

PREVISÃO DA POSIÇÃO ESPACIAL DE TEMPESTADES ELÉTRICAS UTILIZANDO MLP

Mariana Kleina

Programa de Pós Graduação em Métodos Numéricos em Engenharia, UFPR
Av. Cel. Francisco H. dos Santos, Centro Politécnico, C.P. 19011, CEP 81531-980, Curitiba/PR
marianakleina1@gmail.com

Luiz Carlos Matioli

Programa de Pós Graduação em Métodos Numéricos em Engenharia, UFPR
Av. Cel. Francisco H. dos Santos, Centro Politécnico, C.P. 19011, CEP 81531-980, Curitiba/PR
matioli@ufpr.br

Eduardo Alvim Leite

Sistema Meteorológico do Paraná (SIMEPAR)
Av. Cel. Francisco H. dos Santos, Centro Politécnico, C.P. 19100, CEP 81531-980, Curitiba/PR
alvim@simepar.br

Alana Renata Ribeiro

Programa de Pós Graduação em Métodos Numéricos em Engenharia, UFPR
Av. Cel. Francisco H. dos Santos, Centro Politécnico, C.P. 19011, CEP 81531-980, Curitiba/PR
alanar89@gmail.com

RESUMO

Descargas atmosféricas são fenômenos naturais responsáveis por diversos danos terrestres, tais como incêndios, perturbações na rede elétrica e até morte de pessoas. Quando organizadas dão origem as chamadas tempestades elétricas. Prever, mesmo que a curto prazo, para onde estas tempestades se deslocam, pode auxiliar em tomadas de decisões a fim de minimizar danos nos mais diversos setores. Para tal previsão do deslocamento do núcleo de tempestades elétricas é utilizada a rede neural Perceptron Multi-Camadas (MLP), tendo como preditores as posições espaciais dos centros das tempestades nos últimos trinta minutos. Os resultados indicam que a metodologia empregada é promissora, encorajando futuramente a previsão de outras variáveis das tempestades elétricas.

PALAVRAS CHAVE. Tempestades Elétricas, Previsão a Curto Prazo, Perceptron Multi-Camadas.

Área Principal: MH, PM, OA.

ABSTRACT

Lightning strokes are natural events responsible for several land damage, such as fire, disturbances of the electrical network and even death of people. When organized, they generate electrical storms. To predict, even in the short term, where are these storms moving to, may help in decision making in order to minimize damage in various sectors. For such prediction of the electrical storms centers displacement, the Multi-Layer Perceptron (MLP) neural network is used, having as predictors the spatial positions of the storms centers on the last thirty minutes. The

results indicate that the methodology used is promising, encouraging forecast of other variables of electrical storms on the future.

KEYWORDS. Electrical Storms, Short Term Forecast, Multi-Layer Perceptron.

Main Area: MH, PM, OA.



1. Introdução

Descargas atmosféricas despertam interesse de estudo nas mais diversas áreas pois envolvem muitos riscos ao planeta, tais como: mortes, danos ambientais e inúmeros prejuízos materiais. No Brasil esse interesse é ainda maior devido a alta incidência de descargas, em virtude de suas grandes dimensões e seu clima predominantemente tropical. Anualmente 57 milhões de descargas atingem o território brasileiro, causando um prejuízo na ordem de 1 bilhão de reais para a economia do país, além de 100 vítimas fatais, segundo o INPE (2015). É de grande valia saber se, em determinada localidade e em um tempo específico, descargas atmosféricas são capazes de provocar qualquer tipo de ameaça, podendo serem tomadas medidas preventivas a fim de minimizar danos maiores.

Tempestade elétrica, assim denominada nesta pesquisa, é um evento meteorológico formado por um conjunto de descargas atmosféricas organizadas espacial e temporalmente. Rastrear e prever o comportamento destas tempestades utilizando informações de descargas é o principal objetivo desta pesquisa.

Nos últimos anos, diversos trabalhos com a finalidade de monitorar e prever tempestades utilizando diversas variáveis meteorológicas foram desenvolvidos. O algoritmo TITAN (*Thunderstorm Identification, Tracking, Analysis, and Nowcasting*) utiliza dados de radar para identificar, rastrear e prever a curto prazo células de tempestades (Dixon e Wiener, 1993). Nos trabalhos de Steinacker *et al.* (2000) e Bonelli e Marcacci (2008) são utilizados dados de radar em conjunto com descargas atmosféricas para melhorar a qualidade dos resultados, visto que é sugerido forte correlação entre descargas e precipitação (Strauss *et al.*, 2013). No trabalho de Betz *et al.* (2008), somente dados de descargas atmosféricas são usados para identificar células de tempestades.

Na maioria dos trabalhos de monitoramento (os citados anteriormente são exemplos) células de tempestades são identificadas por regiões de alta densidade de variáveis, como descargas ou taxa de precipitação. Em relação à previsão, esta é geralmente feita utilizando extrapolação de dados, a qual tem bom desempenho se a tempestade é bem comportada. Nesta pesquisa, a identificação e rastreamento de tempestades elétricas é feita por meio clusterização de descargas e a previsão da posição do centro da tempestade é realizada através da aplicação de redes neurais.

2. Região de Estudo e Dados

A região selecionada para o estudo foi a área que abrange a linha de transmissão que leva a energia produzida na hidrelétrica de Itaipu para a proximidade do centro de consumo da região Sudeste do Brasil, conhecida como LT 765 kV (Figura 1). Foi escolhida como região piloto para esta pesquisa devido a grande vulnerabilidade desta linha às condições climáticas do Brasil, visto que cerca de 70% dos desligamentos na transmissão e 40% na distribuição de energia são ocasionados por descargas (INPE,2015). Este sistema é composto de três linhas de transmissão entre as subestações de Foz do Iguaçu (PR) e Tijuco Preto (SP), cada uma com extensão de aproximadamente 900 quilômetros. É o sistema de transmissão de tensão mais elevada existente no Brasil, desde que começou sua operação em 1986 (Itaipu, 2015).

De 1999 a 2012 foram registradas 29344598 descargas nuvem-solo na região que engloba a linha LT 765 kV, provenientes da Rede Integrada Nacional de Detecção de Descargas Atmosféricas (RINDAT), mantidos pelo SIMEPAR em cooperação com Furnas, CEMIG e INPE. Estes registros compreendem a posição geográfica (longitude/latitude) e data das descargas atmosféricas.

3. Metodologia

Esta pesquisa se divide em duas etapas primordiais, com metodologias diferentes: a primeira etapa consiste na identificação e rastreamento das tempestades elétricas que será realizada através da clusterização de descargas atmosféricas com o método *Convergent Data Sharpening*, já a segunda etapa consiste em realizar a previsão dos núcleos ou centros das tempestades monitoradas, aplicando-se a rede neural supervisionada MLP. A seguir, as duas metodologias com suas respectivas formas de aplicação serão apresentadas.

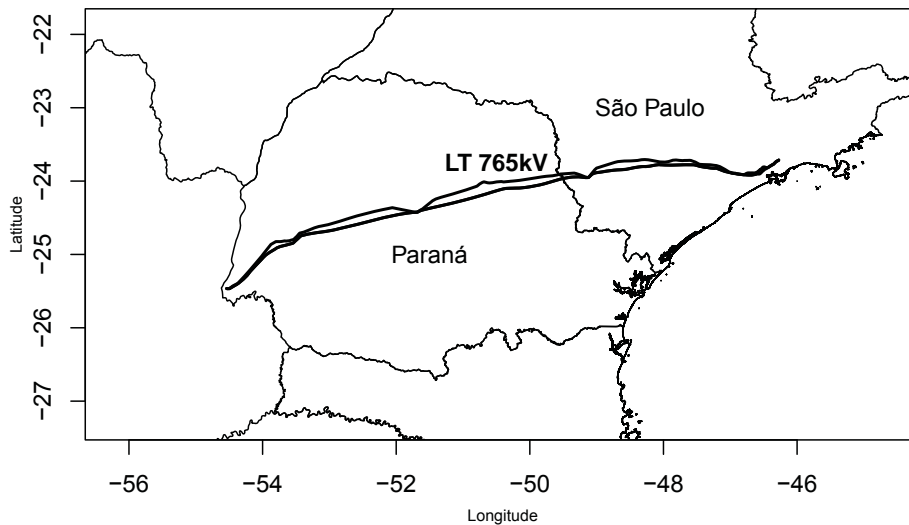


Figura 1: Região de estudo que abrange a linha LT 765 kV.

3.1. Identificação e Monitoramento das Tempestades Elétricas

3.1.1. Método de Clusterização: *Convergente Data Sharpening*

O método *Data Sharpening* foi originalmente introduzido por Choi e Hall (1999) com o intuito de reduzir os vies na estimação de densidades. Tendo os dados originais x_1, x_2, \dots, x_n , uma estimativa para a densidade desconhecida é dada por

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n k_h(x_i - x) \quad (1)$$

em que k_h representa uma função densidade de probabilidade simétrica, conhecida com função *kernel*, com parâmetro h chamado de largura da banda.

Por regressão constante local de x em x , Choi e Hall (1999) obtiveram os dados ajustados $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n$, onde

$$\hat{x}_j = \frac{\sum_{i=1}^n k_h(x_i - x_j) x_i}{\sum_{i=1}^n k_h(x_i - x_j)}$$

e assim a densidade estimada é feita sobre os dados ajustados, isto é, substituir x_i na equação 1 por \hat{x}_i , onde esse ajuste gera menor viés do que $\hat{f}(x)$.

A equação 1 tende a subestimar densidades nos picos e superestimar nos vales (Woolford e Braun, 2006), e este ajuste faz com que cada dado original se mova para mais perto de modos locais. Woolford e Braun (2006) propuseram iterar o algoritmo de perturbação dos dados com o objetivo de indicar os modos locais encontrados como sendo os centros de clusters de dados com aderência a eles. Os autores citados anteriormente provam que para h fixo e qualquer vetor inicial de observações x_0 , o algoritmo *Data Sharpening* de Choi e Hall (1999) converge para um único vetor \hat{x} . A este método deu-se o nome de *Convergent Data Sharpening*. Diversas funções *kernel* podem ser utilizadas, porém a mais comum é a função gaussiana, a qual será utilizada nesta pesquisa.

Portanto, o método *Convergent Data Sharpening* será utilizado nesta pesquisa como método de clusterização de descargas atmosféricas (temporal e espacialmente) para identificar as tempestades elétricas que se desenvolveram na região e período estudados. Este método de clusterização tem a vantagem de não ser necessário a informação do número de clusters a priori, porém este número é altamente influenciado pela escolha do parâmetro h .

O *software* utilizado nesta pesquisa foi o R, o qual possui implementado o método *Convergent Data Sharpening* através da função *sharp2dB* do pacote *CHsharp* (R, 2012).

3.1.2. Aplicação

O reconhecimento das tempestades elétricas foi realizado por meio da clusterização das descargas pela posição espacial (longitude/latitude) através do método *Convergent Data Sharpening* onde o conjunto de descargas no período de uma hora é selecionado e clusterizado. Após a clusterização desse conjunto, a nova janela temporal considerada é a anterior com dez minutos eliminados no início e dez minutos adicionados no final. O centro de cada cluster é calculado como a média das posições das descargas que o compõe. O valor do parâmetro h foi escolhido por meio de testes experimentais e inspeção visual, sendo utilizado o valor de 0.1° (cerca de 10 quilômetros).

Assim, retratos temporais são criados, onde uma mesma descarga encontra-se em seis clusters diferentes. Então é preciso fazer a conexão entre estes retratos temporais para que sejam identificados como tempestades elétricas. Clusters de janelas subsequentes são unidos se a velocidade de deslocamento de um núcleo para outro não ultrapassa um limite máximo, estipulado em 40 km/10min. Com isso, o processo de identificação e rastreamento das tempestades elétricas é finalizado.

Com os 14 anos de dados na região piloto, foram identificadas 636012 tempestades elétricas. A figura 2 ilustra trajetórias de tempestades: 54 tempestades elétricas identificadas das 17:00 às 19:00 horas no dia 06/01/2011 (Figura 2a) e imagem ampliada do trajeto de uma tempestade (Figura 2b).

3.2. Previsão da Posição Espacial das Tempestades Elétricas

3.2.1. Perceptron Multi-Camadas

As redes Perceptron Multi-Camadas consistem de um conjunto de Perceptrons arranjados em diversas camadas e operam com treinamento supervisionado e um algoritmo de retropropagação (*backpropagation*) que se baseia na regra de aprendizagem pela correção do erro. A rede consiste em um passo para frente (a propagação) e um passo para trás (a retropropagação). Inicialmente um padrão de entrada é aplicado aos neurônios da rede e seu efeito se propaga pelas camadas. Então, a saída é produzida com a resposta real da rede. Na propagação, os pesos não se alteram e durante a retropropagação os pesos são ajustados de acordo com a regra de correção do erro. De forma específica, a resposta da rede é subtraída da resposta desejada para produzir um sinal de erro. Portanto, na retropropagação os pesos são ajustados fazendo com que a resposta real da rede se aproxime da resposta desejada.

A figura 3 representa a arquitetura de uma rede MLP em que um neurônio qualquer está ligado a todos os neurônios da camada anterior, isto é, a rede é totalmente conectada. A seguir é apresentado um algoritmo genérico da rede neural MLP.

Algoritmo:

Passo 1: Inicialize aleatoriamente os pesos da rede neural com valores pequenos.

Passo 2: Apresente um padrão para a camada de entrada da rede neural.

Passo 3: Propague o padrão de entrada pelas camadas ocultas da rede neural para assim calcular a saída da sua função de ativação.

Passo 4: A diferença entre as saídas desejada e da função de ativação será usada para calcular o erro de ativação da rede.

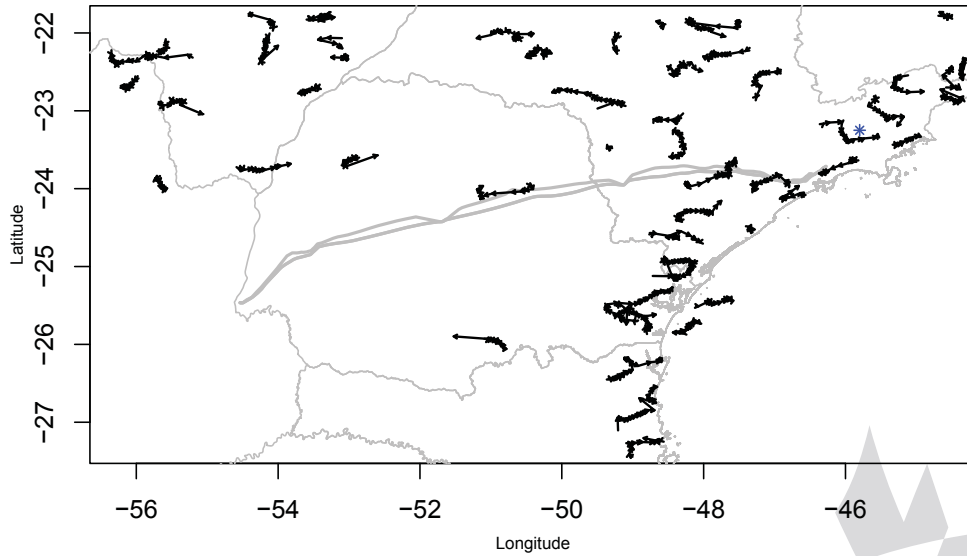
Passo 5: Ajuste os pesos por meio do neurônio de saída da rede a fim de diminuir o erro de ativação para o padrão de entrada apresentado.

Passo 6: Propague o valor do erro de volta (retropropagação) para cada neurônio da camada oculta, considerando a proporção de sua contribuição no cálculo do erro de ativação da rede.

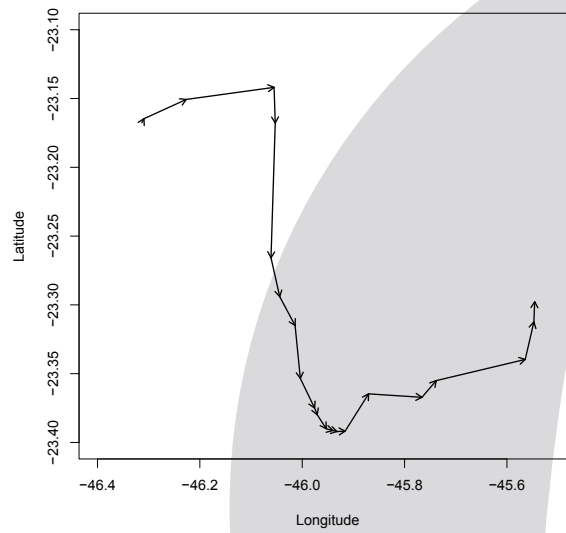
Passo 7: Ajuste os pesos, propagando o erro pelos neurônios da camada oculta a fim de reduzir a sua colaboração para o erro calculado para o padrão de entrada apresentado à rede.

Passo 8: Repetir os passos 2 a 7 para cada padrão de entrada do conjunto de dados apresentado à rede.

Passo 9: Repetir o passo 8 até que a rede esteja treinada.



(a) 54 tempestades elétricas.



(b) Caminho percorrido por uma tempestade.

Figura 2: A figura 2a ilustra as trajetórias de 54 tempestades elétricas identificadas no dia 06/01/2011 das 17:00 às 19:00 horas; a figura 2b é a aproximação da imagem para a tempestade marcada com o asterisco na figura 2a, ilustrando a trajetória desta tempestade que durou 200 minutos.

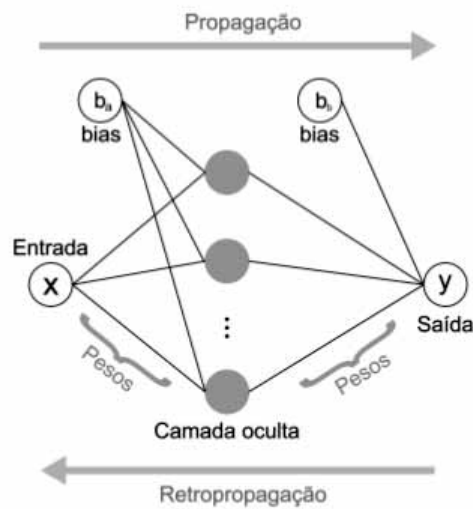


Figura 3: Arquitetura de uma rede MLP totalmente conectada, com uma camada oculta e uma camada de saída. Fonte: Santos (2014).

Todo o embasamento matemático sobre a rede MLP pode ser encontrado em Haykin (2001), Siqueira (2014) e Santos (2014). No *software* R, a rede MLP está disponível na biblioteca *RSNNS*.

3.2.2. Aplicação

A rede neural MLP será utilizada para prever a posição das tempestades elétricas (identificadas e rastreadas na etapa descrita anteriormente) uma hora à frente, dado que a tempestade tenha um passado mínimo. Como entrada para a MLP foram criados os preditores da seguinte forma: as posições espaciais (longitude/latitude) nos tempos -30, -20, -10 e 0, e no tempo +60 para supervisão da rede. Portanto, para cada tempestade que durou no mínimo 90 minutos, foi criada a seguinte entrada para a rede:

$$lon[-30] \quad lat[-30] \quad lon[-20] \quad lat[-20] \quad lon[-10] \quad lat[-10] \quad lon[0] \quad lat[0] \quad lon[+60] \quad lat[+60]$$

onde os 8 primeiros elementos representam o passado ou histórico da tempestade, usados no aprendizado da rede e os 2 últimos elementos representam o futuro da tempestade, os quais a rede utilizará como resposta. Para uma tempestade que durou mais que 90 minutos, todas as possíveis entradas para a rede foram criadas. Por exemplo, se uma tempestade durou 100 minutos, isto é:

$$\begin{pmatrix} lon \\ lat \end{pmatrix}_1 \quad \begin{pmatrix} lon \\ lat \end{pmatrix}_2 \quad \begin{pmatrix} lon \\ lat \end{pmatrix}_3 \quad \begin{pmatrix} lon \\ lat \end{pmatrix}_4 \quad \begin{pmatrix} lon \\ lat \end{pmatrix}_5 \quad \begin{pmatrix} lon \\ lat \end{pmatrix}_6 \quad \begin{pmatrix} lon \\ lat \end{pmatrix}_7 \quad \begin{pmatrix} lon \\ lat \end{pmatrix}_8 \quad \begin{pmatrix} lon \\ lat \end{pmatrix}_9 \quad \begin{pmatrix} lon \\ lat \end{pmatrix}_{10} \quad \begin{pmatrix} lon \\ lat \end{pmatrix}_{11}$$

onde 1 representa a origem e 11 o final da tempestade, duas entradas desta tempestade são formadas:

$$\begin{array}{ccccc} \overbrace{lon_1 \quad lat_1}^{-30} & \overbrace{lon_2 \quad lat_2}^{-20} & \overbrace{lon_3 \quad lat_3}^{-10} & \overbrace{lon_4 \quad lat_4}^0 & \overbrace{lon_{10} \quad lat_{10}}^{+60} \\ \overbrace{lon_2 \quad lat_2} & \overbrace{lon_3 \quad lat_3} & \overbrace{lon_4 \quad lat_4} & \overbrace{lon_5 \quad lat_5} & \overbrace{lon_{11} \quad lat_{11}} \end{array}$$

Com as 636012 tempestades elétricas identificadas e monitoradas na etapa anterior, foram geradas 1329240 entradas para a rede neural conforme descrito anteriormente. A desvantagem no processo de criação destes preditores é que somente tempestades com um histórico de no mínimo 100 minutos puderam ser utilizadas, assim tempestades com pouco passado não foram aproveitadas no aprendizado da rede.

Os dados antes de serem apresentados a rede foram normalizados pela regra Min-Max. As configurações da rede MLP para o problema de previsão da posição espacial de tempestades

elétricas foram: 8 neurônios na camada de entrada, uma camada oculta com 40 neurônios (número escolhido a partir de testes empíricos) e 2 neurônios na camada de saída; função de ativação logística para as camadas oculta e de saída.

Para aprendizado da rede foram utilizados 70% dos dados e para validação 25%. Os outros 5% de dados inexplorados foram usados para aplicação da rede já treinada e validada, cujos resultados serão utilizados para avaliar o desempenho da rede.

4. Resultados e Discussões

Nesta seção serão apresentados os resultados obtidos com a aplicação da metodologia descrita anteriormente, isto é, a previsão a curto prazo (uma hora à frente) do deslocamento de tempestades elétricas identificadas sobre a região da linha LT 765 kV utilizando a rede neural MLP.

Observando os resultados encontrados, nota-se a duas situações distintas (Figura 4): a superestimação (figura à direita) e a subestimação (figura à esquerda) da posição prevista em comparação com a posição real. As setas pretas representam o passado da tempestade que foi apresentado a rede treinada (a origem da primeira seta representa a posição da tempestade no tempo -30, a origem da segunda seta no tempo -20, a origem da terceira seta no tempo -10 e o final da terceira seta no tempo 0), a seta azul representa o deslocamento real da tempestade e a seta vermelha representa o deslocamento previsto pela MLP da posição da tempestade no tempo +60. Utilizando a técnica Viés (Santos, 2014), que indica a direção média dos desvios de um conjunto a partir de um conjunto de valores observados (valores positivos apontam superestimativas e negativos subestimativas), obteve-se o valor de 0.0051 (longitude e latitude consideradas conjuntamente), indicando uma discreta superestimação das previsões geradas pela rede.

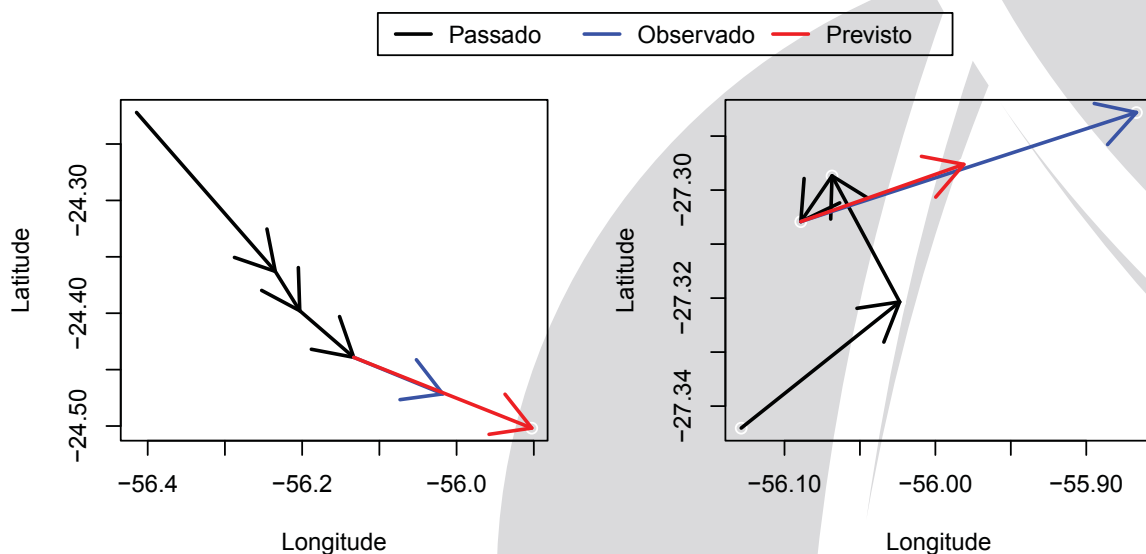


Figura 4: Superestimação e subestimação da previsão gerada pela rede.

A figura 5 apresenta alguns resultados em que a rede conseguiu prever bem a posição real da tempestade. Conforme esperado, o desempenho da rede é melhor nas situações em que o passado é bem comportado, isto é, trajetória aproximadamente retilínea.

Com o intuito de avaliar os resultados da rede neural, foram utilizados dois índices de comparação (Santos, 2014): o Coeficiente de Correlação de Pearson ρ , que mede a dependência linear entre os valores previsto e observado, e a Raiz do Erro Médio Quadrático (REM), que indica a magnitude média do erro.

Foram calculados dois coeficientes de correlação, um para longitude e outro para latitude, e transformando estes dois índices em apenas um, obteve-se o valor de 0.99, significando que os

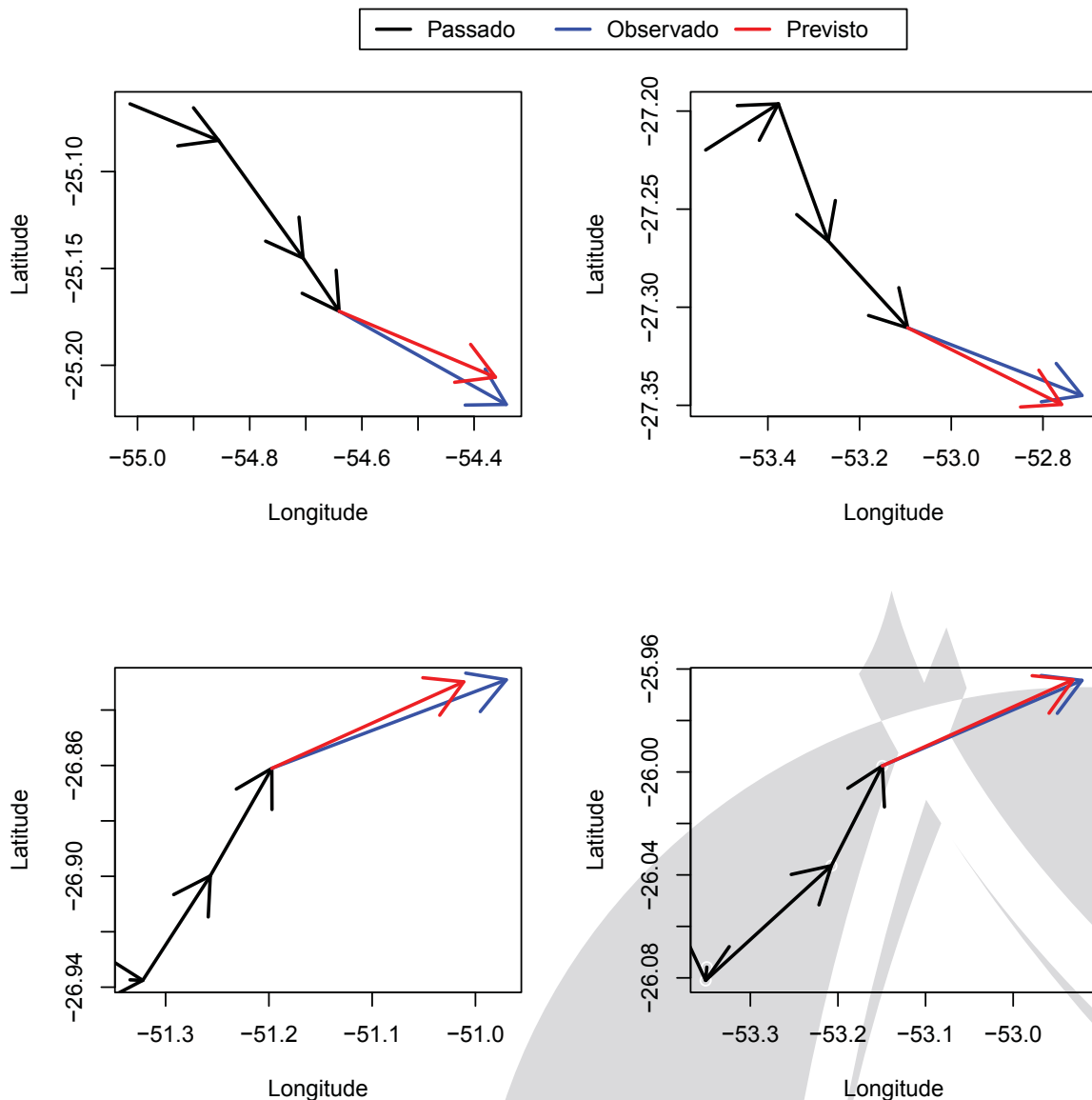


Figura 5: Exemplos do resultado da rede MLP na previsão da posição de tempestades elétricas.

valores previstos e observados são altamente correlacionados. Da mesma forma, o índice REMQ foi calculado, sendo encontrado o valor 0.18, significando quem, em média, o resultado previsto pela rede erra em 0.18° (ou equivalentemente 18 quilômetros) do real valor observado.

5. Conclusões

Nesta pesquisa fez-se uso de uma rede neural do tipo Perceptron Multi-Camadas, a fim de que esta fosse capaz de prever a posição uma hora à frente de tempestades elétricas, conhecido um passado de 30 minutos dela. Antes porém da aplicação da rede, fez uso de um método de clusterização para identificar e rastrear as trajetórias destas tempestades elétricas.

O desempenho da rede foi bastante satisfatório, visto que os índices de comparação entre dados previstos e observados mostraram a eficácia da MLP para o problema de previsão da posição espacial de tempestades elétricas. O coeficiente de correlação atingiu valor próximo do máximo e o REMQ mostrou que há um erro médio de 18 quilômetros aproximadamente na previsão. Tendo em vista a dificuldade deste tipo de previsão, é um erro considerado aceitável, considerando também

todas as incertezas do processo de identificação e monitoramento das tempestades.

Para treinamento, validação e análise de desempenho da rede, somente tempestades com duração de 90 minutos no mínimo foram utilizadas, visto que foram usados dados históricos nesta pesquisa. Porém, pensando na previsão das trajetórias de tempestades em tempo real, que é a ideia futura para continuação desta pesquisa, as tempestades precisam possuir esse passado de meia hora para aplicação da rede. Infelizmente, a rede neural apresentada não pode ser aplicada em tempestades que estão no seu estágio inicial, sem um histórico de 30 minutos. Contudo, tendo disponível algum histórico da trajetória da tempestade, outras técnicas de previsão poderão ser aplicadas, tais como a extrapolação de dados.

Em tempo real, tendo o conhecimento da posição de uma tempestade elétrica uma hora à frente na região da linha LT 765 kV, diversas medidas operacionais podem ser tomadas se uma tempestade está se aproximando de torres de transmissão ou subestações, por exemplo.

Analisando os resultados da previsão da posição espacial das tempestades elétricas pela rede neural, desperta-se um interesse em prever outras variáveis meteorológicas de tempestades elétricas, tais como o número de descargas e o pico de corrente médio, por exemplo.

Referências

Betz, H. D., Schmidt, K., Oettinger, W. P. e Montag, B. (2008), Cell-tracking with lightning data from LINET, *Advances in Geosciences*, 17, 55-61.

Bonelli, P. e Marcacci, P. (2008), Thunderstorm nowcasting by means of lightning and radar data: algorithms and application in northern Italy, *Natural Hazards and Earth System Sciences*, 8, 1187-1198.

Choi, E. e Hall, P. (1999), Data sharpening as a prelude to density estimation, *Biometrika*, 86, 941-947.

Dixon, M. e Wiener, G. (1993), TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting — A Radar-based Methodology, *Journal of Atmospheric and Oceanic Technology*, 10, 785-797.

Haykin, S., *Redes Neurais: princípios e prática*, Bookman, Porto Alegre, 2001.

INPE – Instituto Nacional de Pesquisas Espaciais (2014), *Sistema Elétrico*, disponível em: <http://www.inpe.br/webelat/homepage/menu/infor/relampagos.e.efeitos/sistema.eletrico.php>, acessado em: 15/12/2014.

INPE – Instituto Nacional de Pesquisas Espaciais (2015), *Brasil, “país dos raios”*, disponível em: http://www.inpe.br/webelat/elatMidia/arquivos/20130130173223Ambiente%20Energia_30%2001%202013.pdf, acessado em: 10/04/2015.

ITAIPU (2015), *Integração ao Sistema Brasileiro*, disponível em:

<http://www.itaipu.gov.br/energia/integracao-ao-sistema-brasileiro>, acessado em: 20/02/2015.

R Core Team (2012), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

Santos, T. N., *Redes Neurais Artificiais e Relação ZR Aplicadas à Estimativa de Chuva*, Dissertação de Mestrado – Universidade Federal do Paraná, 2014.

Siqueira, P. H. (2014), *Metaheurísticas e Aplicações*, Notas de aula, Universidade Federal do Paraná.

Steinacker, R., Dorninger, M., Wolfelmaier, F. e Krennert, T. (2000), Automatic Tracking of Convective Cells and Cell Complexes from Lightning and Radar Data, *Meteorology and Atmospheric Physics*, 72, 101-110.

Strauss, C., Rosa, M. B. e Stephany, S. (2013), Spatio-temporal clustering and density estimation of lightning data for tracking of convective events, *Atmospheric Research*, 134, 87-99.

Woolford, D. G. e Braun, W. J. (2006), Convergent data sharpening for identification and tracking of spatial and temporal centers of lightning activity, *Environmetrics*, 18, 461-479.