

Avaliação da influência do ENOS sobre início da estação chuvosa via modelos semiparamétricos de sobrevivência

Aline de Holanda Nunes Maia – Embrapa Meio Ambiente¹

Ricardo Antônio Almeida Pazianotto - Embrapa Meio Ambiente²

Caio Augusto dos Santos Coelho – CPTEC/INPE³

Resumo: Características da estação chuvosa tais como tempo até o início (T) e sua duração são importantes para o planejamento agrícola, especialmente em regiões com alta variabilidade interanual da estação chuvosa. A variável T é influenciada pelo fenômeno El Niño/Oscilação Sul (ENSO) cuja intensidade pode ser mensurada por anomalias da temperatura da superfície do mar (ATSM). Neste trabalho, quantificamos a influência de ATSM sobre T em Guarimiranga e Juazeiro do Norte (Ceará), utilizando uma abordagem inovadora, a análise de sobrevivência semiparamétrica (ASSP). Para definição do início da estação em cada ano, foi utilizado um critério baseado no total de chuva acumulada em dois dias (20 mm) consecutivos e ausência de veranicos superiores a 10 dias nos trinta dias seguintes. O modelo de Cox foi utilizado para quantificar a influência de ATSM sobre T e estimação de distribuições condicionais. A ASSP apresenta vantagens em relação aos modelos de regressão de componentes principais, entre elas: a) modelagem das distribuições de probabilidade ao invés de valores esperados (médias); b) não há necessidade de pressupor normalidade para T e c) não considera relações lineares entre os preditores e T . Quando comparada à regressão logística, a ASSP é naturalmente mais adequada, uma vez permite a estimação de distribuições condicionais ao invés de compor funções de sobrevivência a partir de vários modelos logísticos, um método complexo e desnecessário para esse tipo de problema. Nos dois locais avaliados, a fase quente do ENOS (El Niño) implicou em aumento do risco de início tardio da estação chuvosa.

Palavras-chave: modelo de regressão de Cox, previsão climática sazonal, avaliação quantitativa de risco.

1. Introdução

Características da estação chuvosa, tais como tempo até o início (T) e sua duração são importantes para o planejamento agrícola, especialmente em regiões com alta variabilidade interanual da precipitação

¹ Agônoma, Mestre em Estatística, Doutora em Agronomia, Laboratório de Geotecnologias e Métodos Quantitativos. E-mail: ahmaia@cnpma.embrapa.br

² Matemático, Mestre em Biofísica Molecular, Laboratório de Geotecnologias e Métodos Quantitativos.

³ Doutor em Meteorologia, Centro de Previsão de Tempo e Estudos Climáticos.

pluvimétrica. Diversos estudos demonstram a influência do fenômeno de larga escala El Niño/Oscilação Sul (ENSO) sobre T em muitas regiões do globo, entre elas o nordeste e sul do Brasil e norte da Austrália (RAUSCHER et al, LO et al, 2007). O gerenciamento de riscos climáticos associados ao início tardio da estação chuvosa requer informação sobre estados futuros de T , geralmente representadas por funções de distribuição de probabilidade acumulada (CDF, $P[T \leq t]$) ou por suas funções complementares ($P[T > t]$), ditas funções probabilidade de exceder (FPE). Uma variedade de métodos estatísticos tem sido utilizada para estimação de FPE , incluindo regressão logística e métodos não paramétricos de suavização de dados. MAIA et al (2010), propuseram o uso de modelos de regressão de Cox para previsões climáticas sazonais, apresentando estudos de caso para o início da estação chuva em Cairns (norte da Austrália) e total de chuva acumulado durante a estação chuvosa em Quixeramobim (Ceará).

Neste trabalho, avaliamos a influência de anomalias da temperatura da superfície do mar (TSM) sobre o tempo até o início da estação chuvosa em Guaramiranga e Juazeiro do Norte, no estado do Ceará, utilizando uma abordagem inovadora, a análise de sobrevivência não paramétrica. Para definição do início da estação chuvosa em cada ano, foi utilizado um critério baseado na ocorrência de chuva acumulada em dois dias consecutivos superior a 20mm e ausência de veranicos superiores a dez dias nos trinta dias consecutivos (SANSIGOLO, 1989).

Neste trabalho, são apresentados dois estudos de caso nos quais o modelo de regressão semiparamétrico de Cox foi usado para investigar a influência do ENOS, mensurada via anomalia média da TSM no período de outubro a dezembro ($ANOM_OND$), sobre o início da estação chuvosa em Guaramiranga e Juazeiro do Norte (Ceará, Brasil). O modelo de Cox foi utilizado para quantificar a influência do ENOS sobre T e posterior estimação de distribuições condicionais para geração de informações para avaliação probabilística de risco. O objetivo da apresentação desses estudos é demonstrar o potencial do método proposto como ferramenta de avaliação quantitativa de riscos climáticos, útil para auxílio à tomada de decisão em agricultura.

2. Dados e métodos

Foram utilizadas as seguintes séries históricas: anomalias mensais da temperatura da superfície do mar (TSM) na região Niño 3.4 do pacífico Equatorial (www.cpc.noa.gov/data/indices/sstoi.indices) e série histórica de dados diários (1974-2011) de chuva de estações dos municípios de Guaramiranga (04°15'48"S, -38° 55' 59") e Juazeiro do Norte (-07°12'47"S, -39° 18' 55"W), obtidos da Fundação Cearense de Meteorologia, Funceme (www.funceme.br/DEPAM/download/postos/81.txt).

Para cada local e ano, foram calculadas anomalias médias da TSM para o período de outubro a dezembro ($ANOM_OND$) do ano anterior à estação chuvosa considerada. O tempo até o início da estação chuvosa

(T) foi calculado para cada ano, seguindo critério definido por Sansigolo (1989): o primeiro dia após a data de referência (1 de janeiro), com 20 mm acumulados em um ou dois dias consecutivos, condicionado à não ocorrência de veranicos maiores que dez dias nos trinta dias subsequentes.

A influência do ENOS sobre T , em cada local, foi quantificada via modelo de regressão de Cox (Cox, 1972), utilizando $ANOM_OND$ como preditor. Nos exemplos apresentados neste trabalho temos o seguinte modelo para função probabilidade de excedência (FPE):

$$FPE(t_i, x_i) = P(T > t_i / X = x_i) = [FPE_0(t_i)] \exp(x_i \cdot \beta) \quad (\text{Equação 1})$$

Onde t_i é o tempo até a ocorrência do início da estação chuvosa no local avaliado (Guaramiranga ou Juazeiro do Norte) num particular ano i ; x_i é o valor médio da anomalia da TSM na região Niño 3.4 do Pacífico equatorial nos meses de outubro a dezembro ($ANOM_OND$) no ano ($i-1$) e $FPE_0(t_i)$, a função de sobrevivência obtida quando x_i tem valor nulo; β é o parâmetro do modelo que quantifica a influência das anomalias da TSM sobre T .

3. Resultados e Discussão

As estimativas dos parâmetros do modelo de Cox (Equação 1) com respectivos erros-padrão, razões de riscos e níveis de significância nominais (valores p) do teste de Wald são apresentadas Tabela 1. Na Figura 1 são apresentadas estimativas das distribuições condicionais de T para alguns valores do preditor : -1,5 (fase fria), 0 (fase neutra) e 1,5 (fase quente). Os tempos medianos até o início da estação chuvosa (dias após 1º de janeiro) para os valores -1,5 (fase fria) e 1,5 (fase quente) da $ANOM_OND$ em Guaramiranga foram 7 e 27 dias, respectivamente; para Juazeiro as medianas foram 12 dias para a fase fria e 24 dias para a fase quente. Os resultados indicam efeito pronunciado do ENOS: a fase fria (La Niña) favorece o início antecipado da estação.

Tabela 1: Influência do preditor $ANOM_OND$ (período de outubro a dezembro) sobre riscos associados ao início da estação chuvosa em Guaramiranga e Juazeiro do Norte (Ceará, Brasil), quantificadas via modelo de regressão de Cox.

Local	Estimativa de β	Erro Padrão	Estimativa da razão de riscos ($\exp(\beta)$)	χ^2	Valor p*
Guaramiranga	-0,3816	0,1879	0,683	4,12	0,0423
Juazeiro do Norte	-0,2651	0,1651	0,767	2,58	0,1083

* Nível de significância nominal do teste qui-quadrado (χ^2) de Wald para as hipóteses $\beta = 0$.

Os valores p derivados dos testes de χ^2 de Wald indicam forte evidência de efeito do preditor $ANOM_OND$ sobre o início da estação chuvosa em Guaramiranga (p= 0.0423) e moderada evidência para Juazeiro do Norte (p=0.1083). As estimativas de β (Equação 1) indicam que $ANOM_OND$ afeta T de forma mais intensa em Guaramiranga que em Juazeiro do Norte (Tabela 1). Analisando as estimativas das razões de

riscos (Tabela 1), observamos que em Guaramiranga, para cada acréscimo unitário no valor da $ANOM_OND$, a taxa de falha instantânea ($h(t)=exp(\beta)$) é multiplicada por 0,683 o que corresponde a uma redução de 31,7% . Para Juazeiro, a redução em $h(t)$ é menos intensa: 23,3 % para cada acréscimo unitário do preditor. A probabilidade de T ser inferior à mediana histórica (19 dias) em Guaramiranga para $ANOM_OND=-1.5$ (fase fria) é 2,12 vezes superior à $Prob(T<19)$ para $ANOM_OND=1.5$ (fase quente). Para Juazeiro do Norte (mediana=18 dias), a razão entre essas probabilidades é de 1,70.

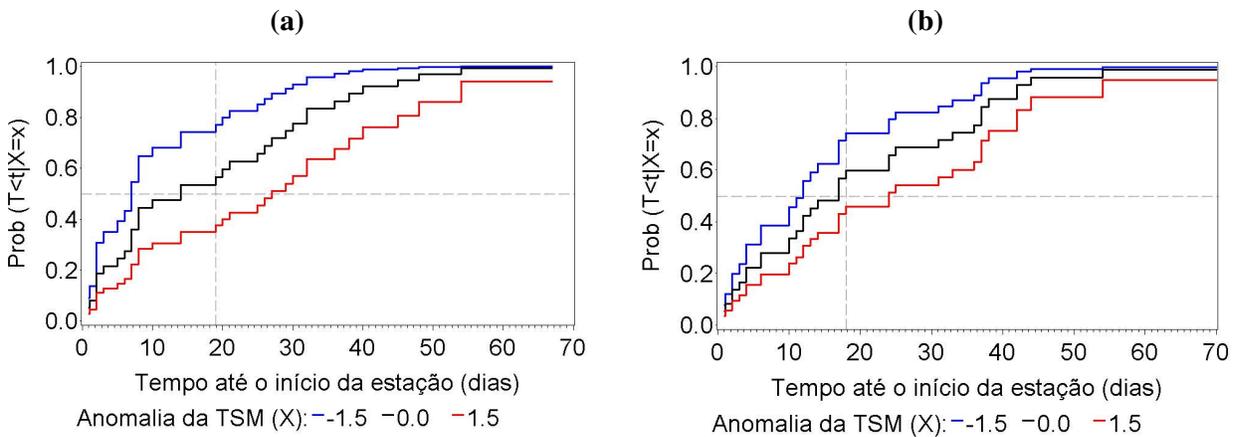


Figura 1: Distribuições condicionais ($P[T \leq t | X=x]$) para o tempo até o início da estação chuvosa (T) em dois locais no estado do Ceará, estimadas via modelo de Cox, mostrando a influência do preditor anomalia da TSM na região Nino 3.4 no período de outubro a dezembro ($ANOM_OND$): (a) Guaramiranga e (b) Juazeiro do Norte. Linhas verticais pontilhadas indicam a data mediana de início da estação chuvosa. O tempo até o início da estação chuvosa e o número de dias entre 1° de janeiro e a data em que foi atingido o critério definido em SANSIGOLO (1989).

Essa abordagem apresenta as seguintes vantagens em relação ao uso tradicional de modelos de regressão de componentes principais: a) modelagem das distribuições de probabilidade ao invés de valores esperados (médias); b) não há necessidade de pressupor normalidade para T e c) não considera relações lineares entre T e os preditores. Quando comparada ao método de regressão logística utilizado por LO et al (2008), a abordagem de sobrevivência é naturalmente mais adequada, uma vez que considera múltiplos valores de t em $P(T > t | X=x)$. A regressão logística estima essa probabilidade para um valor fixo de t . LO et al (2008) ajustaram diferentes logísticas, variando t , para compor uma função de sobrevivência, um processo complexo desnecessário para esse tipo de problema.

4. Considerações finais

Modelos probabilísticos de previsão sazonal que utilizam preditores derivados de estados do oceano e da atmosfera (anomalias de temperaturas da superfície do mar – TSM, índice de oscilação sul, IOS)

constituem alternativas promissoras para auxílio à tomada de decisão, em escalas locais e regionais, sob a hipótese de estabilidade da relação entre variável resposta e preditores em cenários de mudanças climáticas. A análise de sobrevivência, tradicionalmente utilizada para modelagem de variáveis que mensuram o tempo até a ocorrência de um evento de interesse, a partir de uma origem previamente especificada, em investigações na área médica, engenharia e ciências sociais, pode ser de grande utilidade para avaliação probabilística de riscos climáticos, especialmente para variáveis como tempo até o início e duração da estação chuvosa. Para variáveis positivas que não constituem tempos de falha, os métodos de análise de sobrevivência podem ainda ser utilizados com o simples propósito de estimar distribuições condicionais, mesmo que a taxa de falha instantânea ($h(t)=exp(\beta)$) não seja interpretável em tais contextos. Modelos de sobrevivência, tanto paramétricos como semiparamétricos (modelo de Cox), são úteis para avaliar a influência de preditores (índices climáticos) sobre riscos de interesse, para estimar *CDF* ou *FPE* para combinações específicas de preditores e suas respectivas incertezas, além de fornecer informação sobre riscos relativos, de grande valor para tomadores de decisão. O modelo de Cox utilizado neste trabalho considera a pressuposição de riscos proporcionais. No entanto, essa classe de modelos pode facilmente generalizada para acomodar situações onde essa pressuposição não é satisfeita. Abordagens similares utilizando novos modelos paramétricos extremamente flexíveis estão também sendo investigadas com o objetivo de contribuir com alternativas viáveis e de fácil uso para previsão probabilística sazonal.

Agradecimentos

Agradecemos ao Dr. Holger Meinke, diretor do TIAR (Tasmanian Institute of Agricultural Research) pelo incentivo à investigação sobre novas abordagens para avaliação quantitativa de riscos climáticos.

Referências

COX, D.R, 1972: Regression Models and Life-Tables (with Discussion), **Journal of the Royal Statistical Society**, Series B, v. 34, p. 187–220.

LO F.; M. WHEELER; H. MEINKE; A. DONALD, 2007. Probabilistic forecasts of the onset of the North Australian wet season. **Monthly Weather Review**, v. 135(10), p. 3506–3520.

MAIA, A. H. N.; H. MEINKE, S. LENNOX; R. C. STONE, 2007: Inferential, non-parametric statistics to assess quality of probabilistic forecast systems. **Monthly Weather Review**, v. 135(2), p. 351-362.

RAUSCHER, S. A.; SETH, A.; LIEBMANN, B.; QIAN, J-H; S J CAMARGO, 2007. Regional Climate Model Simulated Timing and Character of Seasonal Rains in South America. **Monthly Weather Review**, v. 135(7), p. 2642–2657.

SANSINGOLO, A. S.,1989: Variabilidade Interanual da Estação Chuvosa em São Paulo. **Climanálise**, v.4(9), p. 40-43.