

UTILIZAÇÃO E ANÁLISE DE REDES NEURAIS ARTIFICIAIS COM PARÂMETROS CLIMATOLÓGICOS APLICADO A DADOS DE PRECIPITAÇÃO

Camila Bermond Ruezzene¹

Carlos Eduardo Nunes Medina Martinez²

Renato Billia de Miranda³

Adriano Rogerio Bruno Tech⁴

Frederico Fábio Mauad⁵

¹Doutoranda em Ciências da Engenharia Ambiental, USP/Campus São Carlos/SP, camila.ruezzene@gmail.com

²Mestrando pelo Programa de Pós-Graduação em Ciências da Engenharia Ambiental, USP/Campus São Carlos/SP, carlosenmm@gmail.com

³Professor Doutor do Programa de Pós-Graduação em Ciências da Engenharia Ambiental, USP/Campus São Carlos/SP, rbm1706@gmail.com

^{4,5}Professor Associado do Programa de Pós-Graduação em Ciências da Engenharia Ambiental, USP/Campus São Carlos/SP, adriano.tech@usp.br; mauadffm@gmail.com

Introdução

As ocorrências das precipitações são derivadas de fenômenos atmosféricos que se interagem com a superfície terrestre por meio de trocas de calor e umidade. A precipitação pluvial possui grande atuação nas condições do ambiente, pois representa o principal componente de entrada do balanço hídrico, contribuindo assim com a recarga de corpos d'água (GOUVEA et al., 2018).

O conhecimento do regime pluviométrico e da variabilidade da precipitação se torna fundamental nas análises ambientais e de cenários socioeconômicos, já que eventos de grande intensidade ou de escassez podem gerar impactos negativos nas atividades do homem e, consequentemente, no desenvolvimento de uma região. Estudos de precipitação são realizados por meio de simulações numéricas e geração de modelos que permitem a realização de projeções e avaliação de impactos relacionados a processos erosivos, produção agrícola e eventos climáticos, sendo também aplicados na gestão e gerenciamento dos recursos hídricos (KLEIBER et al., 2012; SANTOS et al., 2018).

Há uma grande limitação quanto ao conhecimento aprofundado da distribuição dos dados climatológicos que influenciam diretamente na precipitação devido à complexidade física de seus processos. É de grande relevância que se obtenha resultados satisfatórios de projeções e possíveis anomalias de chuva que possa estar ocorrendo ao longo do

tempo. Dessa forma, as Redes Neurais Artificiais (RNAs) apresentam grande potencial para compreensão da variabilidade espaço/temporal e modelagem de vários processos meteorológicos e hidrológicos existentes em bacias hidrográficas, por ser uma técnica complexa de classificação e reconhecimento de padrões (MELLO et al., 2017; WANDERLEY et al., 2014).

A RNA é uma ferramenta que simula o sistema neurológico do ser humano. Operando em camadas e neurônios, ela recebe do usuário uma ou mais séries de parâmetros que são denominadas como entrada da rede e uma única série de parâmetros denominado como alvo da rede. A rede atribui pesos a cada tipo de dado de entrada para obter uma saída que é, então, comparado com o valor do alvo, ou seja, a saída esperada. Este processo é repetido diversas vezes sendo que, a cada interação, a rede altera os valores dos pesos atribuídos em busca de uma correlação cada vez mais precisa entre os valores de saída e o alvo (NKUNA & ODIYO, 2011; HEMA & KANT, 2017).

Esses novos métodos de análise através das redes neurais vem sendo cada vez mais utilizados para modelagem climáticas, sendo essas ferramentas estatísticas não lineares que tornam as análises mais eficazes para relacionar as complexas entradas e saídas de simulações, gerando assim, séries consistentes e confiáveis para os estudos mais detalhados do clima (COUTINHO et al., 2018; MAQSOOD & ABRAHAM, 2007).

Desta maneira, a partir da classificação da intensidade da precipitação, o objetivo deste estudo é criar redes neurais artificiais para o preenchimento de falha em dados mensais de chuva através de parâmetros climatológicos, utilizando dados de 2010 a 2018 de uma estação climatológica localizada no município de Itirapina-SP.

Material e Métodos

A área de estudo localiza-se no município de Itirapina-SP com uma população de 15.524 habitantes e conta com uma área territorial de 564,603 km² (IBGE, 2010). Geograficamente, as regiões situam-se na porção central do Estado de São Paulo (Figura 1).

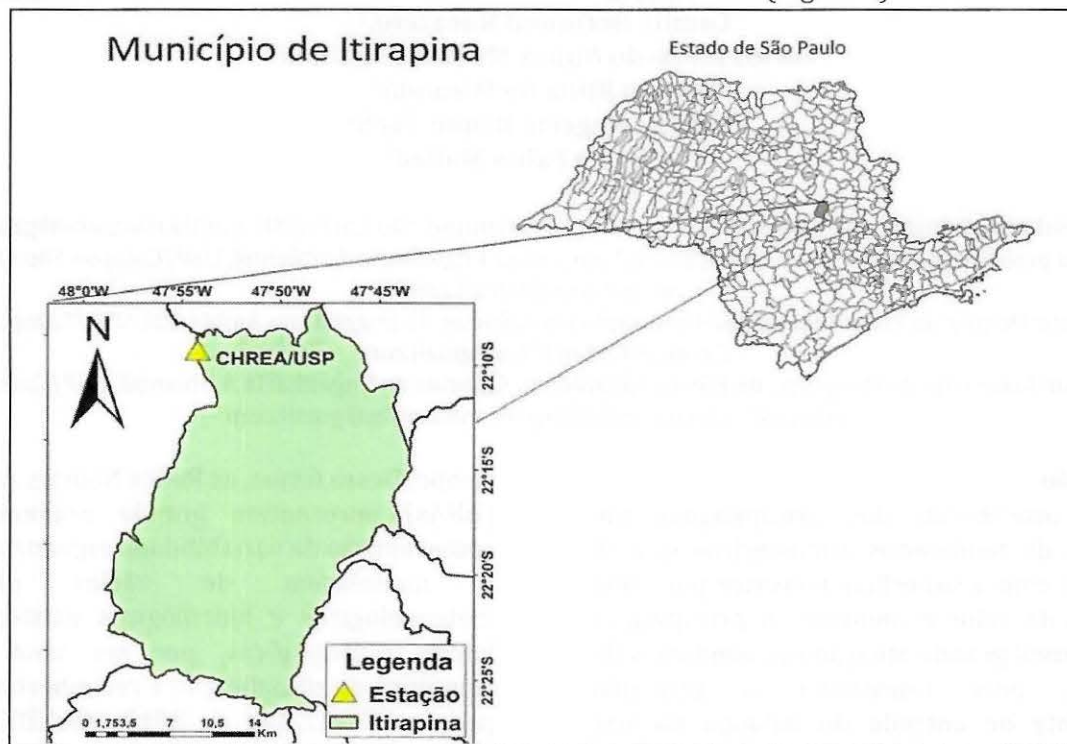


Figura 1. Mapa de localização da área de estudo. Fonte: Adaptado de IBGE (2009).

O clima predominante da região, de acordo com a classificação de Köppen é CWA, definido por clima temperado, de invernos secos e verões chuvosos. Esta região central tem uma sazonalidade com estação úmida entre os meses de outubro a março, de intensidade concentrada principalmente nos meses de dezembro a fevereiro, e a estação seca, correspondente aos meses de abril a setembro, apresentando uma precipitação média anual de 1493mm (MACHADO & MATTOS, 2001). A temperatura média é de 22°C, porém, variando entre mais amenas e altas, conforme a época do ano (CEPAGRI, 2015; PEEL et al., 2007; PERUSSI & VECCHIA, 2016). Os solos da região são, basicamente arenosos e em direção a Serra Geral encontra-se maior presença de solos em argila (ZANCHETTA & PINHEIRO, 2007).

Neste estudo, foram utilizados os dados da estação meteorológica do Centro de Recursos Hídricos e Estudos Ambientais (CRHEA) da Escola de Engenharia de São Carlos (EESC) da Universidade de São Paulo (USP). Essa estação foi instalada no ano de 1972 e a escolha desta foi devido a ela possuir dados primários e consistidos. Para a criação da rede foi definido uma série onde não havia nenhuma falha nos parâmetros climatológicos diários, sendo compreendido entre os anos de 2010 a 2018 e utilizado como variáveis de entrada temperaturas mínima, máxima e média (°C), velocidade do vento a 0,5 m de altura (m/s), umidade relativa (%) e como parâmetro de saída a precipitação (mm) como pode ser observado na Figura 2.

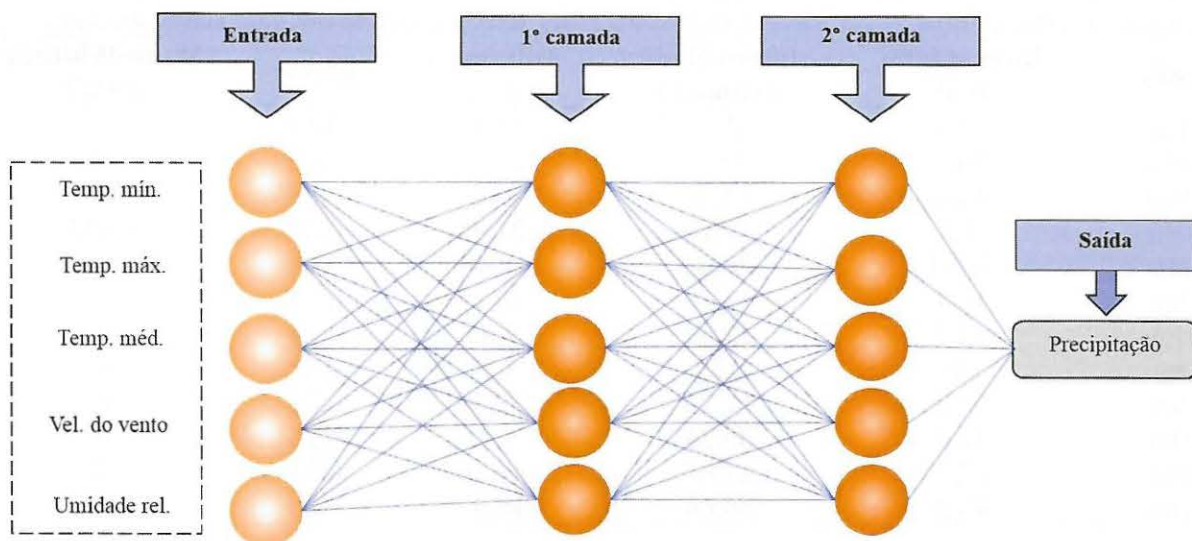


Figura 2. Arquitetura da rede neural.

As redes foram treinadas no software MATLAB versão R2015a utilizando a rede *feed-forward backpropagation*, com função de treinamento Levenberg-Marquardt, a função de aprendizado adaptativo *Gradient descent with momentum weight and bias*, a função de performance *mean-squared error*. A estrutura da rede foi implementada com 5 neurônios de entrada, duas camadas intermediárias com 10 neurônios e, uma camada de saída com um único neurônio, sendo a função de transferência utilizada a *tansig*.

Inicialmente, foi criada uma única rede que recebeu os parâmetros da série diária dos 9 anos, entretanto, após os testes iniciais foram criadas trinta e seis (36) redes para melhor modelagem dos dados e, conseqüentemente, melhor desempenho da rede. Os dados foram divididos tanto por mês quanto por classificação de intensidade da chuva, adaptada de Reichardt (1986), sendo a baixa menor que 2,5mm; a média maior ou igual a 2,5mm e menor que 7,5mm; e a alta maior ou igual a 7,5mm. Após as redes terem calculado os valores esperados de precipitação para cada intensidade nos respectivos meses, foram feitas três análises, a diferença entre a precipitação acumulada real e a estimada, o coeficiente de determinação e o teste não paramétrico de Mann-Witney proposto por Sneyers (1975). Foram utilizados valores acumulados de precipitação para cada intensidade correspondentes aos 12 meses e, após, realizou-se a diferença entre o acumulado real e o estimado.

Foi utilizado o coeficiente de determinação para avaliar a qualidade do ajuste

de modelo indicando o quanto ele foi capaz de explicar os dados de referência, sendo que valores mais próximos de 1 indicam melhores modelos criados, e segundo recomendações de Pruski et al. (2004) e Barbosa et al. (2005) adotou-se como critério mínimo, valores obtidos de R^2 acima de 0,7. Foram calculados os coeficientes para as 36 redes criadas comparando-se os valores de precipitação diários reais com os estimados.

Foi realizado o teste não paramétrico de Mann-Witney com nível de significância de 0,05 para testar se houve diferenças significativas entre o valor estimado pelas redes neurais e o dado real de referência da precipitação. Ao encontrar p-valores maiores que o nível de significância, aceita-se H_0 onde os dados provêm de amostras iguais, e rejeita-se a H_1 , onde os dados provêm de amostras diferentes, em que a distribuição das duas amostras são as mesmas. Onde o Teste de hipótese para H_0 : os dados provêm de amostras iguais e H_1 : os dados provêm de amostras diferentes.

Resultados e Discussão

Observa-se nas Tabelas 1 a 3 os valores mensais de precipitação de 2010 a 2018 correspondentes a cada rede criada com as suas respectivas estimativas classificadas em baixa ($\leq 2,5$ mm), média ($>2,5$ e $<7,5$ mm) e alta ($\geq 7,5$ mm) intensidade de chuvas comparando as diferenças percentuais com os valores reais, assim como, os valores correspondentes ao R^2 e os p-valores do teste de Mann-Whitney para cada mês.

Tabela 1. Rede neural mensal para baixa intensidade de chuvas. Precipitação acumulada real e estimada, diferença percentual entre os valores estimados e a reais, R^2 e p-valor do teste de Mann-Whitney

Mês	Intensidade real	Intensidade estimada	Diferença (%)	R^2 (%)	Mann-Whitney (MW)
Jan	49,5	41,38	16,40	0,49	0
Fev	30,9	23,51	23,92	0,63	0,018
Mar	45,8	54,32	-18,60	0,33	0
Abr	22	17,18	21,91	0,60	0,598
Mai	770,1	789,16	-2,48	0,43	0
Jun	16,4	11,79	28,17	0,73	0
Jul	13,7	12,63	7,81	0,74	0
Ago	8,8	8,51	3,30	0,82	0
Set	15,3	11,11	27,39	0,71	0
Out	32,2	30,04	6,71	0,56	0
Nov	32,2	17,87	44,50	0,51	0
Dez	45,3	36,66	19,07	0,76	0

É possível observar pela Tabela 1, que as redes para baixa intensidade superestimam os meses de março e maio com diferenças de -18,60 e -2,48%, respectivamente, e as demais redes subestimam os valores, encontrando menores diferenças para os meses de julho e agosto, e maiores diferenças para novembro e junho.

Por meio do coeficiente de determinação obtido pelo teste, dentre as três classificações, a de baixa intensidade é a que apresenta, no geral, menores valores de R^2 . Os dois melhores resultados são os obtidos para meses de agosto e dezembro, e os dois menores são para os meses de março e maio.

Pode-se verificar, por meio do teste de Mann-Whitney com nível de significância de 0,05, a diferença significativa entre os valores estimados e o real. Desta forma, para baixa intensidade, apenas o mês de abril apresenta estimativas significativamente iguais aos dados reais.

Conforme encontrado por Silva e Jardim (2017), a transição entre o período seco e úmido representa uma maior dispersão nos dados acumulados de chuva, por apresentarem períodos com baixa precipitação e valores acima da normalidade, resultando em

anomalias que ocorrem de forma desproporcional entre as estações avaliadas o que pode dificultar as estimativas. Como pôde ser verificado nesse estudo onde não foram encontradas boas estimativas para baixas precipitações, sendo indicado utilizar a rede neural criada para preencher os dados correspondentes ao mês de abril.

Souza e Galvani (2017), ao analisarem a precipitação de 1968 a 1998 de 16 postos localizados na bacia do rio Jacaré Guaçu que compõe parte do município de Itirapina, afirmam que essa região apresenta uma grande variabilidade espacial e temporal das chuvas, conforme verifica-se para áreas localizadas nos trópicos. Desta forma, conjunturas de efeito orográfico da região podem ter comprometido as estimativas de baixa intensidade, não sendo possível estimar de forma coerente por meio das redes neurais artificiais.

Na Tabela 2 são apresentados os dados da rede neural para média intensidade correspondente a cada mês e os seus respectivos valores acumulados mensais, diferença percentual, coeficiente de determinação (R^2) e p-valor do teste de Mann-Whitney, todos referentes aos dados reais e estimados de precipitação.

Tabela 2. Rede neural mensal para média intensidade de chuvas. Precipitação acumulada real e estimada, diferença percentual entre os valores estimados e a reais, R^2 e p-valor do teste de Mann-Whitney

Mês	Intensidade real	Intensidade estimada	Diferença (%)	R^2 (%)	Mann-Whitney (MW)
Jan	174,1	172,96	0,65	0,85	0,991
Fev	154,63	153,8	0,54	0,94	0,825
Mar	145,4	148,79	-2,33	0,83	0,714
Abr	66	65,68	0,48	0,98	1
Mai	42,5	41,87	1,48	0,95	0,958
Jun	23,6	22,61	4,19	0,53	1
Jul	40,6	39,69	2,24	0,87	1
Ago	51,2	49,88	2,58	0,98	0,948
Set	59,7	54,1	9,38	0,50	0,436
Out	106,7	107,61	-0,85	0,99	0,982
Nov	124,2	117,31	5,55	0,92	0,464
Dez	156,9	161,08	-2,66	0,68	0,720

As redes de média intensidade de chuva (Tabela 2) apresentam resultados satisfatórios com poucas diferenças percentuais entre os meses, sendo encontrados maiores diferenças para os meses de setembro e novembro, e menores para abril e fevereiro. Com relação ao R^2 , seis dos doze meses apresentam valores superiores a 0,9, e para o teste de Mann-Whitney, todos os meses apontam estimativas significativas indicando que não existe

diferença significativa entre os dados reais e estimados.

Na Tabela 3 pode ser observado as redes neurais criadas para alta intensidade correspondente a cada mês, com os respectivos valores acumulados mensais, diferença percentual, coeficiente de determinação (R^2) e p-valor do teste de Mann-Whitney, todos referentes aos dados reais e estimados de precipitação.

Tabela 3. Rede neural mensal para alta intensidade de chuvas. Precipitação acumulada real e estimada, diferença percentual entre os valores estimados e a reais, R^2 e p-valor do teste de Mann-Whitney

Mês	Intensidade real	Intensidade estimada	Diferença (%)	R^2 (%)	Mann-Whitney (MW)
Jan	2599,7	2690,68	-3,50	0,58	0,810
Fev	1303,3	1276,23	2,08	0,96	0,726
Mar	1390	1409,96	-1,44	0,67	0,630
Abr	735,5	731,69	0,52	0,91	0,945
Mai	710,5	736,83	-3,71	0,99	0,792
Jun	437,8	448,79	-2,51	0,91	0,862
Jul	109,2	111,56	-2,16	0,98	0,958
Ago	145,23	160,3	-10,38	0,68	1
Set	538,8	548,42	-1,79	0,37	0,352
Out	955,11	97,8	-2,06	0,87	0,802
Nov	1290,3	1285,4	0,38	0,79	0,868
Dez	1691,5	949,9	43,84	0,20	0

Para alta intensidade de chuvas, o mês de dezembro apresenta a maior diferença entre os valores estimados e reais, enquanto novembro aponta a menor diferença. Os meses de dezembro e setembro exibem os menores valores de R^2 , enquanto os meses de maio e julho são próximos de 1. Observa-se um p-valor

igual a 1 para o mês de agosto, mas não se obteve resultado satisfatório para o R^2 . Para o teste de Mann-Whitney todos os meses, exceto para dezembro, indicam que não existe diferença significativa entre os dados reais e estimados.

Para todas as análises realizadas, as classificações de média e alta intensidade mostram os melhores resultados. As de baixa intensidade apresentam diversos resultados insatisfatórios, que podem ser devidos aos parâmetros de entrada não conseguirem representar de maneira adequada, no período seco, os dias em que há pouca ou nenhuma precipitação.

O fato de os dados diários de chuva terem uma alta variabilidade espaço/temporal como observado também nos estudos de Barreto et al. (2012), tal fato causa um impacto negativo nas RNA devido à dificuldade de atribuição de pesos dos parâmetros de entrada. Este problema foi suprido por meio da classificação mensal da intensidade, a qual permitiu com que as redes conseguissem encontrar padrões e ajustassem os pesos, o que ocasionou em resultados mais precisos devido a menores variações de dados (WANDERLEY et al., 2014).

Os resultados de classificação de período encontrados neste trabalho são similares ao estudo de Depiné et al. (2014), que utilizaram as redes neurais e dividiram as séries de precipitação em três diferentes alternativas, sendo a primeira com 365 dias de dados, a segunda com séries curtas de 12 a 30 dias e a terceira semelhante a segunda, mas com o acréscimo apenas de análises dos períodos sazonais. O estudo destes autores obteve melhores resultados do preenchimento na terceira alternativa em todos os postos analisados, reproduzindo de forma eficiente os picos de verão.

Para valores de precipitação acima de 2,5mm observa-se valores de R^2 acima de 0,9 e p-valores acima de 0,05, o que indica tanto uma ótima relação entre os dados estimados e os dados reais, quanto a hipótese de que os resultados provêm de amostras iguais em cerca de metade dos meses para cada classificação, sendo possível utilizar estas redes criadas para o preenchimento de falha em dados de precipitação.

Pode-se observar neste estudo que a partir da utilização de dados diários de temperatura mínima, média e máxima, velocidade do vento e umidade relativa foram obtidos altos índices de correlação com a precipitação a partir da aplicação de redes neurais artificiais. Isto foi atingido devido a criação de uma rede para cada mês e para cada intensidade de chuva.

Conclusão

Para valores de precipitação acima de 2,5mm foram encontrados valores de R^2 acima de 0,9 e p-valores acima de 0,05, o que indicou tanto uma ótima correlação entre os dados estimados e os dados reais quanto a hipótese de que os resultados provêm de amostras iguais em cerca de metade dos meses para cada classificação.

Comprovou-se neste estudo que a partir da utilização de dados diários de temperatura mínima, média e máxima, velocidade do vento e umidade relativa foram obtidos altos índices de correlação com a precipitação a partir da aplicação de redes neurais artificiais. Isto foi atingido devido a criação de uma rede para cada mês e para cada intensidade de chuva.

Esse trabalho trouxe contribuições através da criação de RNAs para estudo de preenchimento de falha de dados de precipitação a partir de outras variáveis e para o entendimento da relação dos parâmetros climatológicos com a precipitação, sendo possível esta metodologia ser aplicada em outras regiões através das análises empregadas.

Referências

- BARBOSA, S. E. da S.; BARBOSA JÚNIOR, A. R.; SILVA, G. QUEIROZ da; CAMPOS, E. N. B.; RODRIGUES, V. de C. Geração de modelos de regionalização de vazões máximas, médias de longo período e mínimas de sete dias para a Bacia do Rio do Carmo, Minas Gerais. Engenharia Sanitária e Ambiental, v.10, n.1, p.64-71, 2005.
- BARRETO, H. B. F.; SANTOS, W. de O.; CRUZ, C. M. Análise da distribuição da precipitação pluviométrica média anual no estado do Ceará. Revista Verde, v.7, n.2, p.122-128, 2012.
- CEPAGRI. Centro de Pesquisa Meteorológicas e Climáticas Aplicadas à Agricultura. Clima dos municípios paulistas. 2015. Disponível em: http://www.cpa.unicamp.br/outras-informacoes/clima_muni_560.html. Acesso em: 15 de abril 2019.
- COUTINHO, E. R.; SILVA, R. M.; MADEIRA, J. G. F.; COUTINHO, P. R. de O. dos S.; BOLOY, R. A. M.; DELGADO, A. R. S. Application of artificial neural networks (ANNs) in the gap filling of meteorological time series. Revista Brasileira de Meteorologia, v.33, n.2, p.317-328, 2018.
- GOUVEA, R. L.; CAMPOS, C. C.; MENEZES, J. T.; MOREIRA, G. F. Análise de frequência de precipitação e caracterização de anos secos e chuvosos para a bacia do Rio Itajaí. Revista

Brasileira de Climatologia, v.14, n.22, p.309-323, 2018.

HEMA, N.; KANT, K. Reconstructing missing hourly real-time precipitation data using a novel intermittent sliding window period technique for automatic weather station data. *Journal of Meteorological Research*, v.31, n.4, p.774-790, 2017.

IBGE. Instituto Brasileiro de Geografia e Estatística. 2010. Disponível em: <https://cidades.ibge.gov.br/brasil/sp/itirapina/panorama>. Acesso em: 10 abril 2019.

KLEIBER, W.; KATZ, R. W.; RAJAGOPALAN, B. Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. *Water Resources Research*, v.48, n.1, p.1-17, 2012.

MACHADO, R. E.; MATTOS, A. Construção e instalação de um lisímetro com sistema de drenagem. *Revista Brasileira de Agrometeorologia*, v.9, n.1 p.147-151, 2001.

MAQSOOD, I.; ABRAHAM, A. Weather analysis using ensemble of connectionist learning paradigms. *Applied Soft Computing*, v.7, n.3, p.995-1004, 2007.

MATLAB. Matlab R2015a, 2018. Disponível em: <https://atendimento-prod.sti.usp.br/otrs/public.pl?Action=PublicFAQZoom;ItemID=316>. Acesso em: 26 fevereiro 2019.

MELLO, Y. R.; KOHLS, W.; OLIVEIRA, T. M. N. Uso de diferentes métodos para o preenchimento de falhas em estações pluviométricas. *Boletim de Geografia*, v.35, n.1, p.112-121, 2017.

NKUNA, T. R.; ODIYO, J. O. Filling of missing rainfall data in Luvuvhu River Catchment using artificial neural networks. *Physics and Chemistry of the Earth, Parts a/b/c*, v.36, n.14-15, p.830-835, 2011.

PEEL, M. C.; FINLAYSON, B. L.; MCMAHON, T. A. Updated world map of the Köppen- Geiger

climate classification. *Hydrology and Earth System Sciences*, v.11, n.5, p.1633-1644, 2007.

PERUSSI, R.; VECCHIA, F. A. da S. Análise do comportamento térmico de um sistema de cobertura verde utilizando plataformas de teste. *Pesquisa em Arquitetura e Construção*, v.7, n.4 p.225-236, 2016.

PRUSKI, F. F.; PEREIRA, S. B.; NOVAES, L. F.; SILVA, D. D.; RAMOS, M. M. Precipitação média anual e vazão específica média de longa duração, na Bacia do São Francisco. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.8, n.3, p.247-253, 2004.

REICHARDT, K. A Água em Sistemas Agrícolas. 1. ed. São Paulo: Manole. 1986. 188p.

SANTOS, R. S.; VIERA, P. D.; EVANGELISTA, D. H. R.; OLIVEIRA, L. J. C.; NONATO, D. Caracterização de extremos mensais de precipitação em Cacoal (RO). *Revista Brasileira de Climatologia*, v.22, p.267-280, 2018.

SILVA, A. A. F.; JARDIM, C. H. Aplicação de técnicas de preenchimento de falhas de dados de pluviosidade mensal e anual para o noroeste do estado de Minas Gerais - Brasil. *Geografias Artigos Científicos*, v.15, n.2, p.83-106, 2017.

SNEYERS, R. Sur l'analyse statistique des séries d'observations. Ed. Genève: Organisation Météorologique Mondiale, 1975. 192p.

WANDERLEY, H. S.; AMORIM, R. F. C.; CARVALHO, F. O. Interpolação espacial de dados médios mensais pluviométricos com redes neurais artificiais. *Revista Brasileira de Meteorologia*, v.29, n.3, p.389-396, 2014.

ZANCHETTA, D.; PINHEIRO, L. de S. Análise biofísica dos processos envolvidos na invasão biológica de sementes de *Pinus elliottii* na Estação Ecológica de Itirapina - SP e alternativas de manejo. *Climatologia e Estudos de Paisagem*, v.2, n.1, p.72-89, 2007.