

Sistema de vigilância para detecção de interação espaço-tempo de eventos pontuais

Taynãna C. Simões, Renato M. Assunção

Departamento de Estatística – Universidade Federal de Minas Gerais (UFMG)
Caixa Postal: 702 – 31270-901 – Belo Horizonte – MG – Brasil
tay_estatistica@yahoo.com.br, assuncao@est.ufmg.br

Abstract. *We introduce a surveillance system to monitor point events located in geographical space and in time. The aim is to detect the space-time clusters as soon they start emerge by triggering an alarm. In this article, the main objective is to adapt a technique proposed by Rogerson (2001), aiming at identifying the eventual cluster that would be the principal cause the alarm going off. Additionally, we believe that our technique isolate events that don't belong to the cluster but that contribute to the alarm warning.*

Resumo. *Propomos um sistema de vigilância para o monitoramento de eventos pontuais, a fim de detectar rapidamente a formação de conglomerados espaço-tempo, assim que estes começam a surgir, disparando um alarme. Neste artigo, o objetivo principal é adaptar a técnica apresentada por Rogerson (2001), com o intuito de identificar o cluster eventualmente existente e que seria a principal causa do disparo do alarme. Além disso, acredita-se que a técnica proposta permita isolar eventos que não pertençam ao eventual cluster, mas que contribuam para fazer o alarme soar.*

1. Introdução

Detectar aglomerados espaço-tempo de forma rápida, eficiente e em tempo real é uma necessidade em várias áreas do conhecimento. Na área da saúde principalmente, técnicas e ferramentas com este fim têm uma importância particular, pois permitem que intervenções ou ações de controle sejam realizadas de forma eficiente.

O sistema de vigilância, proposto neste artigo, consiste no monitoramento de um conjunto de eventos pontuais, a fim de detectar rapidamente a formação de conglomerados espaço-temporais, assim que estes começam a surgir. O sistema é um método estatístico prospectivo, no qual é feita análise repetida de dados acumulados seqüencialmente ao longo do tempo, com o objetivo de detectar rapidamente uma mudança relevante que ocorra na série de dados.

Utilizando a estatística de detecção de conglomerados espaço-tempo de Knox e métodos de soma acumulada (CUSUM), Rogerson (2001) propõe um sistema de vigilância que detecte clusters ativos (existentes no momento da análise). No entanto, dois problemas verificados nesta técnica são: 1 - eventos que disparam o alarme podem

não pertencer ao eventual cluster; 2 – a técnica é capaz de detectar a presença de clusters localizados, mas não identifica estes clusters, dando sua posição e extensão no espaço e tempo.

Desta forma, nosso trabalho procura adaptar a técnica apresentada por Rogerson (2001), com o intuito de identificar o cluster eventualmente existente e que seria a principal causa do disparo do alarme. Acredita-se ainda que a técnica proposta permita isolar eventos que não pertençam ao cluster, mas que contribuam para fazer o alarme soar. A idéia é monitorar os novos eventos não através de uma Soma Acumulada, mas através de superfícies acumuladas. De forma resumida, a cada novo evento, é calculado um escore de interação espaço-temporal localizado e este escore é distribuído no espaço através de uma densidade de kernel, gerando uma superfície na região do evento. Estas superfícies são acumuladas e geram uma saliência pronunciada em torno de um eventual cluster.

2. Vigilância Estatística

Um sistema de vigilância monitora mudanças quando novos eventos tornam-se disponíveis, no decorrer do estudo. A vigilância estatística significa um monitoramento de um processo estocástico $X = \{X(t); t=1, 2, \dots\}$ com o objetivo de detectar uma mudança importante no processo, em um tempo desconhecido τ , tão rápida e precisamente possível.

A cada instante de tempo s , deve-se discriminar entre dois estados no sistema monitorado: sob-controle e fora-de-controle. Para que isso ocorra, utilizam-se os eventos acumulados até s (denotado por X_s) para formar conjuntos alarme. Se X_s pertencerem ao conjunto alarme, há uma indicação de que o processo está no estado fora-de-controle e um alarme é soado. Usualmente, isso é feito usando uma função alarme $p(X_s)$ e um limite de controle h . O tempo de um alarme t_A é escrito como

$$t_A = \min\{s, p(X_s) > h\}$$

Diferentes tipos de medidas são utilizados para avaliar um sistema de vigilância, caracterizando seu comportamento quando o processo está sob-controle e fora-de-controle. A distribuição de um alarme falso, por exemplo, é freqüentemente resumida pelo número médio de eventos até que o alarme soe, dado que o processo está sob-controle (ARL_0).

$$ARL_0 = E[t_A | \tau = \infty]$$

onde τ é o tempo verdadeiro de mudança no processo e que é desconhecido na prática.

Outra medida normalmente utilizada é a probabilidade de um alarme falso:

$$P(t_A < \tau) = \sum_{t=1}^{\infty} P(\tau = t)P(t_A < \tau | \tau = t)$$

Geralmente, a distribuição de probabilidade $P(\tau = t)$ para τ não é conhecida e isto dificulta o uso desta última medida.

Quando um sistema de vigilância é avaliado, deve-se encarar um *trade-off* entre alarmes falsos e tempos de espera curtos para observar um alarme verdadeiro.

3. Sistema de Vigilância – Rogerson (2001)

Rogerson (2001) propõe um sistema de vigilância que combina métodos de soma acumulada (CUSUM) com uma estatística de detecção de conglomerados espaço-tempo para um conjunto de dados pontuais (Teste de Knox).

3.1. Teste de Knox

O teste de Knox é um teste baseado na contagem do número de pares de eventos que ocorrem dentro de intervalos críticos pré-especificados de tempo (T) e distância (D).

Dados n pontos localizados no tempo e no espaço, existem $n(n-1)/2$ pares distintos que podem ser divididos em uma matriz 2×2 (Tabela 1), indicando os pares que estão próximos no espaço ou próximos no tempo.

Tabela 1. Tabela 2×2 para o cálculo do Teste de Knox (Os valores indicados por • são obtidos a partir dos demais).

Situação	Próximos no espaço	Não próx. no espaço	Total
Próximos no tempo	n_{st}	•	n_t
Não próx. no tempo	•	•	•
Total	n_s	•	$n(n-1)/2$

Sejam n_s os pares de eventos observados que são próximos no espaço (separados por uma distância menor ou igual a D), n_t os pares de eventos que são próximos no tempo (separados no tempo por menos que o intervalo crítico T), e n_{st} os pares de eventos que são próximos em ambos, espaço e tempo. A estatística de teste N_{st} deve ser comparada com o seu valor esperado, sob a hipótese nula H_0 , de que não há interação espaço-tempo. Caso o número de eventos próximos no espaço e no tempo exceda substancialmente este valor esperado, há uma indicação da existência de interação espaço-tempo.

Esta estatística testa a presença de interação de forma global, sem localizar os clusters, e sem preocupar-se em verificar se os clusters ainda estão presentes ou se já deixaram de existir. Para resolver estes dois problemas, Rogerson (2001) propôs o uso das estatísticas locais de Knox.

3.2. Estatística Local de Knox

Sejam $n_s(i)$ o número de eventos que são próximos ao i -ésimo evento no espaço, $n_t(i)$ o número de eventos que são próximos ao i -ésimo evento no tempo e $n_{st}(i)$ o número de eventos que são próximos simultaneamente no espaço e tempo ao evento i .

Para encontrar a distribuição de $N_{st}(i)$, sob a hipótese nula de não interação espaço-tempo, usa-se que cada permutação aleatória dos índices dos tempos, mantendo as posições espaciais fixas, é igualmente provável. Seja $n_t^j(i)$ o número de pontos que estão próximos no tempo do evento i quando a este é associado o j -ésimo valor do

tempo. Mostra-se que a distribuição de $N_{st}(i)$ é hipergeométrica com parâmetros $n-1$, $n_s(i)$ e $n_t^j(i)$. Considerando todos os valores possíveis de tempo para o i -ésimo evento, mostra-se que $N_{st}(i)$ é uma soma ponderada de hipergeométricas. Pode-se mostrar também que:

$$E\{N_{st}(i)\} = \frac{2n_t n_s(i)}{n(n-1)}$$

$$Var\{N_{st}(i)\} = \left[\sum_{j=1}^n (n_t^j(i))^2 \right] \frac{n_s(i)}{n(n-1)^2} \left[\frac{n-1-n_s(i)}{n-2} + n_s(i) \right]$$

$$+ \left[\frac{2n_t n_s(i)}{n(n-1)} \frac{n-1-n_s(i)}{n-2} - \left(\frac{2n_t n_s(i)}{n(n-1)} \right)^2 \right]$$

Padroniza-se $N_{st}(i)$, resultando na seguinte estatística escore z_i ajustada:

$$z_i = \frac{n_{st}(i) - E\{N_{st}(i)\} - 0.5}{\sqrt{Var\{N_{st}(i)\}}}$$

A variável z_i possui distribuição aproximadamente normal, com média zero e variância 1 (denotada por $N(0,1)$), caso o processo esteja sob controle. Caso exista um cluster, espera-se que z_i tenha distribuição com média maior que zero, se o evento i fizer parte do cluster. A distribuição de z_i tende a permanecer aproximadamente igual a $N(0,1)$ para aqueles eventos que não fazem parte do cluster.

Suponha que $i-1$ eventos já tenham ocorrido. Pode-se comparar o valor da estatística de Knox após o caso i (denotado K_i) com o valor que seria esperado sob a hipótese nula e condicionado ao valor da estatística de Knox após o evento $i-1$ e valores observados de $n_s(i)$ e $\{n_t^j(i); j=1, \dots, i\}$:

$$z_i^* = \frac{K_i - E\{K_i | K_{i-1}, n_s(i), n_t(i)\} - 0.5}{\sqrt{Var\{K_i | K_{i-1}, n_s(i), n_t(i)\}}}$$

A informação em z_i^* , obtida nesta comparação de K_i com sua esperança condicional, serve para medir a contribuição do evento i na estatística de Knox. Embora o escore z_i^* seja interessante por causa dessas propriedades, ele é difícil de ser calculado, pois os dois momentos $E\{K_i | K_{i-1}, \dots, K_1\}$ e $Var\{K_i | K_{i-1}, \dots, K_1\}$ não possuem fórmulas analíticas. Por isso, o escore z_i acaba sendo preferido na prática. É possível mostrar que $z_i = K_i - K_{i-1}$.

3.3. Métodos de Soma Acumulada (CUSUM)

Os métodos de soma acumulada são muito usados para monitorar processos industriais com o objetivo de detectar rapidamente uma mudança indesejável no processo. Esses métodos são comuns no controle de um processo estatístico seqüencial de uma variável para potenciais desvios da média esperada.

Seja X_i o valor de uma variável aleatória com distribuição $N(\mu, \sigma^2)$, associada ao i -ésimo evento. Assumindo que não há correlação na série de eventos, a soma acumulada até o evento i é dada através da relação:

$$S_i = \max(0, S_{i-1} + x_i - \mu - k\sigma); S_0 = 0$$

Assim, a soma acumulada acumula desvios da média que excedem k desvios padrão, detectando rapidamente qualquer mudança no valor médio de X_i . Um sinal de "fora-de-controle" (houve uma mudança na média do processo) é soado no primeiro tempo i , tal que S_i exceda algum nível de decisão predeterminado h .

Os parâmetros h e k são expressos em termos do desvio-padrão dos eventos. A escolha de h depende do número médio de eventos até que ocorra uma mudança, sob a hipótese de que o processo esteja sob controle (ARL_0). O valor de ARL_0 é fixado pelo usuário e espera-se que ele seja grande. Um bom procedimento de vigilância terá ARL_1 pequeno, onde ARL_1 é o número médio de eventos até o alarme soar a partir de uma mudança do processo.

Em Rogerson (2001), a estatística de soma acumulada correspondente ao evento i é dada por:

$$S_i = \max(0, S_{i-1} + z_i - k); S_0 = 0$$

O limiar h a partir do qual o alarme deve soar é dado pela expressão derivada por Siegmund (1985):

$$ARL_0 \approx 2\{\exp(h+1.166) - h - 2.166\}$$

Para ilustrar o comportamento da técnica proposta por Rogerson (2001), foram gerados 100 pontos com coordenadas $(x_i; y_i; t_i)$ no espaço e tempo. Os primeiros 80 eventos têm coordenadas espaciais x e y geradas de uma distribuição uniforme entre os valores 0 e 1 (denotada $U(0;1)$) e instantes de tempo em que o evento ocorreu t gerados de uma distribuição uniforme entre os valores 0 e 10 (denotada $U(0;10)$). Assim, os eventos estão distribuídos de forma aleatória dentro de um cubo de base $(0;1)^2$ e altura 10.

Os outros 20 eventos foram gerados com coordenadas espaciais x e y de uma distribuição uniforme entre os valores 0.5 e 0.6 (denotada $U(0.5;0.6)$) e com instantes de tempo t gerados de uma distribuição uniforme entre os valores 5 e 6 (denotada $U(5;6)$), a fim de formar um cluster nesta região. Os eventos foram ordenados de acordo com o tempo de ocorrência e rotulados por $i = 1, \dots, 100$ de forma que $t_i < t_{i+1}$. Para calcular as estatísticas locais de Knox, usamos os parâmetros críticos $D = 0.1$ e $T = 1.0$. A escolha dos valores dos parâmetros D e T é arbitrária e neste exemplo foram escolhidos como 10% da amplitude dos eventos gerados no espaço e no tempo.

Na Figura 1 observa-se o resultado do método, compreendendo um gráfico dos eventos (denotados por id) versus a soma acumulada (denotada por S_i), mostrando como a soma acumulada pode ser monitorada ao longo da ocorrência dos eventos.

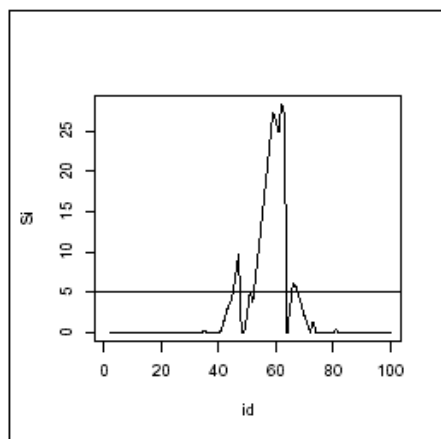


Figure 1. Visualização do resultado fornecido pelo método do Rogerson (2001) para detecção de interações espaço-tempo.

Através da Figura 1, verifica-se ainda que o limiar h a partir do qual o alarme soa tem seu valor próximo a 5 (reta horizontal paralela ao eixo das abcissas) e que é ultrapassado por S_i pela primeira vez no 45º-ésimo evento. O cluster de tamanho 20 contém os eventos de número 40-47, 50, 51, 54-59, 62, 65, 66. Como o primeiro alarme foi verificado no evento de número 45, nota-se que o mesmo foi soado por um evento que pertence ao cluster. Após o 45º-ésimo evento, a Soma Acumulada retorna a zero e volta a ultrapassar o limiar h nos eventos de número 50, 53-62, 65 e 66.

4. Problemas com a metodologia de Rogerson (2001)

Como visto anteriormente, Rogerson (2001) propõe um sistema de vigilância que detecte clusters ativos (vivos) através do monitoramento de uma quantidade S_i , que vai sendo acumulada através de uma soma, à medida que um novo evento i se torna disponível na análise. Caso esta soma S_i ultrapasse um limiar predeterminado h , o alarme deve ser soado. No entanto, dois problemas podem ser verificados nesta técnica: 1 - eventos que disparam o alarme podem não pertencer ao eventual cluster; 2 - a técnica avalia a presença de clusters, mas não os identifica, dando sua posição e extensão no espaço e tempo.

No primeiro problema, verifica-se que pode ser que a soma acumulada S_i esteja tão próxima do limiar h , que um evento com $z_i > 0$ e que não pertence ao cluster pode fazer o alarme soar. Numa situação como esta, provavelmente o alarme já estava na iminência de ser soado e esta contribuição positiva para S_i faz o alarme disparar. Isto quer dizer que a localização do evento disparado do alarme não pode ser usada como um identificador da posição do cluster. Esta situação pode acontecer naturalmente, pois os escores $z_i \sim N(0,1)$ de eventos que não pertençam ao cluster podem exceder k desvios padrões da média ocasionalmente.

Desta forma, nosso objetivo principal é adaptar a técnica apresentada por Rogerson (2001), com o intuito de detectar o cluster eventualmente existente e que teria disparado o alarme, através da visualização da posição em que este foi formado. A técnica proposta neste artigo também permite isolar eventos que não pertencem ao cluster, mas que contribuem para fazer o alarme soar.

A idéia proposta é monitorar os novos eventos não através de uma soma acumulada, mas através de superfícies acumuladas, que utilizam além do método CUSUM, técnicas de estimação de densidade por kernel.

Suponha que o i -ésimo evento ocorra na posição (x_i, y_i, t_i) no espaço-tempo e que z_i seja o escore de interação espaço-tempo (estatística local de Knox) calculado para este evento. Um valor muito positivo de z_i é indicativo de um excesso de eventos que ocorreram pouco tempo antes e que estão geograficamente próximos do evento i . Nós distribuimos este valor de z_i em torno da posição (x_i, y_i) do i -ésimo evento, usando uma superfície de kernel bi-dimensional.

Uma função de kernel bidimensional é obtida a partir de funções de densidade de probabilidade $K^*(x, y) \geq 0$, tais que $\iint K^*(x, y) dx dy = 1$, com máximo na origem $(0,0)$, radialmente simétrica e decrescendo suavemente à medida que nos afastamos da origem. Duas funções muito utilizadas são a função de kernel gaussiana e a função de kernel quártica. A função de kernel gaussiana é definida como

$$K^*(x, y) = \left(\frac{1}{\sqrt{2\pi}} \right) \exp\left(-\frac{1}{2}(x^2 + y^2) \right)$$

O gráfico dessa função está representado na Figura 2.

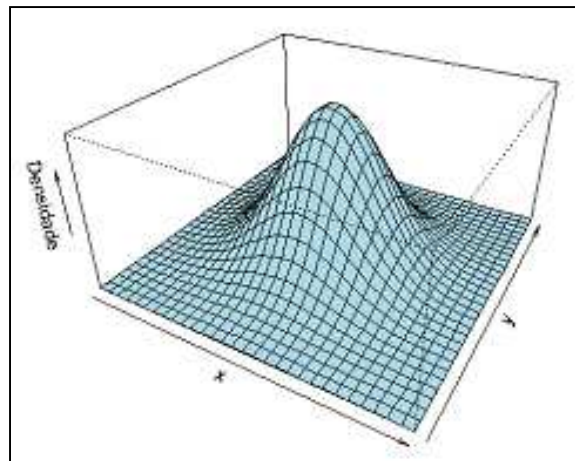


Figure 2. Superfície de kernel usando função gaussiana.

A função de kernel quártica é definida como

$$K^*(x, y) = \begin{cases} \frac{3}{\pi} (1 - (x^2 + y^2))^2 & \text{se } x^2 + y^2 < 1 \\ 0 & , \text{ c.c.} \end{cases}$$

As funções de kernel modificam as funções $K^*(x, y)$ deslocando-as para um novo centro e mudando sua concavidade na origem, com um parâmetro chamado de largura de banda e denotado por λ . Por exemplo, a função de kernel gaussiana $K^*(x, y)$, centrada na posição (x_i, y_i) do i -ésimo evento e com largura de banda λ é dada por

$$K(x, y) = \frac{1}{\lambda} K^* \left[\frac{x - x_i}{\lambda}, \frac{y - y_i}{\lambda} \right] = \frac{1}{\sqrt{2\pi}\lambda} \exp \left\{ -\frac{1}{2} \left(\frac{(x - x_i)^2 + (y - y_i)^2}{\lambda} \right) \right\}$$

De maneira geral, dada uma função base $K^*(x, y)$ com as propriedades descritas, a função de kernel centrada em (x_i, y_i) é dada por

$$K_i(x, y) = \frac{1}{\lambda} K^* \left[\frac{x - x_i}{\lambda}, \frac{y - y_i}{\lambda} \right]$$

O efeito do parâmetro λ é crítico. Um valor muito pequeno irá gerar uma superfície muito descontínua e se for grande demais, a superfície poderá ficar extremamente alisada, mascarando sua forma verdadeira.

Nós usamos a função de kernel, espalhando o escore z_i em torno da localização espacial do i -ésimo evento:

$$w_i(x, y) = z_i K_i(x, y) = \frac{z_i}{\lambda} K^* \left(\frac{x - x_i}{\lambda}, \frac{y - y_i}{\lambda} \right) \quad (1)$$

Observe que, $w_i(x_i, y_i) = \frac{z_i K^*(0,0)}{\lambda}$ e que $w_i(x, y)$ converge para 0 à medida que a posição (x, y) afasta-se de (x_i, y_i) .

Note também que $z_i = \iint w_i(x, y) dx dy$, mostrando que, de fato, estamos distribuindo o valor z_i no espaço usando a função $w_i(x, y)$.

A cada i -ésimo evento, a superfície é calculada iterativamente como:

$$S_{i+1}(x, y) = \max\{0, S_i(x, y) + z_i K_i\{x, y\}\} = \max\{0, S_i(x, y) + w_i(x, y)\}$$

se $i \geq 1$ enquanto que $S_0(x, y) = 0$ para toda posição (x, y) .

De forma geral, as técnicas de soma acumulada e superfícies acumuladas são bem similares, salvo o fato de a primeira estar somando simplesmente os escores z_i e a segunda estar multiplicando uma superfície de kernel a este escore, que permite identificar a localização do eventual cluster.

Resumindo, a técnica de soma acumulada é reformulada para ser aplicada a somas de superfícies (superfícies acumuladas). A partir desta técnica, a superfície $i+1$ seria formada por muitas curvas, de forma que na posição em que estivesse o cluster, se formaria uma curva muito maior, resultante das contribuições dos kernels de cada um dos eventos. A técnica de superfícies acumuladas pode ser visualizada, de forma bem esquematizada, através da Figura 3.

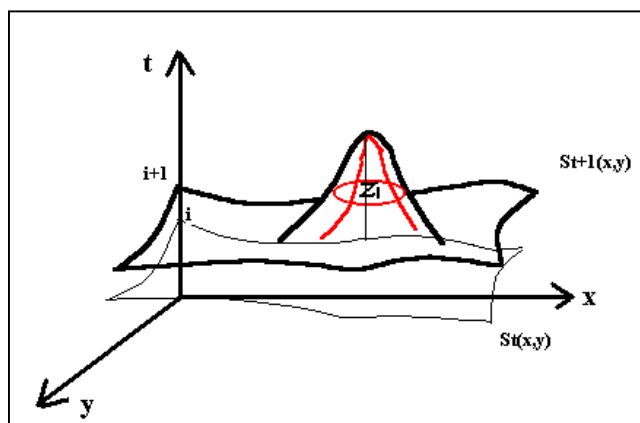


Figure 3. Visualização da técnica de superfícies acumuladas, para detecção de clusters espaço-tempo.

Verifica-se, desta forma, que o problema de o alarme ser soado ocasionalmente por eventos que não pertencem ao cluster é solucionado, no sentido que tais eventos contribuirão com uma superfície de kernel na sua localização de origem, não influenciando efetivamente na superfície formada por elementos do cluster.

A superfície formada após os n eventos terem contribuído para a soma deve aparecer com várias protuberâncias resultantes das superfícies de kernel formadas por cada evento, de forma que no intervalo de tempo em que se encontra o cluster, observasse-ia uma protuberância muito maior se comparada às demais, permitindo assim a localização espacial do cluster.

5. Simulando a Superfície Acumulada no software R

Para se ter uma idéia do comportamento da superfície acumulada, no caso bidimensional, foram gerados 100 pontos com coordenadas (x_i, y_i, t_i) . Como antes, 80 eventos possuem coordenadas x e y de uma distribuição uniforme $U(0;1)$ e instantes de

tempo t de uma uniforme $U(0;10)$. Os 20 eventos restantes foram gerados, com coordenadas x e y de uma uniforme $U(0.5;0.6)$ e tempo t de uma uniforme $U(5;6)$, de forma que formassem um cluster nesta região, uma vez que estão próximos no espaço e no tempo. Os parâmetros críticos para o cálculo das estatísticas de Knox são de $D=0.1$ e $T=1.0$. O limiar h utilizado foi obtido através da fórmula de Siegmund (1985), usando um $ARL_0 = -n / \log(1-p)$, onde p é uma taxa aceitável, predeterminada de alarmes falsos e foi escolhida como 10%. A fórmula para o cálculo de ARL_0 foi obtida, supondo que o número de eventos observados até que um alarme soe sob H_0 , tenha aproximadamente distribuição exponencial.

Os gráficos da Figura 4 mostram as técnicas do Rogerson (2001) e de superfícies acumuladas, sob a hipótese nula de não interação espaço-tempo. Foram gerados apenas os 100 eventos com coordenadas (x,y) de uma distribuição uniforme $U(0,1)$ e instantes de tempo t de uma distribuição uniforme $U(0,10)$, ou seja, nenhum cluster foi criado. O gráfico de somas acumuladas mostra que S_i sai algumas vezes do valor zero, mas não chega a atingir o limiar $h \cong 5$, que nem é mostrado por efeito de escala. Desta forma, o alarme não é soado, como era esperado. No gráfico de superfícies acumuladas, vemos apenas algumas elevações resultantes da contribuição das somas das superfícies de kernel dos eventos anteriores.

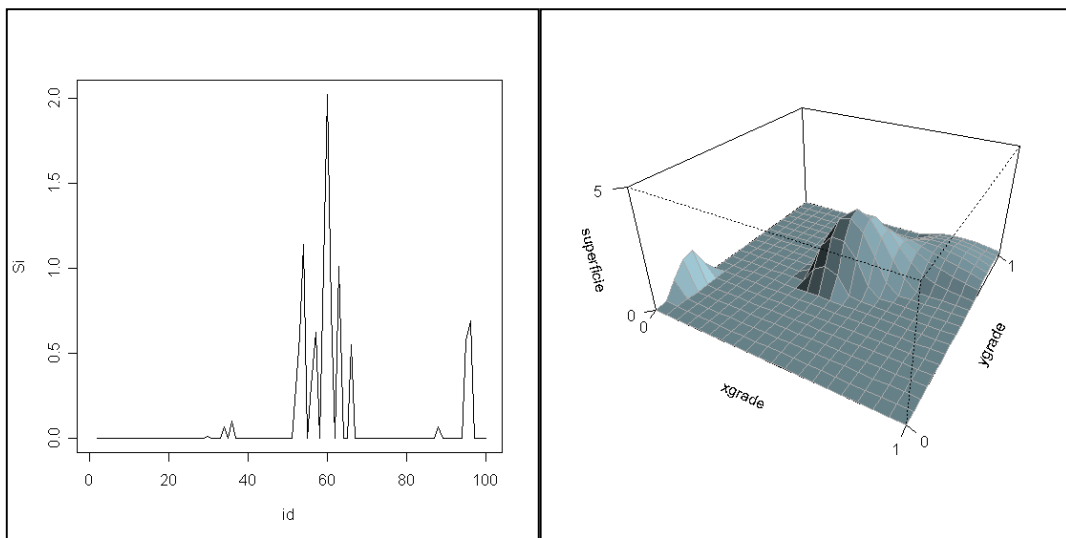


Figure 4. Gráficos de soma acumulada (Rogerson (2001)) e de superfícies acumuladas, sob H_0 .

Na Figura 5, foi criado um cluster de tamanho 20, como descrito no início do exemplo. O cluster gerado contém os eventos de número 37- 49, 54, 55, 57 - 59, 61, 63. No gráfico de somas acumuladas S_i , o limiar h é ultrapassado pela primeira vez na ocorrência do 41º-ésimo evento, ou seja, o alarme é soado quase no início do cluster. O gráfico de superfícies acumuladas ilustra o método no momento da ocorrência do 41º-ésimo evento que acionou o alarme no método do Rogerson (2001). Observamos uma elevação significativa em relação às demais regiões, levando a considerar a presença de um cluster na região de maior saliência.

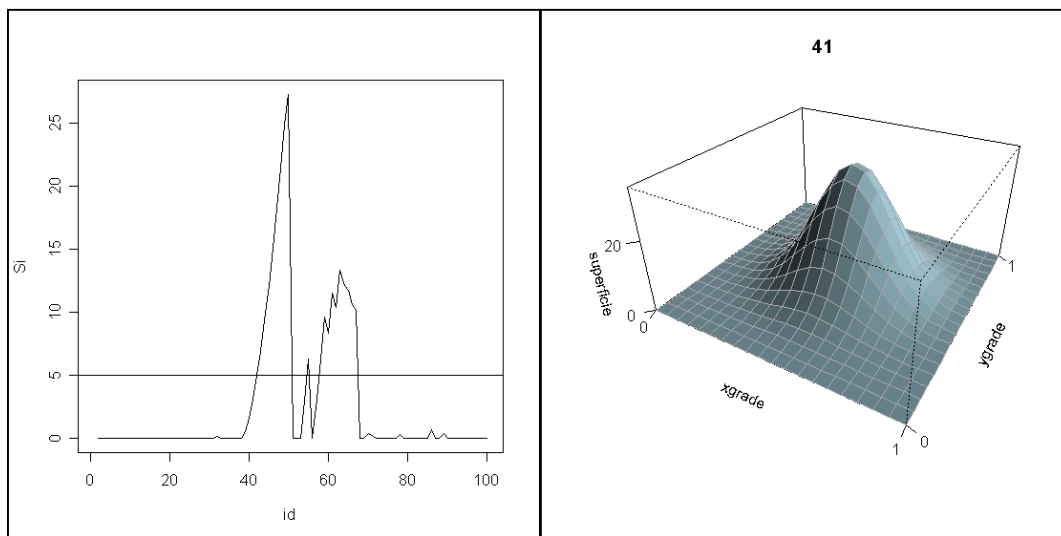


Figure 5. Gráficos de soma acumulada (Rogerson (2001)) e de superfícies acumuladas, com a existência de um cluster.

De forma geral, as Figuras 6 e 7 fornecem um mecanismo de visualização do cluster, no momento em que o alarme é soado no sistema do Rogerson (2001). Na figura 6 vemos a posição do cubo compreendendo os eventos do cluster verdadeiro (gerado) na superfície formada no momento de ocorrência do 41º-ésimo evento. Na Figura 7, o primeiro gráfico mostra cortes da superfície (níveis de contorno) e o segundo é um gráfico de imagem, no qual valores maiores da superfície (cores mais fortes) indicam que naquele local existe uma saliência maior se comparada às outras regiões, indicando a localização mais provável do eventual cluster. Além disso, ressaltamos o fato de que em ambos os gráficos da Figura 7, os maiores valores da superfície tenham aparecido entre as posições 0.5 e 0.6 dos eixos x e y, posições estas em que o cluster verdadeiro foi formado.

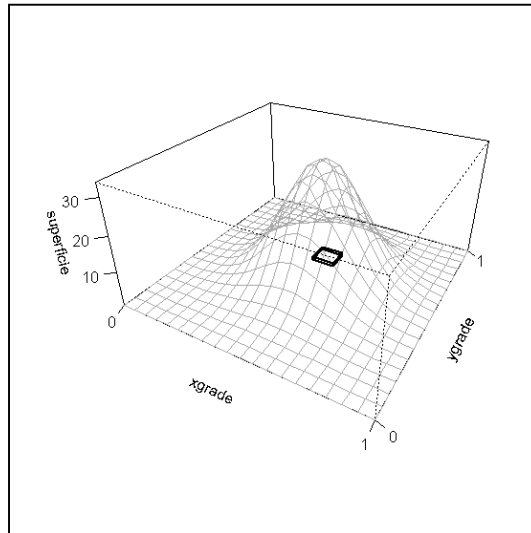


Figure 6. Gráficos da superfície acumulada, mostrando a posição do cubo representante do cluster gerado.

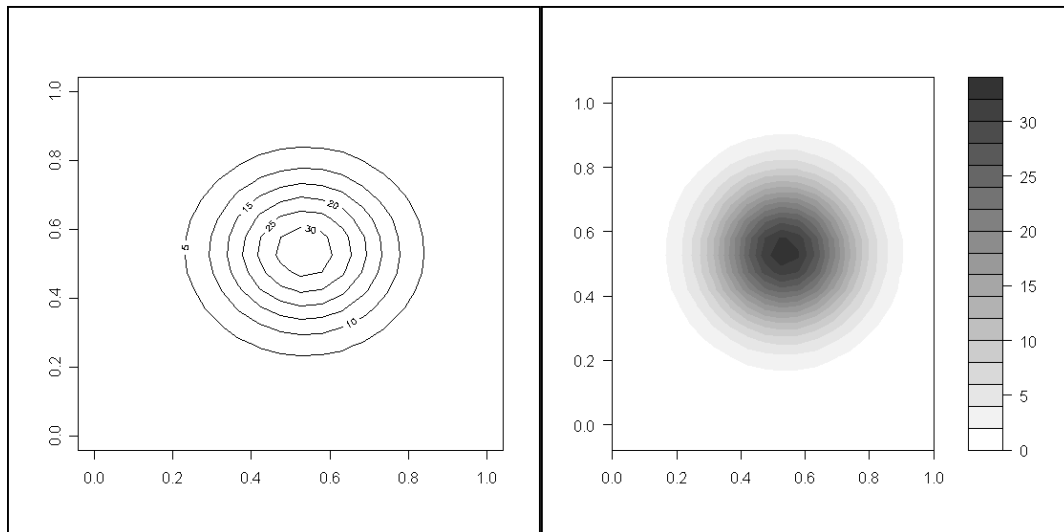


Figure 7. Gráficos de contorno e gráfico de imagem formados no momento do 41º-ésimo evento.

Apesar da técnica gráfica mostrada dar base a uma boa visualização da posição do cluster, é de interesse encontrar os valores aproximados de suas coordenadas x e y no espaço. Para isso, a superfície que contém o ponto de máximo é seccionada por níveis de contorno. Seccionando a superfície que tem o maior pico, obtemos o conjunto de coordenadas (x,y) tais que $S_i(x,y) = c$, onde c é uma constante que deve ser escolhida de forma a obter as coordenadas mais próximas das coordenadas no cluster.

A Figura 8 ilustra o impacto da escolha da constante c . No primeiro gráfico, o valor de c foi escolhido muito próximo do plano, levando à superestimação do tamanho do cluster, uma vez que superfícies vizinhas também são atingidas. No segundo gráfico, a escolha da constante c foi apropriada, dado que o corte foi feito apenas na superfície que contém o valor máximo e, portanto a região mais provável de conter os eventos do cluster.

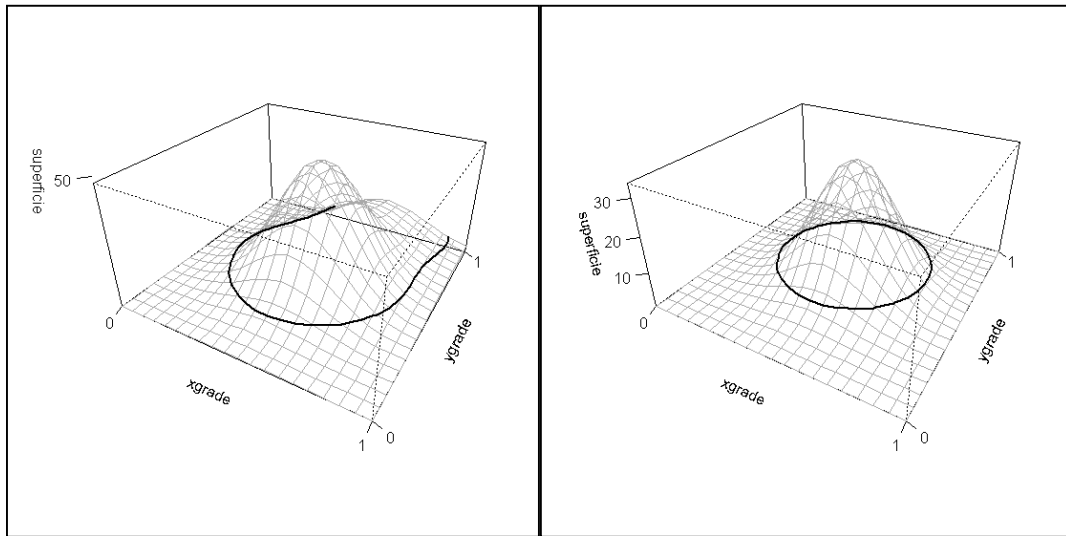


Figure 8. Gráficos de superfícies acumuladas, com escolhas diferentes da constante c para o corte das superfícies.

Uma maneira de avaliar o método de monitoramento é verificando se o alarme foi soado muito antes/depois do cluster verdadeiro ter surgido e através da probabilidade de alarmes falsos, sob H_0 . Para testar a primeira situação, foram feitas 1000 simulações do mesmo cenário acima, utilizando o limiar proposto por Rogerson (2001). Apesar de termos conhecimento de que este limiar é próprio para o uso de séries de observações e não de superfícies, utilizaremos o mesmo, em ambos os métodos, para deixar clara a necessidade de obtermos um limiar diferente quando trabalhamos com superfícies.

Seja d_i o tempo entre o primeiro alarme e o início do cluster. Se $d_i > 0$ é possível calcular o número esperado de eventos até que um alarme ocorra, dado que o processo esteja fora de controle, ou seja, dado que existe um cluster (ARL_1).

A Figura 9 apresenta a distribuição dos d_i 's nos métodos do Rogerson (2001) e de superfícies acumuladas. Note que a amplitude dos valores no primeiro nos dá uma indicação de que o tempo de espera por um alarme verdadeiro é maior no método do Rogerson que no segundo método.

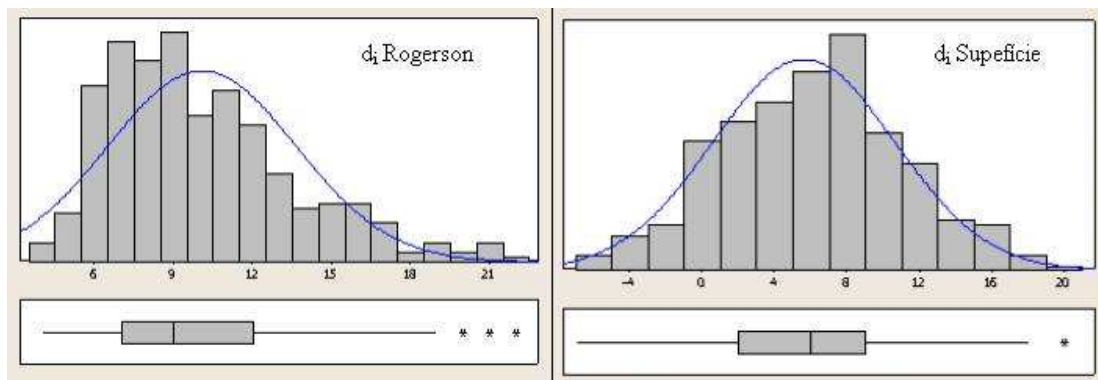


Figure 9. Histogramas dos d_i 's nos dois métodos.

Esta colocação pode ser confirmada através da Tabela 2, que mostra além de informações descritivas dos d_i 's nos métodos do Rogerson (2001) e de superfícies acumuladas, tais como a média, o desvio-padrão, o mínimo (Min), máximo (Max), 1º, 2º (Mediana) e 3º Quartis, os tempos médios de espera por um alarme verdadeiro dado que exista um cluster, em ambos os métodos.

Tabela 2. Informações relevantes obtidas nas simulações (k=0,5).

Informações sobre os d_i 's (k=0,5)								
Método	Média	Desvio-padrão	Min	1ºQuartil	Mediana	3ºQuartil	Max	ARL ₁
Rogerson	10,066	3,580	4,000	7,000	9,000	12,000	22,000	0,000
Superfície	5,603	5,048	-7,000	2,000	6,000	9,000	20,000	0,766
Teste de Mann-Whitney: $w = 140074,5$, $p\text{-valor} = 0,000$								

Através da Tabela 2, verifica-se que, em geral, o alarme soa mais rápido no método de superfícies acumuladas, uma vez que o tempo médio de espera até que um alarme verdadeiro soe, dado que exista interação espaço-tempo (ARL₁), é menor que no Rogerson (2001). Além disso, esta diferença de tempos de espera dos d_i 's positivos, entre os dois métodos, é estatisticamente significativa uma vez que o teste de *Mann-Whitney* foi significativo, ao nível de 5%.

No entanto, outra maneira de avaliar um sistema de vigilância é através da probabilidade de alarmes falsos, sob H_0 (sem cluster). Esta probabilidade é calculada como sendo a razão entre o número de vezes que o alarme soa dentro o total de simulações.

A Tabela 3 mostra os resultados da probabilidade de alarmes falsos sob H_0 para 100 simulações do método de somas acumuladas - Rogerson e do método de superfícies acumuladas, utilizando o mesmo limiar do Rogerson (2001) – Superfície e mais dois limites empiricamente observados: Lim1 e Lim2.

Lim1 foi obtido da seguinte maneira: para uma dada simulação, era calculado o valor máximo que a superfície atinga, a cada evento observado. Assim, para uma única simulação, existiam 100 valores de máximos, correspondentes às etapas de formação das superfícies a cada evento acrescentado. Lim1 é o percentil 90 destes máximos. A probabilidade de alarmes falsos é calculada através da razão entre o número de Lim1's diferentes de zero dentre as 100 simulações.

No Lim2, para uma única simulação, era calculado o percentil 90 da superfície em cada etapa de sua formação. Desta forma, existem 100 valores de percentil 90, dos quais é calculada a média aritmética. Isto é feito para cada uma das 100 simulações, permitindo calcular a probabilidade de alarmes falsos como a razão do número de Lim2's que são diferentes de zero, sobre o total de 100 simulações.

Tabela 3. Informações da probabilidade de alarmes falsos.

Método	P(Alarmes Falsos)
Rogerson	0
Superfície	0,77
Lim1	0,87
Lim2	0,97

Através da Tabela 3, vemos que a probabilidade de alarmes falsos sob H_0 para o método de superfícies acumuladas, utilizando o limiar do Rogerson (2001) de 0.77 é muito maior se comparado ao método de somas acumuladas que é zero. No entanto, este resultado é esperado, devido ao fato de estarmos usando um limiar apropriado para séries de observações e não para superfícies, como dito anteriormente. Além disso, o método do Rogerson (2001) apresenta um valor muito abaixo que o esperado de 10% de alarmes falsos. Acreditamos que estes resultados possam ser devido ao fato do limiar abordado por Rogerson (2001) ser inadequado nas duas técnicas.

Para os limites Lim1 e Lim2, verificamos que a probabilidade é muito alta, no entanto, este resultado pode ser devido ao fato do procedimento ter sido repetido poucas vezes. Assim, um número maior de iterações deve ser simulado a fim de avaliar principalmente o comportamento de Lim1, que acreditamos ser o mais adequado dentre os limites sugeridos, para o método de superfícies acumuladas.

A intensidade de pontos no cenário gerado e o número de eventos por unidade de volume dentro do cilindro em estudo pode, em muitos casos, ser baixa ou mesmo nula. Assim, uma alternativa seria reduzir o parâmetro k , utilizado na soma acumulada, para zero, não deixando assim, que a Soma S_i se torne zero tão frequentemente. Desta forma, fazendo o ajuste $k = 0$, fará com que a soma S_i não tenha tendência a diminuir, possibilitando um alarme mais eficaz, no sentido da detecção mais rápida do cluster.

Estudos ainda não publicados, realizados no LESTE–UFMG (Laboratório de Estatística Espacial), mostram que em cenários onde a intensidade de pontos por volume é baixa, a influência do k é grande, trazendo subestimação de falsos alarmes no método e prejudicando os resultados devido ao baixo valor dos S_i 's. Desta forma, a Tabela 4 mostra os resultados descritivos dos d_i 's, utilizando $k = 0$, bem como as probabilidades de alarmes falsos sob H_0 e os ARL_1 's, nos métodos do Rogerson (2001) e de superfícies acumuladas.

Tabela 4. Informações relevantes obtidas nas simulações ($k = 0,0$).

Método	Informações sobre os d_i 's ($k=0,0$)							
	Média	esvio-padrã	Min	1ºQuartil	Mediana	3ºQuartil	Max	ARL_1
Rogerson	7,646	2,854	2,000	5,000	7,000	9,000	18,000	0,000
Superfície	5,463	5,095	-7,000	2,000	5,500	9,000	20,000	0,834
Teste de Mann-Whitney: $w = 123847,5$, p -valor = 0,0001								

De uma maneira geral, a Tabela 4 mostra os mesmos resultados que a Tabela 2, ou seja, o método de superfícies acumuladas apresenta tempo médio de eventos até uma mudança verdadeira, dado que exista um cluster, menor que no método do Rogerson (2001).

6. Conclusões

Verifica-se que o método proposto de superfícies acumuladas, com base no sistema de vigilância do Rogerson (2001), dá a indicação de ser um bom identificador do cluster, uma vez que detecta a posição espacial do mesmo, através da visualização das representações gráficas das superfícies. Contribuição esta de grande relevância, uma vez que, na prática, a localização do cluster possa levar a várias tomadas de decisão importantes.

No entanto, a suposição de que o novo método forneça um número menor de alarmes falsos deve ser apurada mais profundamente, dado que por simulações prévias, tenha sido verificado que o limiar proposto por Rogerson (2001) e outros empíricos, não sejam apropriados. Acredita-se que um limiar adequado possa ser encontrado, utilizando-se técnicas de teoria de valor extremo, em particular, de máximo de superfícies aleatórias (campos gaussianos). Portanto, os esforços futuros serão direcionados na procura de um limiar ótimo para a nova técnica proposta.

7. Referências

- Rogerson, P.A.(2001) “Monitoring point patterns for the development of space-time clusters”. *Jornal Royal Statistical Society* (2001) 164, Part 1, 87-96. University at Bualo, USA.
- Siegmund, D., O..(1985) “Sequential Analysis: Tests and Condence Intervals”. New York: Springer.
- Härdle,W..(1990) “Smoothing Techniques”. Louvain-La-Neuve.
- Sonesson, C.; Bock, D..(2002) “A review and discussion of prospective statistical surveillance in public health”. Göteborg University, Sweden. *Jornal Royal Statistical Society* (2003) 166, Part 1, pp 5-21.
- Diggle, P., J..(1983) “Statistical Analysis of Spatial Point Patterns”. Academic Press Inc. Londres.
- Frisen, M..(2003) “Statistical surveillance. Optimality and methods”. *International Statistical Review*, 71, 403-434.
- Knox, E., G..(1964) “The detection of space-time interactions”.. *Appl. Statist.* 13, 25-29.
- Montgomery, D., C..(2000) “Introduction to Statistical Quality Control”.. 4th Edition, New York : John Wiley, 2000.