

Rede neural artificial para detecção de sobremortalidade atribuível à cólera no Ceará

Use of an artificial neural network for detecting excess deaths due to cholera in Ceará, Brazil

Maria Lúcia F Penna

Escola Nacional de Saúde Pública. Departamento de Endemias Samuel Pessoa. Rio de Janeiro, RJ, Brasil

Descritores

Redes neurais (computação). Séries de tempo. Previsões. Cólera, epidemiologia. Vigilância epidemiológica.

Resumo

Objetivo

Avaliar as redes neurais recorrentes enquanto técnica preditiva para séries temporais em saúde.

Métodos

O estudo foi realizado durante uma epidemia de cólera ocorrida no Estado do Ceará, em 1993 e 1994, a partir da sobremortalidade tendo como causa básica as infecções intestinais mal definidas (CID-9). O número mensal de óbitos por essa causa, referente ao período de 1979 a 1995 no Estado do Ceará, foram obtidos do Sistema de Informação de Mortalidade (SIM) do Ministério da Saúde. Estruturou-se uma rede com dois neurônios na camada de entrada, 12 na camada oculta, um neurônio na camada de saída e um na camada de memória. Todas as funções de ativação eram a função logística. O treinamento foi realizado pelo método de *backpropagation*, com taxa de aprendizado de 0,01 e momentum de 0,9, com dados de janeiro de 1979 a junho de 1991. O critério para fim do treinamento foi atingir 22.000 epochs. Compararam-se os resultados com os de um modelo de regressão binomial negativa.

Resultados

A predição da rede neural a médio prazo foi adequada, em dezembro de 1993 e novembro e dezembro de 1994. O número de óbitos registrados foi superior ao limite do intervalo de confiança. Já o modelo regressivo detectou sobremortalidade a partir de março de 1992.

Conclusões

A rede neural se mostrou capaz de predição, principalmente no início do período, como também ao detectar uma alteração concomitante e posterior à ocorrência da epidemia de cólera. No entanto, foi menos precisa do que o modelo de regressão binomial, que se mostrou mais sensível para detectar aberrações concomitantes à circulação da cólera.

Keywords

Neural networks (computer). Time series. Forecasting. Cholera, epidemiology. Epidemiologic surveillance.

Abstract

Objective

To evaluate recurrent neural networks as a predictive technique for time-series in the health field.

Methods

The study was carried out during a cholera epidemic which took place in 1993 and

Correspondência para/ Correspondence to:

Maria Lúcia F Penna
Escola Nacional de Saúde Pública
Departamento de Endemias Samuel Pessoa
Rua Leopoldo Bulhões, 1480 Térreo
21041-210 Rio de Janeiro, RJ, Brasil
E-mail: mlpenna@ensp.fiocruz.br

Recebido em 22/10/2002. Reapresentado em 28/11/2003. Aprovado em 19/1/2004.

1994 in the state of Ceará, northeastern Brazil, and was based on excess deaths having 'poorly defined intestinal infections' as the underlying cause (ICD-9). The monthly number of deaths with due to this cause between 1979 and 1995 in the state of Ceará was obtained from the Ministry of Health's Mortality Information System (SIM). A network comprising two neurons in the input layer, twelve in the hidden layer, one in the output layer, and one in the memory layer was trained by backpropagation using the first 150 observations, with 0.01 learning rate and 0.9 momentum. Training was ended after 22,000 epochs. We compare the results with those of a negative binomial regression.

Results

ANN forecasting was adequate. Excessive mortality (number of deaths above the upper limit of the confidence interval) was detected in December 1993 and October/November 1994. However, negative binomial regression detected excess mortality from March 1992 onwards.

Conclusions

The artificial neural network showed good predictive ability, especially in the initial period, and was able to detect alterations concomitant and a subsequent to the cholera epidemic. However, it was less precise than the binomial regression model, which was more sensitive to abnormal data concomitant with cholera circulation.

INTRODUÇÃO

A predição de eventos em vigilância epidemiológica tem como objetivo a projeção de necessidades futuras para a saúde pública, ou a detecção de uma alteração de comportamento de uma série temporal, indicando a ocorrência de excesso de casos ou óbitos. As técnicas utilizadas incluem o ajuste a funções, os modelos ARIMA e a regressão de Poisson.^{1,4,19,16}

Um dos objetivos da vigilância epidemiológica é a detecção de epidemias, seja pela análise de dados de notificação de doenças, seja pela análise do uso dos serviços de saúde ou da mortalidade. Do ponto de vista estatístico, detectar uma epidemia é detectar a presença de valores aberrantes em uma série histórica. A detecção de aberrações com base na notificação dos últimos cinco anos¹⁸ é utilizada em alguns países na prática rotineira da vigilância epidemiológica, como é o caso dos EUA. Já quando se trata de detecção de sobremortalidade, geralmente são utilizados períodos maiores de tempo.

A sobremortalidade foi introduzida na vigilância epidemiológica na avaliação do impacto das epidemias de influenza,⁵ devido à dificuldade de classificar os óbitos como atribuídos à doença, já que a maioria era decorrente de complicações como a pneumonia. Esse conceito de excesso de mortes também foi utilizado para avaliar o impacto de episódios de calor, poluição atmosférica e epidemias aparentemente benignas.

A vigilância epidemiológica no Brasil tem maior variabilidade operacional do que nos países desenvolvidos, resultando em baixa precisão (maior va-

riância) dos dados e menor sensibilidade do sistema de notificação de detecção de novas doenças e da ocorrência de epidemias. Assim, a utilização de técnicas de sobremortalidade poderia ser útil para o aumento de sensibilidade do sistema de vigilância, uma vez que a cobertura do sistema de mortalidade é mais estável do que o de notificação, e a série histórica disponível, mais longa.

No entanto, os métodos estatísticos tradicionalmente utilizados exigem conhecimento estatístico aprofundado para seleção e avaliação dos modelos, dificultando sua aplicação descentralizada.^{3,17} As redes neurais artificiais (RNA) têm sobre essas técnicas a vantagem de poder serem aplicadas a várias séries históricas em seqüência, sem prévio diagnóstico de seu comportamento, obtendo bons resultados.^{12,21}

A maioria das séries de interesse em saúde pública tem sua maior variabilidade atribuível à tendência e à sazonalidade, e um comportamento não linear, além de apresentar ciclos não regulares. A vantagem da utilização das redes neurais como um aproximador de funções não lineares é que essa técnica vem tendo sucesso na análise de séries em que o conhecimento matemático do processo estocástico que a gera é desconhecido e difícil de ser racionalizado.^{2,10}

As redes neurais se desenvolveram inicialmente como uma estratégia de simular os processos mentais humanos, como reconhecimento de imagens e sons, e depois como instrumento tecnológico eficiente para inúmeras tarefas.

Uma rede neural é composta de neurônios ou nós e

suas conexões, que podem ser classificadas como entradas e saídas em relação a cada neurônio individual. O nó ou neurônio corresponde à etapa que executa o processamento matemático. Esse processamento consiste em dois passos: o primeiro é a soma ponderada das entradas ($\sum W_i X_i$) e o segundo, a aplicação de uma função de ativação a essa soma, gerando a saída do neurônio, que pode se constituir em entrada para outros neurônios. Essa função é, em geral, a função logística ou a tangente hiperbólica. Essas funções têm forma sigmoide, apresentando variações muito pequenas para valores de x extremos, o que simula a saturação de um neurônio biológico quando os estímulos de entrada são muito grandes.

As RNA podem ser descritas como uma estratégia de modelagem matemática de problemas, concebidos como sistemas com entradas e saídas. Ao contrário de outras estratégias de modelagem, não é necessário conhecer a relação matemática entre as entradas e saídas. Assim, ao contrário da regressão múltipla, por exemplo, não é preciso propor uma função para o modelo,¹¹ uma vez que algumas redes neurais são aproximadoras universais de funções. A rede solucionaria qualquer problema que ela possa representar.

O objetivo do presente estudo é avaliar a adequação das redes neurais recorrentes enquanto técnica preditiva para séries temporais em saúde.

MÉTODOS

O estudo foi realizado durante uma epidemia de cólera ocorrida no Estado do Ceará, em 1993 e 1994, a partir da sobremortalidade tendo como causa básica as infecções intestinais mal definidas (CID-9).

A sobremortalidade foi atribuída a casos letais de cólera não diagnosticados.⁸ O Ceará foi escolhido por ter apresentado durante dois anos consecutivos as maiores taxas anuais de incidência de cólera já registradas no País, 346,19 e 302,74 por 100.000 habitantes em 1993 e 1994, respectivamente.

O número mensal de óbitos ocorridos com causa básica "infecções intestinais mal definidas" (CID-9) de 1979 a 1995, no Sistema de Informação de Mortalidade (SIM) do Ministério da Saúde.

Como técnica estatística para comparação, foi utilizada a regressão binomial negativa, bastante comum em dados tipo conta, quando a regressão de Poisson não se mostrou adequada.¹⁵

Estruturou-se uma rede recorrente com dois neurônios na camada de entrada, correspondentes ao ano

e mês, doze neurônios na camada oculta, um na camada de saída e outro na camada de memória. A conexão recorrente liga a camada de saída à camada de memória, que por sua vez é ligada à camada oculta. Todas as funções de ativação eram a função logística. O treinamento foi realizado pelo método de *back propagation*, com taxa de aprendizado de 0,01 e momentum de 0,9. O critério para fim do treinamento foi atingir 22.000 epochs. Utilizou-se o programa NeuroShell[®].²³

Para o treinamento da rede e ajustamento dos modelos foram utilizados os dados de janeiro de 1979 a junho de 1991, correspondendo a 150 observações. A escolha visou a garantir a ausência de circulação do *Vibrio colerae* no período utilizado para o ajuste, já que o primeiro caso notificado no Estado do Ceará foi em fevereiro de 1992. Pode ter havido casos nos meses imediatamente anteriores, e não detectados.

Foram preditos dados a partir da rede para agosto de 1991 a dezembro de 1995, correspondendo a 54 meses. Optou-se por não prolongar a extrapolação para garantir a homogeneidade da série, devido à mudança para a CID-10 em janeiro de 1996. O intervalo de confiança do modelo implícito da RNA foi estimado a partir da distribuição dos resíduos do período utilizado no treinamento, assumindo que apresentam distribuição normal de média zero. Utilizou a técnica de *bootstrap* com 2.000 amostras para avaliar a adequação dos parâmetros diretamente estimados dos 150 resíduos.

Os dados foram também ajustados a um modelo de regressão binomial negativa, sendo os óbitos função da ordem de ocorrência, representando o mês-ano, e do mês de ocorrência, transformado em uma variável *dummy*, representando o componente sazonal, após a verificação de excesso de dispersão (*overdispersion*) em uma regressão de Poisson. Utilizou-se o programa Stata.¹⁹

RESULTADOS

Na Tabela 1 encontra-se o resultado da regressão binomial negativa. O mês de junho foi retirado do modelo por colinearidade, ou seja, junho, segundo o modelo ajustado, é o mês base para a definição dos demais parâmetros, que expressam o componente sazonal. Os meses de julho a novembro não apresentaram significância estatística, mas foram mantidos para garantir a validade do modelo.

Os parâmetros relativos aos resíduos da rede neural são apresentados na Tabela 2, onde se observa que os intervalos de confiança da média incluem o zero.

Tabela 1 - Resultado da regressão binomial negativa: parâmetros e significância estatística

| Série | Coef. | Erro-padrão | z | p>z | Intervalo de confiança 95% | |
|-------|----------|-------------|--------|-------|----------------------------|------------|
| | | | | | L inferior | L superior |
| Ordem | -0,00893 | 0,000523 | -17,06 | 0,000 | -0,00995 | -0,0079 |
| Jan | 0,550417 | 0,10777 | 5,11 | 0,000 | 0,339193 | 0,761642 |
| Fev | 0,788315 | 0,107209 | 7,35 | 0,000 | 0,578189 | 0,99844 |
| Mar | 0,919921 | 0,10693 | 8,6 | 0,000 | 0,710342 | 1,129501 |
| Abr | 0,724294 | 0,107387 | 6,74 | 0,000 | 0,513819 | 0,934768 |
| Mai | 0,371552 | 0,108558 | 3,42 | 0,001 | 0,158782 | 0,584322 |
| Jul | -0,12263 | 0,112787 | -1,09 | 0,277 | -0,34368 | 0,098431 |
| Ago | -0,03607 | 0,112405 | -0,32 | 0,748 | -0,25638 | 0,184244 |
| Set | -0,18182 | 0,11329 | -1,6 | 0,109 | -0,40387 | 0,040219 |
| Out | -0,07642 | 0,112647 | -0,68 | 0,498 | -0,2972 | 0,144366 |
| Nov | 0,047475 | 0,112008 | 0,42 | 0,672 | -0,17206 | 0,267007 |
| Dez | 0,236944 | 0,111021 | 2,13 | 0,033 | 0,019347 | 0,454541 |
| Cons | 4,949957 | 0,087621 | 56,49 | 0,000 | 4,778223 | 5,121691 |
| Alpha | 0,061099 | 0,008092 | | | 0,04713 | 0,079207 |

Teste da razão de verossimilhança para $\alpha=0$: $\chi^2_1=851.78$ $p<=0.000$

Razão de verossimilhança $\chi^2_{12}=250.16$ $p<0.0000$

Log verossimilhança = -690.65416

Tabela 2 - Parâmetros dos resíduos dos 150 pontos utilizados no treinamento da rede neural.

| Parâmetro | Valor observado | IC do parâmetro tradicional | Parâmetro nas amostras do bootstrap | IC estimado pelo bootstrap |
|---------------|-----------------|-----------------------------|-------------------------------------|----------------------------|
| Média | 0,57234 | -4,3530:5,4976 | 0,3766 | -4,198:5,343 |
| Mediana | -0,67900 | -5,048:3,310 | — | -4,964:3,606 |
| Desvio-padrão | 30,527 | | | 25,23:35,83 |

A Figura 1 apresenta as estimativas dos dois modelos no período utilizado para o ajuste. Observa-se uma boa concordância entre os dois, com coeficiente de correlação de Pearson de 0,95. A diferença entre as duas estimativas teve como média 0,544740, com intervalo de confiança a 95% de -2,94971 a 4,039190. A Figura 2 mostra as estimativas dos dois modelos extrapoladas para julho de 1991 a dezembro de 1995. Nesse caso existe menor concordância entre as estimativas, com coeficiente de correlação de Pearson de 0,92, sendo a diferença com média de -6,56195 e intervalo de confiança a 95%, de -13.9088 a 0,784901. A estimativa da RNA é maior do que a da regressão e mais próxima do observado de dezembro a maio, estação de maior ocorrência de óbitos, sendo a estimativa da regressão maior e mais próxima do observado nos demais meses.

A Figura 3 apresenta o limite superior do intervalo de confiança das duas estimativas, os dados observados e a ocorrência de cólera, de julho de 1991 a dezembro de 1995. O modelo regressivo detectou sobremortalidade em março-abril de 1992, logo após a detecção dos primeiros casos de cólera no Estado, em fevereiro, e também em outubro-novembro de 1992. A partir de fevereiro de 1993, todos os pontos estão acima do limite superior do modelo regressivo, com exceção de agosto e setembro de 1994. Considerando a sobremortalidade como a diferença entre o número de óbitos observados e o limite

superior do intervalo de confiança, a sobremortalidade definida pela regressão foi de 68 em 1992, 266 em 93, 285 em 94 e 205 óbitos em 95.

A rede neural detectou sobremortalidade de cinco óbitos apenas em dezembro de 1993, mês anterior ao maior pico da epidemia de cólera, e em novembro e dezembro de 1994, a estação climática seguinte, com 17 óbitos.

DISCUSSÃO

Os resultados obtidos apontam para a ocorrência de excesso de mortalidade no Estado do Ceará e para a possibilidade de sua detecção por meio dos dados

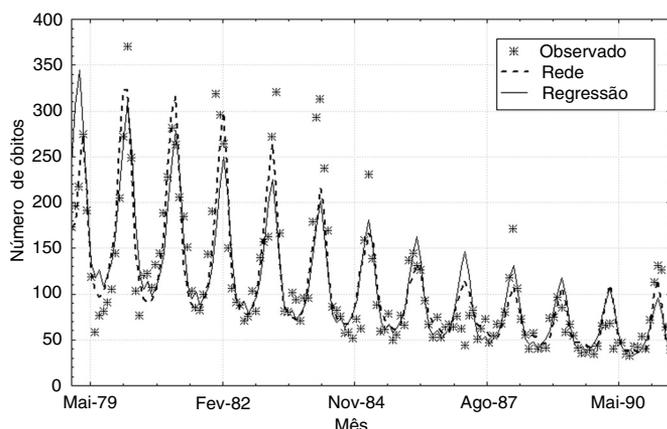


Figura 1 – Observado e estimado pelos modelos regressão binomial negativa e rede neural. Janeiro de 1979 a junho de 1991.

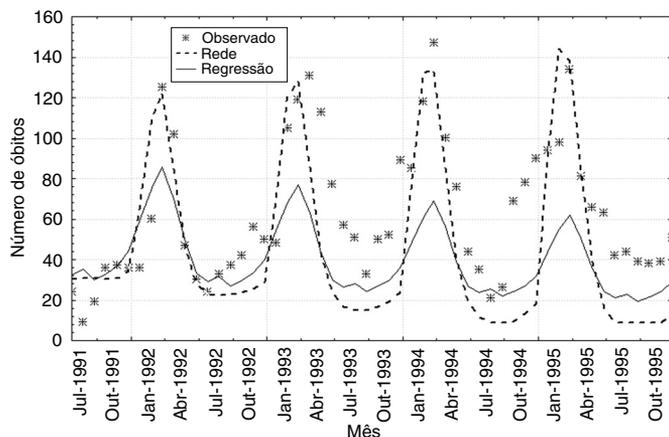


Figura 2 – Óbitos observados e estimados por extrapolação regressão binomial negativa e rede neural. Julho de 1991 a dezembro de 1995.

mensais de óbitos. Durante o período estudado, o Estado registrou 217 óbitos por cólera, sendo 19 em 1992, 89 em 1993, 104 em 1994 e cinco em 1995. A rede neural subestimou o excesso de óbitos, mas há dúvidas quanto à estimativa da regressão binomial para 1995, de magnitude ainda considerável, o que pode se dever ao longo período de extrapolação. Prolongar a série para um período imediatamente posterior sem ocorrência de cólera poderia esclarecer esse ponto, mas infelizmente a mudança para a CID-10 não garantiria a homogeneidade da série, introduzindo um provável viés e dificultando a interpretação dos achados.

Note-se que na estação com maior número de óbitos por infecções intestinais mal definidas, a estimativa pontual da rede neural foi mais próxima do observado do que a da regressão binomial negativa no período de extrapolação (Figura 2). O mesmo não ocorre no período utilizado para o ajuste dos modelos (Figura 1), o que sugere que a extrapolação com a regressão binomial para intervalos mais longos é menos confiável do que a da rede neural. Essa possível característica da rede neural, no entanto, não traz vantagens, uma vez que existe baixa precisão da estimativa intervalar.

As estimativas dos dois modelos tiveram boa concordância, apontando para uma adequação do uso de redes neurais artificiais em séries na área de saúde. No entanto, encarando-se mais pragmaticamente, como é usual em econometria, a escolha entre diferentes estratégias de predição deve considerar como a melhor aquela que funciona, prediz corretamente. Em vigilância epidemiológica, o interesse da predição é a detecção de valores aberrantes,²⁰ e não a maior proximidade entre o observado e o predito, como acontece quando se trata de predição de preços, por

exemplo. No presente exemplo, a regressão binomial negativa se mostrou mais adequada para a detecção de sobremortalidade durante a circulação da cólera, por possuir uma variância menor.

A diferença entre a regressão binomial negativa e a de Poisson está na estimativa da variância, que incorpora um parâmetro de hiper-dispersão (*over-dispersion*) – alfa. Ela só foi utilizada, porque os resíduos após o ajuste de uma regressão de Poisson mostravam uma maior dispersão do que a correspondente distribuição. Quando esse fator é igual a zero, a distribuição binomial negativa se reduz à distribuição de Poisson.

Diferentes estratégias têm sido utilizadas para estimar o erro no caso de predição por meio de redes neurais.² A utilizada no presente estudo assume que, teoricamente, um modelo exato para a série temporal pode ser encontrado, mas devido a erros de medida e à influência de fatores incontroláveis e desconhecidos existe um erro residual produzido aleatoriamente. A rede neural representa um modelo quase ótimo e é uma área a ser melhor explorada e desenvolvida.

As redes mais utilizadas são as não recorrentes ou *feed-forward*, onde cada neurônio de uma determinada camada se relaciona com todos os neurônios das camadas adjacentes, mas não com os da mesma camada. O processamento ocorre sempre no sentido da entrada em direção à saída da rede. Já as redes neurais recorrentes, como utilizada no presente trabalho, podem aprender seqüências; portanto, são a melhor escolha para dados de séries temporais. Enquanto redes com conexões padrão respondem a uma dada entrada sempre com a mesma saída, uma rede

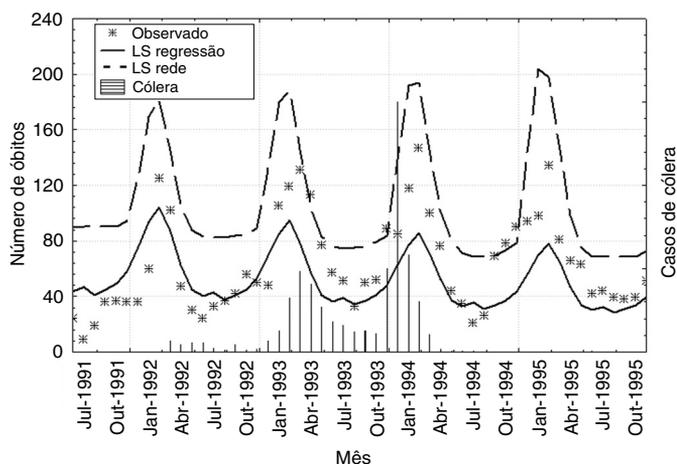


Figura 3 – Limite superior do intervalo de confiança estimado regressão binomial negativa e rede neural. Julho de 1991 a dezembro de 1995.

recorrente pode responder à mesma entrada com diferentes saídas em diferentes momentos, dependendo da entrada que foi apresentada anteriormente.

Ao incorporar um neurônio na camada de memória, a rede incorpora um componente auto regressivo,²² além dos componentes de tendência e sazonalidade representados pelos dois neurônios de entrada, ano e mês. O componente auto regressivo faz com que os valores preditos para um determinado momento tenham influência sobre a predição seguinte, conferindo aos valores recentes maior influência. Esse fato não gerou divergências entre os modelos, apesar da regressão binomial negativa não considerar a auto correlação no tempo.

A principal dificuldade no uso das redes neurais é a pouca familiaridade dos pesquisadores em geral, já que se trata de método recente em relação aos demais métodos estatísticos.¹⁴ O critério de seleção entre diferentes redes é pragmático, ou seja, é escolhida aquela que atinge os objetivos esperados. Também não está assegurada sempre a reprodutibilidade dos resultados, já que a cada treinamento os pesos iniciais são aleatórios, o que pode conduzir a diferentes áreas da superfície de erro. Existem ainda diferentes critérios de convergência, que podem resultar em diferentes mínimos locais. Esses fatos contribuem para uma certa insegurança no uso desse instrumento, que pode ser superada pela realização de vários treinamentos e observação da distribuição de seus resultados.

Outro problema citado é o super treinamento, onde

a rede captura relações quantitativas do ruído dos dados, prejudicando sua generalização (validade externa).⁹ Há a alternativa de se comparar redes com diferentes tempos de treinamento.^{6,7}

No presente estudo, a rede apresentada foi a primeira ajustada, não acontecendo diferenças importantes nas redes subseqüentes. Pode-se dizer que o critério de convergência foi o “tempo” de aprendizado, já que foi determinado pelo número de vezes que todos os dados foram apresentados ao treinamento. Utilizou-se como conjunto de treinamento os dados do período anterior à introdução da cólera no País, sem separar um conjunto para calibração, pois o tamanho do conjunto de treinamento tem grandes implicações em relação à generalização. Avaliou-se que o método de calibração (*cross-validation*), na verdade, não compensa a diminuição do conjunto de treinamento.

Nesse exemplo, a rede neural mostrou-se como um instrumento pouco sensível se comparado à regressão binomial negativa. Sua principal vantagem é que exige menos treinamento estatístico para sua aplicação. É claramente a técnica de mais fácil aplicação, já que não implica reconhecimento de modelos segundo o comportamento da série, nem avaliação do modelo ajustado. Os presentes resultados apontam que as redes neurais apresentam aspectos promissores quanto à sua utilização em vigilância epidemiológica. Porém, há ainda a necessidade de aprofundamento teórico quanto ao comportamento estatístico dos resíduos da rede, de modo a permitir uma maior precisão das estimativas.

REFERÊNCIAS

1. Alves MT, Silva AAM, Nemes MIB, Brito GO. Tendência da incidência e da mortalidade por Aids no Maranhão, 1985 a 1998. *Rev Saúde Pública* 2003;37:177-82.
2. Castiglione F. Forecasting price increments using an artificial neural network. *Adv Complex Systems* 2001;4:45-56.
3. Chatfield C. The analysis of time series. 4th ed. London: Chapman & Hall; 1994.
4. Choi K, Thacher SB. An evaluation of influenza mortality surveillance, 1962-1979. *Am J Epidemiol* 1981;113:215-22.
5. Collins SD, Lehman J. Trends and epiemics of influenza and pneumonia, 1918-1951. *Public Health Rep* 1951;66:1487-505.
6. Duh M, Walker AM, Pagano M, Kronlund K. Prediction and cross-validation of neural networks versus logistic regression: using hepatic disorders as an example. *Am J Epidemiol* 1998;147:407-12.
7. Duh M, Walker AM, Ayanian JZ. Epidemiologic interpretation of artificial neural networks. *Am J Epidemiol* 1998;147:1112-9.
8. Gerolamo M. Cólera no Brasil: a sétima pandemia [tese de doutorado]. Rio de Janeiro: Instituto de Medicina Social da UERJ; 2002.
9. Gorni AA. The application of neural networks in the modeling of plate rolling processes. *JOM-e* [serial online] 1997; 49. Available from: URL: <http://www.tms.org/pubs/journals/JOM/9704/Gorni/Gorni-9704.html> [2003 nov 12]
10. Haykins S. Neural Networks, a comprehensive foundation. 2nd ed. New Jersey: Prentice Hall; 1999.

11. Haydon GH, Jalan R, Ala-Korpela M, Hiltunen Y, Hanley J, Jarvis LM, Ludlum CA, Hayes PC. Prediction of cirrhosis in patients with chronic hepatitis C infection by artificial neural network analysis of virus and clinical factors. *J Viral Hepat* 1998;5:255-64.
12. Joo CN, Koo JY, Yu MJ. Application of short-term water demand prediction model to Seoul. *J Water Sci Technol* 2002;46:255-61.
13. Kao JJ, Huang SS. Forecasts using neural network versus Box-Jenkins methodology for ambient air quality monitoring data. *J Air Waste Manag Assoc* 2000;50:219-26.
14. Kattan MW, Hess KR, Beck JR. Experiments to determine whether recursive partitioning (CART) or an artificial neural network overcomes theoretical limitations of Cox proportional hazards regression. *Comput Biomed Res* 1998;31:363-73.
15. Lawless JF. Negative binomial and mixed Poisson regression. *Can J Stat* 1987;15:209-25.
16. Simonsen L, Clark MJ, Stroup DF, Williamson GD, Arden NH, Cox NJ. A method for timely assessment of influenza associated mortality in the United States. *Epidemiology* 1977;8:390-5.
17. StatSoft Inc. STATISTICA for Windows [Computer program manual]. Tulsa; 1998.
18. Stroup D, Wharton M, Kafadar K, Dean AG. Evaluation of a method for detecting aberrations in public health surveillance data. *Am J Epidemiol* 1993;137:373-80.
19. StataCorp. Stata statistical software: release 7.0. College Station, TX; 2000.
20. Teush SM, Churchill RE. *Principles and practice of public health surveillance*. New York: Oxford University Press; 1994.
21. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol* 1996;49:1225-31.
22. Wasserman PD. *Neural computing: theory and practice*. New York: Van Nostrand Reinhold; 1989.
23. Ward System Group Inc. NeuroShell 2, release 4 [computer program manual]. Frederick; 1998.