XXIII SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS
ABRHidro
Associação Brasileira de Recursos Hídricos

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

Camila Bermond Ruezzene [1]; Carlos Eduardo Nunes Medina Martinez [2]; Renato Billia de Miranda[3]; Adriano Rogerio Bruno Tech [4]; & Frederico Fábio Mauad [5]

RESUMO – A precipitação é um fenômeno climático de fundamental importância devido ao papel da água com os seres vivos e pelos seus eventos extremos trazerem impactos devastadores. Há uma deficiência nos estudos climatológicos quanto a distribuição das grandezas que influenciam na precipitação por conta da complexidade física envolvida. Para tentar sanar este problema, as Redes Neurais Artificiais (RNAs) se mostram como uma ótima ferramenta devido à sua capacidade preditiva e dos diversos testes que são realizados a cada segundo. O objetivo do trabalho foi correlacionar os dados climatológicos diários de 2010 a 2018 do Centro de Recursos Hídricos e Estudos Ambientais, localizado em Itirapina-SP, divididos por mês em baixa, média e alta intensidade de chuva com os valores de precipitação. Foram feitas análises da diferença dos valores acumulados de precipitação, coeficiente de determinação e teste não paramétrico de Wilcoxon. Para classificações de média e alta intensidade, foram encontrados valores de R^2 acima de 0,9 e p-valores acima de 0,05 para quase metade dos meses. Constatou-se que as RNAs são ferramentas com enorme potencial para o entendimento das grandezas que influenciam na precipitação, mas é necessária a divisão dos dados em séries com menor variabilidade para a obtenção de um bom resultado.

ABSTRACT– Rainfall is a climatic phenomenon of fundamental importance due to the water role with living beings and its extreme events that cause devastating impacts. There is a deficiency in climatological studies regarding the distribution of quantities that influence precipitation due to the physical complexity involved. To address this problem, Artificial Neural Networks (ANNs) are a great tool because of their predictive capacity and the many tests that are performed every second. The goal of this paper was to correlate the daily climatological data of the Center for Water Resources and Environmental Studies, located in Itirapina-SP, divided by month in low, medium and high rainfall intensity with precipitation values. Analyses were made on the difference between the cumulative values of precipitation, coefficient of determination and nonparametric Wilcoxon test. For medium and high intensity classifications, R^2 values above 0.9 and p-values above 0.05 were found for almost half of the months. It was found that ANNs are tools with great potential for understanding the magnitudes that influence precipitation, but it is necessary to divide the data into series with less variability in order to obtain good results.

Palavras-Chave – Precipitação, Redes Neurais Artificiais, Estudos Climáticos.

SYSNO PROD ACERVO EESC

1) Universidade de São Paulo: Av. Trab. São Carlense, 400 - Parque Arnold Schimidt, São Carlos – SP, camila.ruezzene@gmail.com
2) Universidade de São Paulo: Av. Trab. São Carlense, 400 - Parque Arnold Schimidt, São Carlos – SP, carlosenmm@gmail.com
3) Universidade de São Paulo: Av. Trab. São Carlense, 400 - Parque Arnold Schimidt, São Carlos – SP, rbm1706@gmail.com
4) Universidade de São Paulo: R. Duque de Caxias, 225 - Jardim Elite, Pirassununga – SP, adriano.tech@usp.br
5) Universidade de São Paulo: Av. Trab. São Carlense, 400 - Parque Arnold Schimidt, São Carlos – SP, mauadffm@gmail.com

1. 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; KATZ; RAJAGOPALAN, 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. Desta 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; KOLHS; OLIVEIRA, 2017; WANDERLEY; AMORIM; CARVALHO, 2014).

As RNAs são métodos que vem sendo cada vez mais utilizados para modelagem climáticas sendo estas 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, através da classificação da intensidade precipitação, o objetivo deste estudo foi criar redes neurais artificiais que correlacionassem os parâmetros climatológicos com dados de chuva de 2010 a 2018 de uma estação climatológica localizada no município de Itirapina-SP.

2. MATERIAL E MÉTODOS

2.1. Área de estudo

O presente estudo localiza-se no Município de Itirapina com uma população de 15.524 habitantes do último censo realizada em 2010 e uma área territorial de 564,603 km² (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA- IBGE, 2010). Geograficamente, as regiões situam-se na porção central de São Paulo (Figura 1).

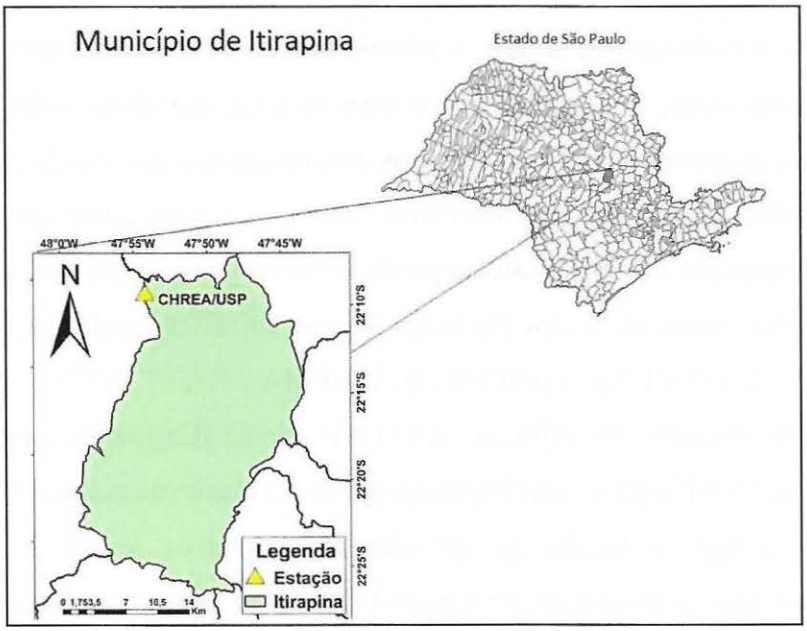

FIGURA 1. Mapa de localização da área de estudo.

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 de 1493 mm, anualmente (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; FINLAYSON; MCMAHON, 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).

2.2. Redes neurais artificiais

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 utilizadas, sendo compreendido 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) e umidade relativa (%) e o como parâmetro de saída a precipitação (mm).

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 e ODIYO, 2011; HEMA e KANT, 2017).

As redes foram treinadas no software MATLAB versão R2015a da empresa MathWorks sendo que o tipo de rede foi a *Feed-forward backpropagation*, a função de treinamento foi Levenberg-Marquardt *backpropagation*, a função de aprendizado adaptativo foi *Gradient descent with momentum weight and bias*, a função de performance foi *mean-squared error*, foram utilizados 2 camadas e 10 neurônios e a função de transferência foi *tan-sigmoid*.

Inicialmente, foi criada uma única rede que recebia os parâmetros da série diária dos nove anos. Como os resultados não se mostraram satisfatórios, foram feitos vários testes, alterando tanto as variáveis que eram inseridas como entrada da rede quanto o período analisado. Os resultados apresentaram certa melhoria quando foram definidos os cinco parâmetros de entrada já mencionados e quando deixamos de criar uma única rede para os valores diários do período de 2010 a 2018 e passamos a criar doze redes com os valores diários para o mesmo período, mas divididas por mês. Como os resultados ainda estavam longe do desejado, foi proposto fazer uma nova divisão que mostrou resultados muito acima do esperados: criaram-se trinta e seis redes dividindo os dados tanto por mês quanto por classificação de intensidade da chuva, adaptada de Reicharrdt (1986), sendo a baixa menor que 2,5 mm; a média maior ou igual a 2,5 mm e menor que 7,5 mm; e a alta maior ou igual a 7,5 mm.

2.3. Testes estatísticos

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 Wilcoxon.

Foram utilizados valores acumulados de precipitação para cada intensidade correspondentes à aos doze meses e após realizamos a diferença entre o acumulado real e estimado

O coeficiente de determinação avalia a qualidade do ajuste de modelo indicando o quanto ele foi capaz de explicar os dados de referência. Foram calculados os coeficientes para as 36 redes criadas comparando os valores de precipitação diários reais com os estimados.

Foi realizado o teste não paramétrico de Wilcoxon com nível de significância de 0,05 para testar se havia diferenças significativas entre o valor estimado pelas redes neurais e o dado real de referência da precipitação. Ao encontrar p-valores menores que o nível de significância, aceita-se H0 e rejeita-se a H1, em que a distribuição das duas amostras são as mesmas.

Teste de hipótese:

H0: os dados provêm de amostras iguais.

H1: os dados provêm de amostras diferentes.

3. RESULTADOS E DISCUSSÕES

Pode-se observar nas tabelas 1, 2 e 3 os valores mensais de precipitação de 2010 a 2018 correspondentes a cada rede criada com as suas respectivas estimativas classificadas em baixa, média e alta intensidade de chuvas comparando as diferenças percentuais com os valores reais, assim como, o valores correspondentes ao R^2 e o p-valor do teste de Wilcoxon para cada mês.

As redes para baixa intensidade superestimaram os meses de março e maio com diferenças de -18,60% e -2,48%, respectivamente, e as demais redes subestimaram os valores, encontrando menores diferenças para os meses de julho e outubro e maiores diferenças para novembro e junho.

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

Pôde-se testar, por meio do teste de Wilcoxon com nível de significância de 0,05, a diferença significativa entre o valor estimado e o real. Desta forma, para baixa intensidade, apenas os meses fevereiro, março e junho apresentaram estimativas significativamente iguais aos dados reais.

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 Wilcoxon.

Mês	Intensidade real	Intensidade estimada	Diferença (%)	R^2	P-valor teste Wilcoxon
Jan	49.5	41.38	16.40%	0,49	0,033
Fev	30.9	23.51	23.92%	0,63	0,65
Mar	45.8	54.32	-18.60%	0,33	0,65
Abr	22	17.18	21.91%	0,60	< 0,0001
Maio	770.1	789.16	-2.48%	0,43	< 0,0001
Jun	16.4	11.79	28.17%	0,73	0,058
Jul	13.7	12.63	7.81%	0,74	< 0,0001
Ago	8.8	8.51	3.30%	0,82	< 0,0001
Set	15.3	11.11	27.39%	0,71	<0,0001
Out	32.2	30.04	6.71%	0,56	<0,0001
Nov	32.2	17.87	44.50%	0,51	0
Dez	45.3	36.66	19.07%	0,76	0,002

As redes de média intensidade apresentaram resultados satisfatórios com poucas diferenças percentuais entre os meses, sendo encontrado maiores diferenças para os meses de setembro e novembro e menores para abril e fevereiro. Com relação ao R^2, seis dos doze meses apresentaram valores superiores a 0,9. Para o teste de Wilcoxon, todos os meses apontaram estimativas significativas, com exceção do mês de novembro.

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 Wilcoxon.

Mês	Intensidade real	Intensidade estimada	Diferença (%)	R^2	P-valor teste Wilcoxon
Jan	174.1	172.96	0.65%	0,85	0,754
Fev	154.63	153.8	0.54%	0,94	0,282
Mar	145.4	148.79	-2.33%	0,83	0,742
Abr	66	65.68	0.48%	0,98	0,978
Maio	42.5	41.87	1.48%	0,95	0,742
Jun	23.6	22.61	4.19%	0,53	1
Jul	40.6	39.69	2.24%	0,87	0,922
Ago	51.2	49.88	2.58%	0,98	0,102
Set	59.7	54.1	9.38%	0,50	0,092

Out	106.7	107.61	-0.85%	0,99	0,823
Nov	124.2	117.31	5.55%	0,92	0,001
Dez	156.9	161.08	-2.66%	0,68	0,28

Por fim, para alta intensidade, o mês de dezembro apresentou a maior diferença entre os valores estimados e reais enquanto maio apontou a menor diferença. Os meses de dezembro e setembro exibiram os menores valores de correlação enquanto os meses de maio e julho foram próximos de 1. Foi encontrado um p-valor igual a 1 para o mês de agosto, mas não foi obtido resultado satisfatório para o R^2, assim como uma grande diferença entre os valores reais e estimados.

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 Wilcoxon.

Mês	Intensidade real	Intensidade estimada	Diferença (%)	R^2	P-valor teste Wilcoxon
Jan	2599.7	2690.68	-3.50%	0,58	0,875
Fev	1303.3	1276.23	2.08%	0,96	0,114
Mar	1390	1409.96	-1.44%	0,67	0,419
Abr	735.5	731.69	0.52%	0,91	0,185
Maio	710.5	736.83	-3.71%	0,99	0,005
Jun	437.8	448.79	-2.51%	0,91	0,099
Jul	109.2	111.56	-2.16%	0,98	0,641
Ago	145.23	160.3	-10.38%	0,68	1
Set	538.8	548.42	-1.79%	0,37	0,2
Out	955.11	974.8	-2.06%	0,87	0,779
Nov	1290.3	1285.4	0.38%	0,79	0,846
Dez	1691.5	949.9	43.84%	0,20	<0,0001

Para todas as análises realizadas, as classificações de média e alta intensidade mostraram os melhores resultados. As de baixa intensidade apresentaram 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 causa um impacto negativo nas RNAs devido à dificuldade de atribuição de pesos dos parâmetros de entrada. Esse 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; AMORIM; CARVALHO, 2014).

Os resultados de classificação de período encontrados neste trabalho foram similares ao estudo de Depiné et al. (2014) que utilizou as redes neurais e dividiu 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 obteve melhores resultados do preenchimento na terceira alternativa em todos os postos analisados, reproduzindo de forma eficiente os picos de verão.

4. CONCLUSÃO

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

Em diversos casos, dados diários de temperatura mínima, média e máxima, velocidade do vento e umidade relativa tiveram altos índices de correlação com a precipitação a partir da aplicação de redes neurais artificias. Isso foi atingido devido a criação de uma rede para cada mês e para cada intensidade de chuva.

Esse trabalho traz contribuições através da criação de RNAs para os possíveis estudos de preenchimento de falhas de precipitação a partir de outras variáveis e para o entendimento da correlação de parâmetros climáticos com a precipitação, na espera de orientar estudos para outras regiões através das análises e metodologias empregadas.

REFERÊNCIAS

CENTRO DE PESQUISA METERIOLÓGICAS E CLIMÁTICAS APLICADAS À AGRICULTURA. Clima dos municípios paulistas. Disponível em:<http://www.cpa.unicamp.br/outras-informacoes/clima_muni_560.html>. Acesso em: 15 de abr. 2019.

COUTINHO, E. R. et al. (2018). "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.

GOUVEA, R. L. et al. (2018) *"Análise de frequência de precipitação e caracterização de anos secos e chuvosos para a bacia do Rio Itajaí"*. Revista Brasileira de Climatologia, ano 14, v. 22, p. 309-323.

HEMA, N.; KANT, K. (2017). 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.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Censo 2010. Disponível em:<https://cidades.ibge.gov.br/brasil/sp/itirapina/panorama>. Acesso em: 10 abr. 2019.

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

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

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

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

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

NKUNA, T. R.; ODIYO, J. O. (2011). *"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.

PEEL, M. C.; FINLAYSON, B. L.; MCMAHON, T. A. (2007). "*Updated world map of the Köppen-Geiger climate classification*". Hydrology and Earth System Sciences, v. 11, n. 5, p.1633-1644.

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

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

SANTOS, R. S. et al. (2018). "*Caracterização de extremos mensais de precipitação em Cacoal (RO)*". Revista Brasileira de Climatologia, ano 14, v. 22, p. 267-280.

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

ZANCHETTA, D.; PINHEIRO, L. de S. (2007). "*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.