

INTERPOLAÇÃO ESPACIAL DE DADOS MÉDIOS MENSAIS PLUVIOMÉTRICOS COM REDES NEURAIS ARTIFICIAIS

HENDERSON SILVA WANDERLEY¹, RICARDO FERREIRA CARLOS DE AMORIM², FREDE OLIVEIRA DE CARVALHO³

¹Universidade Federal Rural do Rio de Janeiro (UFRRJ/IF/DCA), Rio de Janeiro, RJ, Brasil

²Universidade Federal de Alagoas (UFAL/ICAT), Maceió, AL, Brasil

³Universidade Federal de Alagoas (UFAL/DEG), Maceió, AL, Brasil

henderson@ufrj.br, rfcamorim@gmail.com, fredecarvalho@yahoo.com.br

Recebido Abril de 2013- Aceito Outubro de 2013

RESUMO

A falta de informação quanto à distribuição da precipitação é um sério obstáculo para se compreender e modelar sua variabilidade, surgindo assim a necessidade de se obter informações para regiões que não apresentam estações de medição ou que apresentem falhas em seu banco de dados por meio da interpolação. Desta forma, o objetivo deste estudo consiste em utilizar Redes Neurais Artificiais (RNA's), propondo diferentes procedimentos para sua utilização, na interpolação espacial de dados pluviométricos no Estado de Alagoas. Para o estudo foram utilizadas 245 estações pluviométricas localizadas nos Estados de Alagoas e Pernambuco, das quais se usou as informações de latitude, longitude, altitude e precipitação das estações próxima à estação base que se desejou estimar a precipitação, como parâmetros de entrada das redes. A utilização de RNA's, no preenchimento de falhas de dados pluviométricos, mostrou diferença estatística em apenas um procedimento adotado pelas redes. As estimativas realizadas para o mês de novembro apresentou resultados mais coerentes com os observados nas estações bases, devido a menor variabilidade espacial da precipitação neste mês. **Palavra-Chave:** Preenchimento de falhas, geoestatística, variograma, variabilidade espacial.

ABSTRACT: SPATIAL INTERPOLATION OF MONTHLY RAINFALL DATA USING ARTIFICIAL NEURAL NETWORKS

The lack of precipitation data distribution is an obstacle to understand and model its variability, thus to obtain information for regions which do not have meteorological stations or have missing data in your database by interpolation is needed. Thus, the objective of this study consists in using Artificial Neural Networks (ANN's), proposing different procedures for its utilization for spatial interpolation of rainfall data over the Alagoas State. For this study we used 245 rainfall stations located in the States of Alagoas and Pernambuco, from which the information of latitude, longitude, altitude and rainfall near the interpolated station were used as input parameters of the neural networks. Using ANN's in the estimation of lacked rainfall data, showed statistical difference in only one procedure adopted by neural networks. The estimates for the November month showed more consistency with those observed in the base stations, due to lower spatial variability of rainfall during this month.

Keyword: Infilling missing, goestatistical, variograms, spatial variability.

1. INTRODUÇÃO

A falta de informação quanto à distribuição de dados climatológicos é uma limitação frequente para se compreender e modelar a variabilidade espacial e/ou temporal de vários processos meteorológicos e hidrológicos existentes em uma bacia hidrográfica. As variáveis utilizadas nestes processos podem apresentar erros diversos em sua medição, nas quais, provavelmente, há existência de falhas, além de ser um erro comum, é sem sombra de dúvidas, o que requer mais atenção em estudos com séries temporais.

À disponibilidade de uma série de dados contínua e sem falhas nos dados observados, é fundamental para se estabelecer e caracterizar o clima de uma região. A compreensão das condições climáticas reinante em uma localidade, só é possível mediante análise de séries de dados consistentes, as quais são essenciais para observar se o clima está apresentando alguma mudança em decorrência da intensificação do ciclo hidrológico, em função do aumento da temperatura.

Dois abordagens distintas para tratamento de dados que apresentam falhas são destacadas por Piazza et al. (2011). Uma que consiste em utilizar apenas os registros contínuos, descartando as informações anteriores ou posteriores às falhas, enquanto que a outra sugere em ignorar as lacunas, admitindo que os dados sejam uma série de registro contínuo. Embora essas abordagens sejam utilizadas no meio acadêmico, o descarte de dados reduz o período de informação, impossibilitando que algumas inferências estatísticas sejam realizadas com precisão.

Por outro lado, o uso de uma série de dados que apresente falha pode levar a conclusões errôneas, exibindo padrões temporais e espaciais diferentes dos que seriam observados sem a existência de lacunas. No entanto, a nomenclatura atual da estatística de dados ausentes incide em descartar os dados que não sejam pareados, isto consiste na rejeição do banco de dados em que uma das observações esteja ausente e não permite seu pareamento com as demais colunas de observação de diferentes estações.

No entanto, se a análise dos dados que apresentam falhas persistir, uma alternativa válida é o preenchimento das informações ausentes, nos bancos de dados das estações de medições, mediante a interpolação. Essa técnica possibilita a construção de um novo conjunto de dados a partir de um conjunto discreto de dados pontuais previamente conhecidos, permitindo a análise para regiões que não dispõem de registros, em decorrência da ausência de estação de medição, ou que apresentem falhas em seu bando de dados.

Uma questão relevante no processo de preenchimento de falhas ou de interpolação é que a substituição de dados faltantes por valores oriundos de inferências diversas pode alterar a estrutura estatística das amostras da variável em análise.

No entanto, para ambas as situações mencionadas, estimar os dados inexistentes de variáveis como a precipitação ainda é uma tarefa complexa devido à variabilidade espaço-temporal das chuvas e dos processos físicos inerentes ao ciclo hidrológico (Teegavarapu et al., 2012).

A reconstrução de séries temporais incompletas e a interpolação espacial tem sido objeto de estudo de inúmeros trabalhos científicos, onde várias técnicas para estimar valores ausentes foram desenvolvidas. Em vários destes estudos utilizam-se os métodos clássicos determinísticos e estocásticos, comumente apresentados na literatura como o inverso da distância, interpolação não linear e os diversos procedimentos de krigagem (Wanderley et al., 2012). No entanto, a aplicabilidade da Inteligência Artificial, como método preditor de dados inexistentes, vem apresentando resultados satisfatórios (Teegavarapu et al., 2009).

Apesar dos Sistemas Inteligentes serem uma ferramenta importante e eficiente no processo de predição de dados pluviométricos com falhas, a aplicação de redes neurais artificiais (RNA's), até pouco tempo era escassa (Teegavarapu e Chandramouli, 2005). As redes neurais artificiais podem ser entendidas como um conjunto de técnicas computacionais, inspiradas na estrutura intrincada de aprendizagem do cérebro humano, o qual adquire conhecimento por meio de experiências. Assim, as RNA's têm sido amplamente utilizadas para modelar uma série de processos meteorológicos e hidrológicos, sendo o método da inteligência artificial mais utilizado para a predição de dados ausentes.

As redes neurais artificiais por apresentar-se como uma ferramenta estatística não linear, se mostra eficiente no aprendizado de relações complexas entre entradas e saídas, como os processos de simulação precipitação-vazão, precipitação-precipitação, previsão de enchentes, previsão de influxo do reservatório, estimativa e o transporte de sedimentos, qualidade da água, bem como a simulação do impacto das alterações climáticas sobre sistemas de recursos hídricos, meteorológico e agrícola (Maqsood e Abraham, 2007; Ozgur et al., 2009).

As redes neurais se sobressaem dos demais métodos preditivos por serem capazes de generalizar as relações existentes entre os dados a partir de um pequeno subconjunto de informações, permanecendo robustas na presença de ruído existente nas séries de entrada das redes (Nkuna e Odiyo, 2011). Por ter sua aplicação recente na construção de séries temporais, a comparação dos diferentes tipos de redes neurais para o preenchimento de falha em dados de precipitação e temperatura foi realizado por Coulibaly e Évora (2007), os quais destacaram as redes de múltiplas camadas (MLP) como uma das mais eficazes para o preenchimento de falhas de precipitação.

Embora mostre resultados satisfatórios, o desenvolvimento de novas abordagens para reconstrução de dados pluviométricos

com RNA's ainda é o objetivo de estudos devido à precipitação apresentar maior variabilidade, quando comparadas a outros elementos meteorológicos (Kim e Paxhepsky, 2010). Desta forma, o objetivo deste estudo consiste em utilizar Redes Neurais Artificiais (RNA's), propondo diferentes procedimentos para a utilização das RNA's para a interpolação espacial de dados de pluviométricos no estado de Alagoas.

2. MATERIAIS E MÉTODOS

Para esse estudo foram utilizados dados médios mensais de 245 estações pluviométricas distribuídas espacialmente nos estados de Alagoas e Pernambuco, provenientes do banco de dados da Agência Nacional de Águas (ANA), compreendidos entre os anos de 1965 a 1980. Destas estações, 63 estão localizadas no estado de Alagoas e as demais no estado de Pernambuco. Das estações localizadas em Alagoas, cinco foram consideradas como estação base e retiradas do conjunto de dados e classificadas em: A (para a estação Limoeiro de Anadia), B (Mar Vermelho), C (Sebastião Ferreira), D (Batalha) e E (Colônia Pindorama) (Figura 1). As informações geográficas das estações utilizadas para a interpolação espacial e sua respectiva precipitação média mensal, medidas nas estações são apresentadas na Tabela 1.

Para as estações bases foram admitidas a existência de possíveis falhas, onde se buscou preencher as mesmas com o uso das redes neurais artificiais (RNA's). Este procedimento foi realizado para os meses de maio e novembro, meses de máxima e mínima precipitação em Alagoas. Em seguida, os dados estimados pelas RNA's foram confrontados com os respectivos valores medidos nas estações bases, onde foram admitidas falhas. O coeficiente de determinação (r^2), a raiz do erro quadrático médio (REQM), o erro absoluto médio (EAM) e o erro percentual (EP) foram utilizados para avaliar o desempenho e o erro na estimativa dos dados pluviométricos pelas RNA's, Equações 1 a 4, respectivamente:

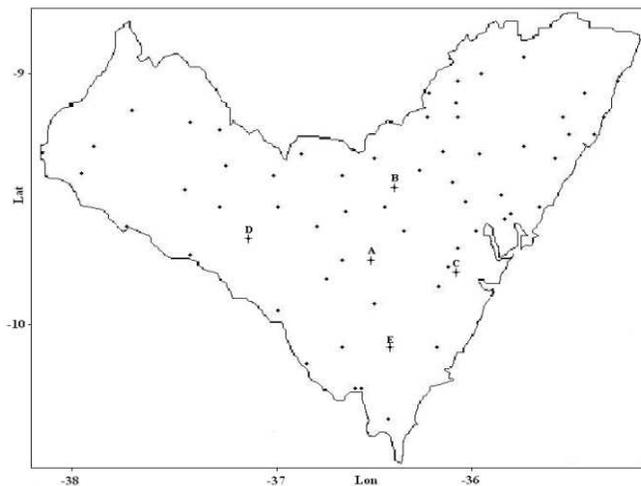


Figura 1 - Distribuição espacial das estações no Estado de Alagoas e a localização das estações bases.

$$r^2 = \frac{\sum_{i=1}^n (O_i - \bar{P}_i)^2}{\sum_{i=1}^n (P_i - \bar{P})^2} \tag{1}$$

$$REQM = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \tag{2}$$

$$EAM = \frac{\sum_{i=1}^n |O_i - P_i|}{n} \tag{3}$$

$$EP = \frac{(O_i - P_i)}{O_i} \cdot 100 \tag{4}$$

em que: O_i são os valores observados nas estações de medição, P_i os valores preditos pelas RNA's, \bar{P}_i os valores preditos pelas RNA's médios e n o número de observações

Para comparar se houve mudança entre os valores medidos nas estações bases e os estimados pelas redes neurais, foi realizado o teste t de Student para verificar se há diferença significativa entre as médias. A hipótese nula (H_0) adotada foi de que as médias mensais antes e depois de serem estimadas pelas

Tabela 1 - Informações geográficas das estações utilizadas para a interpolação espacial e sua respectiva precipitação mensal

Estações	Nome das Estações	Latitude (°)	Longitude (°)	Altitude (m)	Precipitação (mm)	
					Maior	Novembro
A	Limoeiro de Anadia	-9,75'	-36,50	150	168,60	18,79
B	Mar Vermelho	-9,45'	-36,38'	620	204,85	18,89
C	Sebastião Ferreira	-9,80'	-36,06'	0,0	104,05	3,92
D	Batalha	-9,66'	-37,13'	120	104,04	15,72
E	Colônia Pindorama	-10,11'	-36,40'	110	264,55	35,76

redes neurais são iguais, ou seja, a diferença entre elas deve ser supostamente igual à zero. A hipótese de nulidade é rejeitada quando o valor de t calculado é maior, em valor absoluto, ao valor crítico tabelado, a um determinado nível de significância α_0 , sendo adotado neste estudo $\alpha_0 = 0,05\%$.

As redes neurais artificiais apresentam-se como um sistema paralelo distribuído, composto por unidades de processamentos, que calculam funções não lineares, adequadas para processos complexos, como o de estimar dados pluviométricos. Para estimar os dados de precipitação foram utilizadas redes neurais do tipo Perceptrons multicamada (MLP). Este modelo de rede é constituída por uma camada de entrada, compostas pelos neurônios de entrada, por pelo menos uma camada oculta de neurônios computacionais artificiais, e uma camada de saída (Figura 2).

em que: σ é o desvio padrão, μ a média, w os pesos, m a última variável de entrada, Q a saída e f a função gaussiana

Cinco procedimentos distintos foram admitidos como entrada das redes neurais, sendo do primeiro ao quarto procedimento, utilizadas às informações das estações localizadas no estado de Alagoas. Nos procedimentos adotados, as redes neurais tiveram como variáveis de entrada para os processos de treinamento, teste e validação as informações geográficas das estações de medição como latitude, longitude, altitude e precipitação das estações dentro da área de dependência espacial da estação base, a qual foi determinada por análise geoestatístico. Para o quinto procedimento foram utilizadas as estações localizadas nos estados de Alagoas e Pernambuco. Com o aumento do número de estações espera-se melhorar a capacidade de aprendizagem das redes neurais, por existir uma maior quantidade de dados para o período de validação, teste e treinamento.

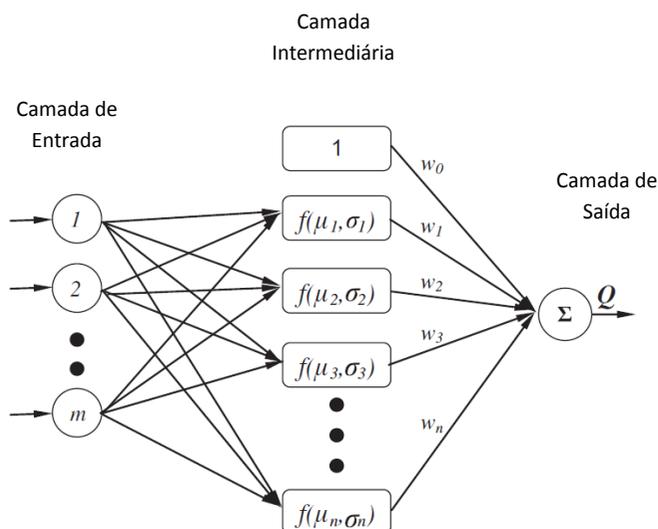


Figura 2 - Configuração de uma rede neural.

Para o primeiro procedimento (I) as redes neurais tiveram como entrada as informações de latitude e longitude das estações de medição dentro da região de dependência espacial, para o período de treinamento das redes, processo pelo o qual as redes neurais buscam encontrar a saída desejada. Para a validação, verificação do desempenho da rede em função do treinamento, e teste foram utilizadas as informações das estações fora de região de dependência espacial. Esse procedimento foi adotado para os demais processos, onde apenas foram alteradas as variáveis de entrada.

O procedimento (II) teve como variáveis de entrada as informações das estações referentes à latitude, longitude e altitude. No procedimento (III), os dados de latitude, longitude e precipitação foram utilizados. O procedimento (IV) utilizou todos os dados de entrada dos demais procedimentos, a saber: latitude, longitude, altitude e precipitação. No procedimento (V) novamente foram usadas as informações de latitude e longitude das estações de medição. No entanto, nesta abordagem foram utilizadas as informações das estações localizadas nos estados de Alagoas e Pernambuco.

A definição das estações dentro da região de dependência espacial da estação base foi obtida com a utilização de técnica geoestatística por meio da análise variográfica. O uso do variograma permite por meio de análise simples de seus parâmetros, patamar, efeito pepita e alcance, definir a região de correlação espacial da variável em estudo, sendo essa ferramenta o traço marcante das características geoestatísticas.

A região de dependência espacial, utilizada para a interpolação da precipitação, foi definida por meio do parâmetro alcance do variograma, o qual define a distância dentro da qual, as amostras apresentam-se correlacionadas espacialmente. Neste estudo foram utilizados os modelos isotrópicos experimentais de variograma do tipo esférico e exponencial, ambos com patamar.

A análise variográfica definiu um alcance de 72 km para o mês de maio, obtido pelo melhor ajuste do variograma esférico e de 80 km para o mês de novembro, decorrente do ajuste do modelo exponencial. Pois, a literatura recomenda que o variograma exponencial utilize apenas 95% do valor alcance (Shao et al., 2007). Desse modo, foi adotada região de dependência espacial de 70 km, padronizada para ambos os meses. A Figura 3 exemplifica a região de dependência espacial obtida para a estação A (Limoeiro de Anadia), onde todas as estações, dentro dessa região de dependência espacial, foram utilizadas pela rede neural para estimar a precipitação da estação A, localizada no centro da circunferência.

3. RESULTADOS E DISCUSSÕES

A interpolação dos dados de precipitação para o estado de Alagoas com a utilização de redes neurais artificiais possibilitou

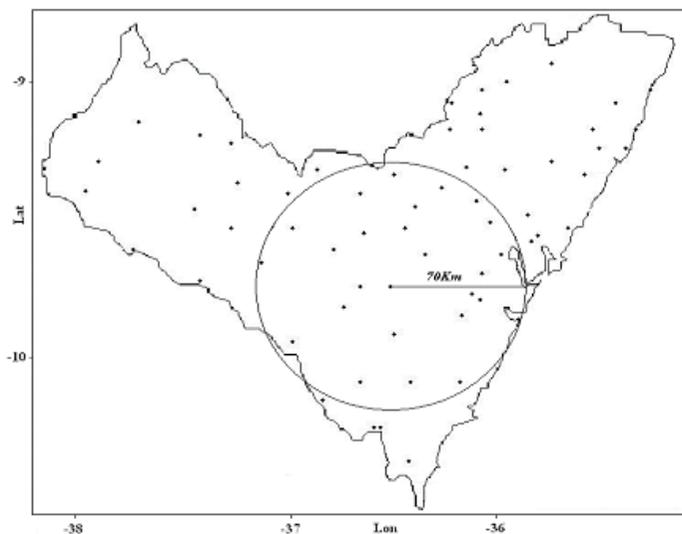


Figura 3 - Região de dependência espacial da estação A.

estimar dados de precipitação em função das informações geográficas e da precipitação observada em estações próximas. A caracterização das chuvas em função das coordenadas geográficas surge como uma alternativa para minimizar erros na interpolação, onde qualquer melhoria na estimativa de precipitação pode ser considerada significativa, uma vez que a precipitação é um insumo fundamental que governa a resposta do sistema hidro meteorológico. Essa aplicação já apresentou resultados satisfatórios na estimativa da precipitação realizada por Mello e Silva (2009) em função da latitude, longitude e altitude.

Neste estudo foi constatado que os dados estimados pelas redes neurais só mostraram diferença estatisticamente significativa na média, entre os dados medidos e estimados, para o mês de maio e para os valores de precipitação estimados pelo procedimento V. Para os demais procedimentos não foi constatada diferença na média entre os valores medidos e estimados. Isso demonstra que a metodologia adotada pode ser

promissora para regiões remotas que não apresentam estações de medição ou que apresentam falhas em seu bando de dados, as quais podem ser estimadas por meio dessas informações.

A principal razão para esses resultados é a capacidade das redes neurais de utilizar as informações geográficas e pluviométricas de outras estações, e incorporar a variabilidade espacial existente no processo de distribuição da precipitação, uma vez que, a análise da variabilidade espacial da precipitação pluviométrica com a altitude apresenta forte dependência espacial e estão diretamente correlacionadas (Fenille e Cardim, 2007).

Os dados pluviométricos estimados pelas redes neurais para o mês de maio, para as estações bases, bem como os valores de precipitação medidos nestas estações, podem ser observados na Tabela 2. Os valores estimados variaram em função do mês em estudo e dos procedimentos utilizados pelas redes neurais. Os resultados obtidos, em sua maioria, estão próximos dos medidos nas estações bases, onde o erro percentual das estimativas variaram de EP = 0,04 a 30,52%, REQM = 36 mm e o EAM = 25,21 mm para o mês de maio.

O menor erro percentual obtido na estimativa foi EP = 0,04%, para a estação C, estimado pelo procedimento III, o qual demonstrou erro absoluto de 0,04 mm entre o valor medido na estação C 104,04 mm e o estimado pela rede neural 104,01 mm. As outras abordagens utilizadas, também apresentaram erro percentual pequeno, como EP = 0,54 e 1,83% obtidos pelo procedimento I para a precipitação das estações C e E. Para o procedimento II, EP = 2,14 e 3,02%, respectivamente para as estações B e D; e EP = 0,39, 2,69 e 0,05% para as estações B, C e D, respectivamente, obtidos pelo procedimento V.

No estado de Alagoas o mês de maio é caracterizado por apresentar a maior variabilidade mensal na distribuição espacial da precipitação, bem como os maiores índices pluviométricos, podendo apresentar precipitação superior a 300 mm na região litorânea e inferior a 100 mm no sertão do Estado. Essa variabilidade pode ter influenciado os resultados obtidos para este período, uma vez que, para esse mês, a amplitude

Tabela 2 - Precipitação estimada pelas RNA's para o mês de maio.

Procedimentos	Estações				
	A	B	C	D	E
	168,60	204,85	104,05	104,04	264,55
I	191,36*	143,00*	103,49*	119,80*	259,72*
II	208,19*	209,23*	72,30*	100,90*	214,32*
III	118,20*	195,16*	104,01*	89,74*	185,10*
IV	213,47*	208,06*	75,55*	87,38*	243,01*
V	117,20*	205,65*	101,25*	104,09*	183,80*

*Precipitação estimada com Redes Neurais Artificiais (mm).

observada, entre os dados medidos nas estações utilizadas para a estimativa da precipitação, foi elevada. Segundo Chiang et al. (2007) a variabilidade espacial da precipitação pode ser causadora de grande incerteza na estimativa de dados.

A utilização das informações geográficas e da precipitação das estações próximas proporcionou coeficientes de determinação elevados e significativos, mostrando que foi possível modelar adequadamente totais precipitados em função de suas coordenadas. Os resultados encontrados neste estudo apresentaram coeficientes de determinação satisfatórios, os quais foram obtidos no processo de interpolação da precipitação pelas redes neurais, variando de $r^2 = 0,72$ a $0,99$ para os meses em estudo, demonstrando ajuste aceitável entre a precipitação estimada pelas RNA's e a observada nas estações meteorológicas de medição.

Em estudo como o de Kim e Paxhepsky (2010), o coeficiente de determinação que apresentou melhor ajuste foi de $r^2 = 0,81$, obtido na interpolação de dados pluviométricos por redes neurais. Teegavarapu e Chandramouli (2005) encontraram coeficiente de correlação de $r = 0,75$ na tentativa de estimar dados de precipitação, e Teegavarapu et al. (2009) obtiveram coeficientes de correlação que variaram de $r = 0,75$ e $0,82$ para $r = 0,95$ e $0,97$, em função da proximidade das estações utilizadas para a interpolação da estação base, apresentando coeficientes maiores nas estações mais próximas da estação base. Vale ressaltar que coeficientes de determinação baixos e pouco significativos poderão influenciar no ajuste dos modelos estatísticos preditores, reduzindo sua dependência espacial.

Os erros observados na interpolação da precipitação com redes neurais, tanto para a raiz do erro quadrático médio (REQM), como para o erro absoluto médio (EAM), apresentaram índices variados. Para o mês de maio foram observadas as maiores dispersões para os dados estimados, com RMSE variando de 29,89 a 41,63 mm, obtidos pelo procedimento IV e V, respectivamente. Para novembro RMSE variou de 0,67 a 4,38, sendo o menor valor decorrente do procedimento II. Para o mês de maio, coeficiente de determinação que apresentou melhor ajuste $r^2 = 0,85$, decorrente do procedimento IV, com raiz do erro quadrático médio de $REQM = 29,89$ mm. O erro absoluto médio variou de $EAM = 21,15$ a $30,77$, com o menor erro para o procedimento V.

Essa maior discrepância entre os dados observado e estimado, para o mês de maio, foi decorrente da maior variabilidade espacial apresentada neste mês. Como as estações selecionadas para usar o método de interpolação da precipitação localizam-se em todas as microrregiões do estado de Alagoas, observou-se a existência de diferença superior a 160 mm, entre as chuvas observadas nas estações D, localizada na zona da mata alagoana e a estação E, situada no sertão. Este é um fator importante que mostra

como ocorre a distribuição espacial dos dados, que podem proporcionar melhoria nos valores estimados pelos modelos. Se a distribuição da precipitação, para uma região, apresenta-se de forma homogênea, a probabilidade dos dados estimados apresentarem essa mesma distribuição é maior, ocorrendo menores erros na estimativa.

O incremento do número de estação utilizado no procedimento V, não apresentou uma redução significativa no erro absoluto obtido na interpolação das estações bases A e D, os quais apresentaram erro percentual de 30,49 e 30,52%, para o mês de maio. No entanto, o erro percentual do procedimento V, obtido para os dois meses em estudo e para as cinco estações, foi inferior a 5% do valor medido nas estações base, para 6 das 10 tentativas de se estimar os dados pluviométricos.

Este resultado demonstra que o uso de longas séries de dados assegura uma melhor generalização temporal e espacial dos dados, reduzindo o risco de excesso na construção de funções não linear para avaliar a estrutura da correlação espacial das observações, melhorando a precisão do método de estimativa. Deste modo, ao aumentar o número de estações ou quantidade de dados, aumentou-se também a capacidade de aprendizagem da rede. Estudos como os de Teegavarapu e Chandramouli (2005) e Teegavarapu et al. (2009), encontraram redução significativa do erro absoluto quando o número de estações utilizadas, para estimar a precipitação no período de treinamento e testes aumentou.

Para o mês de novembro a interpolação simulou resultados mais precisos do que o mês de maio, com maiores coeficientes de determinação $r^2 = 0,89$ a $0,99$ e $REQM = 2,55$ mm e o $EAM = 2,06$ mm. Para este mês a interpolação para a estação C, mediante a utilização do procedimento V, obteve 100% de acerto entre a precipitação medida na estação base e a interpolada pela rede neural (Tabela 3). Para o mês de novembro o procedimento II, foi o que apresentou a melhor estimativa da precipitação, com $REQM = 0,67$ mm e $EAM = 1,22$.

O erro percentual obtido pelos procedimentos II foi $EP = 12,08; 3,18; 1,02; 4,39$ e $6,99\%$, respectivamente para as estações bases A a E. Embora esse procedimento apresente resultados condizentes aos medidos nas estações bases, não minoram a interpolação obtida pelos outros processos interpolativos utilizados pelas redes neurais, os quais apresentaram erro percentual inferiores ao procedimento II para algumas estações. Destes resultados destaca-se o erro percentual obtido para a estação base A de $EP = 0,37\%$, onde o valor medido corresponde a 18,79 mm e o estimado a 18,72 mm; e $EP = 1,46\%$ para a estação D, com uma diferença de 0,23 mm, ambos obtidos pelo procedimento I, ou $EP = 5,92\%$ para a estação D e $1,47\%$ para a estação A, obtidos pelos procedimentos III e IV, respectivamente.

Tabela 3 - Precipitação estimada pelas RNA's para o mês de novembro.

Procedimentos	Estações				
	A	B	C	D	E
	18,79	18,89	3,92	15,72	35,76
I	18,72*	13,10*	3,48*	15,49*	32,34*
II	21,06*	19,49*	3,88*	16,41*	38,26*
III	12,70*	21,00*	2,93*	14,79*	28,19*
IV	18,42*	21,01*	5,45*	13,47*	30,70*
V	19,55*	21,63*	3,92*	15,24*	33,05*

* Precipitação estimada com Redes Neurais Artificiais (mm).

Na estimativa de dados com falhas de precipitação com redes neurais realizada por Marzano et al. (2006), o EAM variaram entre 4,4 a 4,9 mm, já os obtidos por Teegavarapu e Chandramouli (2005) encontraram-se entre AME=0,10 e 58 mm. Feng e Hong (2008) encontraram erro percentual na estimativa da precipitação de 0,44 e 5,8%, próximo ao demonstrado neste estudo. Estes resultados mostram coerência e eficácia da metodologia empregada para se interpolar a precipitação. Entretanto, é bom ressaltar que o erro preditivo de dados pluviométricos, quando comparado com os erros obtidos por outros elementos meteorológicos, apresentam valores superiores para a precipitação, como demonstrando por Xia et al. (1999).

4. CONCLUSÕES

Os valores de precipitação estimados pelas Redes Neurais Artificiais (RNA's) variaram em função do procedimento utilizado para teste, treinamento e validação das redes. No entanto, a utilização de redes neurais para o preenchimento de falhas de dados pluviométricos mostrou diferença estatística apenas para o procedimento V, quando este procedimento foi utilizada pelas redes para a interpolação realizada no mês de maio.

Os dados de precipitação estimada para o mês de novembro apresentou resultados mais precisos do que os estimados para o mês de maio. Esse resultado é decorrente da menor variabilidade espacial da precipitação apresentada no mês de novembro. Para o mês de maio, coeficiente de determinação variou de $r^2 = 0,72$ a $0,85$, apresentou melhor ajuste para o procedimento IV, com raiz do erro quadrático médio de REQM = 29,89 mm, também para esse procedimento, sendo que este apresentou a melhor estimativa para o mês de maio. O erro absoluto médio variou de EAM = 21,15 a 30,77, com o menor erro para o procedimento I.

Para o mês de novembro, o coeficiente de determinação obtido apresentou valor de $r^2 = 0,99$, com REQM = 0,67 mm

e EAM = 1,22, decorrente da utilização do procedimento II, o que apresentou a melhor estimativa dos dados pluviométricos para novembro. O procedimento V apresentou uma precisão de 100% na interpolação da precipitação da estação C, ou seja, o valor medido na estação e o estimado pela rede neural foram idênticos.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- CHIANG, Y. M.; CHANG, F. J.; JOU, B. J. D.; LIN, P. F. Dynamic ANN for precipitation estimation and forecasting from radar observations. **Journal of Hydrology**, v. 334, p. 250-261. 2007.
- COULIBALY, P.; EVORA, N. D. Comparison of neural network methods for infilling missing daily weather records. **Journal of Hydrology**, v. 341, p. 27-41. 2007.
- FENG, L.; HONG, W. On hydrologic calculation using artificial neural networks. **Applied Mathematics Letters**, v. 21, p. 453-458. 2008.
- FENILLE, M. C.; CARDIM, M. Correlação espacial entre a altitude e as precipitações pluviométricas no Estado de São Paulo, **Geociências**, v. 26, n. 2, p. 135-141, 2007.
- KIM, J. W.; PACHEPSKY, Y. A. Reconstructing missing daily precipitation data using regression trees and artificial neural networks for SWAT streamflow simulation. **Journal of Hydrology**, v. 394, p. 305-314. 2010.
- MAQSOOD, I.; ABRAHAM, A. Weather analysis using ensemble of connectionist learning paradigms. **Applied Soft Computing**, v. 7, p. 995-1004. 2007.
- MARZANO, F. S.; FIONDA, E.; CIOTTI, P. Neural-network approach to ground-based passive microwave estimation of precipitation intensity and extinction. **Journal of Hydrology**, v. 328, p. 121-131. 2006.
- MELLO, C. R. e SILVA, A. M. Modelagem estatística da precipitação mensal e anual e no período seco para o estado de Minas Gerais. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 13, n. 1, p. 68-74. 2009.

- 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**, v. 36, p. 830-835. 2011.
- OZGUR, N. B.; KOYUNCU, M.; YAZICI, A. An intelligent fuzzy object-oriented database framework for video database applications. **Fuzzy Sets and Systems**, v. 160, p. 2253-2274. 2009.
- PIAZZA, A.; CONTI, F.; NOTO, L. V.; VIOLA, F.; LOGGIA, G. Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy. **International Journal of Applied Earth Observation and Geoinformation**, v. 13, p. 396-408. 2011.
- SHAO, X.; STEIN, M.; CHING, J. Statistical comparisons of methods for interpolating the output of a numerical air quality model. **Journal of Statistical Planning and Inference**, v. 137, n.7-1, p. 2277-2293. 2007.
- TEEGAVARAPU, R. S. V.; CHANDRAMOULI V. Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. **Journal of Hydrology**, v. 312, p. 191-206. 2005.
- TEEGAVARAPU, R. S. V.; MESKELE, T.; PATHAK, C. S. Geo-spatial grid-based transformations of precipitation estimates using spatial interpolation methods. **Computers & Geosciences**, v. 40, p. 28-39. 2012.
- TEEGAVARAPU, R. S. V.; TUFAIL, M.; ORMSBEE, L. Optimal functional forms for estimation of missing precipitation records. **Journal of Hydrology**, v. 374, p. 106-115. 2009.
- WANDERLEY, H. S.; AMORIM, R. S. C.; CARVALHO, F. O. Variabilidade espacial e preenchimento de falhas de dados pluviométricos para o estado de Alagoas. **Revista Brasileira de Meteorologia**, v.27, p. 347-354, 2012.
- XIA, Y.; FABIAN, P.; STOHL, A.; WINTERHALTER, M. Forest climatology: estimation of missing values for Bavaria, Germany. **Agricultural and Forest Meteorology**, v. 96, p. 131-144. 1999.