adotada serão apresentados. Por fim, são mostrados os resultados de previsão alcançados e então sugeridos estudos futuros que dêem continuidade à idéia proposta.

## SUMÁRIO

	<u>Pág.</u>
<b>LISTA DE FIGURAS .....</b>	<b>4</b>
<b>1 INTRODUÇÃO.....</b>	<b>5</b>
<b>2 MATERIAL E MÉTODOS .....</b>	<b>9</b>
2.1 Relâmpagos.....	9
2.2 Modelo de Mesoescala WRF.....	12
2.3 Mineração de Dados.....	14
2.3.1 Banco de Dados.....	16
2.3.2 Rede Neural Artificial.....	18
<b>3 CONCLUSÕES .....</b>	<b>20</b>
<b>4 REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>21</b>

## LISTA DE FIGURAS

	<b><u>Pág.</u></b>
Figura 01 – Diferentes tipos de relâmpagos: descarga para o ar (DA), relâmpago intranuvem (IN), nuvem-solo positivo (NS+), nuvem-solo negativo (NS-) e solo-nuvem (SN). ....	9
Figura 02 – Rede Brasileira de Detecção de Descargas Atmosféricas. ....	11
Figura 03 – Etapas do processo de simulação no modelo WRF.....	13
Figura 04 – Mapa de eficiência de detecção da BrasilDAT.....	16
Figura 05 – Domínios de integração do modelo WRF.....	17
Figura 06 – Variação diurna do número total de raios negativos e positivos na região Sudeste para o período de 1999-2004. ....	18

## 1 INTRODUÇÃO

Relâmpagos existem em nosso planeta desde os tempos mais remotos, como parte de um complexo sistema que conhecemos pelo nome de natureza. Apesar de sua beleza e imponência, este fenômeno natural representa uma ameaça potencial para atividades importantes como, por exemplo, defesa civil, aviação, telecomunicações, transmissão e distribuição de energia elétrica. Segundo Pinto Jr. e Pinto (2008), os raios causam no Brasil cerca de cem mortes por ano e entre 300 e 400 pessoas feridas, além de prejuízos da ordem de um bilhão de reais. Estima-se que aproximadamente 150 milhões de relâmpagos ocorram em nosso país; destes, 50 milhões atingem o solo.

Identificar, com antecedência de várias horas, condições favoráveis à ocorrência de descargas atmosféricas é essencial para a veiculação adequada de alertas e antecipar a adoção de estratégias que eliminem ou minimizem o impacto negativo deste fenômeno meteorológico. Discutir estratégias de previsão, visando à implementação operacional, implica: conhecer melhor os ambientes atmosféricos propícios ao desenvolvimento de tempestades severas; estudar procedimentos que maximizem a extração de informação relevante dos dados observados e de modelos de mesoescala que reconheçam esses ambientes.

Poucos, ou até mesmo raros, são os estudos na tentativa de desenvolver e relacionar ferramentas de previsão de descargas atmosféricas. Frankel et al. (1991) construíram e treinaram uma rede neural artificial, com variáveis como vento, campo elétrico e o próprio dado de raio na entrada, para gerar mapas de probabilidade da ocorrência de eventos severos, especialmente, raios sobre o Complexo da Estação da Força Aérea do Cabo Canaveral/ Centro Espacial Kennedy, Flórida. Dada a natureza preliminar do estudo, os resultados mostraram-se promissores. Nagae et al. (2000) utilizaram uma rede neural acoplada à lógica fuzzy para a previsão da atividade elétrica sobre a região de

Tóquio, Japão, durante os verões de 1995 e 1997. Os autores conseguiram prever de forma significativa a ocorrência de relâmpagos de uma hora à frente, com resultados satisfatórios quando comparados aos dados observacionais.

Zepka (2005), a partir de correlações entre dados de descargas atmosféricas e variáveis meteorológicas simuladas no modelo de mesoescala MM5, elaborou um estudo inicial para o desenvolvimento de um previsor de descargas atmosféricas usando redes neurais artificiais. O sistema de previsão é proposto e, então, testado em parte da região Sudeste do Brasil. Os resultados alcançados, apesar de promissores, são restritos ao caso de previsão de curtíssimo prazo.

Frisbie et al. (2009), a fim de prever a localização e a frequência de ocorrência de relâmpagos na cidade de *Grand Junction*, Colorado, propuseram o *Lightning Potential Index* (LPI), calculado a partir de uma equação empírica que leva em conta parâmetros de instabilidade e umidade obtidos de modelo meteorológico. O índice é dividido em quatro níveis de risco: baixo, moderado, alto e extremamente alto. Apesar de ser um bom indicador da região onde o relâmpago possa acontecer, a atividade elétrica é subestimada em áreas de forte movimento ascensional, uma vez que variáveis meteorológicas relacionadas a mecanismos dinâmicos de levantamento não são computadas no cálculo do LPI.

Modelos numéricos de mesoescala (com espaçamento horizontal de grade tipicamente entre 10 e 30 km) podem ser utilizados para o cálculo de parâmetros relacionados à atividade convectiva (NASCIMENTO, 2005). Esta estratégia apresenta a vantagem de não requerer a execução de modelos numéricos em escala de tempestade (espaçamento horizontal de grade em torno de ou menor que 5 km), onde células convectivas são resolvidas explicitamente e uma maior capacidade computacional é indispensável (NASCIMENTO, 2005).

O modelo WRF (*Weather Research and Forecasting*) é o sistema de previsão numérica do tempo em mesoescala da próxima geração, desenvolvido e constantemente atualizado para atender tanto às necessidades das previsões operacionais como das pesquisas atmosféricas (WANG ET AL., 2004). Ele foi concebido de forma a promover uma estreita relação entre operação e pesquisa, em prol da qualidade da previsão meteorológica para a sociedade em geral. Um dos aspectos do desenvolvimento do sistema WRF, que tem sido ressaltado como de fundamental importância, é sua utilização extensiva, mediante testes em diferentes regiões do globo com o emprego de diferentes opções de parametrizações físicas, bem como, avaliações da integridade do código numérico propriamente dito (YAMASAKI ET AL., 2006).

Fazer uso de parâmetros meteorológicos de saída do modelo WRF, na busca por reconhecimento de ambientes favoráveis à ocorrência de descargas atmosféricas, é uma estratégia interessante. Entretanto, talvez um problema a ser enfrentado com a coleta automatizada de dados provenientes de simulações numéricas seja, devido ao volume de informações, a extração de conhecimento útil, necessário para fazer inferências, obter conclusões ou comprovar hipóteses. O conhecimento sobre fenômenos relativos a estes dados permitirá justamente análises mais complexas como a própria previsão de descargas atmosféricas. Nesse caso, o processo geral de descoberta de conhecimentos úteis previamente desconhecidos a partir de grandes bancos de dados possa acontecer pelo uso de técnicas matemáticas apropriadas na transformação de dados em padrões.

O objetivo deste trabalho é mostrar resultados preliminares de um estudo inicial, a partir de correlações entre dados de descargas atmosféricas e campos meteorológicos simulados no modelo WRF, para o desenvolvimento de um previsor de descargas atmosféricas aplicado à região Sudeste do Brasil usando redes neurais artificiais.

A fim de fundamentar a discussão acima, a estrutura do texto que segue está disposta da seguinte forma: o item 2 tratará do material e métodos utilizados neste trabalho, começando por uma revisão das características e dos tipos de relâmpagos e uma breve descrição do modelo WRF. No fim deste item, uma abordagem conceitual de ferramentas matemáticas é introduzida em mineração de dados, com a apresentação do banco de dados e da arquitetura adotada de redes neurais artificiais. Os resultados preliminares do estudo proposto são então discutidos no item 3, assim como as perspectivas de trabalhos futuros.

## 2 MATERIAL E MÉTODOS

### 2.1 Relâmpagos

Em termos de nomenclatura, considera-se o relâmpago como um fenômeno físico completo, desde seu início na nuvem até a sua total extinção. A denominação raio é específica para relâmpagos do tipo nuvem-solo (PINTO JR., 2005), comentados a seguir. Relâmpagos são descargas elétricas de grande intensidade e comprimento que se originam na camada atmosférica mais importante do ponto de vista meteorológico, a troposfera. Possuem um canal ionizado, caracterizado ou não por várias ramificações e que se estende a longas distâncias verticais e horizontais. Esse canal atua como uma efetiva antena de irradiação de ondas eletromagnéticas que se propagam ao redor do planeta através da atmosfera, gerando vários fenômenos não somente próximos aos locais de ocorrência, como também a grandes distâncias, desde a superfície da Terra até as camadas mais altas da atmosfera (RAKOV E UMAN, 2003).

Existem diversos tipos de relâmpagos classificados de acordo com o local onde se originam e terminam (Figura 01).

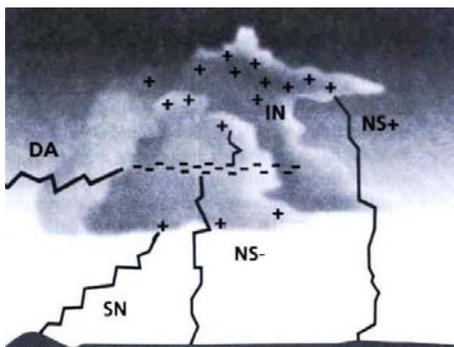


Figura 01 – Diferentes tipos de relâmpagos: descarga para o ar (DA), relâmpago intranuvem (IN), nuvem-solo positivo (NS+), nuvem-solo negativo (NS-) e solo-nuvem (SN).

FONTE: Pinto Jr. (2005).

Os relâmpagos mais freqüentes começam e terminam dentro da nuvem e são chamados intranuvem (IN). Globalmente, eles representam cerca de 70% do número total de relâmpagos e isso se deve ao fato de a capacidade isolante do ar diminuir com a altura em função da diminuição da densidade do ar e devido à maior proximidade de centros de carga de polaridades opostas. Relâmpagos mais raros que terminam na atmosfera são as descargas no ar (DA). Os outros tipos de relâmpagos conectam-se ao solo e são os nuvem-solo (NS) e solo-nuvem (SN). Dependendo da carga efetiva que é transferida para o solo, eles ainda podem ser separados em positivos e negativos. O mais comum dos relâmpagos NS é o nuvem-solo negativo (NS-), observado em cerca de 90% dos casos. Os restantes 10% são basicamente nuvem-solo positivos (NS+). Já os relâmpagos solo-nuvem são os mais raros de todos, porém eles tendem a ocorrer mais freqüentemente a partir de lugares muito altos como topo de montanhas e grandes edificações (PINTO JR. E PINTO, 2000).

Os raios são os mais estudados e, conseqüentemente, melhor compreendidos, não só devido a sua relativa facilidade de observação em relação aos relâmpagos intranuvem como também ao seu caráter destrutivo (RAKOV E UMAN, 2003; PINTO JR, 2005; NACCARATO, 2006). Duram cerca de um quarto de segundo, embora valores variando desde um décimo de segundo a dois segundos têm sido registrados. Durante esse tempo, percorrem trajetórias na atmosfera de até dezenas de quilômetros, com a corrente elétrica variando desde centenas de ampères até centenas de quiloampères (PINTO JR., 2005). Ao conectar no solo, a carga transferida varia de algumas dezenas de coulombs até centenas de coulombs (UMAN, 1987; RAKOV E UMAN, 2003).

Os atuais sistemas de detecção de relâmpagos são baseados na medida da radiação eletromagnética emitida pelas descargas a partir de um conjunto de sensores dispostos adequadamente em áreas de até centenas de milhares de quilômetros quadrados, com o objetivo de localizar precisamente os raios para um determinado fim (PINTO JR., 2005).



Figura 02 – Rede Brasileira de Detecção de Descargas Atmosféricas.

FONTE: Naccarato e Pinto Jr. (2009).

No Brasil, a rede BrasilDAT (Rede Brasileira de Detecção de Descargas Atmosféricas) abrange cerca de 60% do país (PINTO JR. ET AL., 2006), sendo composta por 47 sensores, que operam na frequência de LF (*Low Frequency*), instalados em 12 estados brasileiros: RS, SC, PR, SP, MS, RJ, ES, MG, GO, TO, MA, e PA (os dois últimos cobertos parcialmente), como mostrado acima na Figura 02. Os sistemas de detecção que trabalham na faixa de frequência de LF (30-300 kHz) são eficientes para a detecção de relâmpagos NS com uma resolução espacial média de 1 km e uma eficiência de detecção de aproximadamente 80 a 90% (CUMMINS ET AL., 1998; NACCARATO, 2006). No Brasil, a eficiência da BrasilDAT é de 87,4% para raios e 54,7% para

descargas de retorno, medidas no Vale do Paraíba (BALLAROTTI ET AL., 2006; PINTO JR. ET AL., 2007). Esses valores variam de região para região, dependendo do tipo de sensor utilizado, da geografia do lugar e da disposição dos sensores entre si.

## **2.2 Modelo de Mesoescala WRF**

O modelo WRF é um sistema de modelagem numérica da atmosfera, destinado tanto para fins de pesquisa quanto operacionais, em diferentes escalas espaciais, desde alguns metros até milhares de quilômetros. O seu desenvolvimento contínuo é fruto da colaboração das seguintes instituições e agências governamentais: *National Center for Atmospheric Research (NCAR)*, *National Oceanic and Atmospheric Administration (NOAA)*, *National Centers for Environmental Prediction (NCEP)*, *Forecast Systems Laboratory (FSL)*, *Air Force Weather Agency (AFWA)*, *Naval Research Laboratory*, *Oklahoma University* e *Federal Aviation Administration (FAA)* (SKAMAROCK ET AL., 2005). O esforço de aperfeiçoamento do modelo WRF é, para a comunidade científica, um desafio a melhor compreensão dos fenômenos atmosféricos no que diz respeito à capacidade de modelá-los corretamente. Concebido para ser uma ferramenta de topo na arte da simulação atmosférica, flexível, portátil e eficiente em variadas plataformas computacionais, o sistema WRF é de domínio público e disponibilizado gratuitamente. Oferece uma diversidade de opções físicas, além de sistemas avançados de assimilação de dados que estão sendo desenvolvidos e testados (SKAMAROCK ET AL., 2005).

Em aplicações reais, seja na previsão do tempo, seja na simulação do tempo passado, o WRF, como modelo de área limitada, necessita de condições de fronteira laterais que representem, dentro do possível, o estado real da atmosfera vizinha ao longo do tempo de integração das equações dinâmicas. Essas condições são fornecidas por um modelo de previsão numérica do tempo em escala global. O estado inicial deve ser definido pelos valores de

análise nos vértices da própria grade desse modelo. Os modelos globais são capazes de fornecer previsões para vários dias num prazo suficientemente curto a fim de possibilitar o avanço das previsões dos modelos de área limitada. As condições de inicialização do WRF são fornecidas pelo modelo global GFS (*Global Forecast System*), rodado quatro vezes ao dia (00, 06, 12 e 18 UT) com horizonte de simulação de 384 horas, resolução espacial de 0,5, 1 ou 2,5 graus de latitude e longitude e resolução temporal de três horas. Os dados gerados são disponibilizados na internet pelo NCEP.

A Figura 03 ilustra resumidamente as etapas do processo de simulação no modelo WRF.

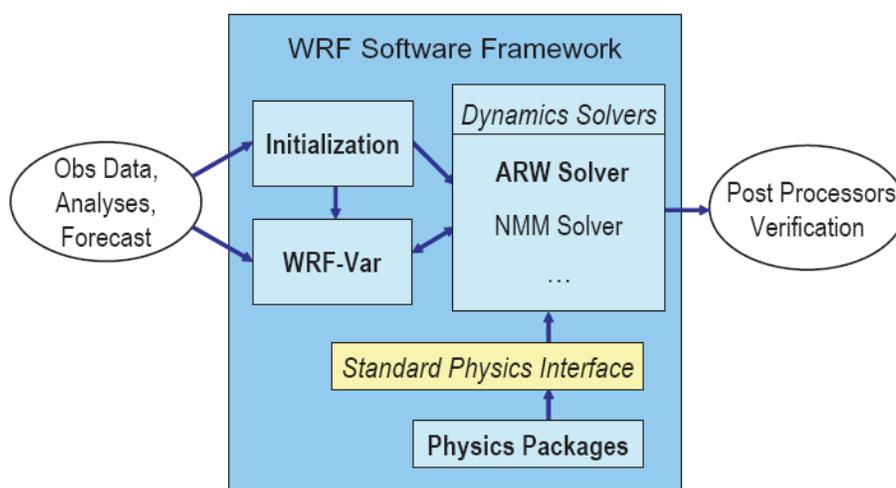


Figura 03 – Etapas do processo de simulação no modelo WRF.

FONTE: Skamarock et al. (2005).

O pacote WRFSI (*WRF Standard Inicialization*) é formado por um conjunto de programas que transformam dados terrestres (topografia, uso do solo, cobertura vegetal, etc.) e dados meteorológicos (gerados pelos modelos globais) em arquivos de condições de contorno e condições iniciais. O primeiro passo executado nesta etapa é a definição da grade (tipo de projeção, localização, número de pontos e espaçamentos) para que sejam interpolados

os campos estáticos que descreverão o domínio. Em seguida, os dados meteorológicos fornecidos pelo modelo global (geralmente em formato *GRidded Binary* – GRIB) são interpolados horizontalmente, no sistema de coordenadas da grade, e verticalmente, no sistema de coordenadas  $\eta$  (eta). Como resultado, são gerados arquivos binários, no formato netCDF (*network Common Data Form*), com as condições iniciais e de contorno que serão usadas pelo modelo WRF. Após a simulação, os arquivos de saída do modelo WRF são convertidos do formato netCDF para o formato binário do GrADS (Sistema de Análise e Exibição de Grade) para a visualização das variáveis de interesse.

### **2.3 Mineração de Dados**

Mineração de dados é o nome dado ao conjunto de técnicas que permitem a extração de conhecimentos a partir de bancos de dados (SUMATHI E SIVANANDAM, 2006). A descoberta de conhecimentos em bancos de dados (*knowledge discovery in databases* ou KDD) é o processo não-trivial de identificação de padrões válidos, novos, compreensíveis e potencialmente úteis (FAYYAD ET AL., 1996). KDD pode ser visto como um processo iterativo composto das seguintes etapas: compreensão do domínio da aplicação, do conhecimento existente e objetivos do usuário final; escolha de subconjunto dos dados; limpeza e pré-processamento dos dados; redução e projeção dos dados; escolha da tarefa de mineração dos dados; mineração dos dados; interpretação dos resultados e consolidação do conhecimento obtido (FAYYAD ET AL., 1996).

O modelo WRF oferece aos usuários mais de uma centena de variáveis meteorológicas representativas do comportamento da atmosfera. Escolher aleatoriamente parâmetros preditores em potencial da ocorrência de descargas atmosféricas é uma tarefa bastante complexa, visto as inúmeras correlações que podem ser inferidas. Contudo, sabe-se que as condições primariamente

necessárias para a formação de tempestades são o perfil térmico da atmosfera, umidade e movimento ascensional, o que já limita de forma considerável o conjunto de escolhas. Mesmo assim, as possibilidades ainda são muitas e a aplicação de uma técnica de mineração de dados é uma alternativa.

Várias técnicas e algoritmos de transformação de dados em padrões podem ser usados e sua escolha deve ser feita considerando o tipo de dado e o tipo de informação associada ao padrão que se deseja obter. É possível ter técnicas híbridas ou que são usadas de forma distinta em diferentes passos do processo de KDD. Cada uma destas técnicas tem vários algoritmos, com variantes e diferentes parâmetros.

As Redes Neurais Artificiais (RNAs ou NNs, de *Neural Networks*), que são simulações simplificadas do comportamento de neurônios reais, representam um método de mineração de dados (BERSON E SMITH, 1997). RNAs de diversos tipos podem ser usadas como classificadores supervisionados, sendo mais comuns os perceptrons (neurônios artificiais) em múltiplas camadas (FAUSETT, 1994). Um único neurônio é capaz de separar duas classes, contanto que as mesmas sejam linearmente separáveis por um único plano. Para separar mais classes e possibilitar a separação não linear (através da combinação de separadores lineares), deve-se usar um conjunto de neurônios em camadas. As ativações dos neurônios são propagadas nas camadas da rede; seu treinamento é feito apresentando-se valores de dados e classes esperadas e ajustando-se os pesos de cada neurônio até que a rede esteja treinada adequadamente. Tradicionalmente, o número de neurônios na camada de entrada é igual ao número de atributos dos registros e o número de neurônios na camada de saída é igual ao número de classes. Redes neurais de múltiplas camadas de perceptrons conseguem separar classes com distribuições complexas, contanto que um número adequado de neurônios em camadas escondidas seja usado (FAUSETT, 1994).

### 2.3.1 Banco de Dados

O banco de dados deste trabalho é composto por dados de relâmpagos nuvem-solo provenientes da rede BrasilDAT e por parâmetros meteorológicos obtidos a partir de simulações numéricas com o modelo de mesoescala WRF, para um período de seis meses de dados: outubro de 2005 a março de 2006.

A Figura 04 mostra o mapa de eficiência de detecção de descargas atmosféricas da BrasilDAT. Com base nele, a região de interesse foi definida como parte do Sudeste do Brasil, onde a eficiência de detecção é de aproximadamente 90% (NACCARATO E PINTO JR., 2009).

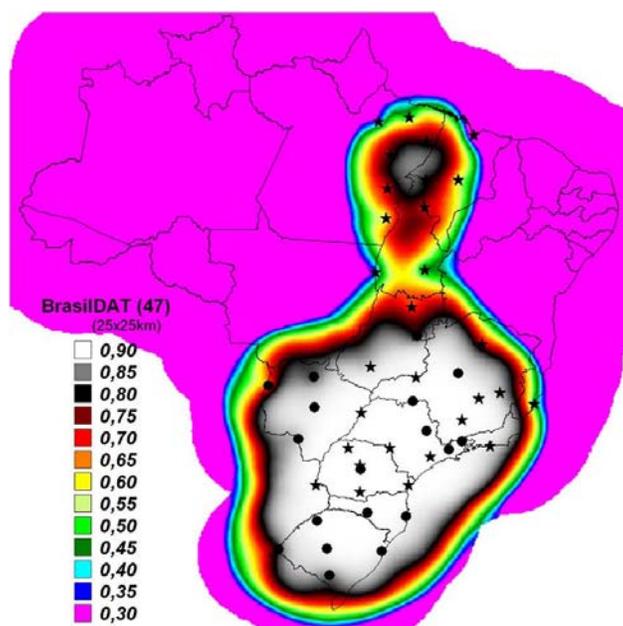


Figura 04 – Mapa de eficiência de detecção da BrasilDAT.

FONTE: Naccarato e Pinto Jr. (2009).

O modelo WRF foi configurado com dois domínios de integração, conforme mostra a Figura 05: o domínio d01 com resolução horizontal de 30 km (35° S a 14° S de latitude; 62° O a 34° O de longitude) e o domínio d02, aninhado ao domínio maior, com resolução horizontal de 10 km (25° S a 21° S de latitude;

50° O a 44° O de longitude). Simulações numéricas de 21 horas foram realizadas para cada dia, empregando os prognósticos numéricos das previsões globais a cada três horas com resolução de 1 grau do sistema de previsão global NCEP/GFS. Somente os parâmetros meteorológicos obtidos do domínio de alta resolução (d02), cobrindo parte do estado de São Paulo, sul de Minas Gerais e extremo nordeste do Paraná, foram de fato utilizados para a previsão das descargas atmosféricas.

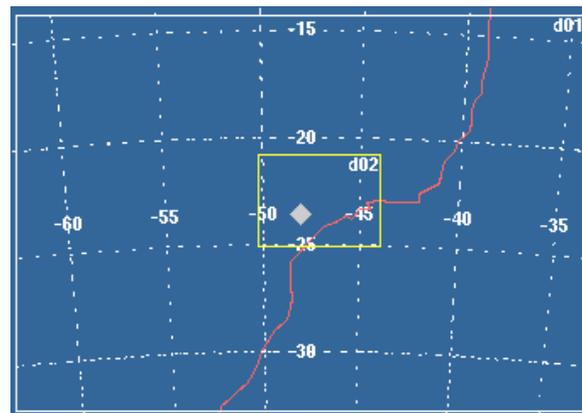


Figura 05 – Domínios de integração (d01, d02) do modelo WRF.

A Figura 06 ilustra a variação diurna (com a hora do dia) do número de raios (soma total do número de eventos) separados por polaridade para 6 anos de dados (1999-2004) na região Sudeste, mostrando o já conhecido pico de atividade no período da tarde (acompanhando o ciclo diário de insolação), por volta das 15-16h LT, horário em que se cria condições propícias de instabilidade na atmosfera (NACCARATO, 2006). É possível identificar que o máximo de raios negativos ocorre 1-2h antes do máximo de positivos. Isso é condizente com o fato dos raios positivos ocorrerem predominantemente na fase final das tempestades convectivas, cuja duração média é cerca de 1h (MACGORMAN E RUST, 1998; RAKOV E UMAN, 2003). Portanto, neste trabalho, o período da tarde é considerado como aquele mais importante no

reconhecimento de padrões atmosféricos relevantes à ocorrência de atividade elétrica.

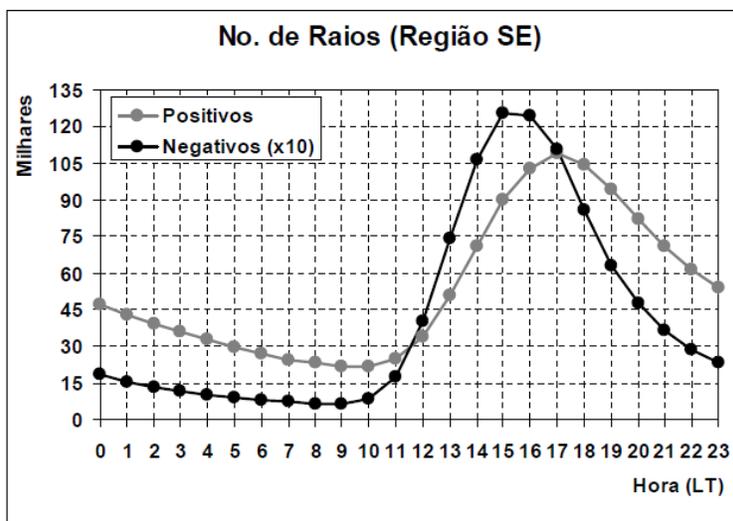


Figura 06 – Variação diurna do número total de raios negativos e positivos na região Sudeste para o período de 1999-2004.

FONTE: Naccarato (2006).

### 2.3.2 Rede Neural Artificial

A idéia deste trabalho é, a partir de determinadas variáveis meteorológicas simuladas no modelo WRF, alimentar a entrada de uma rede neural artificial, obtendo como saída a ocorrência de descargas atmosféricas para um futuro próximo. Neste caso, a RNA pode ser vista como uma função de transferência, relacionando algumas entradas com uma saída. A dinâmica real da atmosfera que relaciona as entradas escolhidas com a saída é bastante complexa e, evidentemente, desconhecida. O esforço consiste em justamente fazer com que uma RNA aprenda esta dinâmica, de forma a prever como a atmosfera comportar-se-á um tempo depois.

Para uma RNA ter um caráter preditivo com real aplicabilidade prática, parece razoável assumir que o simulador é confiável ou, ao menos, confere resultados numéricos com erros que se mantêm dentro de limites aceitáveis, e que

diferentes valores numéricos para um conjunto de variáveis de entrada induzem diferentes saídas. Outra importante condição é que as variáveis escolhidas para servirem de entrada no sistema neural de previsão devam ser de fato representativas do fenômeno.

A arquitetura de rede adotada foi uma rede *backpropagation (multilayer, feedforward, fully connected)*, sendo utilizados o *backpropagation* com momentum (BEALE E JACKSON, 1990; FAUSETT, 1994) como regra de treinamento e uma função de ativação sigmoïdal com algoritmo genético.

O conjunto de entrada da RNA foi composto pelo número de raios que aconteceu no período da manhã de cada dia e campos meteorológicos de CAPE (Energia Potencial Convectiva Disponível) (MONCRIEFF E GREEN, 1972), CINE (Energia de Inibição Convectiva) (COLBY, 1984), ILEV (Índice de Instabilidade por Levantamento) (GALWAY, 1956) e umidade relativa, obtidos a partir do modelo WRF para o período da tarde destes mesmos dias. A única camada escondida apresentou quatro neurônios e a camada de saída, apenas um, representando a quantidade de descargas atmosféricas que ocorreu no período da tarde para cada dia. O treinamento da rede deu-se de forma dinâmica nos três testes realizados com diferentes conjuntos de treino (outubro de 2005, outubro e novembro de 2005, outubro a dezembro de 2005). Por exemplo, considerando inicialmente o mês completo de outubro de 2005 para prever o primeiro dia de novembro de 2005, o conjunto de treino sempre se manteve com 31 dias, com os dias previstos seguintes substituindo os dias treinados anteriores. Da mesma forma ocorreu com os demais treinamentos de dois e três meses, que se mantiveram sempre com 61 e 92 dias, respectivamente.

A seguir serão apresentados os resultados preliminares das RNAs treinadas. O comportamento previsto das descargas atmosféricas para a tarde de cada dia do período de estudo foi obtido na forma de um índice, classificado da seguinte maneira: 0 para baixa, 1 para média e 2 para alta atividade de raios.

### **3 CONCLUSÕES**

Dependendo da quantidade de casos apresentada à rede como padrões de treinamento, diferentes foram os conjuntos que restaram para testar a eficiência do previsor neural de descargas atmosféricas. A melhor taxa de acerto, 64%, foi alcançada quando utilizado como conjunto de treino os meses de outubro a dezembro de 2005. Praticamente insignificantes foram as diferenças entre as taxas de acerto da RNA quando treinada com um e dois meses: 60 e 63%, respectivamente. Os resultados iniciais não se mostraram satisfatórios, indicando a necessidade de dar seqüência ao estudo proposto.

Fundamentalmente, seriam necessários mais dados de descargas atmosféricas de outros anos aqui não analisados, a fim de ampliar a quantidade de padrões apresentados à RNA na fase de treinamento e de aumentar o número de testes com casos não treinados. Tratando-se do conjunto de entradas, diferentes variáveis meteorológicas que possivelmente estabelecessem melhores correlações com a ocorrência das descargas atmosféricas poderiam ser avaliadas em um estudo futuro, assim como outro tipo de discretização para a quantidade de raios.

#### 4 REFERÊNCIAS BIBLIOGRÁFICAS

BALLAROTTI, M. G.; SABA, M. M. F.; PINTO JÚNIOR, O. A new performance evaluation of the Brazilian Lightning Location System (RINDAT) based on high-speed camera observations of natural negative ground flashes. In: INTERNATIONAL LIGHTNING DETECTION CONFERENCE, 19. (ILDC), , Tucson, Arizona. **Proceedings...** 2006. On-line. (INPE-14369-PRE/9456). Disponível em: <<http://urlib.net/sid.inpe.br/mtc-m17@80/2006/11.29.16.09>>. Acesso em: 15 jul. 2009.

BEALE, R.; JACKSON, T. **Neural computing: an introduction**. Massachusetts: MIT Press, 1990.

BERSON, A.; SMITH, S. J. **Data warehousing, data mining and OLAP**. McGraw-Hill, 1997. 612 p.

COLBY, JR., F. P. Convective inhibition as a predictor of convection during AVE-SESAME II. **Mon. Weather Rev.**, v. 112, 2239-2252 p., 1984.

CUMMINS, K. L.; MURPHY, M. J.; BARDO, E. A.; HISCOX, W. L.; PYLE, R. B.; PIFE, A. E. A combined TOA/MDF technology upgrade of the U.S. National Lightning Detection Network. **Journal of Geophysical Research**, v. 103, n. D8, 9035-9044 p., 1998.

FAUSETT, L. **Fundamentals of neural networks: Architecture, Algorithms and Applications**. 1<sup>st</sup> ed. Upper Saddle River: Prentice-Hall, 1994. 461 p. ISBN 0-133-34186-0.

FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; UTHURUSAMY, R. **Advances in knowledge discovery and data mining**. 1. ed. MIT Press, 1996. 611 p.

FRANKEL, D.; SHILLER, I.; DRAPER, J. S.; BARNES, A. A. JR. Use of neural network to predict lightning at Kennedy Space Center. In: SEATTLE INTERNATIONAL JOINT CONFERENCE, 1991, Seattle. **Electronic Proceedings...** Seattle: IEEE, 1991. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=00155197>>. Acesso em: 15 jul 2009.

FRISBIE, P. R.; COLTON, J. D.; PRINGLE, J. R.; DANIELS, J. A.; RAMEY JR., J. D.; MEYERS, M. P. Lightning prediction by WFO Grand Junction using Model Data and Graphical Forecast Editor Smart Tools. In: CONFERENCE ON THE METEOROLOGICAL APPLICATIONS OF LIGHTNING DATA, 4., 2009, Phoenix. **Electronic Proceedings...** Phoenix: MAS, 2009. Disponível em:

[http://ams.confex.com/ams/89annual/techprogram/paper\\_149101.htm](http://ams.confex.com/ams/89annual/techprogram/paper_149101.htm).

Acesso em: 15 jul 2009.

GALWAY, J. G. The lifted index as a predictor of latent instability. **Bull. Amer. Meteor. Soc.**, v. 43, 528-529 p., 1956.

MACGORMAN, D. R.; RUST, W. D. **The electrical nature of storms**. New York: Oxford University Press, 1998. 422 p.

MONCRIEFF, M. W.; GREEN, J. S. A. The propagation of steady convective overturning in shear. **Quart. J. R. Met. Soc.**, v. 98, 336-352 p., 1972.

NACCARATO, K. P. **Análise das características dos relâmpagos na região sudeste do Brasil**. 2005. 362 p. (INPE-14083-TDI/1069). Tese (Doutorado em Geofísica Espacial) – INPE, São José dos Campos, 2006.

NACCARATO, K. P.; PINTO JR, O. Improvements in the detection efficiency model for the Brazilian lightning detection network BrasilDAT. **Atmospheric Research**, n. 91, v. 2, 546-563 p., 2009.

NAGAE, Y.; OKUMURA, K; SUZUKI, T.; KAWAMURA, T.; MIYAKE, Y.; TAKAHASHI, S. Prediction of lightning activities by using fuzzy-neural network. In: INTERNATIONAL CONFERENCE ON LIGHTNING PROTECTION. **Proceedings...** Patras:University of Patras, 2000. p. 155-160.

NASCIMENTO, E. L. Previsão de tempestades severas utilizando-se parâmetros convectivos e modelos de mesoescala: uma estratégia operacional adotável no Brasil? **Revista Brasileira de Meteorologia**, v. 20, n. 1, 121-140 p., 2005.

PINTO JR., O. **A arte da guerra contra os raios**. São Paulo: Oficina de textos, 2005. 80 p. ISBN 85-86238-40-6.

PINTO JR, O.; PINTO, I. R. C. A. **Tempestades e relâmpagos no Brasil**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 2000. 196 p., 306366 cm. ISBN 85-901719-1-4.

PINTO JR, O.; PINTO, I. R. C. A. **Relâmpagos**. São Paulo: Brasiliense, 2008. 110 p. ISBN 978-85-11-00112-9.

PINTO JR., O.; NACCARATO, K. P.; PINTO, I. R. C. A.; FERNANDES, W. A.; NETO, O. P. Monthly distribution of cloud-to-ground lightning flashes as observed by lightning location systems. **Geophysical Research Letters**, v. 33, L09811, doi:10.1029/2006GL026081, 2006.

PINTO JR., O.; PINTO, I. R. C. A.; NACCARATO, K. P. Maximum cloud-to-ground lightning flash densities observed by lightning location systems in the tropical region: A review. **Atmospheric Research**, v. 84, 189-200 p., 2007.

RAKOV, V. A.; UMAN, M. A. **Lightning: Physics and Effects**. Cambridge: Cambridge University, 2003. 687 p. ISBN 0-521-58327-6.  
Electrodynamics. H. Volland, CRC Press, v. 1, p. 61-92, 1995.

SKAMAROCK, W. C.; KLEMP, J. B.; DUDHIA, J.; GILL, D. O.; BARKER, D. M.; WANG, W.; POWERS, J. G. **A Description of the Advanced Research WRF Version 2**. NCAR/Tech. Notes, 2005.

SUMATHI, S.; SIVANANDAM, S. N. **Introduction to data mining and its applications**. Studies in Computational Intelligence. 1<sup>st</sup> ed. Springer, 2006. 828 p.

UMAN, M. A. **The lightning discharge**. 1<sup>a</sup> ed. New York: Academy Press, 1987. 377 p.

WANG, W., BARKER, D., BRUYERE, C., DUDHIA, J., GILL, D., MICHALAKES, J. **WRF Modeling System User's Guide**. MMM Division/NCAR/Tech. Notes, 2004.

YAMASAKI, Y; MACHADO, A. M. S. P. T.; PICADO, A. T. S.; ORGAZ, M. L. D. M. (2006) Diagnósticos de Previsões Numéricas do WRF. In: Jornadas Científicas de la AME, 29/ Encuentro Hispano-Luso, 7, 2006. Disponível em: <<http://www.ame-web.org/JORNADAS/P14-trabajo%20Yamasaki%20et%20al-1.pdf>>. Acesso em: 15 jul 2009.

ZEPKA, G. S. **Estudo para o desenvolvimento de um previsor de descargas elétricas atmosféricas aplicado à região costeira do estado do Rio de Janeiro**. 2005. 149 p. Dissertação (Mestrado em Engenharia Oceânica) – Fundação Universidade Federal do Rio Grande, Rio Grande, 2005.