

Nº 23, jul/98, p.1-7

REDES NEURAIS PARA PREDIÇÃO DE VARIÁVEIS CLIMÁTICAS RELEVANTES PARA A AGRICULTURA¹

José Dalton Cruz Pessoa²
Luciano Freire³

Este trabalho utilizou a base de dados da estação climatológica da (ESALQ) Escola Superior de Agricultura Luiz de Queiroz, campus da Universidade de São Paulo em Piracicaba (SP) para projetar, treinar e testar algumas redes neurais para predição da temperatura média e precipitação diárias. As redes *feedforward*, treinadas com *backpropagation*, foram comparadas em três categorias: i) predição da temperatura usando valores anteriores (janelas) e a sazonalidade; ii) predição da precipitação usando valores anteriores e a sazonalidade; iii) predição da precipitação a partir de um conjunto de séries (umidade relativa, insolação, temperatura média e precipitação), com janela de 1 dia e sazonalidade. Nas três categorias o tamanho da camada intermediária da rede também foi avaliado a partir do erro quadrático médio e da distribuição do erro relativo. As topologias de melhor desempenho são apresentadas e avaliada a utilização da sazonalidade para predição a partir de múltiplas séries.

A estação climatológica da ESALQ registra dados climatológicos diários desde 1917. Nem todas as séries começaram a ser medidas simultaneamente e todas apresentam falhas de diferentes durações e freqüências. Atualmente a estação fornece valores de temperatura, velocidade e direção do vento, precipitação, umidade relativa do ar e evaporação, o que requer um suporte técnico instrumental e uma boa logística para manter a periodicidade das medidas. Seus dados servem à pesquisa realizada nas escolas de agricultura e no Cena/USP (Centro Energia Nuclear na Agricultura), à comunidade científica e à agricultura da região, cuja produção depende fortemente das condições climáticas, pois afetam a hidratação das plantas, a fotossíntese, o florescimento, o manejo do solo, e na soma de todos os fatores: a produtividade.

O desenvolvimento de técnicas para estimar as medidas faltantes na base de dados amenizaria o rigor da logística de uma estação climatológica, e a previsão dos valores futuros reduziria os riscos da atividade agrícola. Dentre as possíveis abordagens para este problema, as redes neurais artificiais são uma técnica atraente, mas metodologicamente incompleta, em que a determinação de uma

¹ Projeto Embrapa SEP 12.098.800

² dalton@cnpdia.embrapa.br Embrapa Instrumentação CP 741 São Carlos - SP, Brasil.

³ luciano_freire@hotmail.com Embrapa Instrumentação CP 741 São Carlos - SP, Brasil.

CT/23, CNPDIA, jul/98, p.2

topologia eficiente depende muito da tentativa-e-erro. Alguns trabalhos sugerem soluções locais, como por exemplo (Fernandes, 1995), que desenvolveu uma metodologia baseada na decomposição de uma série temporal (sazonalidade, tendência, ciclo e irregularidade) para a construção da camada de entrada de redes do tipo *feedforward* com algoritmo de aprendizado *backpropagation*. Neste trabalho ele comprova a eficiência da decomposição para modelamento da camada de entrada e compara os resultados obtidos através do modelo estatístico UCM (Unobservable Components Model). Outro exemplo é o trabalho de (Abelém, 1994) que acentuou alguns cuidados a serem observados na construção de uma rede, como o pré-processamento dos dados e a escolha de parâmetros que influenciam no treinamento. Ele também utilizou o modelo de *windowing* para descobrir regularidades das séries no treinamento das redes que se mostraram eficazes quando comparadas com o modelo estatístico de Box-Jenkins.

Como guias de aplicação genérica para o planejamento de uma rede neural preditiva pode-se dizer:

a) A camada de entrada pode ser constituída:

- por valores passados da série
- por valores passados da série, referentes ao mesmo mês mas de anos anteriores, visando incorporar informações da tendência ou ciclos da série.
- por componentes de outras séries, preferencialmente correlacionadas com a série a ser prevista.

b) Com grande frequência a camada intermediária é determinada de maneira empírica, mas pode ser auxiliada pela seguinte premissa: "A rede que melhor generaliza os dados é a menor rede apta a realizar o treinamento dos dados"

Outra abordagem, mais promissora, é exemplificada por Zhang et al. (1997). Esses autores construíram um sistema especialista baseado na teoria dos grupos de redes neurais para previsão de chuva que, testado no National Oceanic and Atmospheric Administration (US Department of Commerce), reduziu os erros de previsão de 30% para 10%.

Este estudo introdutório pretende avaliar o comportamento das séries temporais da base de dados da estação climatológica da ESALQ utilizando redes neurais artificiais. As ocorrências de trechos contínuos de dados são identificados e utilizados no treinamento e validação de diversas redes inicialmente planejadas para predição de valores futuros a partir de uma ou mais séries. O efeito da sazonalidade no desempenho das previsões também é avaliado.

Procedimento de Construção e Avaliação das Redes

Todas as redes neurais usadas neste trabalho foram construídas e treinadas utilizando os recursos do simulador SNNS (Stuttgart Neural Network Simulator), versão 4.1 (1995) instalado em uma estação de trabalho AlphaStation, 300Mhz, rodando Unix (Dec OSF/1). As redes, do tipo *feedforward*, foram treinadas com 630 a 712 casos durante 1000 ciclos com o algoritmo *backpropagation*. O Erro Quadrático Médio (MSE) quantificou os erros de validação e treinamento e os Fatores de Análise 10 e 100 (FA10, FA100) ajudaram a avaliar a eficiência das redes. Estes fatores foram calculados da seguinte forma:

1- Calcula-se o erro relativo: subtrai-se o valor esperado (real) do valor previsto, divide-se pelo valor esperado e multiplica-se por 100.

CT/23, CNPDIA, jul/98, p.3

2- Seja 'a' o número de ocorrências de cada erro relativo e 'A' a soma de 'a' até 10%.

3- Seja 'B' o número total de ocorrências de erros percentuais

4- O Fator de Análise até 10% é calculado como:

$$FA10 = A / B$$

5- Se 'C' for calculado como a soma dos valores de 'a' a partir de 100%, o Fator de Análise a partir de 100% é calculado como:

$$FA100 = C / B$$

O trabalho empregou a base de dados do posto agrometeorológico da ESALQ em Piracicaba, São Paulo (Latitude : 22 42' 30" sul, Longitude: 47 38' 30" oeste, Altitude : 546 m). Foram escolhidos os dados de 94 a 95 para treinamento e os de 97 para validação, por formarem conjuntos relativamente extensos de todas as séries relevantes. As séries originais foram usadas sem nenhum tratamento prévio, exceto a normalização. A sazonalidade de 12 meses, a inexistência de tendência e os altos valores do ruído foram observados graficamente.

Três ensaios foram realizados para testar a eficácia das redes na previsão *single-step* da temperatura média e precipitação diárias: i) predição da temperatura usando valores anteriores (janelas) e a sazonalidade; ii) predição da precipitação usando valores anteriores e a sazonalidade; iii) predição da precipitação a partir de um conjunto de séries (umidade relativa, insolação, temperatura média e precipitação), com janela de 1 dia, e da sazonalidade.

As redes neurais foram criadas com uma camada de entrada, uma intermediária, e uma de saída. A camada de saída era composta por apenas uma unidade, enquanto a camada intermediária poderia possuir 1,2,3,5,7,9,12,15,18 e 21 unidades e a quantidade de unidades de entrada dependia do ensaio.

Previsão dos Valores da Temperatura Média

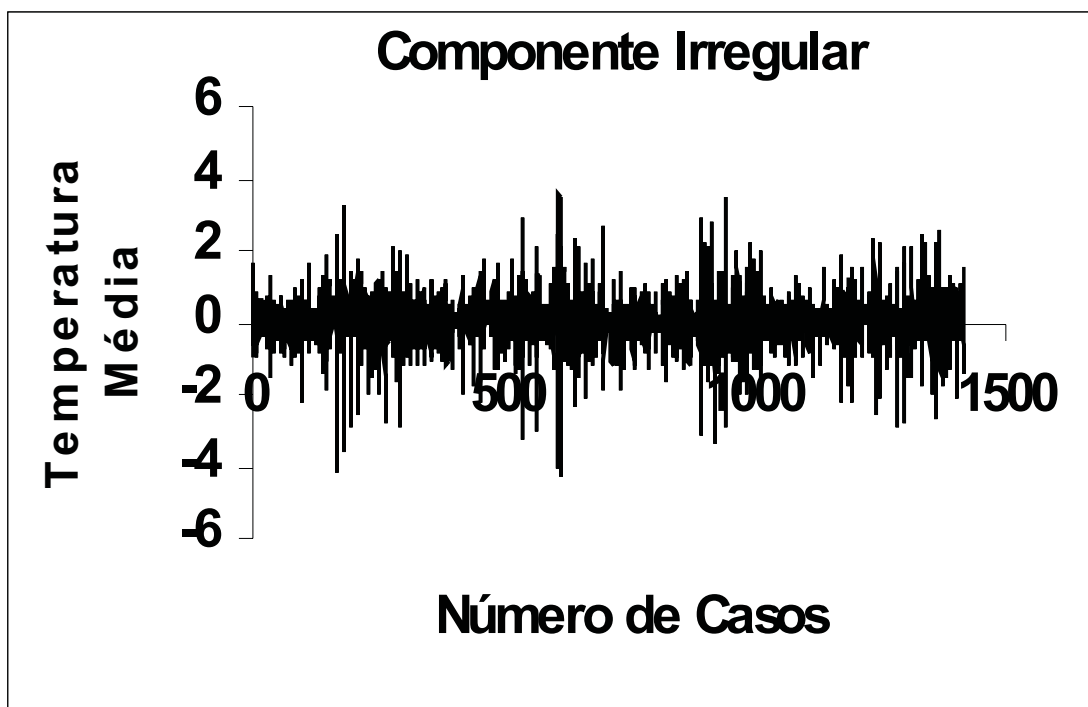


Figura 1 - gráfico da componente 'irregularidade' da série de temperatura média

CT/23, CNPDIA, jul/98, p.4

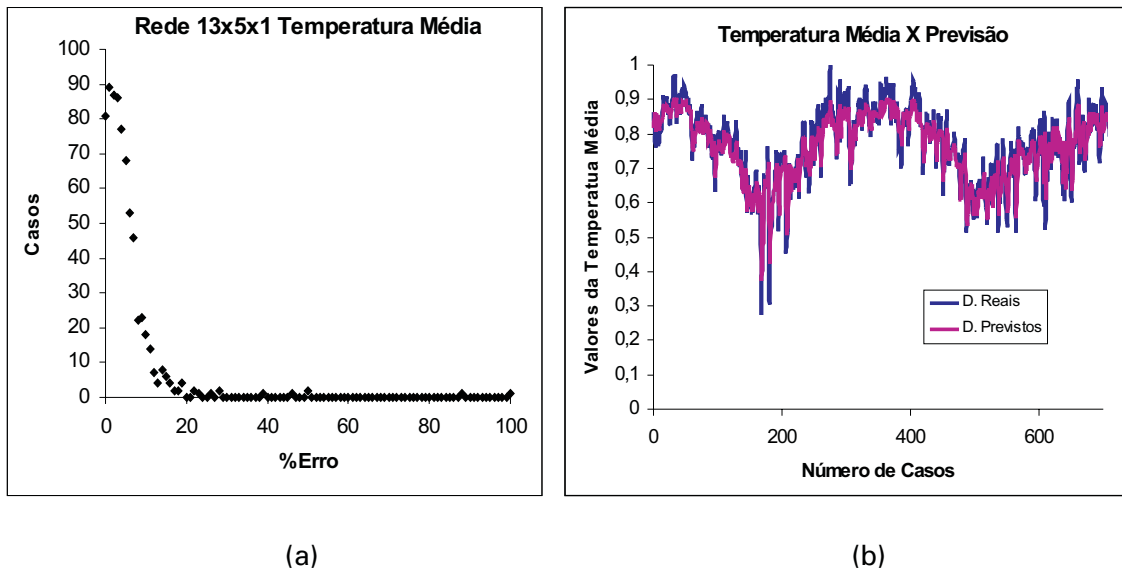


Figura 2 (a) Distribuição do erro relativo da rede com melhor desempenho (maior FA10) para previsão da temperatura média. Ela é formada por 13 unidades de entrada (12 para sazonalidade e 1 dia de janela) e 5 unidades intermediárias. (b) Comparação entre os valores originais de temperatura média e os valores obtidos utilizando uma rede neural (13x5x1) para previsão dos valores de temperatura média.

A série da precipitação é caracterizada por grandes variações de amplitude entre dias consecutivos criando, na estação chuvosa, um padrão formado por um conjunto de dias com precipitação intensa seguido de dias com pouca ou nenhuma chuva, ou ainda um conjunto de dias com precipitação nula interrompido por um dia chuvoso. Na estação seca são encontrados longos períodos com precipitação zero. Outra característica marcante dessa série é a sazonalidade, sincronizada com o comportamento anual das estações.

A previsão de valores futuros utilizando a própria série foi feita para um conjunto de redes neurais que possuíam na camada de entrada 12 unidades, representando a sazonalidade na forma de um número binário que indicava a qual mês do ano aquele valor ou valores se referiam, e um conjunto de unidades representado valores passados, compondo janelas de 1,3,5 ou 7 dias. Neste conjunto de redes observa-se uma grande dispersão do erro relativo (Fig 1). As redes não conseguiram obter um desempenho satisfatório, nem mesmo com o aumento da janela (FA10 variou entre 32,01% e 37,59%).

Na previsão a partir de múltiplas séries a insolação não apresentou qualquer tipo de influência na previsão da precipitação o que, possivelmente, se deve à baixa correlação entre a série de precipitação e a de insolação. A informação sobre sazonalidade da precipitação, proposta em outro trabalho (Fernandes, 1995) para melhorar a previsão de redes univariadas, quando acrescida aos valores anteriores de temperatura média e umidade relativa também melhorou o desempenho de redes multivariadas (FA10 = 41,99%). A introdução de valores históricos da precipitação na camada de entrada provocou uma diminuição drástica no desempenho, pois os valores de precipitação aumentam a irregularidade dos dados de entrada dificultando a previsão da rede, anterior ao ponto de tornar esta a rede com o pior desempenho na previsão de precipitação.

CT/23, CNPDIA, jul/98, p.5

Previsão dos Valores de Precipitação

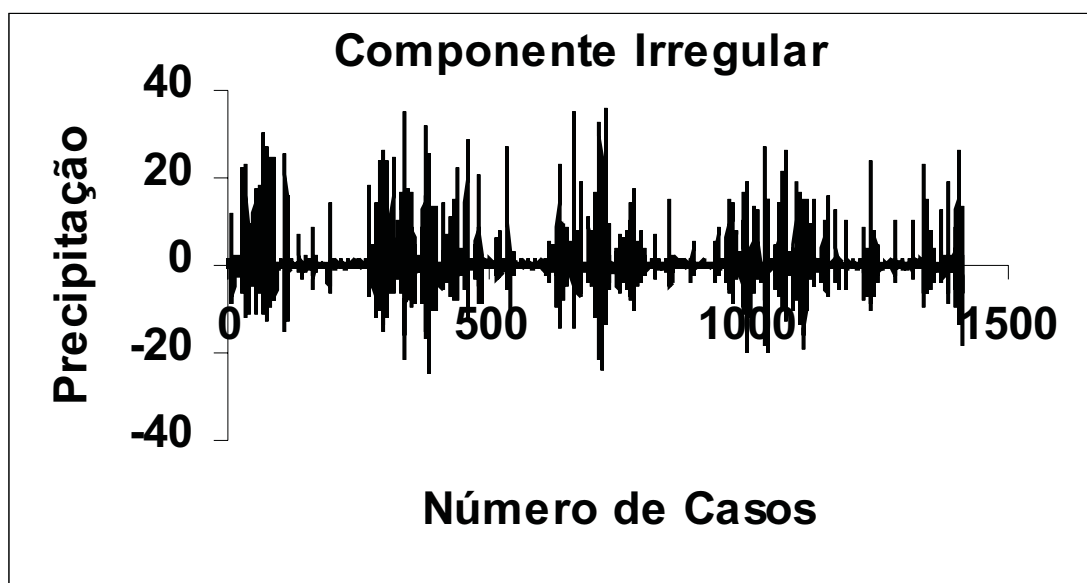
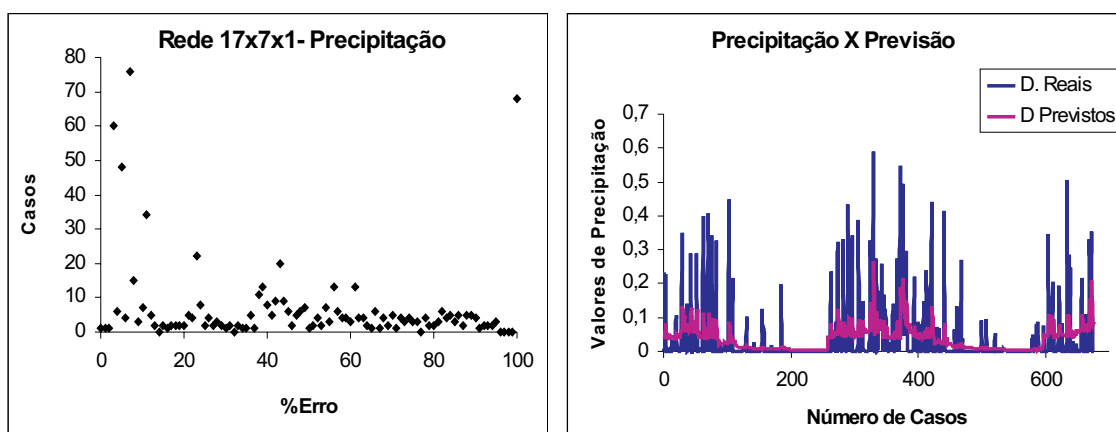


Figura 3 - gráfico da componente irregularidade da série de precipitação



(a)

(b)

Figura 4 (a) Distribuição do erro relativo da rede com melhor desempenho (maior FA10) para previsão da precipitação a partir da sazonalidade (12 nós) e uma janela de 5 dias, (b) Comparação entre os valores originais de precipitação e os valores obtidos utilizando uma rede neural (17x7x1) para previsão dos valores de precipitação

Dentre as estudadas a temperatura média é a série com menor variação diária. O cálculo da média produz uma suavização dos valores da série, o que melhora o desempenho da rede.

As redes utilizadas para fazer a previsão dos valores de temperatura média diários tinham na camada de entrada um conjunto de unidades representando as informações sobre a sazonalidade da série e o valores históricos, com janelas de 3,5 e 7 dias.

CT/23, CNPDIA, jul/98, p.6

Tabela1 - características das melhores redes para previsão da temperatura média e da precipitação

Rede Neural	Entrada	MSE	FA10	FA100
Previsão de Precipitação Utilizando Múltiplas Séries				
2x7x1	Umd rel Temp med	0.01716	0.3174	0.1671
3x9x1	Insol, Umd rel, Temp med	0.01703	0.3174	0.1629
14x3x1	Umid rel, Saz, Temp med	0.01590	0.4199	0.1306
15x2x1	Umid rel, Saz, Precip, Temp med	0.02135	0.1938	0.1910
Previsão de Precipitação Utilizando a Própria Série de Precipitação				
17x7x1	Precip Janela 5 d Saz	0.00681	0.3279	0.1004
Previsão de Temperatura Média Utilizando a Própria Série de Temperatura Média				
13x5x1	Temp méd Janela 1 d Saz	0.00279	0.9116	0.0014

Metodologia de Avaliação da Eficiência

A análise de eficiência a partir dos Fatores de Análise (FA10 e FA100) da distribuição dos erros relativos (fig. 1,2 e 3) estabelece de forma pragmática a aplicabilidade dos resultados. Para a temperatura, por exemplo, 91,16% das previsões estão erradas em menos de 11%, o que equivale a aproximadamente 1 ou 2 graus centígrados. Em condições climáticas favoráveis erros desta ordem não afetam a maioria dos processos agrícolas, mas se a cultura estiver sob estresse térmico uma variação de 1 ou 2 graus pode determinar uma perda significativa para o produtor.

O fato de ter-se treinado as redes com a temperatura média e não com temperatura máxima ou mínima, dilui um pouco o efeito destas conclusões, mas apresenta a vantagem de criar um contraponto aos resultados para a precipitação, cuja série apresenta uma componente irregular muito maior. Aparentemente o gráfico da fig. 1 mostra outros efeitos desta irregularidade: a maior dispersão dos erros, uma baixa taxa de acerto (FA10 = 32,79%) e uma elevada taxa de erros relativos a partir de 100% (FA100 = 10,04%) . Com este desempenho é pouco provável a aplicação desta metodologia em um problema prático

A fig. 2 mostra a dispersão do erro para uma rede multivariada, treinada a partir da umidade relativa, sazonalidade, precipitação e temperatura média. Com o gráfico um pouco menos disperso que o da fig. 1, apresenta porém um maior FA100 (19,10%) e um menor FA10 (19,82%), indicando que a abordagem

CT/23, CNPDIA, jul/98, p.7

multivariada apresentou resultados piores que a univariada na previsão da precipitação. A menor dispersão talvez possa ser atribuída às informações introduzidas pelas séries de entrada, que não foram suficientes para se sobrepor ao efeito da componente irregular. As correlações cruzadas entre a umidade relativa, a temperatura média e a precipitação não são importantes, mas são maiores que para as outras séries da base de dados.

Conclusões

A sazonalidade melhorou a eficiência das redes multivariadas.

Nos casos estudados o aumento da janela não melhorou significativamente as predições.

A utilização das séries com uma grande componente irregular na camada de entrada reduziu a eficiência das redes.

Para previsão dos índices pluviométricos a abordagem univariada apresentou os melhores resultados, mas não o suficiente.

As previsões de temperatura média foram as melhores obtidas neste trabalho, com perspectivas de tornar-se útil em aplicações agropecuárias.

Referências

- ABELÉM, A.J.K. **Redes neurais artificiais na previsão de séries temporais**. Rio de Janeiro: PUC/RJ, 1994. 100p. Dissertação mestrado.
- FERNANDES, L.G.L. **Utilização de redes neurais na análise de previsão de séries temporais**. Porto Alegre: UFRGS, 1995. 76p. Dissertação mestrado.
- ZHANG, M.; FULCHER, J.; SCOFIELD, R. A. Rainfall estimation using artificial neural network group. **Neurocomputing**, New York, v.16, p.97-115, 1997.
- SNNS: Stuttgart neural network simulator: user manual, version 4.1; <http://www.rhrk.uni-kl.de/WWW-aix/Dokumentations/SNNSinfor/UserManual/UserManual.html>. 1995.