

# Sistema de alarme ótimo aplicado na ferrugem do café

## Optimal alarm system applied in coffee rust

Luciene Resende Gonçalves<sup>1\*</sup>; Thelma Sáfadi<sup>2</sup>;  
Anderson Castro Soares de Oliveira<sup>3</sup>

### Resumo

---

Os sistemas de alarme ótimos têm tido grande utilidade na detecção e advertência de ocorrência de catástrofes. Essa metodologia foi aplicada via modelo TARSO com ajuste bayesiano, servindo como mais um mecanismo de previsão para a doença ferrugem do café. A cultura cafeeira é muito suscetível ao ataque dessa doença, existindo registros de incidência na maioria das lavouras cultivadas. Pesquisas envolvendo esse fator limitante para a produção são intensas e frequentes, indicando os fatores ambientais como responsáveis pelo alastramento de epidemias, que não ocorrem se esses fatores não forem favoráveis. O tipo de ajuste utilizado por meio das probabilidades a posteriori, permitem que o sistema seja atualizado a cada ponto do tempo. A metodologia foi aplicada à série de índice de ferrugem na presença da série de temperatura média sendo possível verificar que o alarme resultou em uma detecção de catástrofe alta em pontos nos quais a catástrofe ainda não ocorreu e em detecções baixas se o ponto já estava no estado de catástrofe.

**Palavras-chave:** Ferrugem, inferência bayesiana, modelo TARSO, limiar

### Abstract

---

Alarm systems have very great utility in detecting and warning of catastrophes. This methodology was applied via TARSO model with Bayesian estimation, serving as a forecasting mechanism for coffee rust disease. The coffee culture is very susceptible to this disease causing several records of incidence in most cultivated crops. Researches involving this limiting factor for production are intense and frequent, indicating environmental factors as responsible for the epidemics spread, which does not occur if these factors are not favorable. The fitting type used by the a posteriori probability, allows the system to be updated each time point. The methodology was applied to the rust index series in the presence of the average temperature series. Thus, it is possible to verify the alarm resulted or in a high catastrophe detection in points at which the catastrophe has not occurred, or in the low detections if the point was already in the catastrophe state.

**Key words:** Rust, bayesian inference, TARSO model, threshold

---

<sup>1</sup> Prof<sup>a</sup> do Instituto de Ciências Sociais Aplicadas, Universidade Federal de Alfenas, UNIFAL, Campus Varginha, Varginha, MG. E-mail: [luciene-rg@bol.com.br](mailto:luciene-rg@bol.com.br)

<sup>2</sup> Prof<sup>a</sup> do Dept<sup>o</sup> de Ciências Exatas, Universidade Federal de Lavras, UFLA, Lavras, MG. E-mail: [thelsafadi@hotmail.com](mailto:thelsafadi@hotmail.com)

<sup>3</sup> Prof. do Dept<sup>o</sup> de Estatística, Universidade Federal de Cuiabá, Cuiabá, MT. E-mail: [andersoneso@gmail.com](mailto:andersoneso@gmail.com)

\* Autor para correspondência

## Introdução

Os sistemas de alarme ótimos constituem uma metodologia potencial em descobrir quando uma variedade de processos aleatórios cruzarão um certo nível crítico  $u$ . Na terminologia de Lindgren (1985), esse fato é chamado de “catástrofe” e é usado como terminologia mais atraente.

Exemplos de catástrofes para Monteiro, Pereira e Scotto (2008) e Svensson, Holst e Lindquist (1996) abrangem inundações de planícies, falta de energia elétrica, itens faltosos em um estoque, predição do crescimento em média do nível do mar e frequências de inundações devidas ao aquecimento global decorrentes do efeito estufa, avaliação dos impactos na saúde em razão da poluição do ar, bem como diversas outras situações.

A maneira mais simples para construir um sistema de alarme, também chamado de preditor ótimo de catástrofe, é por meio da utilização de um preditor simples. No entanto, Lindgren (1985) argumenta que tais alarmes simples não funcionam bem na habilidade de detectar os eventos, localizá-los precisamente no tempo, e produzir tão poucos alarmes falsos quanto possível.

Dessa forma, levando em consideração tais ideias, Lindgren (1975, 1980, 1985) e Maré (1980) estabeleceram os princípios ótimos dos cruzamentos de níveis e correspondentemente de sistemas de alarme ótimos para processos aleatórios. A seguir são elencados alguns trabalhos desses autores bem como de outros.

Maré (1980) desenvolveu um sistema de alarme ótimo e aplicou o método a processos gaussianos. Para ele, um sistema de alarme tem uma alta habilidade de detectar catástrofes se a proporção de catástrofes detectadas dentre todas é próxima de um. Se, ao mesmo tempo, a proporção de alarmes corretos dentre todos os alarmes é tão alta quanto possível, então o sistema de alarme é dito ser ótimo.

Lindgren (1985) utilizou um processo estacionário e gaussiano para predizer o cruzamento

futuro de um nível  $u$  investigando critérios para um preditor com bom nível de cruzamento no sentido de maximizar a probabilidade de detecção do alarme.

Turkman e Turkman (1990) sugeriram uma versão preditiva bayesiana para processos autorregressivos de ordem 1,  $AR(1)$ , apesar de não terem feito nenhuma tentativa para solucionar o problema computacionalmente. O trabalho desses autores foi uma tentativa para superar a desvantagem dos sistemas de alarme introduzidos por Lindgren (1980) e Maré (1980) os quais ignoram as variações amostrais dos parâmetros do modelo, além das probabilidades a posteriori dos eventos não serem atualizadas a cada ponto do tempo.

Svensson, Holst e Lindquist (1996) apresentaram um preditor ótimo de cruzamentos de níveis e catástrofes para processos autorregressivos e médias móveis e investigaram o desempenho do preditor. Para eles, um preditor de catástrofe ótimo é o preditor que dá um número mínimo de alarmes falsos para uma probabilidade de detecção fixada. Utilizaram as características operacionais como uma ferramenta de avaliação para comparar e construir os preditores, isto é, usando a probabilidade de detecção de uma catástrofe para o preditor. Uma condição explícita para o preditor ótimo de catástrofe baseado na predição linear de valores futuros do processo é dado e comparado com um preditor simples de catástrofe, os quais alarmam quando os valores preditos do processo excedem um dado nível. Seus resultados indicaram que o preditor ótimo de catástrofe é melhor que o preditor simples. Svensson e Holst (1997) aplicaram essas ideias para desenvolver um sistema de alarme ótimo para predizer os níveis de altura das águas no mar Báltico.

Antunes, Amaral-Turkman e Turkman (2003) estenderam os resultados de Turkman e Turkman (1990) para processos autorregressivos de ordem  $p$ ,  $AR(p)$ , mostrando como as características do alarme podem ser obtidas numericamente. Esses autores nomearam o sistema de alarme de Lindgren

(1980) e Maré (1980) de previsão de eventos “off-line”, em que o sistema é construído para eventos incondicionais, supostos terem a mesma probabilidade em todo tempo devido à suposição de estacionariedade. Em contrapartida estabeleceram os sistemas de alarme chamados previsão de eventos “on-line”, para os quais as probabilidades a posteriori são atualizadas a cada ponto temporal.

Monteiro, Pereira e Scotto (2008) desenvolveram um sistema de alarme para prever se um processo de contagem ultrapassa superiormente um certo nível  $u$  sempre que esse fato é predito. Concluíram que a técnica leva a eventos preditores ótimos no sentido de fornecer o menor número de alarmes falsos para uma determinada probabilidade de alarme condicionada a valores passados. Obtiveram um número alto de alarmes falsos e sugeriram, como solução futura para este problema, incluir no modelo informações externas adicionais, ou ainda, considerar a probabilidade do evento de interesse variando no tempo.

Costa, Scotto e Pereira (2010) desenvolveram um sistema de alarme para prever se uma série financeira temporal ajustada via modelo FIAPARCH cruza superiormente ou inferiormente algum nível particular  $u$  e dá um alarme sempre que o cruzamento é predito. Fizeram uma aplicação a um conjunto de dados referentes aos retornos diários da bolsa de valores de São Paulo. Concluíram que o sistema de predição “on-line” exibe um comportamento adaptativo, já que, a informação disponível é integrada à experiência informativa o que faz o sistema adaptar-se produzindo um número mínimo de alarmes falsos.

O objetivo, neste artigo, será aplicar o sistema de alarme ótimo ao processo TARSO o qual foi desenvolvido sob o enfoque bayesiano por Sáfadi e Morettin (2001). A importância desse método está na possibilidade de contribuir com mais um mecanismo de previsão para a doença que mais ocorre nas lavouras cafeeiras brasileiras, a ferrugem.

Para isto, além desta introdução, o artigo

é organizado como segue. Em Material são apresentados os dados, enquanto, nos Métodos é apresentada a metodologia do alarme de acordo com Antunes, Amaral-Turkman e Turkman (2003) e o sistema de alarme ótimo para processos TARSO. Em Resultados e Discussão, é feita a aplicação à série de índices de ferrugem na presença da variável climática temperatura média. Em seguida são apresentadas as Conclusões, Agradecimentos e Referências Bibliográficas.

## Material e Métodos

### Material

Os dados utilizados na construção do sistema de alarme ótimo para a ferrugem do café referem-se ao acompanhamento mensal da incidência de ferrugem no cafeeiro ( $Y_t$ ) em (%) e temperatura média ( $Z_t$ ) em ( $^{\circ}C$ ). Esses dados foram coletados por Japiassú et al. (2007) no período de outubro de 1999 a agosto de 2007, na fazenda experimental da Fundação Procafé, localizada em Varginha, MG. As observações mensais de incidência da ferrugem e temperatura média para o período de setembro de 2007 a fevereiro de 2011 foram obtidas no site da Fundação Procafé (2011).

### Métodos

#### Metodologia do alarme

Seja  $Y_t$  um processo estocástico discreto no espaço paramétrico  $\theta \subset \mathfrak{R}^k$  com  $k$  fixado. A sequência de tempo  $\{t = p+1, \dots, t-1, t+1, \dots\}$  é dividida em três seções:  $\{p+1, \dots, t-p\}$  - passado,  $\{t-p+1, \dots, t\}$  - presente e  $\{t+1, \dots\}$  - futuro. Essas seções, para algum  $p > 0$ , originam três conjuntos distintos:

$D_t = \{Y_{p+1}, \dots, Y_{t-p}\}$  – dados ou experimento informativo;

$Y_2 = \{Y_{t-p+1}, \dots, Y_t\}$  – experimento presente;

$Y_3 = \{Y_{t+1}, \dots\}$  – experimento futuro.

Termos como catástrofe e região de alarme são comuns em eventos de predição e, por isso, são dados em forma de definição.

**Definição 1** (ANTUNES; AMARAL-TURKMAN; TURKMAN, 2003). O evento de interesse  $C_{t,j}$ , ou catástrofe, é qualquer evento na  $\sigma$ -álgebra gerada por  $\mathbf{Y}_3$ . Um evento preditor ou alarme para o evento de interesse  $C_{t,j}$  é qualquer evento na  $\sigma$ -álgebra gerada por  $\mathbf{Y}_2$ .

$C_{t,j}$  é o evento cruzamento superior  $C_{t,j} = \{Y_{t+j-1} \leq u < Y_{t+j}\}$  para algum inteiro positivo  $j$  e para algum número real  $u$ .

**Definição 2** (ANTUNES; AMARAL-TURKMAN; TURKMAN, 2003). Um alarme é dado, em um instante  $t$ , para um evento futuro  $C_{t,j}$ , se o valor observado de  $\mathbf{Y}_2$  pertence ao evento preditor. Se o evento de interesse ocorre quando um alarme é dado, então diz-se que o evento é detectado. Se o evento de interesse não ocorre quando um alarme é dado, então diz-se que um alarme falso foi dado.

As probabilidades preditivas para quaisquer eventos  $C_{t,j}$  e  $A_{t,j}$  gerados, respectivamente, por  $\mathbf{Y}_3$  e  $\mathbf{Y}_2$  são dadas por

$$P(A_{t,j} | D_t) = \int_{\theta} P(A_{t,j} | \theta, D_t) p(\theta, D_t) d\theta$$

$$P(C_{t,j} | A_{t,j}, D_t) = \frac{\int_{\theta} P(A_{t,j} \cap C_{t,j} | \theta, D_t) p(\theta, D_t) d\theta}{\int_{\theta} P(A_{t,j} | \theta, D_t) p(\theta, D_t) d\theta}.$$

**Definição 3** (ANTUNES; AMARAL-TURKMAN; TURKMAN, 2003). Um evento  $A_{t,j} \in \sigma_{\mathbf{Y}_2}$  é uma região de alarme ou evento preditor para o evento  $C_{t,j}$  em  $\mathbf{Y}_3$  se, sempre que  $\mathbf{Y}_2 \in A_{t,j}$ , é provável que o evento  $C_{t,j}$  ocorrerá.

**Definição 4** (ANTUNES; AMARAL-TURKMAN; TURKMAN, 2003). A região

de alarme  $A_{t,j} \in \sigma_{\mathbf{Y}_2}$  tem tamanho  $\alpha_{t,j}$  e probabilidade de detecção  $\gamma_{t,j}$  se

$$\alpha_{t,j} = P(A_{t,j} | D_t)$$

$$\gamma_{t,j} = P(C_{t,j} | A_{t,j}, D_t).$$

$\alpha_{t,j}$  pode ser visto como a proporção de tempo gasto no estado do alarme, ao passo que  $\gamma_{t,j}$  é a probabilidade de predizer corretamente um evento  $j$  – passos à frente.

**Definição 5** (ANTUNES; AMARAL-TURKMAN; TURKMAN, 2003). A região de alarme  $A_{t,j}$  é ótima de tamanho  $\alpha_{t,j}$  se

$$P(C_{t,j} | A_{t,j}, D_t) = \sup_{B \in \sigma_{\mathbf{Y}_2}} P(C_{t,j} | B, D_t) \quad (1)$$

em que o supremo é tomado sobre todos os conjuntos tal que  $P(B | D_t) = \alpha_{t,j}$ .

A região de alarme  $A_{t,j}$  é ótima se, entre todas as regiões com o mesmo tamanho de alarme, ela tem a mais alta probabilidade de detecção. Ou ainda, a região de alarme é ótima se ela fornece o menor número de alarmes falsos para uma dada probabilidade de detecção (ANTUNES; AMARAL-TURKMAN; TURKMAN, 2003).

**Definição 6** (ANTUNES; AMARAL-TURKMAN; TURKMAN, 2003). Um sistema de alarme ótimo de tamanho  $\alpha_{t,j}$  é uma família de regiões de alarme  $A_{t,j}$  que satisfazem a Expressão (1).

**Lema 1** (LINDGREN, 1985): Sejam  $p(\mathbf{y}_2 | D_t)$  e  $p(\mathbf{y}_2 | C_{t,j}, D_t)$  as densidades preditivas de  $\mathbf{Y}_2$  e  $\mathbf{Y}_2$  condicionado ao evento  $C_{t,j}$ , respectivamente. O sistema de alarme  $A_{t,j}$  com

$$A_{t,j} = \left\{ \mathbf{y}_2 \in \mathfrak{R}^p : \frac{p(\mathbf{y}_2 | C_{t,j}, D_t)}{p(\mathbf{y}_2 | D_t)} \geq k_{t,j} \right\}$$

ou equivalentemente,

$$A_{t,j} = \left\{ \mathbf{y}_2 \in \mathfrak{R}^p : \frac{p(C_{t,j} | \mathbf{y}_2, D_t)}{p(C_{t,j} | D_t)} \geq k_{t,j} \right\}, \quad (2)$$

em que  $k_{t,j}$  é tal que  $P(\mathbf{Y}_2 \in A_{t,j} | D_t) = \alpha_{t,j}$ , é ótima de tamanho  $\alpha_{t,j}$ .

As definições acima asseguram que, em qualquer tempo  $t$ , para um dado tamanho de alarme  $\alpha_{t,j}$ ,  $A_{t,j}$  fornece a mais alta probabilidade de detecção. A questão é como escolher o próprio tamanho do alarme. Uma escolha lógica é  $\alpha_{t,j} = P(C_{t,j} | D_t)$ , ou seja, a proporção de tempo gasto na região de alarme é igual à proporção de tempo na região de catástrofe  $C_{t,j}$  (ANTUNES; AMARAL-TURKMAN; TURKMAN, 2003).

A região de alarme (2) pode ainda ser reescrita

$$A_{t,j} = \{y_2 \in \mathfrak{R}^p : p(C_{t,j} | y_2, D_t) \geq \bar{k}\} \quad (3)$$

sendo  $\bar{k} = k_{t,j} P(C_{t,j} | D_t)$ . Dessa forma, a região de alarme depende do valor de  $k_{t,j}$ . A seguir são listadas algumas sugestões para escolha de  $\bar{k}$ .

Antunes, Amaral-Turkman e Turkman (2003) estabeleceram que o valor ótimo de  $\bar{k}$ , denominado aqui por  $\bar{k}^*$ , pode ser obtido fixando o tamanho do alarme  $\alpha_{t,j} = P(C_{t,j} | D_t)$  ou, de maneira alternativa, calculando-se as características do alarme para diferentes valores de  $\bar{k}$  e escolhendo aqueles que resultam na melhor coleção de características operacionais.

Svensson, Holst e Lindquist (1996) sugeriram que  $\bar{k}$  deveria ser escolhido de modo que a probabilidade de alarme correto e a probabilidade de detecção do evento fossem aproximadamente iguais, isto é,  $P(C_{t,j} | A_{t,j}) \cong P(A_{t,j} | C_{t,j})$ .

Com a determinação da região de alarme (3) é possível calcular as características do sistema de alarme dadas por

1. Tamanho do alarme:  $P(A_{t,j} | D_t)$ .
2. Probabilidade de detecção do evento:  $P(A_{t,j} | C_{t,j}, D_t)$ .
3. Probabilidade de alarme correto:  $P(C_{t,j} | A_{t,j}, D_t)$ .
4. Probabilidade de alarme falso:  $P(C_{t,j}^c | A_{t,j}, D_t) = 1 - P(C_{t,j} | A_{t,j}, D_t)$ .
5. Probabilidade de não detecção do evento:  $P(A_{t,j}^c | C_{t,j}, D_t) = 1 - P(A_{t,j} | C_{t,j}, D_t)$ .

### Sistema de alarme ótimo para processos TARSO

Seja  $\{Y_t, t = 1, \dots, n\}$  o processo TARSO(2;1,1,1,1) com dois regimes dado por

$$Y_t = \begin{cases} \phi_{11} Y_{t-1} + \beta_{11} Z_{t-1} + a_t^{(1)}, & \text{se } Z_{t-d} \leq r \\ \phi_{21} Y_{t-1} + \beta_{21} Z_{t-1} + a_t^{(2)}, & \text{se } Z_{t-d} > r, \end{cases} \quad (4)$$

sendo  $a_t^{(i)} \sim N(0, \tau_i^{-1})$ ,  $\gamma_i = (\phi_{i1}, \beta_{i1})'$  e  $\tau_i$  os parâmetros com  $i=1,2$ . Fazendo  $X_{i,t} = (Y_{t-1}, Z_{t-1})'$ , o modelo (4) pode ser reescrito como

$$Y_t = \begin{cases} \gamma_1' X_{1t} + a_t^{(1)}, & \text{se } Z_{t-d} \leq r \\ \gamma_2' X_{2t} + a_t^{(2)}, & \text{se } Z_{t-d} > r \end{cases}$$

Tomando  $j=2$ , o evento de interesse a ser predito – a catástrofe – toma a forma

$$C_{t,2} = \{Y_{t+1} \leq u < Y_{t+2}\}$$

e a região de alarme  $A_{t,2}$  desse evento torna-se

$$A_{t,2} = \{y_t : P(C_{t,2} | y_t, D_t) > k_{t,2} P(C_{t,2} | D_t)\},$$

ou de forma equivalente

$$A_{t,2} = \{y_t : P(C_{t,2} | y_t, D_t) > \bar{k}\} \quad (5)$$

em que  $\bar{k} = k_{t,2} P(C_{t,2} | D_t)$ .

Pela Expressão (5), o primeiro passo na construção do sistema de alarme consiste em calcular a probabilidade  $P(C_{t,2} | y_t, D_t)$ , probabilidade de catástrofe condicional em  $y_t$  e  $D_t$ . Essa probabilidade e a probabilidade de catástrofe condicional em  $D_t$ ,  $P(C_{t,2} | D_t)$ , são dadas por

$$\begin{aligned} P(C_{t,2} | y_t, D_t) &= \int_{-\infty}^u \int_u^{\infty} P(y_{t+2}, y_{t+1} | y_t, D_t) dy_{t+2} dy_{t+1} = \\ &= \int_{-\infty}^u \left[ \int_u^{\infty} P(y_{t+2} | y_{t+1}, y_t, D_t) dy_{t+2} \right] P(y_{t+1} | y_t, D_t) dy_{t+1} = \\ &= \int_{-\infty}^{u_1} \left[ \int_{u_2}^{\infty} P(w_{t+2} | w_{t+1}, w_t, D_t) dw_{t+2} \right] P(w_{t+1} | w_t, D_t) dw_{t+1} = \\ &= \int_{-\infty}^{u_1} P(w_{t+1} | w_t, D_t) [1 - F_v(u_2)] dw_{t+1} \end{aligned} \quad (6)$$

e

$$\begin{aligned}
 P(C_{t,2} | D_t) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(y_{t+2}, y_{t+1} | D_t) P(y_t | D_t) dy_{t+2} dy_{t+1} dy_t = \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} P(y_{t+2} | y_{t+1}, D_t) dy_{t+2} \right] P(y_{t+1} | D_t) dy_{t+1} \\
 &= \int_{-\infty}^{\infty} P(y_t | D_t) dy_t = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(w_{t+2} | w_{t+1}, D_t) dw_{t+2} \left[ \int_{-\infty}^{\infty} P(w_{t+1} | D_t) dw_{t+1} P(w_t | D_t) dw_t = \right. \\
 &= \int_{-\infty}^{\infty} P(w_t | D_t) dw_t \int_{-\infty}^{\infty} P(w_{t+1} | D_t) [1 - F_v(u_2)] dw_{t+1} \quad (7)
 \end{aligned}$$

As variáveis  $w_t, w_{t+1}$  e  $w_{t+2}$  são variáveis transformadas, isto é,  $W_{n+h} = [V^{(h)}]^{-\frac{1}{2}} (Y_{n+h} - E^{(h)})$ .  $F_v$  é a função de distribuição com  $\nu = n_i + 2$  graus de liberdade.

As probabilidades que aparecem nas Expressões (6) e (7) são distribuições preditivas e para qualquer  $h$ , a distribuição preditiva de  $y_{n+h}$  dado  $\mathbf{y}, d$  e  $r$  é uma distribuição t-Student não central com  $\nu = n_i + h$  graus de liberdade, média  $E^{(h)}$  e variância  $V^{(h)}$ . Detalhes da demonstração dessa distribuição podem ser encontrados em Gonçalves (2011).

Dessa forma tem-se,

$$P(y_{n+h} | \mathbf{y}, d, r) \sim St(\nu, E^{(h)}, V^{(h)}).$$

Na notação de matrizes,

$$E^{(h)} = \frac{(\mathbf{B}_i^{(n+h-1)})' (\mathbf{A}_i^{(n+h)})^{-1} X_{in+h}}{(1 - X_{in+h}' \mathbf{A}_i^{(n+h)})^{-1} X_{in+h}} \quad (8)$$

e

$$\begin{aligned}
 V^{(h)} &= \left\{ \frac{C^{(n+h-1)} - (\mathbf{B}_i^{(n+h-1)})' (\mathbf{A}_i^{(n+h)})^{-1} \mathbf{B}_i^{(n+h-1)}}{(1 - X_{in+h}' \mathbf{A}_i^{(n+h)})^{-1} X_{in+h}} \right. \\
 &\quad \left. - \left( \frac{\mathbf{B}_i^{(n+h)} (\mathbf{A}_i^{(n+h)})^{-1} X_{in+h}}{1 - X_{in+h}' (\mathbf{A}_i^{(n+h)})^{-1} X_{in+h}} \right)^2 \right\} \frac{1}{(n_i + h)}. \quad (9)
 \end{aligned}$$

Fazendo  $h = 1$  e  $h = 2$  nas Expressões (8) e (9) tem-se

$$u_1 = [V^{(1)}]^{-\frac{1}{2}} (u - E^{(1)})$$

e

$$u_2 = [V^{(2)}]^{-\frac{1}{2}} (u - E^{(2)}).$$

No desenvolvimento do modelo

$$Y_t = \begin{cases} \phi_{11} Y_{t-1} + \beta_{11} Z_{t-1} + a^{(1)}_t, & \text{se } Z_{t-d} \leq r \\ \phi_{21} Y_{t-1} + \beta_{21} Z_{t-1} + a^{(2)}_t, & \text{se } Z_{t-d} > r, \end{cases}$$

tem-se

$$\mathbf{x}_{n+h} = \begin{pmatrix} Y_{n+h-1} \\ Z_{n+h-1} \end{pmatrix}, \mathbf{B}^{(m)}_i = \begin{pmatrix} \sum_{t=p+1}^m Y_{t-1} Y_t \\ \sum_{t=p+1}^m Z_{t-1} Y_t \end{pmatrix}, C^{(m)} = \left( \sum_{t=p+1}^m Y_t^2 \right) e \mathbf{A}_i^{(m)} = \begin{pmatrix} \sum_{t=p+1}^m Y_{t-1}^2 & \sum_{t=p+1}^m Y_{t-1} Z_{t-1} \\ \sum_{t=p+1}^m Y_{t-1} Z_{t-1} & \sum_{t=p+1}^m Z_{t-1}^2 \end{pmatrix},$$

sendo o número  $m$  de elementos somados nessas expressões substituídos pelas somas que aparecem em  $E^{(h)}$  e  $V^{(h)}$ , Expressões (8) e (9).

Em  $P(C_{t,2} | D_t)$ , as distribuições são condicionadas somente ao passado e  $P(C_{t,2} | D_t)$  é fixa para qualquer  $t$ . As expressões  $E^{(h)}$  e  $V^{(h)}$ , neste caso, são dadas por

$$E^{(h)} = \frac{(\mathbf{B}_i^{(n+h-2)})' (\mathbf{A}_i^{(n+h-1)})^{-1} X_{in+h}}{(1 - X_{in+h}' \mathbf{A}_i^{(n+h-1)})^{-1} X_{in+h}}$$

e

$$\begin{aligned}
 V^{(h)} &= \left\{ \frac{C^{(n+h-2)} - (\mathbf{B}_i^{(n+h-2)})' (\mathbf{A}_i^{(n+h-1)})^{-1} \mathbf{B}_i^{(n+h-2)}}{(1 - X_{in+h}' \mathbf{A}_i^{(n+h-1)})^{-1} X_{in+h}} \right. \\
 &\quad \left. - \left( \frac{\mathbf{B}_i^{(n+h-1)} (\mathbf{A}_i^{(n+h-1)})^{-1} X_{in+h}}{1 - X_{in+h}' (\mathbf{A}_i^{(n+h-1)})^{-1} X_{in+h}} \right)^2 \right\} \frac{1}{(n_i + h)}.
 \end{aligned}$$

A partir do cálculo de  $P(C_{t,2} | y_t, D_t)$  e  $P(C_{t,2} | D_t)$  é possível determinar a região de alarme (5) e calcular as características do sistema de alarme, ou seja, tamanho do alarme, probabilidade de detecção do evento, probabilidade de alarme correto, probabilidade de alarme falso e probabilidade de não detecção do evento.

Um procedimento sugerido por Antunes, Amaral-Turkman e Turkman (2003) para determinação das características operacionais do alarme é dado a seguir:

1. Forme um *grid*  $G = \{y_t \in \mathfrak{R}\}$  de maneira que

$$P[(Y_t) \in G | D_t] \cong 1;$$

2. para cada  $y_t \in G$ , calcule a probabilidade

$$P(C_{t,2} | y_t, D_t);$$

3. para vários valores adequadamente escolhidos de  $\bar{k}$ , forme os conjuntos

$$A_{t,2} = \left\{ y_t \in G : \int_{-\infty}^{u_t} P(w_{t+1} | w_t, D_t) [1 - F_v(u_2)] dw_{t+1} \geq \bar{k} \right\}$$

em que  $\bar{k} = k_{t,2} P(C_{t,2} | D_t)$ ;

4. para cada  $\bar{k}$ , calcule o tamanho  $\alpha_{t,\bar{k}}$  de  $A_{t,2}$  por

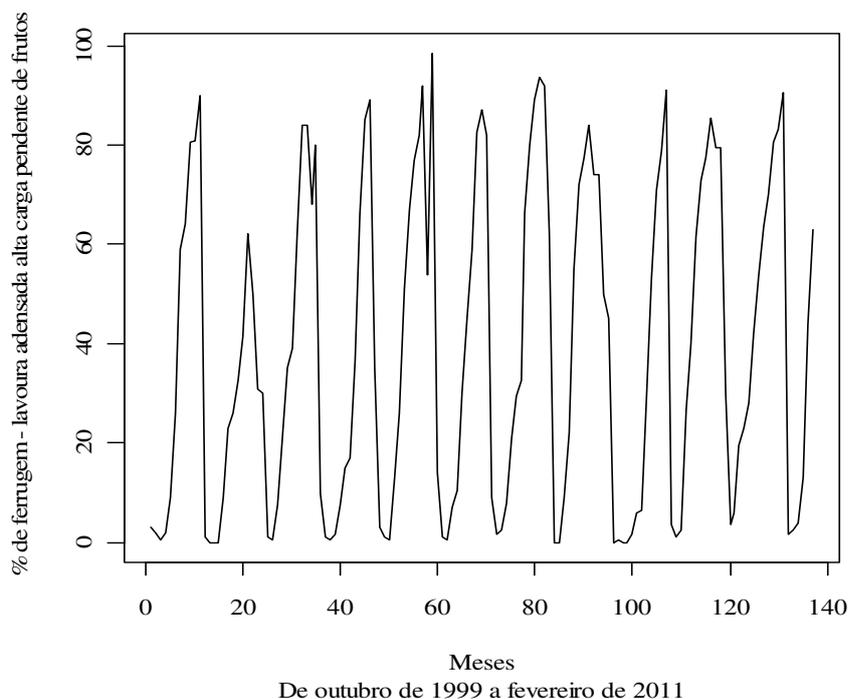
$$\alpha_{t,\bar{k}} = P(Y_t \in A_{t,2} | D_t) = \int_{A_{t,2}} P(y_t | D_t) dy_t.$$

As características operacionais do alarme podem então ser calculadas para cada valor de  $\bar{k}$ , ou simplesmente para o valor ótimo de  $\bar{k}$ ,  $\bar{k}^*$ .

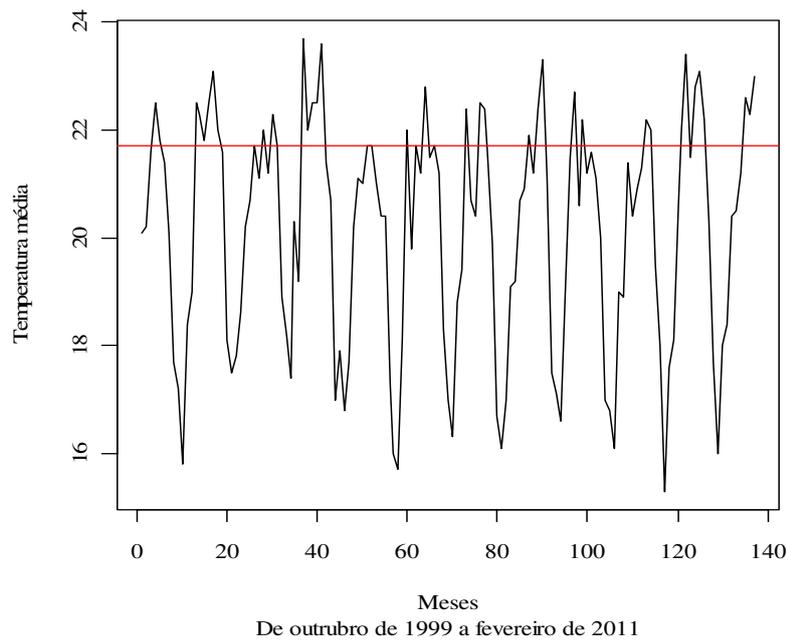
## Resultados e Discussão

Nesta seção serão apresentados os resultados do sistema de alarme ótimo do modelo TARSO (2;1,0,1,1) para a série de incidência da ferrugem do cafeeiro em lavoura adensada com produção alta na presença da variável temperatura média. Nas Figuras 1 e 2 são apresentados, respectivamente, os gráficos dessas séries sendo que a linha vermelha no gráfico de temperatura média indica o que se chama de valor *threshold*  $r$ . O valor *threshold* é também chamado de valor limiar e separa a série em dois regimes.

Figura 1. Índices de ferrugem.



Fonte: Elaboração dos autores.

**Figura 2.** Valores de temperatura média com valor *threshold*  $r = 21,5^{\circ}C$ .

**Fonte:** Elaboração dos autores.

Ambas as séries são sazonais de periodicidade igual a 12 meses e o modelo TARSO foi ajustado utilizando a metodologia bayesiana considerando a priori de Jeffreys,  $d = 1$  e  $r = 21,5^{\circ}C$ . O valor  $r$  de temperatura tomado pertence à faixa considerada ótima no alastramento da doença ferrugem. A equação ajustada do modelo (4) é dada por

$$Y_t = \begin{cases} 0,8391Y_{t-1}, & \text{se } Z_{t-1} \leq 21,5 \\ 1,1115Y_{t-1} + 0,3267Z_{t-1}, & \text{se } Z_{t-1} > 21,5 \end{cases}$$

Nota-se que o índice de ferrugem no instante  $t$  é influenciado apenas pelo seu valor defasado de 1 para temperaturas abaixo do limiar  $r = 21,5^{\circ}C$ , já para valores de temperaturas acima  $r = 21,5^{\circ}C$ , o índice de ferrugem no instante  $t$  é respondido pelo índice do período anterior e também pela temperatura do período anterior. Maiores detalhes desse ajuste podem ser encontrados em Gonçalves (2011).

O evento de interesse é a catástrofe definida pelo cruzamento superior de um nível fixo  $u$  em  $t + 2$  dado por

$$C_{t,2} = \{Y_{t+1} \leq u < Y_{t+2}\} \quad (10),$$

sendo  $u$  tomado como 5%, 10% e 15%. Nesses índices, o nível de ferrugem ainda é considerado baixo e, portanto, eficazes para a inserção de medidas de controle da doença.

Os pontos de estudo das características operacionais do alarme são  $t=12$ ,  $t=39$  e  $t=133$ . Foram escolhidos, dessa forma, de modo a permitir a análise do comportamento do alarme em pontos com poucas e muitas informações passadas e também em pontos nos quais o evento de interesse, a catástrofe, dada pela Expressão (10), é satisfeita e onde não é.

Na Tabela 1 são apresentados os pontos de análise, o nível de cruzamento superior ( $u\%$ ), o valor de  $\bar{k}^*$  e as regiões de alarme ótimas (RA) a partir das quais são obtidas as características do alarme. O valor de  $\bar{k}^*$  é o valor que faz com que  $P(C_{t,2} | D_t) \approx P(A_{t,2} | D_t)$ . Essa aproximação pode ser vista na Tabela 2.

**Tabela 1.** Nível de cruzamento,  $\bar{k}^*$ , e região de alarme dos pontos de cruzamento analisados na construção do alarme.

t	u%	$\bar{k}^*$	RA
12	5	1,3301	[92;100]
39	10	1,4921	[1;5,5]
133	15	0,2845	[1;19,5]

Nota: RA – Região de Alarme.

Fonte: Elaboração dos autores.

Na Tabela 2 são apresentadas as características operacionais do alarme: tamanho do alarme ( $P(A_{t,2} | D_t) = P(A | D)$ ), probabilidade de detecção do evento ( $P(A_{t,2} | C_{t,2}, D_t) = P(A | C, D)$ ), probabilidade de alarme correto ( $P(C_{t,2} | A_{t,2}, D_t) = P(C | A, D)$ ), probabilidade de alarme falso ( $P(C_{t,2}^c | A_{t,2}, D_t) = P(C^c | A, D)$ ) e probabilidade de não detecção do evento ( $P(A_{t,2}^c | C_{t,2}, D_t) = P(A^c | C, D)$ ) para os pontos t=12, t=39 e t=133. O ponto t=12 refere-se a setembro de 2000 no qual nota-se, para o nível de cruzamento u=5% uma probabilidade de detecção do evento e probabilidade de alarme correto em torno de 0%. Dessa forma seus complementares, probabilidade

de alarme falso e não detecção do evento, ficam muito próximos de 100%. O índice de infecção em t=39 refere-se a dezembro de 2002 e considerando um cruzamento de u=10%, a probabilidade de detecção do evento é de 60,70% e a probabilidade de alarme correto é de 52,13%. Neste ponto, as probabilidades de alarme falso e de não detecção do evento são, respectivamente, 47,87% e 39,30%. A observação t=133 refere-se a outubro de 2010 e considerando um nível de cruzamento superior de 5%, a probabilidade de detecção do evento é de 41,04% e a probabilidade de alarme correto é de 41,89%. As probabilidades de alarme falso e de não detecção do evento ficam em torno de 58%.

**Tabela 2.** Características operacionais do alarme em t=12, t=39, e t=133 do processo TARSO (2;1,0,1,1).

u	t	$P(C_{t,2}   D_t)$	$P(A   D)$	$P(A   C, D)$	$P(C   A, D)$	$P(C^c   A, D)$	$P(A^c   C, D)$
5%	12	0,0521	0,0594	0,0010	0,0009	0,9991	0,9990
5%	133	0,2495	0,2444	0,4104	0,4189	0,5811	0,5896
10%	39	0,2141	0,2493	0,6070	0,5213	0,4787	0,3930

Fonte: Elaboração dos autores.

O desempenho do alarme pode ser melhor justificado analisando a Tabela 3 na qual são apresentados os índices de ferrugem nos pontos

considerados e seus dois futuros mais próximos que são os índices que compõem o evento catástrofe dado pela Expressão (10).

**Tabela 3.** Índices de ferrugem em t, t+1 e t+2.

t	$Y_t$	$Y_{t+1}$	$Y_{t+2}$
12	1,0%	0,0%	0%
39	1,5%	7,5%	15%
133	2,5%	4,0%	13%

**Fonte:** Elaboração dos autores.

Analisando a Tabela 3, o desempenho ruim do alarme em t=12 justifica-se pelo fato do evento catástrofe,  $C_{t,2} = \{Y_{t+1} \leq u < Y_{t+2}\}$ , não ser satisfeito para o nível de cruzamento considerado. Porém, qualquer outro nível que se considerasse o alarme teria desempenho bom, já que, na prática, uma catástrofe aqui é rara, pois em setembro o café está sem folhas e está fora do período infeccioso que se estende de dezembro a maio.

As probabilidades encontradas em outubro e dezembro já eram esperadas, isto é, em t=133 e t=39, já que inicia-se o período de infecção que coincide com o início do período chuvoso.

## Conclusões

O sistema de alarme se comporta razoavelmente bem em pontos nos quais a catástrofe ainda não ocorreu. O sistema fornece probabilidades mais altas de alarme correto e detecção em níveis de cruzamento mais altos. O maior número de observações passadas em um ponto parece não favorecer tanto as características do alarme.

## Agradecimentos

À CAPES e a FAPEMIG ( APQ 01714-09) o auxílio financeiro. À Fundação Procafé pela concessão dos dados.

## Referências

ANTUNES, M.; AMARAL-TURKMAN, M. A.; TURKMAN, K. F. F. A bayesian approach to event

prediction. *Journal of Time Series Analysis*, Clevedon, v. 24, n. 6, p. 631-646, dec. 2003.

COSTA, C.; SCOTTO, M. G.; PEREIRA, I. Optimal alarm systems for fiaparch processes. *Revstat – Statistical Journal*, Lisboa, v. 8, n. 1, p. 37-55, 2010.

FUNDAÇÃO PROCAFÉ. *Estação de avisos*: Sul de Minas. 2011. Disponível em: <<http://www.fundacaoprocafe.com.br=estacaoavisos=suldeminas=boletimdeaviso>>. Acesso em: 23 maio 2011.

GONÇALVES, L. R. *Sistema de alarme ótimo para o modelo TARSO com aplicação na ferrugem do café*. 2011. Tese (Doutorado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, MG.

JAPIASSÚ, L. B.; GARCIA, A. W. R.; MIGUEL, A. E.; CARVALHO, C. H. S.; FERREIRA, R. A.; PADILHA, L.; MATIELLO, J. B. Influência da carga pendente, do espaçamento e de fatores climáticos no desenvolvimento da ferrugem do cafeeiro. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 5., 2007, Águas de Lindóia. *Anais...* Brasília: EMBRAPA, 2007. p. 5.

LINDGREN, G. Model processes in nonlinear prediction with application to detection and alarm. *The Annals of Probability*, Hayward, v. 8, n. 4, p. 775-792, 1980.

\_\_\_\_\_. Optimal prediction of level crossing in Gaussian processes and sequences. *The Annals of Probability*, Hayward, v. 13, n. 3, p. 804-824, 1985.

\_\_\_\_\_. Prediction from a random time point. *The Annals of Probability*, Hayward, v. 3, n. 3, p. 412-423, 1975.

MARÉ, J. de. Optimal prediction of catastrophes with applications to gaussian processes. *The Annals of probability*, Hayward, v. 8, n. 4, p. 841-850, 1980.

MONTEIRO, M.; PEREIRA, I.; SCOTTO, M. G. Optimal alarm systems for count processes. *Communications in Statistics-Theory and Methods*, New York, v. 37, n. 19, p. 3054-3076, 2008.

SÁFADI, T.; MORETTIN, P. A. Análise Bayesiana do modelo open loop threshold autoregressive. *Revista Brasileira de Estatística*, Rio de Janeiro, v. 62, n. 217, p. 91-105, 2001.

SVENSSON, A.; HOLST, J.; LINDQUIST, R. Optimal prediction of catastrophes in autoregressive moving-average processes. *Journal of time Series Analysis*, Clevedon, v. 17, n. 5, p. 511-531, oct. 1996.

SVENSSON, A.; HOLST, J. Prediction of high water levels in the Baltic. *Journal of the Turkish Statistical Association*, Ankara, v. 1, n. 1, p. 39-46, 1997.

TURKMAN, M. A. A.; TURKMAN, K. F. Optimal alarm systems for autoregressive processes: a bayesian approach. *Computational Statistics & Data Analysis*, New York, v. 10, n. 3, p. 307-314, dec. 1990.

