

MODELAGEM DAS RAZÕES DIÁRIAS DE BRILHO SOLAR EM SÃO PAULO ATRAVÉS DE CADEIAS DE MARKOV NÃO ESTACIONÁRIAS

Clóvis Angeli Sansigolo
INPE, S.J. Campos, SP

RESUMO

Cadeias de Markov de 4 estados, não estacionárias, são usadas para modelar as razões diárias de brilho solar na cidade de São Paulo. Séries de Fourier descrevem as variações sazonais nas probabilidades de transição. Todas as regressões e testes, baseados em modelos lineares generalizados foram feitas através do software GLIM. Variações sazonais significativas foram detectadas nas probabilidades de transição, especialmente nas categorias extremas. O modelo ajustado provém um resumo conciso dos dados diários, útil nas previsões a curto prazo (24h) e simulações do número de horas de brilho solar.

INTRODUÇÃO

As razões diárias de brilho solar podem ser adequadamente descritas por Cadeias de Markov com probabilidades de transição não estacionárias, pois exibem duas características fundamentais : persistência e sazonalidade.

Fraedrich & Miller (1982) modelaram as durações diárias de brilho solar em Berlim com Cadeias de Markov de 4 estados. Spirke & Ries (1986) ajustaram e testaram Cadeias de Markov de 5 estados com probabilidades de transição não estacionárias ao número de horas de brilho solar em Weistephan, Alemanha.

Todos esses modelos podem ser considerados como lineares generalizados (MLG), análogos aos métodos de regressão linear usados com dados normalmente distribuídos. Além disso, uma grande variedade de modelos podem ser facilmente ajustados através do software GLIM, usado em todas as regressões e testes.

O objetivo deste trabalho é o ajuste e teste de Cadeias de Markov de 4 estados, não estacionárias, às razões diárias de brilho solar na cidade de São Paulo.

METODOLOGIA

Os dados usados neste estudo são os registros diários do número de horas de brilho solar (n) no período de 1933 a 92, obtidos na Estação Meteorológica do IAG/USP, São Paulo. Esse dados foram normalizados pela duração máxima de brilho solar (N), função da latitude do local e declinação do sol, levando a uma nova série de razões de brilho solar (n/N). Esta série foi subdividida em 4 categorias, com aproximadamente mesmas frequências anuais : cat. 1- $[0-0,15[$, cat. 2- $[0,15-0,45[$, cat. 3- $[0,45-0,7[$ e cat. 4- $[0,7-1]$. Estes limites podem ser relacionados à irradiâncias através da equação de Ångstron-Prescott para São Paulo : $Q/Q_0 = 0,3 + 0,54 n/N$.

Numa Cadeia de Markov de 1ª ordem, com 4 estados, a probabilidade de uma dada classe num determinado dia depende somente do estado (categoria) do dia anterior. Os parâmetros a serem estimados são portanto as probabilidades de transição p_{ij} :

$$p_{i,j}(t) = n_{i,j}(t) / [n_{i,1}(t) + n_{i,2}(t) + n_{i,3}(t) + n_{i,4}(t)] \quad (1)$$

onde $i=1,2,3,4$ e $j=1,2,3,4$.

O próximo passo na análise é ajustar uma função às $p_{i,j}(t)$ para modelar a dependência temporal nas probabilidades de transição entre as categorias através do ano. Séries de Fourier finitas foram usadas para levar em conta essas variações sazonais.

$$\ln [p_{i,j}(t') / (1-p_{i,j}(t'))] = a_{i0} + \sum_{k=1}^m a_{ik} \cdot \text{sen}(kt') + b_{ik} \cdot \text{cos}(kt') \quad (2)$$

onde $t = 1,2,3, \dots, 366$ e $t' = \pi (t-183)/183$.

A transformação logística no termo esquerdo da equação (2) assegura que as proporções estimadas estarão sempre compreendidas entre 0 e 1.

O ajuste dessas funções é análogo ao ajuste de equações de regressão para dados normalmente distribuídos. Diferentes modelos são comparados considerando-se a magnitude da diferença entre as somas de quadrados residuais. Com dados não normais a estatística análoga é a Deviância D^2 , dada para a distribuição binomial (Coe & Stern, 1982), por:

$$D^2 = 2 \sum_{t=1}^{366} n_{i,j} \left\{ p_{i,j}(t) \ln [p_{i,j}(t) / \hat{p}_{i,j}(t)] + [1 - p_{i,j}(t)] \ln [(1 - p_{i,j}(t)) / (1 - \hat{p}_{i,j}(t))] \right\} \quad (3)$$

onde $\hat{p}_{i,j}(t)$ é o valor ajustado de $p_{i,j}(t)$.

D^2 é assintoticamente igual à estatística χ^2 de Pearson. Se m parâmetros são estimados e o modelo é o correto, D^2 tem uma distribuição χ^2 com $T-m$ graus de liberdade.

RESULTADOS E DISCUSSÃO

A análise das deviâncias resultante da variação do número de termos ajustados indica que na maioria dos casos 1 ou 2 harmônicos são necessários para descrever as variações sazonais nas probabilidades de transição. Bons ajustes foram obtidos em todos os casos, pois as deviâncias estão próximas aos graus de liberdade.

O problema de se determinar se cadeias de 1ª ordem são adequadas se reduz ao problema de se determinar se 4 curvas (p_1, p_2, p_3 e p_4) se ajustam melhor do que 16 curvas ($p_{i,j}, i=1,4$ e $j=1,4$). A deviância dos 4 conjuntos de dados $p_{1,j}, j=1,4$ em relação à curva ajustada p_1 (4468 com 1424 GL) é comparada com as deviâncias desses mesmos 4 conjuntos de dados em relação às 4 curvas ajustadas $p_{1,j}, j=1,4$ (1572 com 1446 GL). Semelhantemente, as deviâncias dos conjuntos de dados $p_{2,j}, p_{3,j}, p_{4,j}, j=1,4$, em relação à p_2, p_3 e p_4 , são respectivamente, 2414 com 1424 GL, 2409 com 1440 GL e 5328 com 1424 GL, comparadas com as deviâncias sobre as curvas $p_{2,j}, p_{3,j}$ e $p_{4,j}, j=1,4$ (1465 com 1448 GL, 1608 com 1438 GL e 1675 com 1444 GL).

As diferenças nas deviâncias usadas para decidir sobre a complexidade dos modelos são simplesmente testes de razão de verossimilhança (Nelder & Wedderburn, 1972).

As significativas reduções nas deviâncias indicam que, no mínimo cadeias de 1ª ordem são necessárias e se estas se ajustam bem, são preferíveis em relação às de ordem superior, pois o número de parâmetros a ser estimado é menor, logo melhores estimativas são obtidas e o uso subsequente do modelo é mais simples.

Variações sazonais significativas foram detectadas principalmente nas categorias extremas com dias claros e cobertos.

O ganho sobre a climatologia diária (fig. 1) é evidente ($p_{11}=85\%$, $p_{22}=28\%$, $p_{33}=26\%$ e $p_{44}=95\%$). A persistência é maior para as transições entre as mesmas classes, especialmente as extremas (fig.2, 3 e 4).

Este estudo enfatiza a importância de se considerar persistência e sazonalidade nas previsões a curto prazo e simulações do número de horas de brilho solar.

BIBLIOGRAFIA

- COE, R. & STERN, R.D. Fitting Models to Daily Rainfall Data. *J. Appl. Met.* **21**: 1024-1031. 1982
- FRAEDRICH, K. & MÜLLER, K. On Single Station Forecasting : Sunshine and Rainfall Markov Chains. *Beitr. Phys. Atmosph.* **56(1)**: 108-134. 1982.
- NELDER, J. A. & WEDDERBURN, W.M. Generalized Linear Models. *J. Roy. Stat. Soc.* **A-123**: 370-384. 1972.
- SANSIGOLO, C.A. & COSTA, M.M.G. Stochastic Modelling of the Sunshine Daily Ratios at São Paulo. Preprints 12th Conf. Prob. Stat. Atm. Sci. AMS, Toronto. p.188-189. 1991.
- SPIRKE, W. & RIES, H. Non-Stationary Markov Chains for Modelling Daily Radiation Data. *Quart. J. Roy. Met. Soc.* **112**:1219-1229. 1986.

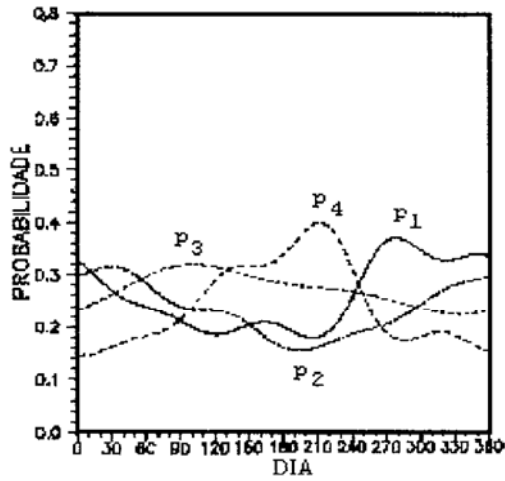


Fig. 1. Curvas ajustadas às probabilidades incondicionais para as categorias 1 (P_1), 2 (P_2), 3 (P_3) e 4 (P_4).

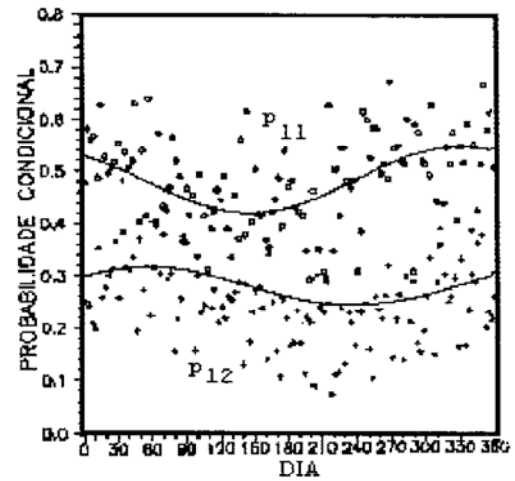


Fig. 2. Dados observados e curvas ajustadas às probabilidades de transição entre as categorias 1 (P_{11}) e 2 (P_{12}).

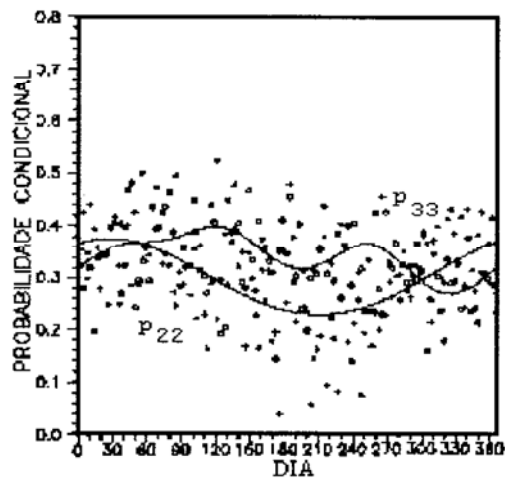


Fig. 3. Dados observados e curvas ajustadas às probabilidades de transição entre as categorias 2 (P_{22}) e 3 (P_{33}).

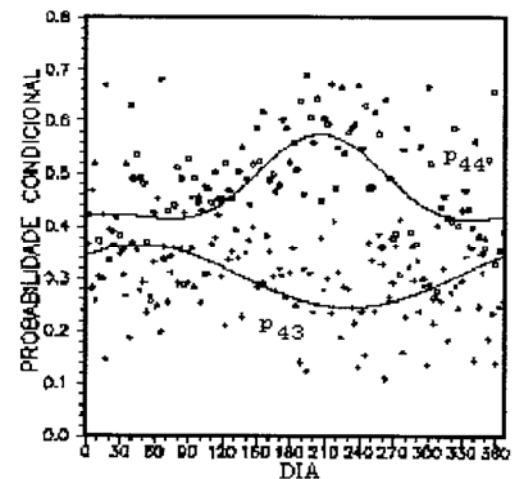


Fig. 4. Dados observados e curvas ajustadas às probabilidades de transição entre as categorias 4 (P_{44}) e 3 (P_{43}).