

Revista Brasileira de Geografia Física



ISSN:1984-2295

Homepage: www.ufpe.br/rbgfe

Redes Neurais Artificiais na Previsão de Queimadas e Incêndios no Pantanal

Hevelyne Henn da Gama Viganó¹, Celso Correia de Souza², Marcia Ferreira Cristaldo³, Leandro de Jesus⁴

¹ Doutoranda em Meio Ambiente e Desenvolvimento Regional, Uniderp MS, Professora Titular Matemática e Estatística, Unidade Acadêmica Campo Grande, Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul, Sede Provisória Rua Treze de Maio, n. 3072, Centro, CEP 79002-356, Campo Grande, Mato Grosso do Sul. (67) 99320-4554. hevelyne.vigano@gmail.com, hevelyne.vigano@ifms.edu.br (autor correspondente). ² Professor Dr. Programa de Pós Graduação Stricto Sensu em Meio Ambiente e Desenvolvimento Regional, Uniderp MS, Rua Alexandre Herculano, n. 1400, Jardim Veraneio, CEP 79037-280, Campo Grande, Mato Grosso do Sul. (67) 3309-6565. csouza939@gmail.com. ³Doutoranda em Meio Ambiente e Desenvolvimento Regional, Uniderp MS, Professora Informática, Unidade Acadêmica Aquidauana, Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul. (87) 3240-1600. marcia.cristaldo@ifms.edu.br. ⁴Doutorando em Meio Ambiente e Desenvolvimento Regional, Uniderp MS, Professor Informática, Unidade Acadêmica Aquidauana, Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul, Rua José Tadao Arima , n. 222, Vila Ycaraí, CEP 79200-000, Aquidauana, Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul, Rua José Tadao Arima , n. 222, Vila Ycaraí, CEP 79200-000, Aquidauana, Mato Grosso do Sul. (67) 3240-1600. leandro.jesus@ifms.edu.br.

Artigo recebido em 05/04/2017 e aceito em 31/09/2017

RESUMO

O bioma pantaneiro é acometido anualmente por um grande número de queimadas e incêndios. Prever esses eventos é de suma importância, uma vez que, prejuízos à fauna e à flora poderiam ser minimizados e catástrofes evitadas. Uma intervenção imediata do poder público na mitigação desses eventos passa, essencialmente, pela previsão do número de focos e das áreas queimadas, e posteriormente, na localização desses focos e na identificação das áreas. Dados precisos sobre as variáveis ambientais, em tempo real, podem ser obtidos através do sensoriamento remoto, aliado aos sistemas de informações geográficas, às técnicas de inteligência artificial e estatística aplicada, favorecendo às tomadas de decisão na previsão. O objetivo deste estudo foi aplicar a técnica de Redes Neurais Artificiais (RNAs) para a previsão dos focos e das áreas queimadas no Pantanal Sul-Mato-Grossense. O centro de previsão de tempo e estudos climáticos do Instituto Nacional de Pesquisas Espaciais (INPE), bem como, o Instituto Nacional de Meteorologia (INMET) possuem bancos de dados do Pantanal, dos quais, as variáveis envolvidas nesse processo foram extraídas. Utilizando as RNAs do tipo *Multilayer Perceptron*, com algoritmo *Retropropagation* de aprendizagem foi possível prever o valor do número de focos com um ajuste de 84,8% utilizando um conjunto de variáveis meteorológicas como preditoras e, de 99,4% usando como preditora, somente a série temporal do número de focos. No entanto, o ajuste passou a ser 90,3% ao se realizar a previsão da área queimada, utilizando as mesmas variáveis, e de 98,6%, usando como preditora, somente os dados da área queimada.

Palavras-chave: previsão de fogo, monitoramento ambiental, inteligência artificial.

Artificial Neural Networks in Prediction of Forest Fires and Burns in the Pantanal

ABSTRACT

On a year basis the Pantanal biome is affected by a large number of fires and fire points. Predicting these events is of paramount importance, since damage to fauna and flora could be minimized and disasters might be avoided. Immediate intervention of the government in mitigating these events essentially depends on the identification and location of fire outbreaks. Accurate and reliable data on environmental variables in real time can be obtained by means of remote sensing coupled with geographic information systems, techniques of artificial intelligence and applied statistics, which would favor the decision-making in foreseeing fire foci and burned area. The aim of this study was to apply the technique of artificial neural networks to predict the fire foci and burned areas in the Pantanal Sul-Mato-Grossense. The center for weather forecast and climate studies of the National Institute for Space Research (INPE) and the National Institute of Meteorology (INMET) afford meteorological database of the Pantanal region, from which the environmental variables involved in this process were extracted. By using Multilayer Perceptron Artificial Neural Networks with Backpropagation algorithm, it was possible to predict the value of the number of fire foci with an adjustment of 84,8% using a set of meteorological variables as predictors and of 99,4% using as a predictor only the time series of the number fire foci. However, the adjustment became 90,3% when the forecast of the burned area, using the same meteorological variables, and 98,6%, using as a predictor, only the time series of the burned area.

Keywords: fire forecasting, environmental monitoring, artificial intelligence.

Introdução

Realizar previsões sobre focos de queimadas e incêndios em tempo real é uma atividade complexa, que depende de um conjunto de informações de satélites, de análise de dados históricos, de estudos sobre a vegetação e composição do solo e de variáveis ambientais. No entanto, as pesquisas têm demonstrado a necessidade de sistemas integrados em tempo real para a geração simultânea de alertas e de ações de combate ao fogo e para que os prejuízos causados à biodiversidade, ao ciclo hidrológico, ao ciclo do carbono na atmosfera, dentre outros danos, sejam minimizados (Artaxo, et al., 2006; Granemann e Carneiro, 2009).

O Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA), possui sob sua competência o Sistema Nacional de Prevenção e Combate aos Incêndios e Queimadas (PREVFOGO), que realiza ações de uso controlado dos recursos naturais renováveis por meio de organizações, implementações operacionalizações de atividades relacionadas à pesquisa, educação, prevenção, controle e combate aos incêndios florestais e queimadas (Ibama, 2009). Queimada pode ser tratada, como prática agropastoril ou florestal, onde o fogo é utilizado de forma controlada, atuando como um fator de produção. Incêndio florestal é todo fogo sem controle que incide sobre qualquer forma de vegetação, podendo ocorrer por fontes naturais ou por negligência humana (Anderson et al., 2005).

A predição de incêndios florestais e queimadas pode ajudar na prevenção evitando possíveis catástrofes. Diferentes modelos têm sido comumente aplicados na predição desses eventos, com a utilização de variáveis meteorológicas e de séries históricas de dados sobre queimadas e incêndios florestais, que podem ser descritos por séries temporais, modelos de regressões lineares (RL), modelos auto-regressivos de médias móveis integradas (ARIMA) e as redes neurais artificiais (RNA).

Em 2005 foi utilizado um modelo de regressão linear (RL) para estimar dados de temperaturas mensais no Estado de Alagoas, para identificar as possíveis áreas susceptíveis à desertificação (Gois et al., 2005).

Almorox et al. (2008) aplicaram uma RL para estimar a radiação solar global de estações meteorológicas da Venezuela. E na Nigéria, segundo Falayi et al. (2008), foram utilizadas regressões lineares múltiplas (RLM) para prever índices de radiações solares globais, associando diferentes variáveis meteorológicas.

Segundo Vieira et al. (2012), no ano de 2011 foi realizada a estimação das médias de temperaturas diárias do estado de Santa Catarina utilizando um modelo de RLM.

Os modelos auto-regressivos ARIMA e SARIMA, esse último incluindo a sazonalidade, vêm sendo aplicados no estado de Mato Grosso para a predição de focos de queimadas e incêndios, associado às variáveis ambientais (Espinosa et al., 2010).

De acordo com Morettin e Toloi (2006), esses modelos ARIMA e SARIMA dependem de uma grande massa de dados históricos que podem gerar erros grosseiros quando submetidos às sazonalidades devido às baixas precisões apresentadas em algumas estimativas e ainda do quantitativo de variáveis a serem consideradas.

A modelagem de sistemas não lineares tem evoluído muito nos últimos tempos e, com elas, as redes neurais artificiais (RNAs). Uma RNA é uma estrutura matemática não-linear que é capaz de, arbitrariamente, representar processos não-lineares que relacionam entradas e saídas de um sistema (Hsu et al., 1995; Haykin, 2001).

As principais vantagens das RNAs na modelagem de queimadas e incêndios florestais são: não requererem conhecimento dos processos físicos causadores do fenômeno; podem ser aplicadas em sistemas sem soluções específicas; possibilitam o treinamento contínuo da rede; não amplificam os erros de medição; permitem uma otimização entre os dados de entrada e saída. As RNAs permitem modelar processos que envolvem séries temporais de sinais de entrada e saída.

Desta forma, o sucesso das RNAs e, muitos campos da ciência e da engenharia, sugeriu sua aplicabilidade também em eventos meteorológicos, especialmente em casos onde o conhecimento dos processos era muito limitado.

As técnicas computacionais de inteligência artificial, com o uso RNAs, foram utilizadas para a previsão de fenômenos como detecção de fumaça e ambientais. incêndios. mapeamentos de florestas. monitoramento de previsão queimadas, análise do uso da terra, dentre outros, no âmbito de ecossistemas para a preservação ambiental, melhorando satisfatoriamente o poder de previsão do modelo (França et al., 2009; Hauser et al., 2012; Acharya et al., 2014).

Segundo Litta et al. (2013), na região de Kolkata na Índia, foi realizada a predição de tempestades usando um modelo de RNA e, no mesmo ano, Coutinho et al. (2016) relataram a obtenção de resultados favoráveis de acertos para

predições de dados meteorológicos usando uma RNA *Multilayer Perceptron* (MLP).

Há ainda destaque do uso das RNAs, por apresentarem vantagens com relação às outras técnicas de análise multivariada de dados, no que tange aos seguintes aspectos: tolerância a falhas; aplicação em tempo real; capacidade de autoadaptação e a rápida resolução de problemas práticos, sem necessidade de definição de listas de regras ou de modelos precisos. Com isso, ao se definir uma rede o objetivo é resolver o problema com a menor estrutura possível, com potencial de em situações que requeiram aplicação classificação de padrões, aproximação de funções e aprendizado em áreas de difíceis modelagens de precisão com a realidade e com frequentes mudanças de ambiente (Haykin, 2001).

Assim o objetivo do presente estudo foi utilizar uma RNA (MLP) para prever o número de focos de incêndios e a respectiva área queimada no

bioma pantaneiro Sul-Mato-Grossense, na presença de variações de níveis de determinadas variáveis ambientais (número de focos e área queimada) e meteorológicas (temperatura do ar e do solo, umidade relativa do ar, quantidade de água no solo, velocidade do vento e radiação solar).

Material e métodos

Área de Estudo

Este estudo foi desenvolvido no bioma do Pantanal brasileiro que, de acordo com o IBGE (2015), possui uma área aproximada de 138.183 km², localizada na bacia do Alto Rio Paraguai, na Região Centro-Oeste brasileira, e abrange parte dos estados do Mato Grosso (MT) (7% da área de MT) e do Mato Grosso do Sul (MS) (25% da área de MS), conforme a Figura 1.

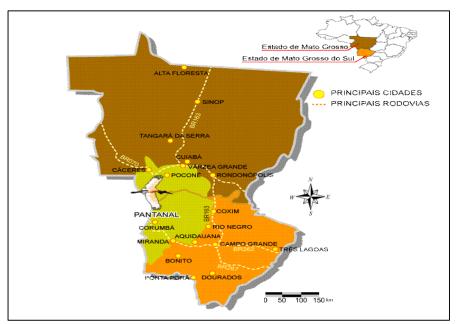


Figura 1. Região do Pantanal brasileiro e abrangência de seus municípios. Fonte: Adaptado de Silva et al. (2013)

A planície pantaneira está posicionada num nível de altitude entre 80 e 150 m, sendo o seu solo predominantemente pouco permeável. As formações de savanas constituem a principal fitofisionomia e, também, possuem formações vegetacionais da savana estépica, além de pequenas áreas de floresta estacional semidecidual e decidual que se interpenetram e se misturam com comunidades de espécies pioneiras de distribuição generalizada.

A base da economia pantaneira é a pecuária bovina de corte, criação extensiva, entretanto, o turismo é a segunda maior fonte de renda da região. Grande parte da cobertura vegetal original foi substituída por lavouras e pastagens e a modificação da paisagem traz aspectos negativos decorrentes das ações antrópicas.

Entre os municípios que constituem o Pantanal Sul-Mato-Grossense, o de maior extensão territorial é Corumbá (MS), com cerca de 65.000

km², sendo 95% de sua área, pertencentes à fisiografia do bioma, ela representa 69% do bioma no Estado (Silva et al., 2013).

A localização geográfica de Corumbá está compreendida entre as coordenadas de 19°00'32'' S e 57°39'10'' W, e altitude de 118 m acima do nível do mar. De acordo com RADAMBRASIL, o Pantanal se divide em três regiões geomorfológicas: Região do Pantanal Mato-Grossense; Região da Depressão do Alto Paraguai e Região da Bodoquena e Morrarias do Urucum-Amolar (Brasil, 1982).

O município de Corumbá possui tipo climático Aw (clima tropical úmido megatérmico), ou seja, clima tropical de altitude, com verões quentes e chuvosos e inverno seco (Soriano, 2007). A cobertura vegetal predominante é o cerrado típico do Pantanal, cerrado parque e cerrado arbóreo aberto. Nos locais mais elevados há áreas de matas e ao Sul do município ocorre a vegetação chaquenha e, nas áreas mais baixas, encontram-se as palmeiras carandá e o paratudal, que trata de formação savânica alagável (Brasil, 1982; Imasul, 2016).

Coleta de Dados

Para a realização deste estudo foram utilizados dados meteorológicos secundários do município de Corumbá, no período de 2005 a 2015, extraídos do banco de dados para ensino e pesquisa do Instituto Nacional de Meteorologia (INMET), da estação de coleta de dados meteorológicos 83552 (19° 00' 36" S e 57° 38' 60" W) e da coleta de plataforma de meteorológicos31949 (19° 01' 19.2" S e 57° 39' 7.2" W), do Centro de Previsão de Tempo e Estudos Climáticos do Instituto Nacional de Pesquisas Espaciais (INPE) (Inmet, 2016; Inpe, 2016).

Utilizou-se para o estudo, as médias compensadas diárias e mensais das variáveis meteorológicas consideradas independentes, que foram calculadas a partir da equação $X=(2.x_{00}+x_{12}+x_{máx}+x_{mín})$, onde X representa a variável em análise, x_{00} e x_{12} são os valores da variável observadas às 00h e 12h do tempo médio de Greenwich (TMG) e os valores extremos $x_{máx}$ e $x_{mín}$ são os registros dos valores máximos e mínimos, respectivamente, dentro do intervalo de 24h de cada variável.

As variáveis meteorológicas selecionadas a partir das correlações com as variáveis dependentes foram: Temperatura Máxima do Ar (°C) (T); Temperatura do Solo (°C) a 100 mm de profundidade (T_S); Contagem de água no Solo (m³) a 200 mm de profundidade (W_S); Umidade Relativa (%) (H); Radiação Solar Acumulada (MJ/m³) (R) e Velocidade do Vento (m/s) (S). O número de focos (N) e a área queimada em quilômetros quadrados (km²) (A), foram as variáveis ambientais dependentes, necessárias para a aplicação do método de predição.

relativos Os dados às variáveis dependentes, (número de focos e área total obtidos queimada), foram através monitoramento de satélites com sensores óticos operando na faixa termal-média de 4 um, que geram imagens que são processadas na Divisão de Satélites e Sistemas Ambientais (DSA) do INPE, e disponibilizadas no Banco de Dados de Queimadas do diretório da mesma instituição. As séries temporais diárias foram obtidas do satélite de referência, cujos dados dos focos detectados podem ser analisados estatisticamente para as mesmas regiões ao longo dos anos, com alta e moderada resolução espectral e com os efeitos atmosféricos já corrigidos.

Os satélites de referência, cujos dados foram extraídos, são: NOAA-12 (de 2005 a 2007) e AQUA M-T (de 2008 a 2015). Esses satélites possuem horário de passagem em um mesmo ponto da Terra, estáveis ao longo dos anos, e mantêm o mesmo padrão de amostragem temporal considerados isentos de falsas detecções devido a reflexos do sol na superfície terrestre. Fornecem, portanto, dados com um contínuo padrão de detecção.

A massa de dados analisada foi constituída de 316.800 dados horários para todo o conjunto de variáveis onde foram considerados 440 registros, sendo 55 registros mensais, (julho, agosto, setembro, outubro e novembro), denominado de período JASON (iniciais dos nomes dos meses em tela), para a série de 2005 a 2015, considerados críticos no que tange a ocorrência de focos de queimadas no bioma pantaneiro.

A escolha do período em análise (JASON) se deu em função das séries históricas de maiores estabilidades disponíveis no INPE e INMET, bem como o período de estiagem da região.

Para as análises de estatística descritiva de Redes Neurais Artificiais foi utilizado o *software* IBM – SPSS.

Funcionamento da Rede Neural Artificial Multilayer Perceptron (RNA – MLP)

As RNAs são modelos computacionais, inspirados no sistema nervoso de seres vivos, que

possuem a capacidade de aquisição e manutenção do conhecimento baseado em informações, e são caracterizadas por neurônios artificiais que são interligados por um grande número de interconexões apresentadas por matrizes de pesos sinápticos (Freeman e Skapura, 1991; Haykin, 2001; Koyács, 2006).

De acordo com Braga et al. (2014), as principais características das RNAs são: i) organização de dados: ii) adaptação do modelo por experiência; iii) capacidade de aprendizado; iv) armazenamento distribuído; v) tolerância a falhas; vi) habilidade de generalização; vii) controle e otimização de sistemas; viii) processamento de sinais e predição. Essas características incentivam a busca por metodologias que possam aprimorar os

processos de previsão de fenômenos meteorológicos, gerando dados satisfatórios próximos dos reais.

O processo de funcionamento das RNAs é composto basicamente por uma arquitetura, uma topologia e um treinamento. A arquitetura define a forma como os neurônios estão dispostos, a topologia define as diferentes formas de composições estruturais assumidas pela rede e o treinamento está associado ao algoritmo de aprendizagem, ou seja, consiste no conjunto de passos ordenados para os ajustes dos pesos, sintonizando as respostas de saída da rede próximas dos valores desejados (Braga, 2014). A Figura 2 mostra o esquema de funcionamento de uma RNA.

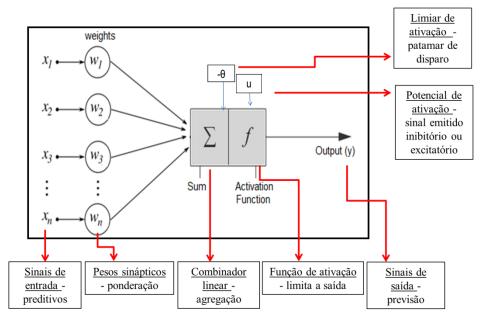


Figura 2. Estrutura de uma RNA MLP. Fonte: Adaptado de Barber et al. (2012)

As RNAs do tipo MLP pertencem à classe de arquitetura *feedforward* (alimentação à frente), e são constituídas de uma camada de entrada de neurônios, uma ou mais camadas ocultas (ou intermediárias) e uma camada de saída. O algoritmo de aprendizado utilizado no treinamento supervisionado da rede MLP é o *backpropagation* ou algoritmo de retropropagação de erros, baseado na regra Delta generalizada (Gomes et al., 2010).

Os modelos de neurônios da RNA, do tipo MLP são não lineares e capazes de gerar qualquer tipo de saída de acordo com a função ativação aplicada. Uma RNA não linear pode ser representada por um modelo de neurônio artificial não linear, onde x_i são os dados de entrada (valores

assumidos pelas variáveis de entrada), w_{ij} são os pesos sinápticos (valores que ponderam as variáveis de entrada), f(.) é a função ativação (limita a saída do neurônio dentro de um intervalo de valores razoáveis de serem assumidos), b_i são os bias (produzem o efeito de diminuir ou aumentar a função ativação), u_i é o potencial de ativação (pode ser inibitório ou excitatório), dado pela combinação linear do somatório dos n-ésimos termos $\sum_{i=1}^{n} (w_{ij} \times x_i + b_i)$ de $j = 1, 2, 3, \cdots$ e y_i é a saída que cada neurônio gera (Haykin, 2001; Freeman 2006; Kovács, 2006).

Para cada camada os pesos (w_{ij}) estão associados às entradas (x_i) , onde a resposta emitida para a próxima camada dependerá do estímulo

gerado pela soma ponderada dos produtos dos pesos de entradas ($w_{ij} \times x_i$) que ativam a função de transferência. O padrão de resposta é gerado quando a camada intermediária gera estímulos para a camada de saída, ou seja, as informações são extraídas da camada intermediária e seus pesos são responsáveis por codificar as características padrões de entrada, permitindo que a rede forneça uma resposta mais detalhada e uma representação mais complexa do problema (Hauser et al., 2012).

treinamento pelo algoritmo retropropagation, ocorre em duas fases, onde cada fase percorre a rede em um sentido. Estas duas fases são chamadas de forward e backward. A fase forward é utilizada para definir a saída da rede para um dado padrão de entradas e a fase backward ajusta os pesos e suas conexões dos sinais que retornam à camada de entrada. O algoritmo de retropropagação de erro pode ser dividido nos seguintes passos: i) selecionar o próximo par do conjunto de treinamento e aplicar o vetor de entrada à rede: ii) calcular a saída da rede: iii) calcular o erro entre a saída da rede e a saída-alvo; iv) ajustar os pesos da rede de maneira a minimizar o erro e; v) repetir do passo 1 até o passo 4 para cada vetor do conjunto de treinamento, até o erro se tornar aceitavelmente baixo para o conjunto inteiro. Pode-se observar que os passos 1 e 2 constituem a etapa de propagação para frente, em que o sinal de entrada é transmitido através da rede da entrada até a saída. Os passos 3 e 4, por sua vez, constituem-se na etapa de retropropagação, na qual o erro calculado é propagado de volta através da rede para ajustar os pesos (Hauser et al., 2012).

A função de entrada para uma RNA_MLP que possui *n* neurônios na camada de saída é dada pela equação 1. A partir dessa equação é possível calcular o ajuste dos pesos na direção contrária ao gradiente do erro (Haykin, 2001).

$$E(t) = \frac{1}{2} \sum_{i=1}^{n} e_i^2(t)$$
 (1)

Onde E(t) é a função de desempenho, propagado para todas as camadas da rede, $e_i(t)$ é o erro originado em cada saída da rede e obtido pela diferença do valor de saída predito e o valor esperado.

O ajuste de pesos do i-ésimo neurônio da camada de saída é dada pela equação 2, que é proporcional à derivada parcial de E(t) em relação a cada um dos pesos da camada de saída w_{ii} .

$$\Delta w_{ij}(t) = -\eta \frac{\partial E(t)}{\partial w_{ij}} \tag{2}$$

Onde η é a taxa de aprendizado do algoritmo de retropropagação que determina a velocidade de mudança dos pesos.

Portanto, de acordo com Haykin (2001), na camada oculta não existe uma resposta desejada para os neurônios e, por isso, deve-se determinar recursivamente um sinal de erro para cada neurônio para que os pesos sejam atualizados. Reescrevendo a equação 2 em termos da função ativação e aplicando a regra da cadeia pode-se redefinir o gradiente local para o neurônio oculto *i* e obter as equações 3 e 4 que expressam a atualização dos pesos da camada oculta pela regra do Delta generalizada.

$$\Delta w_{ij}(t) = -\eta \delta_i(t) h_i(t) \quad (3)$$

$$\Delta w_{ij}(t) = -\eta \delta_i(t) x_j(t) \quad (4)$$

Onde $\Delta w_{ij}(t)$ é o ajuste dos pesos dos neurônios em cada conexão sináptica, $\delta_i(t)$ é o gradiente local originado do produto do erro $e_i(t)$ pela derivada da função ativação $f'(u_i(t))$, η é a taxa de aprendizado do algoritmo de retropropagação que determina a velocidade de mudança dos pesos (quando aparece negativo significa a descida do gradiente no espaço de pesos), $h_j(t)$ é o valor de entrada de uma conexão do neurônio da camada de saída e $x_j(t)$ é a entrada de uma conexão do neurônio da camada de saída com os pesos ajustados, ou seja, com o erro minimizado.

Após o processo de ajuste dos pesos sinápticos a amostra é dividida aleatoriamente em 70% da massa de dados para o treinamento e testes da rede e 30% para a validação.

Estrutura da experimentação para a predição dos dados

Os experimentos foram realizados para oito modelos de previsão com a mesma RNA – MLP, com diferentes estruturas de registros dos dados do período JASON, para a série histórica de 2005 a 2015. Foram testadas como variáveis preditivas o número de focos e a área queimada (em Km²), ambas no intuito de verificar a probabilidade de combustão da região em estudo associada às variáveis meteorológicas e ambientais de entrada. Na Tabela 1 estão dispostas as estruturas

compostas pelas variáveis de entrada e de saída de dados, adotadas para a utilização nas RNAs.

Tabela 1. Estruturas dos dados de entradas e saídas da RNA MLP adotada para cada teste.

Estrutura	Descrição	Dados de entrada – Variáveis Preditivas*	Dados de Saída – Variável Prevista**
1	Previsão dos Focos	N	Média de N
2	Previsão dos Focos	N (2005 a 2014)	N (2015)
3	Previsão dos Focos	T, H e R	N
4	Previsão de Focos	$T, H, R, S, T_S e W_S$	N
5	Previsão da Área Queimada	N	A
6	Previsão da Área Queimada	T, H, R e N	A
7	Previsão da Área Queimada	$T, H, R, S, T_S, W_S e N$	A
8	Previsão da Área Queimada	A (2005 a 2014)	A (2015)

^{*} Onde N é o número de focos; T é a temperatura do ar; H é a umidade relativa do ar; R é a Radiação Solar; S é a velocidade do vento; T_S é a temperatura do solo e W_S é a contagem de água no solo.

Fonte: Própria Autora

É possível observar que as estruturas de um a quatro têm como variável dependente o número de focos (N) e nas estruturas de cinco a oito, as áreas queimadas (A).

Avaliação do desempenho

O desempenho das previsões geradas em cada estrutura foi avaliado através das medidas estatísticas de correlação de Pearson (r), do grau de ajustamento R² entre os dados previstos e os dados observados, análise de variância (ANOVA), erro relativo médio (REM), erro percentual absoluto médio (MAPE) e análise dos resíduos de alguns desses testes.

Foram avaliados também os valores médios, máximos, mínimos, desvio padrão,

covariância, fator de inflação da variância, a tolerância, a homocedasticidade, e significância dos dados reais e dos previstos (Hair et al., 2005).

Configurações das RNAs

A aplicação da RNA – MLP com o algoritmo backpropagation foi realizada com o software IBM-SPSS. As configurações de cada estrutura de dados analisados (com um total de oito estruturas) foram constituídas de apenas uma camada intermediária, a função de ativação adotada para a camada oculta foi a função tangente hiperbólica e, da camada de saída, foi a função identidade. A Tabela 2 dispõe do número de neurônios da camada intermediária de cada estrutura, assim como, as unidades das camadas de entrada e saída das redes e demais informações.

Tabela 2. Informações das configurações das RNAs adotadas para as oito estruturas em análise.

	Cam	ada de entrada	Camada oculta	Camada de saída		
Estrutura	Número de Variáveis variáveis independentes*		Número de neurônios	Número de variáveis	Variável dependente**	
1	1	N	5	1	Média de N	
2	10	N(2005 a 2014)	6	1	N (2015)	
3	3	T, H e R	1	1	N	
4	6	$T, H, R, S, T_S e W_S$	2	1	N	
5	1	N	1	1	A	
6	4	T, H, R e N	4	1	A	
7	7	$T, H, R, S, T_S, W_S e N$	2	1	A	
8	10	A (2005 a 2014)	5	1	A (2015)	

^{**} Onde A é a área queimada em km².

* Onde N é o número de focos; T é a temperatura do ar; H é a umidade relativa do ar; R é a Radiação Solar; S é a velocidade do vento; T_S é a temperatura do solo e W_S é a contagem de água no solo.

** Onde A é a área queimada em km².

Fonte: Própria Autora

O treinamento ocorreu no modo *on line* de aprendizagem, que utiliza padrão-por-padrão, onde os ajustes dos pesos são feitos após a apresentação de cada amostra de treinamento e após a execução do passo de ajuste. A respectiva amostra pode ser descartada não necessitando de grandes espaços de armazenamento. O número máximo de épocas de execução da rede foi de 133 e o método de reescalonamento padronizado das variáveis dependentes e independentes, foi adotado para as oito estruturas.

Resultados e discussão

Para cada estrutura treinada pela RNA – MLP foram realizadas 10 simulações e calculadas as médias dos dados preditos para as análises de desempenho.

A Figura 3 mostra o perfil das variáveis meteorológicas do bioma pantaneiro para o município de Corumbá (MS) do período JASON, para a série de 2005 a 2015.

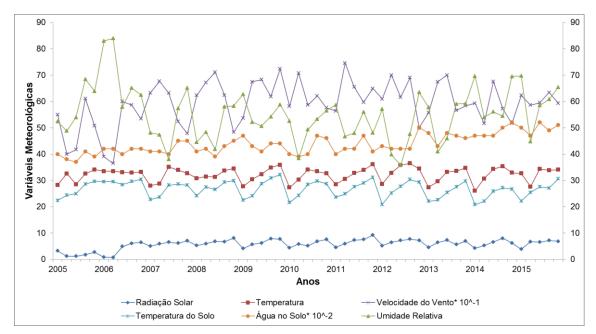


Figura 3. Gráfico do comportamento das variáveis meteorológicas utilizadas como independentes no sistema de previsão, para a série histórica de 2005 a 2015, no período JASON do bioma pantaneiro, no município de Corumbá (MS).

De acordo com as curvas dispostas na figura 3, é possível observar o comportamento das variáveis meteorológicas no decorrer da série temporal para o período JASON. Elas apresentam características cíclicas para os mesmos períodos de estação seca de cada ano, com os dados bem distribuídos e sem a constatação de eventos extremos que pudessem alterar bruscamente os valores das variáveis selecionadas.

Comparando as curvas de temperaturas do ar e do solo, pode-se verificar que as mesmas se encontram em fase, com todos os valores de pontos de máximo, de mínimo, crescimento e

decrescimento. Elas possuem uma diferença de aproximadamente 10°C, entre as temperaturas do ar maior que a do solo. Ambas influenciam no número de focos e na área queimada, pois, aquecem a planta retirando umidade do interior dos vegetais, o que os tornam propícios ao processo de combustão, facilitando a formação de massa seca.

A radiação solar é fortemente influenciada pela presença de nuvens, o que dificulta sua absorção nos vegetais em dias cuja umidade relativa é mais alta.

Ao se comparar as curvas de velocidade do vento com a umidade relativa, nota-se que para os

valores extremos de máximo e mínimo de toda a série, essas variáveis apresentam comportamentos opostos. Valores altos para a umidade relativa significam que o ar fica mais denso, ou seja, a quantidade de partículas de água são maiores, o que dificulta o deslocamento de massas de ar e consequentemente a área queimada será menor,

caso contrário, favorecem a ocorrência de um maior número de focos e o aumento da área queimada.

A Figura 4 mostra a comparação entre o número de focos e a área queimada para a série temporal em análise.

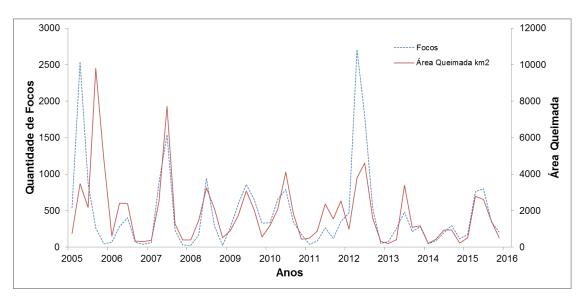


Figura 4. Gráfico do comportamento das variáveis ambientais usadas como dependentes no sistema de previsão, para a série histórica de 2005 a 2015, no período JASON, do bioma pantaneiro no município de Corumbá (MS).

Com base na figura 4, pode-se verificar que as curvas do número de focos e área queimada são diretamente proporcionais, se encontram em fase e quanto maior o número de focos maior é a área queimada, a qual apresentou picos em cerca de quatro vezes maiores do que os focos observados, como por exemplo, nos anos de 2010 e 2014. A velocidade do vento apresentou picos de máximos, quando a umidade relativa foi mínima e as demais variáveis meteorológicas apresentaram valores favoráveis ao processo de combustão, o que justifica a extensão da área queimada nessas proporções, desse mesmo período (figura 3 e 4).

Em geral, os picos dos focos e da área queimada ocorreram para os meses de agosto e setembro, no período de estiagem, com valores médios em torno de 450 focos e 1841 km² de área queimada, conforme exposto na figura 4.

Ao se comparar as figuras 3 e 4, verificase que a temperatura máxima média foi de 32,3°C, onde o número de focos foi maior para as temperaturas mais elevadas, como por exemplo, agosto de 2005, setembro de 2007, agosto e setembro de 2012 e novembro de 2014. Já nos meses de agosto de 2006 e outubro de 2013, mesmo com elevadas temperaturas, houve uma diminuição no número de focos decorrentes da alta umidade relativa do ar. A radiação solar afeta o número de focos e, consequentemente, a área queimada, quando associada às condições de baixa umidade relativa do ar e altas temperaturas. Entretanto, a curva de radiação apresenta defasagem com a curva do número de focos, quando analisada isoladamente.

Destacam-se os anos de 2005, 2007 e 2012, que apresentaram os maiores valores de incidência de focos na série, isso pode ser proveniente do acúmulo de material combustível das plantas dos anos precedentes (2004, 2006 e 2011 respectivamente), em que ocorreram elevadas precipitações e enchentes na região, tornando a vegetação mais fraca e propícia à queima, ocasião em que se pode observar a alta umidade relativa, com média em torno de 60%.

A Tabela 3 apresenta as medidas de média (MED), valor máximo (MAX), valor mínimo (MIN), desvio padrão (DP) e coeficiente de

variação (CV) dos valores das variáveis observadas (dados medidos).

Tabela 3. Estatística dos valores das variáveis observadas, utilizadas nos modelos de previsão, dispostas nas oito diferentes estruturas para o período JASON, de 2005 a 2015.

Dados Observados	MED	MAX	MIN	DP	CV (%)
Radiação Solar	5,72	9,19	0,72	1,94	32,92
Temperatura do Ar	32,21	36,44	26,03	2,65	8,23
Umidade Relativa	55,33	83,97	35,80	10,09	18,24
Velocidade do Vento	6,02	10,79	3,65	1,07	17,77
Temperatura do Solo	26,73	32,22	20,82	3,06	11,45
Contagem de Água no Solo	0,44	0,52	0,37	0,04	9,09
Número de Focos	456	2710	24	55,96	12,27
Área Queimada	1841	9809	199,59	182,94	9,94

Fonte: Própria Autora

De acordo com a tabela 3, é possível observar que os dados das variáveis meteorológicas e ambientais, são homogêneos e as medições não apresentam discrepâncias. O maior coeficiente de variação foi o da radiação solar, cujas medidas são fortemente influenciadas na presença de nuvens, apresentando maior dispersão na série, como constatado nas análises da figura 3.

A Tabela 4 apresenta a análise estatística e a avaliação de desempenho nos modelos de predição (dados previstos), do conjunto de validação dos dados das oito estruturas apresentadas, utilizando como variáveis preditoras ora as variáveis meteorológicas e, ora as variáveis previstas (focos e área queimada).

Tabela 4. Estatística dos valores das variáveis previstas, utilizadas nos modelos de previsão, dispostos nas oito diferentes estruturas para o período JASON, de 2005 a 2015.

Estru - tura	MED	MAX	MIN	DP	\mathbb{R}^2	REM	MAPE (%)	Equação de ajuste linear*
1	466	846	167	28,88	0,995	0,008	2	Y=-11,92+1,05x
2	528	802	208	26,42	0,994	0,006	4	Y=-4,28+1,01x
3	511	1916	98	462,2	0,607	0,406	44	$Y=2,03e^2+0,52x$
4	583	1888	14	425,7	0,848	0,155	32	Y=99,95+0,80x
5	2022	4459	1181	959,5	0,560	0,442	56	$Y=8,38e^2+0,53x$
6	1537	4629	0,2452	123,95	0,650	0,364	35	$Y=3,89e^2+0,69x$
7	1691	5304	163,46	120,19	0,903	0,207	20	$Y=2,1e^2+0,89x$
8	1572	2788	511	196,71	0,986	0,030	3	Y=-7,80+0,97x

*Onde Y corresponde ao valor previsto e x corresponde ao valor observado.

Fonte: Própria Autora

Com base nos resultados da tabela 4, foi possível constatar que os modelos das estruturas de redes neurais artificias dois, quatro, sete e oito apresentaram os resultados mais satisfatórios com os menores erros e um poder de predição de 99,4%, 84,8%, 90,3% e 98,6%, respectivamente. Nos

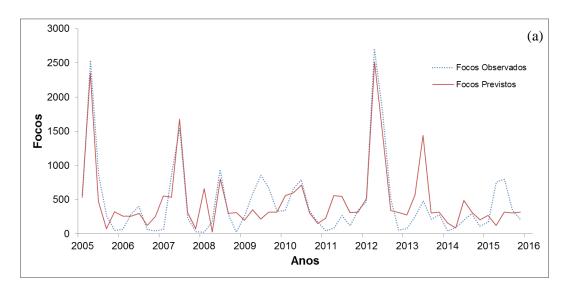
modelos dois e oito foram realizadas as previsões dos focos e da área sem utilizar as variáveis meteorológicas, ou seja, a variável independente foi o número de focos e a área queimada, respectivamente, das séries anteriores e, nas estruturas quatro e sete, todas as variáveis meteorológicas foram consideradas e tiveram, respectivamente, como dados previstos o número de focos e a área queimada.

Com os resultados das estruturas dois e oito, pode-se observar que a técnica de RNA é ótima para a previsão do número de focos e da área queimada e, nesse caso, não havendo a necessidade de se estabelecer a relação dessas variáveis com as

demais variáveis externas ao sistema, tais como as variáveis meteorológicas ou demais causas que possam contribuir para favorecer o processo de combustão, como por exemplo, a ação humana.

Adicionalmente, observa-se que no modelo de estrutura cinco o poder de previsão diminui e os erros de previsão aumentam. Isso se deve ao condicionamento das ações antrópicas que não foram consideradas e as contribuições das demais variáveis meteorológicas que são significativas no fenômeno.

A Figura 5 representa as curvas dos valores observados e previstos para o número de focos (Figura 5a) e para a área (Figura 5b).



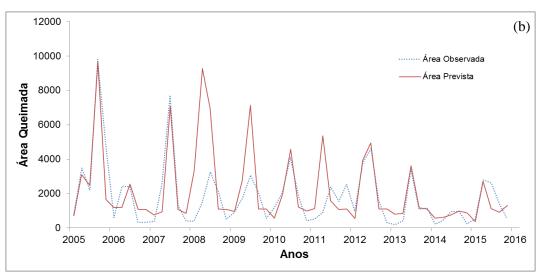


Figura 5. Gráficos dos valores observados e previstos pelas modelagens de RNA das estruturas 4 e 7 respectivamente (Figura 5a e Figura 5b), utilizando como entrada os dados das variáveis meteorológicas selecionadas, para a série histórica de 2005 a 2015 do período JASON, no bioma pantaneiro de Corumbá (MS).

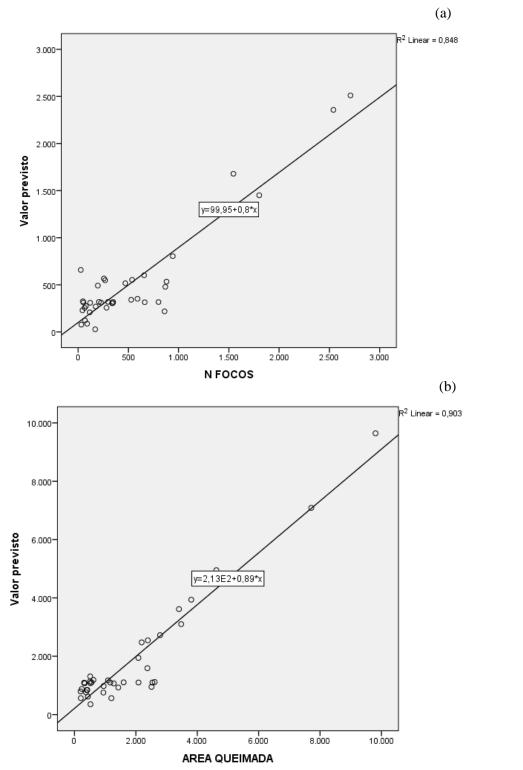
De acordo com os gráficos da figura 5, observa-se que os dados previstos tanto para o

número de focos quanto para a área queimada estão próximos, com vários pontos coincidentes, com

erro relativo médio de 15% para o número de focos e de 3% para a área queimada, o que significa que os dados estão bem ajustados, resultando em uma previsão confiável pela técnica de RNA nas duas estruturas.

Os gráficos de ajuste linear entre os valores observados e os valores previstos das RNAs das

estruturas quatro e sete, dos modelos preditivos a partir das variáveis meteorológicas, estão dispostos na Figura 6 (Figura 6a e Figura 6b, respectivamente), com suas equações de regressão e valores de ajuste \mathbb{R}^2 .



Viganó, H. H. G

Figura 6. Gráficos de ajuste entre os dados previstos e os dados observados na estrutura 4 para o número de focos (Figura 6a) e estrutura 7 para a área queimada (Figura 6b), das RNA utilizadas, com variáveis meteorológicas como preditoras.

É possível observar, a partir da figura 6, que a reta de regressão dos dados observados e previstos, para o número de focos, (figura 6a), possui um ajuste linear menor do que o modelo apresentado para a previsão da área queimada

(figura 6b), pois a dispersão dos dados é maior, o que justifica esse decréscimo de R².

A Figura 7 apresenta os valores residuais dos dados observados e previstos para os dois principais modelos descritos pelas estruturas 4 e 7.

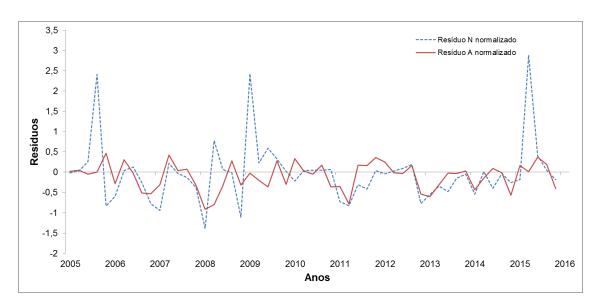


Figura 7. Gráficos dos valores residuais dos dados observados e previstos para o número de focos (estrutura 4) e área queimada (estrutura 7), para a série histórica de 2005 a 2015 do período JASON, no bioma pantaneiro de Corumbá (MS).

Considerando a análise de resíduos (figura 7), pode-se afirmar que a RNA MLP pode ser utilizada para fazer inferência além da amostra de dados utilizados. A previsão da área queimada foi melhor do que a previsão de focos, onde se observa que ocorre uma oscilação menor em torno do zero, apresentando melhor coeficiente de variação. O desempenho geral da técnica, com p-valor < 0,001, para as estruturas quatro e sete foram 68% e 97%, respectivamente, o que confirma a viabilidade de aplicação da RNA para modelos de previsão.

O valor de desempenho para a previsão dos focos (estrutura quatro) é inferior ao da previsão da

área queimada (estrutura sete), o que pode ser corrigido, com a introdução de novas variáveis, que por exemplo, identifiquem as causas das ocorrências dos focos (naturais ou antrópicas).

Conclusões

A técnica de Redes Neurais Artificiais foi aplicada para a previsão do número de focos e da área queimada no bioma pantaneiro e apresentou resultados satisfatórios na avaliação de desempenho, sendo adequado para a previsão utilizando tanto as variáveis meteorológicas como

variáveis preditoras, quanto as variáveis ambientais. Destacaram-se os modelos dois, quatro, sete e oito, em que os erros REM e MAPE ficaram bem ajustados, validando o R² e demonstrando ser uma ótima técnica.

Novos estudos modificando a tipologia da RNA, o número de camadas ocultas e neurônios, a função de ativação de entrada e saída, dentre outros quesitos, podem ser aplicados com a finalidade de aumentar o poder preditivo da técnica RNA e otimizar os resultados de previsão.

Agradecimentos

Os autores agradecem à Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul, (FUNDECT), pelo auxilio financeiro concedido junto ao Edital de Chamada FUNDECT n. 08/2015 de bolsa de Doutorado em Mato Grosso do Sul.

Referências

- Acharya, N., Shrivastava, N.A., Panigrahi, B.K., Mohanty, U.C., 2014. Development of an Articial Neural Network based multi-model ensemble to estimate the Northeast monsoon rainfall over South peninsular India: an application of extreme learning machine. Climate Dynamics 43, 1303-1310.
- Almorox, J., Benito, M., Hontoria, C., 2008. Estimation of Global Solar Radiation in Venezuela. Interciencia 33, 280-293.
- Anderson, L.O., Aragão, L.E.O.C, Lima, A., Shimabukuro, C., 2005. Detecção de cicatrizes de áreas queimadas baseada no modelo linear de mistura espectral e imagens índice de vegetação utilizando dados multitemporais dos sensor MODIS/TERRA no estado do Mato Grosso, Amazônia brasileira. Revista Acta Amazonica 35, 445-456.
- Artaxo, P., Oliveira, P.H., Lara, L.L., Pauliquevis, T.M., Rizzo, L.V., Junior, C.P, 2006. Efeitos climáticos de partículas de aerossóis biogênicos e emitidos em queimadas na Amazônia. Revista Brasileira de Meteorologia 21, 168-189.
- Barber, D., Sollich, P., Saad, D., 2012. Finite Size Effects in on-line learning of multi-layer Neural Network. In: Ellacott, S. W., Mason, J.C., Anderson, I.J., Mathematics of Neural Networks: Models, Algorithms and Applications. Springer, Manchester, pp. 84-88.
- Braga, A.P., Carvalho, A.P.L.F., Ludermir, T.B., 2014. Redes Neurais Artificiais: Teoria e Aplicações, 2 ed. LTC, Rio de Janeiro.

- Brasil. Radambrasil, Ministério das Minas e Energia, 1982. Folha SE. 21. Corumbá e Folha SE. 20. Geologia, Pedologia, Geomorfologia, Vegetação e Uso e Ocupação da Terra, 1 ed. IBGE, Rio de Janeiro.
- Coutinho, E.R., Silva, R.M., Delgado, A.R.S, 2016. Utilização de Técnicas de Inteligência Computacional na Predição de Dados Meteorológicos. Revista Brasileira de Meteorologia 31, 24-36.
- Espinosa, M.M., Prado, S.M., Ghellere, M., 2010. Uso do modelo SARIMA na previsão do número de focos de calor para os meses de junho a outubro no Estado de Mato Grosso. Ciência e Natura 32, 7 21.
- Falayi, E.O., Adepitan, J.O., Rabiu, A.B., 2008. Emprirical models for the correlation of global solar radiation with meteorological data for Iseyin, Nigeria. International Journal of Physical Sciences 3, 210-216.
- França, M.M., Fernandes Filho, E.I., Xavier, B.T.L., 2009. Análise do uso da terra no município de Viçosa-MG mediado por classificações supervisionadas com Redes Neurais Artificiais e MAWVER. Revista Brasileira de Geografia Física 2, 92-101.
- Freeman, J.A., Skapura, D.M., 1991. Neural Networks: Algorithms, Applications, and Programming Techniques, 2 ed. CNS, California.
- Gois, G., Souza, J.L., Silva, P.R.T., Júnior, J.F.O., 2005. Caracterização da desertificação no Estado de Alagoas utilizando variáveis climáticas. Revista Brasileira de Meteorologia 20, 301-314.
- Gomes, L.F.C., Montenegro, S.M.G.L., Valença, M.J.S., 2010. Modelo baseado na Técnica de Redes Neurais para Previsão de Vazões na Bacia do Rio São Francisco. Revista Brasileira de Recursos Hídricos 15, 5-15.
- Granemann, D.C., Carneiro, G.L., 2009. Monitoramento de focos de incêndio e áreas queimadas com utilização de imagens de sensoriamento remoto. Revista de Engenharia e Tecnologia 1, 55-62.
- Hair, J., Anderson, R., Tatham, R., Black, W., 2005. Análise multivariada de dados, 5 ed. Bookman, Porto Alegre.
- Hauser, T., Keats, A., Tarasov, L., 2012. Artificial Neural Network assisted Bayesian Calibration of Climate Models. Climate Dynamics 39, 137-154.
- Haykin, S., 2001. Redes Neurais Princípios e Prática, 2 ed. Artmed Editora S.A., Porto Alegre.

- Hsu, K., Gupta, H.V., Sorooshian, S., 1995. Artificial neural network modeling of the rainfallrunoff process. Water Resources Research 31, 2517-2530.
- IBAMA. Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis, 2009. Manual de operações de prevenções e combate aos incêndios: comportamento do fogo. Brasília. Disponível: http://www.icmbio.gov.br/portal/images/storie s/servicos/sejaumbrigadista.pdf. Acesso: 08 fev. 2016.
- IBGE. Instituto Brasileiro de Geografia e
 Estatística, 2016. Cidades Mato Grosso do Sul
 Corumbá MS. Senso 2015. Disponível:
 http://www.cidades.ibge.gov.br/xtras/perfil.ph
 p?lang=&codmun=500320&search=matogrosso-do-sul|corumba|infograficos: informacoes-completas. Acesso: 08 fev. 2016.
- IMASUL. Instituto de Meio Ambiente de Mato Grosso do Sul, 2016. Monitoramento Ambiental Disponível:

- http://www.imasul.ms.gov.br. Acesso: 21 jan. 2016.
- INMET. Instituto Nacional de Meteorologia, 2016.

 Banco de Dados Meteorológicos para Ensino e
 Pesquisa. Disponível:
 http://www.inmet.gov.br/portal/index.php?r=b
 dmep/bdmep. Acesso: 08 fev. 2016.
- INPE. Instituto Nacional de Pesquisas Espaciais, 2016. Portal do Monitoramento de Queimadas e Incêndios. Disponível: http://www.inpe.br/queimadas. Acesso: 08 fev. 2016.
- Kovács, Z.L., 2006. Redes Neurais Artificiais: Fundamentos e Aplicações, 4 ed. Livraria da Física, São Paulo.
- Litta, A.J., Idicula, S.M., Mohanty, U.C., 2013. Artificial Neural Network Model in Prediction of Meteorological Parameters during Premonsoon Thunderstorms. International Journal of Atmospheric Sciences 10, 1-14.
- Morettin, P.A., Toloi, C.M., 2006. Análise de Séries Temporais, 2 ed. Blucher, São Paulo.
- Silva, M.H.S., Passos, M.M., Sakamoto, A.Y., 2013. As lagoas Salitradas do Pantanal da Nhecolândia: um estudo da paisagem baseado no modelo GTP – Geossistema, Território e Paisagem. Revista Franco-Brasileira de Geografia Confins 19, 1-29.
- Soriano, B.M.A., Pellegrin, L.A., 2007. Monitoramento do número de focos de calor e variáveis meteorológicas observados em 2007 no Pantanal. Embrapa Pantanal, Corumbá. (Comunicado Técnico, 64).
- Vieira, H.J., Back, A.J., Lopes F., Morais, H., 2012. Estimativa das temperaturas médias diárias, diurnas e noturnas a partir das temperaturas horárias. Revista Brasileira de Meteorologia 27, 329-336.