

Título

Modelagem espaço-temporal de contagens de ovos de *Aedes aegypti*

Autores

Wagner Hugo Bonat Departamento de Estatística - DEST, Laboratório de Estatística e Geoinformação - LEG, Universidade Federal do Paraná - UFPR, CEP: 81531-990, Curitiba,Paraná, Brasil, E-mail: *wagner@leg.ufpr.br*

Paulo Justiniano Ribeiro Jr Departamento de Estatística - DEST, Laboratório de Estatística e Geoinformação - LEG, Universidade Federal do Paraná - UFPR, CEP: 81531-990, Curitiba,Paraná, Brasil, E-mail: *paulojus@leg.ufpr.br*

Agradecimentos

Agradecemos ao projeto SAUDAVEL pela disponibilidade dos dados. A doutora Lêda Regis e ao doutor Antonio Miguel Vieira Monteiro pela leitura do artigo preliminar.

Resumo O dengue é um problema de saúde pública em todas as regiões brasileiras. Como não se dispõe de uma vacina efetiva, o elo vulnerável da cadeia epidemiológica é o vetor, o mosquito *Aedes aegypti*. O entendimento das flutuações na população do mosquito é instrumental para reduzir a sua proliferação e a exposição das pessoas à infecção. Neste artigo ilustramos o uso de modelos espaço-temporais para contagens de ovos de *Aedes aegypti*, coletados em ovitrampas em um experimento de campo conduzido em Recife/PE. Uma categoria de modelos espaço-temporais é adotada e revisada, bem como os métodos de inferência a ela associados; Os resultados mostram padrões espaciais e temporais nas contagens com interações espaço-tempo não significativas. O uso desta classe flexível de modelos com inferência computacionalmente eficiente através da aproximação de Laplace aninhada integrada (INLA) possibilitam a incorporação destes modelos em sistemas de vigilância e monitoramento.

Palavras chave Vigilância entomológica, dengue, inferência bayesiana, dados de contagens.

Abstract Dengue is a public health across Brazil whole country. As an effective vaccine is not yet available, the vulnerable part of the epidemiological chain is the vector, the *Aedes aegypti*. Understanding fluctuations on the mosquito population is instrumental to reduce the spreading and exposure of the human population to the infection. The current work illustrates the usage of spatio-temporal models for *Aedes aegypti*'s egg counts collected at ovitraps during an field experiment conduct at Recife/PE. We revise and adopt a category of spatial-temporal models with the associated method of inference. Results shows spatial and temporal patterns on the counts, whereas no spatio-temporal interaction was detected. The flexible class of models and the computationally efficient inference by integrated nested Laplace approximation (INLA) allows for embedding of such models in disease surveillance and monitoring systems.

Keywords Epidemiological surveillance, dengue, Bayesian inference, count data.

1 Introdução

O dengue é uma doença febril aguda, cujo agente etiológico é um vírus do gênero Flavivírus. São conhecidos atualmente quatro sorotipos, antigenicamente distintos: DEN-1, DEN-2, DEN-3 e DEN-4. As manifestações variam de uma síndrome viral, inespecífica e benigna, até um quadro grave e fatal de doença hemorrágica com choque.

O vírus dengue são transmitidos ao homem pela picada do mosquito *Aedes aegypti*, tal mosquito tem hábitos domésticos, pica mais frequentemente durante o dia e tem preferência acentuada por sangue humano (Tauil, 2000).

Durante quase 60 anos, de 1923 a 1982, o Brasil não apresentou registro de casos de dengue em seu território. Porém, desde 1976, o *Aedes aegypti* havia sido reintroduzido no país, a partir de Salvador, Bahia, e estava presente em muitos países vizinhos. Países da América Central, México, Venezuela, Colômbia, Suriname e alguns outros do Caribe já vinham apresentando a doença desde os anos 70.

Até os dias atuais não se dispõe de uma vacina eficaz para uso preventivo contra o dengue, apesar de todos os esforços de pesquisa para o seu desenvolvimento e produção. Enquanto não se puder contar com esta medida de controle, o único elo vulnerável da cadeia epidemiológica é o vetor.

Pensando neste elo vulnerável um dos esforços nacionais foi a elaboração do Projeto SAUDAVEL,¹ o qual pretende contribuir para aumentar a capacidade do setor de saúde no controle de doenças transmissíveis, demonstrando ser necessário desenvolver novos instrumentos para a prática da vigilância entomológica, incorporando aspectos ambientais, identificadores de risco e proteção, além de métodos automáticos e semi-automáticos, que permitam a detecção de surtos e seu acompanhamento no espaço e no tempo (Monteiro et al. 2006).

Com a intensa circulação do vírus do dengue no Brasil a partir da década de 1980, epidemias explosivas têm atingido todas as regiões brasileiras (Braga e Valle, 2007). Em vista disso, também a partir desta década, diversas metodologias para a vigilância do vetor vêm sendo desenvolvidas no país.

Nos programas de controle do dengue, a vigilância entomológica é feita principalmente a partir da coleta de larvas, de acordo com a proposta de Connor e Monroe (1923) para medir a infestação de áreas urbanas pelo *Aedes aegypti*. Esta metodologia consiste em vistoriar os depósitos de água e outros recipientes localizados nas residências e demais imóveis, como borracharias, ferros-velhos, cemitérios, entre outros tipos de imóveis considerados estratégicos, por produzirem grande quantidade de mosquitos adultos, para o cálculo dos índices de infestação predial (IIP) e de Breteal (IB).

A coleta de larvas (ou pesquisa larvária, como é comumente chamada no Brasil) é importante para verificar o impacto das estratégias básicas de controle da doença, dirigidas a eliminação das larvas do vetor. Entretanto, não é um bom indicador para se medir a abundância do adulto, e ineficaz para estimar o risco de transmissão (Braga e Valle, 2007).

Apesar disso, para avaliar a densidade do vetor instalam-se armadilhas de oviposição e armadilhas para coleta de larvas, que visam estimar a atividade de postura. A armadilha de oviposição, também conhecida no Brasil como

¹Sistema de Apoio Unificado para a Detecção e Acompanhamento em Vigilância Epidemiológica (<http://saudavel.dpi.inpe.br>)

“ovitrampa”, é destinada a coleta de ovos. Em um recipiente de cor escura, adere-se um material àspero que permite a fixação dos ovos depositados. Em 1965, iniciou-se o uso de ovitrampas para a vigilância de populações adultas de *Aedes aegypti* (Fay e Eliason, 1965). Posteriormente, ficou demonstrada a superioridade dessas armadilhas em relação a pesquisa larvária, para a verificação da ocorrência do vetor (Fay e Eliason, 1966).

As ovitrampas fornecem dados úteis para a investigação da distribuição espacial e temporal (sazonal) de ovos do mosquito. Dados obtidos com essa metodologia também são usados para verificar o impacto de vários tipos de medidas de controle, que visam a redução do vetor com inseticidas.

O objetivo deste artigo é construir, avaliar e comparar modelos para descrever o padrão espaço-temporal das contagens de ovos de *Aedes aegypti* coletados em ovitrampas. O entendimento do comportamento espaço-temporal das contagens pode contribuir de forma significativa para a elaboração de ações, que busquem evitar a proliferação do mosquito, através de campanhas direcionadas para as áreas de maior infestação, bem como, priorizar épocas do ano em que as contagens sejam maiores. Para isto, será analisado um conjunto de dados de um experimento conduzido pelo “Projeto SAUDAVEL” na cidade de Recife/PE. A idéia é que as construções espaço-temporais aqui apresentadas sirvam de referência para a análise em outras áreas, onde se realize experimentos similares. Os modelos consideram relações espaciais entre as armadilhas, a possibilidade de relação temporal e ainda diversas formas de interação espaço-temporal entre as coletas.

O presente artigo encontra-se dividido da seguinte forma: esta primeira apresenta uma visão geral sobre a dengue no Brasil e o objetivo do artigo. A segunda seção apresenta o conjunto de dados SAUDAVEL/Recife, que será utilizado para o ajuste dos modelos. Na terceira são apresentados os modelos e a forma de inferência. Na seção quatro os principais resultados e fecha-se com uma discussão e principais conclusões obtidas.

2 Conjunto de dados

O conjunto de dados que será analisado corresponde a um experimento desenvolvido pelo projeto SAUDAVEL na cidade de Recife/PE. Neste experimento foram instaladas 464 armadilhas (ovitrampas) para coletar ovos do mosquito *Aedes aegypti*. Estas armadilhas começaram a ser monitoradas em março de 2004. O experimento foi realizado em 7 dos 94 bairros da cidade. Os dados que serão apresentados aqui referem-se apenas ao bairro Brasília Teimosa, por este apresentar uma quantidade expressiva de observações.

A coleta de dados neste bairro teve início em 04/01/2005 e até 15/05/2007 período para o qual os dados estão disponíveis, foram realizadas 2480 observações em 80 armadilhas no período correspondente a 124 semanas. A rede de armadilhas foi instalada de modo a cobrir toda a extensão do bairro, a cada sete dias é feita a contagem de ovos encontrados em cerca de um quarto das armadilhas, assim um ciclo de 28 dias é necessário para que todas as armadilhas sejam monitoradas. Ressalta-se que as coletas são semanais porém apenas 1/4 das armadilhas são observadas em cada semana, tem-se assim quatro grupos de armadilhas, de acordo com a semana em que ela é observada.

Cada armadilha contém três lâminas de material áspero, nas quais a fêmea do mosquito coloca os ovos, quando recolhidas são levadas para um laboratório especializado onde a contagem de ovos é realizada. O departamento de entomologia da FIOCRUZ/PE e os serviços de saúde locais são os coordenadores operacionais e logísticos e responsáveis pela realização do experimento (Monteiro et al. 2006).

Regis et al. (2008) descrevem de forma ampla o experimento SAUDAVEL/Recife, bem como todo o escopo do projeto que visa desenvolver metodologias e tecnologias para o monitoramento de populações de *Aedes aegypti* através de contagens de ovos coletados em ovitrampas.

O banco de dados do SAUDAVEL/Recife, possibilita a construção de diversas covariáveis que podem ser associadas as contagens. Bonat et al. (2009) investigaram o efeito destas covariáveis usando a estrutura dos Modelos Aditivos Generalizados, chegando a algumas significativas, porém os modelos não contemplam a possibilidade de interações espaço-temporais entre as coletas. Bonat (2010) reinvestiga o efeito destas covariáveis com uma estrutura mais geral de modelos de interação espaço-temporal e descarta os efeitos das covariáveis argumentando que o efeito destas podem ser absorvidos pelos efeitos espaciais e temporais do modelo. O autor argumenta também que as covariáveis construídas foram coletadas apenas no início do experimento, sendo que, as condições das residências onde as ovitrampas foram instaladas não são monitoradas junto com as coletas, não se tem nenhuma garantia de que as condições não mudaram durante o experimento. Este fato desacredita a construção de modelos com tais covariáveis.

Neste artigo, não vamos considerar modelos com a inclusão de covariáveis pelos mesmos motivos apresentados em Bonat (2010). Vamos focar na construção de modelos para estruturar apenas os efeitos espaciais e temporais.

Desta forma, o banco de dados utilizado nesta análise conta com 2480 observações divididas em 80 armadilhas (localizações) observadas em 124 datas de coletas.

3 Modelos espaço-temporais

Para a construção dos modelos espaço-temporais, denote por y_{it} a variável resposta na área i ($i = 1, \dots, n$) e no tempo t ($t = 1, \dots, T$). Assuma que a distribuição de y_{it} pertencente a família exponencial, com termo de média μ_{it} e possivelmente um parâmetro de escala ou dispersão ϕ que pode ou não depender do parâmetro de média. Assumindo propriedades de independência condicional a verossimilhança é simplesmente o produto da distribuição de probabilidade atribuída para a resposta. Assim tem-se, $g(\mu_{it}) = \eta_{it}$, onde η_{it} tem a seguinte decomposição

$$\eta_{it} = \alpha + \rho_t + \gamma_t + \varphi_i + \phi_i \quad (1)$$

onde α é o nível médio do processo, ρ_t e γ_t são efeitos temporais, representando desvios da média geral do processo para o tempo t que respectivamente tem e não tem estrutura temporal a priori. Similarmente, φ_i e ϕ_i representam desvios da média geral para a área i que respectivamente possui e não possui estrutura espacial.

A formulação 1 é completada definindo as distribuições a priori para os quatro blocos $\rho = (\rho_1, \dots, \rho_T)^T$, $\gamma = (\gamma_1, \dots, \gamma_T)^T$, $\varphi = (\varphi_1, \dots, \varphi_n)^T$ e $\phi = (\phi_1, \dots, \phi_n)^T$ em η_{it} . Cada priori é assumida como sendo gaussiana multivariada com média zero e matriz de precisão kK , onde k é um escalar desconhecido e K é uma matriz de estrutura conhecida. A matriz de estrutura K vai ser diferente para cada bloco a fim de descrever diferentes suposições a priori sobre o relacionamento entre os parâmetros dentro de cada bloco. Para ρ , assume-se a priori que os efeitos de pontos vizinhos no tempo tendem a ser semelhantes. O mais simples dos modelos dinâmicos é o passeio aleatório (*random walk*) com incrementos gaussianos independentes

$$\pi(\rho|k_\rho) \propto \exp\left(-\frac{k_\rho}{2} \sum_{t=2}^T (\rho_t - \rho_{t-1})^2\right) \quad (2)$$

que tem matriz de estrutura

$$K_\rho = \begin{pmatrix} 1 & -1 & & & & & & \\ -1 & 2 & -1 & & & & & \\ & -1 & 2 & -1 & & & & \\ & & \vdots & \vdots & \vdots & & & \\ & & & -1 & 2 & -1 & & \\ & & & & -1 & 2 & -1 & \\ & & & & & -1 & 1 & \end{pmatrix}$$

ver por exemplo Rue e Held (2005). Para γ , assume-se permutabilidade (*exchangeability*) dos componentes tomando $K_\gamma = I$, a matriz identidade.

Para o bloco estruturado espacialmente φ , escolhe-se uma simples autoregressão Gaussiana; ver, por exemplo, Besag, York e Mollié (1991). Assim, a matriz de estrutura K_φ tem elementos fora da diagonal $k_{ij} = -1$ para áreas geograficamente conectadas (vizinhas) $i \sim j$ e elementos na diagonal k_{ii} igual ao número de áreas, m_i , que são geograficamente contíguas a área i . Todos os outros elementos de K_φ são zero. A priori para φ pode ser escrita como

$$\pi(\varphi|k_\varphi) \propto \exp\left(-\frac{k_\varphi}{2} \sum_{i \sim j} (\varphi_i - \varphi_j)^2\right). \quad (3)$$

Esta priori de campo aleatório Markoviano é o analogo espacial do passeio aleatório sendo assim, é não estacionário. Esta abordagem pode ser estendida introduzindo pesos na formulação a priori (Besag, Your e Mollié, 1991). Finalmente, heterogeneidade espacial não estruturada é acomodada por tomar $K_\phi = I$. Uma representação simbólica do modelo de efeitos principais é dada na figura 1. Em que círculos vazados representam indepência a priori, retângulos representam dependência a priori. Observações no espaço tempo são indicadas por círculos sólidos.

A formulação até aqui, considera efeitos temporais e espaciais separados, e requer uma extensão apropriada para a presença de interações espaço-tempo. Formalmente adiciona-se um parâmetro de interação δ_{it} , $i = 1, \dots, n$ e $t = 1, \dots, T$, o preditor agora é dado por:

$$\eta_{it} = \alpha + \rho_t + \gamma_t + \varphi_i + \phi_i + \delta_{it}. \quad (4)$$

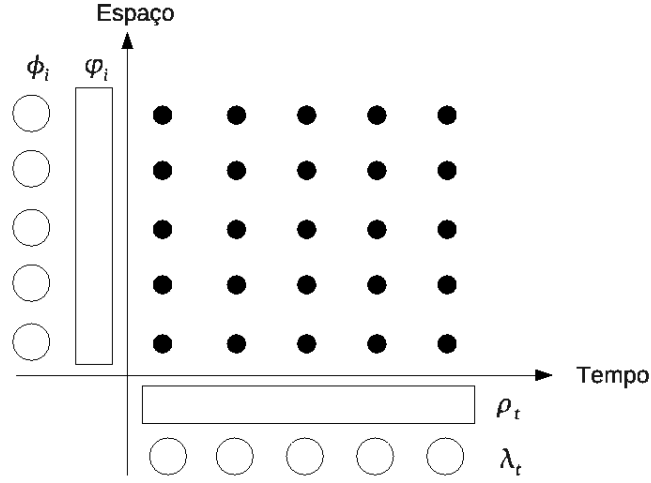


Figure 1: Representação simbólica do modelo de efeito principais. Círculos representam independência a priori, retângulos representam dependência a priori. Observações no tempo espaço são indicadas por círculos sólidos.

O vetor de parâmetros $\delta = (\delta_{11}, \dots, \delta_{nT})^T$ é assumido como Gaussiano com matriz de precisão $k_\delta K_\delta$. Como para os efeitos principais, k_δ é um escalar desconhecido e K_δ é uma matriz de estrutura pré-especificada. Note que se reduz a 1 se todos os $\delta_{ij} = 0$, assim δ captura somente a variação que não pode ser explicada pelos efeitos principais.

Clayton (1996) sugere que K_δ seja especificada como o produto Kronecker das matrizes de estrutura dos efeitos principais que se assume interagir. Esta formulação pode ser vista como o análogo Bayesiano, da modelagem de interações por *tensor products* no contexto de regressão por splines (Stone, 1997). Na formulação proposta acima, $2 * 2 = 4$ combinações são possíveis dependendo de qual dos dois efeitos temporais assume-se interagir com qual dos dois efeitos espaciais. Estes quatro tipos de interação implicam em diferentes priors para o relacionamento entre os δ_{it} , como ilustrado na figura 2. Neste momento, se discutirá cada tipo separadamente, ordenados pelo grau de dependência a priori. A nomenclatura para os tipos de interação, segue da contribuição de Knorr-Held (2000).

- **Interação tipo I** - Se os dois efeitos principais não estruturados γ e ϕ são esperados interagirem, pela proposta de Clayton (2004), tem-se $K_\delta = K_\gamma \otimes K_\phi = I \otimes I$, assim todos os parâmetros de interação δ_{it} são a priori independentes. Este efeito pode ser visto como de covariáveis não observadas para cada pixel (i, t) , que não tem nenhuma estrutura no espaço-tempo.
- **Interação tipo II** - Se combinar o efeito principal *random walk* ρ com o bloco não estruturado ϕ , então cada $\delta_i = (\delta_{i1}, \dots, \delta_{iT})^T$, $i = 1, \dots, n$

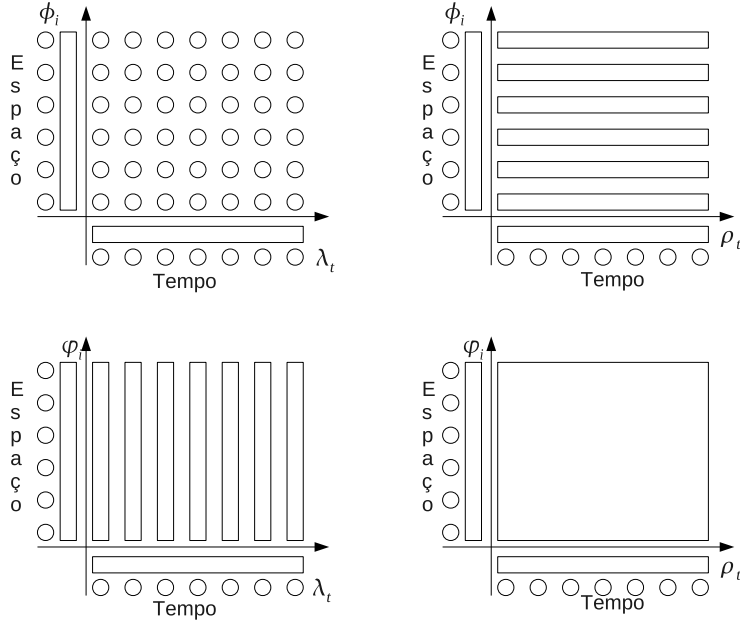


Figure 2: Representação simbólica para os quatro tipos de interação. Círculos representam independência a priori, retângulos representam dependência a priori.

segue um *random walk*, independente para cada área, ou seja, cada área tem a sua própria série temporal, porém não existe dependência espacial entre as séries. O modelo com δ do tipo II vai ser adequado, se as tendências temporais são diferentes para cada área, mas não tem nenhuma estrutura no espaço.

- **Interação tipo III** - Ao assumir, que os efeitos principais γ e φ interagem, então cada $\delta_t = (\delta_{1t}, \dots, \delta_{nt})^T$, $t = 1, \dots, T$, segue uma autoregressão intrínseca (independente). Tal especificação vai ser adequada, se tendências espaciais são diferentes em cada ponto no tempo, porém sem nenhuma estrutura temporal.
- **Interação tipo IV** - Do ponto de vista teórico, a mais interessante forma de interação é com o produto de dois efeitos principais dependentes, o *random walk* ρ e a autoregressão intrínseca φ . Agora δ é completamente dependente sobre o espaço e tempo e já não pode ser fatorada em blocos independentes. A interação do tipo IV é um campo Markoviano, onde não somente vizinhos (primeira ordem) temporais ($\delta_{i,t-1}$ e/ou $\delta_{i,t+1}$) e espaciais ($\delta_{jt}, j \sim i$) entram na condicional completa para δ_{it} , mas também vizinhos de segunda ordem ($\delta_{j,t-1}$ e/ou $\delta_{i,t+1}$), ou seja, vizinhos espaciais de vizinhos temporais ou, equivalentemente, vizinhos temporais de vizinhos espaciais. Esta priori força com que a tendência temporal (em termos de diferenças de primeira ordem) para áreas vizinhas seja parecida. Equivalentemente, esta priori força com que a dependência espacial para pontos

vizinhos no tempo $(t-1, t+1)$ sejam parecidas. Tal modelo a priori vai ser adequado, se tendências temporais são diferentes de área para área, mas são mais parecidas para áreas vizinhas.

A inferência para estes modelos é um desafio para os métodos computacionais atuais. Métodos tradicionais para inferência Bayesiana baseados em simulação, como os métodos de Monte Carlo via Cadeias de Markov (MCMC) são ineficientes devido a grande dimensão do vetor de efeitos aleatórios. Rue, Martino e Chopin (2009) apresentam um novo método para fazer inferência Bayesiana completa em modelos Gaussianos latentes altamente estruturados. A metodologia faz o uso de métodos de integração numérica para obter as distribuições marginais a posteriori de interesse. A abordagem INLA - Integrated Nested Laplace Approximation, supera os problemas com relação a tempo computacional e convergência dos métodos MCMC, tornando-se uma ferramenta de grande valia para a estimação Bayesiana de modelos altamente complexos como os apresentados neste artigo.

A metodologia esta eficientemente implementada em uma biblioteca escrita em C e Fortran chamada GMRFlib (Rue e Follestad, 2002). Além disso, um pacote $R - INLA$ (Martino e Rue, 2009) para o software R é disponível para implementar os modelos diretamente de dentro do R em um formato bastante amigável. Uma revisão detalhada desta metodologia pode ser encontrada em Bonat (2010) e uma revisão com exemplos computacionais é apresentada em Bonat et al. (2012).

Para a especificação do modelo é necessário escolher uma distribuição de probabilidade para as contagens. A escolha mais imediata seria a distribuição Poisson, por esta ser pertencente a família exponencial e ser a distribuição mais usada para contagens. Porém, esta distribuição tem uma forte suposição sobre a forma de relacionamento entre a média e a variância das contagens, assumindo que ambas são iguais. Esta suposição mostrou-se inadequada para a presente situação, já que, análises descritivas iniciais indicaram a presença de superdispersão. Sendo assim, optou-se pela distribuição Binomial Negativa, já que, esta conta com um parâmetro extra que modela a superdispersão. Como função de ligação foi escolhido a função logaritmo. Estas escolhas também foram feitas em Bonat et al. (2009) e Bonat (2010).

Para completar a especificação Bayesiana do modelo, falta designar as distribuições a priori para os hiperparâmetros dos campos Gaussianos. Os hiperparâmetros para o modelo com efeitos principais são k_ρ , k_γ , k_φ e k_ϕ , eles determinam a variação de cada bloco, e precisam ser estimados através dos dados. Adicionalmente, k_δ tem que ser estimado nos modelos com o termo de interação. Seguindo a proposta de Knorr-Held (2000) assume-se para todos estes parâmetros prioris Gama próprias, $k \sim G(a, b)$, para evitar problemas com hiperprioris impróprias. Prioris Gama são computacionalmente convenientes, porque a condicional completa de k vai também ser Gama. Para todas as aplicações neste artigo, hiperprioris Gama de alta dispersão foram escolhidas para todos os blocos com valores $a = 1$ e $b = 0.01$.

4 Resultados

De acordo com a estrutura apresentada na seção 3 pode-se construir diversos modelos competidores. Sendo que cada um representa uma suposição diferente sobre o comportamento espaço-temporal das contagens. A tabela 1 apresenta dez diferentes composições de modelos de acordo com a presença ou ausência dos efeitos espaciais e temporais, bem como, diferentes formas de interação espaço-tempo.

Table 1: Modelos de acordo com a presença ou não de efeitos espaciais e temporais.

| Modelos | γ | ϕ | ρ | φ | δ |
|---------|----------|--------|--------|-----------|--------------------------|
| 1 | X | X | X | X | X |
| 2 | ✓ | ✓ | X | X | X |
| 3 | ✓ | ✓ | X | X | ✓ |
| 4 | X | X | ✓ | ✓ | X |
| 5 | X | X | ✓ | ✓ | ✓ |
| 6 | ✓ | ✓ | ✓ | ✓ | X |
| 7 | ✓ | ✓ | ✓ | ✓ | $\gamma \otimes \phi$ |
| 8 | ✓ | ✓ | ✓ | ✓ | $\rho \otimes \phi$ |
| 9 | ✓ | ✓ | ✓ | ✓ | $\gamma \otimes \varphi$ |
| 10 | ✓ | ✓ | ✓ | ✓ | $\rho \otimes \varphi$ |

O modelo 1 representa a variabilidade total dos dados e foi ajustado apenas para servir de referência para comparações. O modelo 2 leva em consideração efeitos espaciais e temporais, porém assume a priori que tais efeitos não tem nenhuma estrutura. O modelo 3 é o modelo 2 acrescido de um efeito de interação não estruturado.

O modelo 4 assume a priori que os efeitos espaciais e temporais são estruturados, conforme descrito na seção 3. O modelo 5 é o modelo 4 acrescido de um efeito de interação estruturado. O modelo 6 condensa os modelos 2 e 4 dizendo que os efeitos espaciais e temporais podem ser divididos em uma parte estruturada e outra não estruturada.

Os modelos de interação espaço-tempo são diferentes combinações entre os efeitos principais que assume-se interagir a priori. Todos mantêm a mesma estrutura de efeitos principais, ou seja, efeitos temporais e espaciais estruturados e não-estruturados. No modelo 7 assume-se que a interação é do tipo *I* conforme apresentado na seção 3, na sequência tem-se o modelo 8 com a interação do tipo *II*, o modelo 9 com a interação do tipo *III* e por fim o modelo 10 com a interação do tipo *IV*.

Para o ajuste dos modelos foi utilizado a metodologia INLA (Rue, Martino e Chopin, 2009). Para comparação dos modelos usou-se a Deviance Média (DEV) o número esperado de parâmetros (NP), o Critério de Informação da Deviance (DIC) (Spiegelhalter, 2001), o *log score rules* (CPO) (Schrodle e Held, 2011).

A tabela 2 apresenta os critérios de comparação para cada um dos 10 modelos considerados.

Os resultados apresentados na tabela 2, mostram que existe um ganho razoável em termos de DIC quando passa-se do modelo 1 para o 2 e 4, ou seja, a estrutura espaço-tempo contribui para explicar a variabilidade das contagens.

| | DEV | NP | DIC | CPO |
|----|----------|--------|----------|------|
| 1 | 35757.01 | 1.66 | 35758.66 | 8.41 |
| 2 | 34872.84 | 173.16 | 35046.00 | 8.24 |
| 3 | 34850.31 | 199.67 | 35049.99 | 8.22 |
| 4 | 34888.32 | 124.40 | 35012.72 | 8.24 |
| 5 | 34902.35 | 241.59 | 35143.94 | 8.28 |
| 6 | 34888.52 | 125.83 | 35014.36 | 8.23 |
| 7 | 34858.71 | 162.17 | 35020.88 | 8.23 |
| 8 | 34866.62 | 157.82 | 35024.43 | 8.20 |
| 9 | 34905.61 | 221.44 | 35127.05 | 8.27 |
| 10 | 34902.79 | 242.83 | 35145.62 | 8.28 |

Table 2: Modelos ajustados e medidas para comparação.

Pensando na interação espaço-tempo, os modelos 3 e 5 mostram que não se obtêm um ganho razoável em incluir esta estrutura nos modelos.

A partir do modelo 6 os modelos apresentam duas estruturas para representar os efeitos espaciais e temporais, sendo o modelo 6 com apenas a estrutura principal. Os modelos 7 a 10 são as diversas estruturas de interação espaço-tempo. É possível ver claramente que conforme a estrutura vai ficando mais complexa a qualidade do ajuste vai caindo, de acordo com o DIC. Isso se deve porque modelos de maior complexidade tem maior número de parâmetros esperado, e como o DIC penaliza o modelo pela complexidade, praticamente não compensa colocar nenhuma estrutura de interação espaço-tempo.

De acordo com os resultados, pode-se escolher o modelo 4 como o que melhor descreve a estrutura espaço-temporal das contagens de ovos do mosquito *Aedes aegypti*. Este modelo conta com os efeitos principais de espaço e tempo, estruturados. A Figura 4, apresenta as realizações dos efeitos espaciais e temporais, com bandas de credibilidade de 95%. As distribuições a posteriori do intercepto, do parâmetro de superdispersão da Binomial Negativa, e dos parâmetros de precisão dos efeitos espaciais e temporais. O efeito espacial foi ordenado de forma crescente para facilitar a visualização.

De acordo com os resultados apresentados na figura 4, a contagem esperada esta em torno de 1388(1337 – 1440) ovos. O parâmetro de superdispersão da Binomial Negativa, apresentou uma estimativa de 1.5292(1.4440 – 1.6212) indicando a presença de superdispersão. As estimativas para os parâmetros de precisão dos efeitos aleatórios espaciais e temporais, indicam a presença destes, sendo que o efeito temporal é mais pronunciado.

A trajetória do efeito temporal mostra um comportamento cíclico. No ano de 2005 observa-se um pico próximo ao dia 15/03 seguido por um queda brusca e uma tendência crescente até um segundo pico que começa no final de junho e se estende até começo de setembro. A partir deste mês o efeito mostra uma tendência de decaimento até o mês de março de 2006, onde volta a crescer atingindo um pico aproximadamente em maio. Segue um período de oscilação da série em um nível bastante elevado até aproximadamente fevereiro de 2007 quando a série volta a apresentar uma tendência de decrescimento.

O gráfico da trajetória do efeito espacial mostra que existem áreas que apresentam contagens significativamente maiores que a média da área. Tais áreas

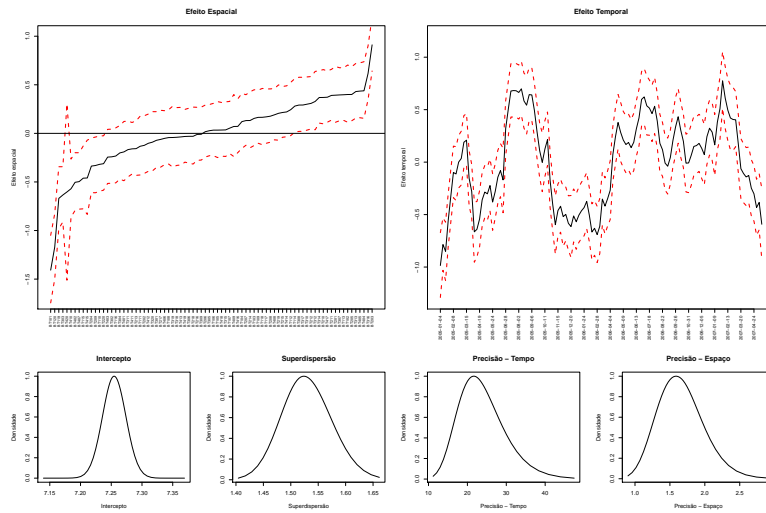


Figure 3: Resultados dos ajuste do modelo 4, usando aproximação de Laplace. Efeito espacial foi ordenado para facilitar a visualização das áreas com contagens significativamente abaixo ou acima da média.

podem ser tidas como áreas quentes, é onde ocorre a maior abundância de ovos. No total 20 armadilhas apresentam efeito significativamente maior que a média geral. Por outro lado, também verifica-se que existem armadilhas que apresentam contagens significativamente menores que a média geral, são um total de 13. Para as restantes o efeito não foi significativo.

As áreas onde o efeito espacial é significativamente superior a média geral, podem ser vistas como pontos quentes que devem receber maior atenção com relação as atividades de prevenção. Importante também ressaltar que como este efeito não muda com o tempo, já que, não houve interações significativas, estes pontos são fixos e devem estar representando o efeito de covariáveis não medidas pelo experimento que apresentam padrão espacial. O mapa apresentado na figura ?? representa tais resultados.

De forma geral, os resultados mostram cinco pequenos aglomerados de duas a cinco de áreas de alto risco. Estas são as áreas prioritárias para intervenção no combate ao mosquito. As áreas de baixo risco aparecem sem um padrão definido ao longo de todo o mapa.

Conclusões

Mostramos a aplicação de modelos espaço-temporais, para determinar a distribuição espaço-temporal de contagens de ovos do mosquito *Aedes aegypti* coletados em ovitrampas, na cidade de Recife/PE. O uso desta classe de modelos é ainda reduzido, pois trata-se de modelos altamente estruturados que apresentam muitos desafios para o procedimento de inferência. Neste artigo foi utilizada a metodologia INLA, para fazer inferência Bayesiana completa. Tal metodologia mostrou-se bastante satisfatória, permitindo o ajuste de diversos modelos altamente estruturados em um tempo computacional aceitável. Isto permitiu a

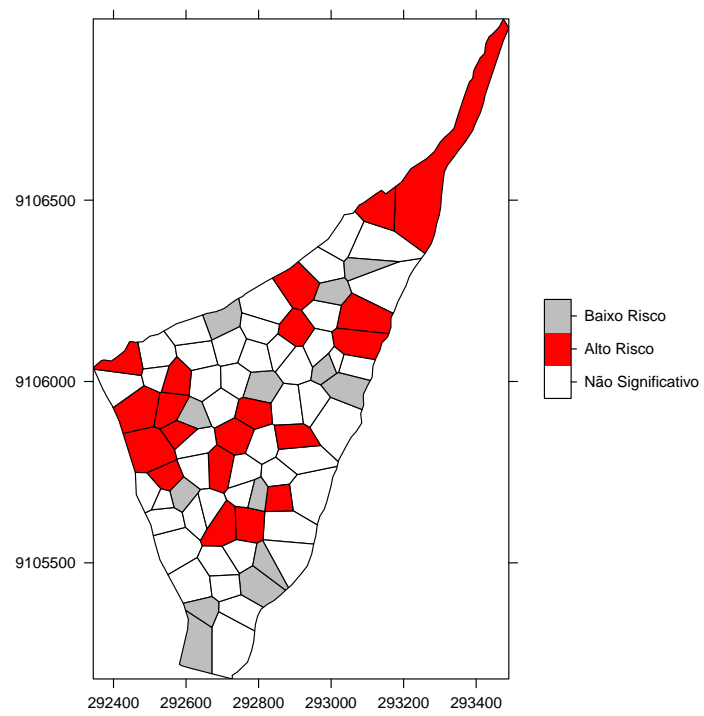


Figure 4: Mapa do efeito espacial, indicando áreas de alto e baixo risco.

comparação de diversos modelos, cada qual testando diferentes hipóteses sobre a realidade do estudo. Tal metodologia pode ser incorporadas a ambientes de monitoramento e vigilância como descritos em Regis et. al. (2008).

De forma geral, os resultados dos modelos indicam que as contagens de ovos de *Aedes aegypti* apresentam dependência espacial o que também verificou-se em análises de outras regiões de Recife, não reportadas aqui. Esta dependência pode estar capturando efeitos de potenciais covariáveis não consideradas que apresentem padrões espaciais na área o que deve orientar estudos posteriores. Este efeito também permitiu identificar áreas onde a ocorrência de ovos é significativamente maior que a média geral da área de estudo. Tal informação pode ser usada para delimitar áreas prioritárias para as ações públicas que objetivam eliminar os mosquitos. Além disso, estas áreas devem se tornar o foco de campanhas educativas, para que a própria população residente evite manter ou criar locais adequados para a proliferação do mosquito.

O efeito temporal mostra maior ocorrência de ovos do mosquito entre outubro e março. Nestes meses a ocorrência de chuvas intermitentes é comum na região, o que pode aumentar o número de locais com água acumulada, que favorecem a proliferação do mosquito. Novamente em tais períodos as campanhas preventivas podem ser intensificadas.

Os modelos propostos exploram diversas possibilidades para investigar a existência de interação espaço-tempo, porém pelos critérios utilizados nenhum efeito de interação foi detectado no caso examinado. A não existência da interação significativa, indica que as áreas quentes são as mesmas independentes do período do ano. O mesmo pode ser dito para o efeito temporal, quando as contagens aumentam devido as condições climáticas que mudam durante o ano, a mudança é comum para todas as áreas.

As contagens altas apresentam superdispersão que foi considerada ao adotar-se distribuição Binomial Negativa. É importante lembrar que esta distribuição pode ser vista como um modelo Poisson acrescido de um efeito aleatório Gama, em cada observação. Desta forma, conjecturamos que efeitos de interação espaço-tempo possam ser capturados por este efeito tornando-se não identificáveis, o que deve ser investigado em trabalhos futuros.

References

- BESAG, J. ; YORK, J. ; MOLLIÉ, A. *Bayesian image restoration with two applications in spatial statistics*. Annals of Institute of Statistical Mathematics, v.43, p.1-59, 1991.
- BRAGA, I. A.; VALLE, D. *Aedes aegypti: vigilância, monitoramento da resistência e alternativas de controle no Brasil*. Epidemiologia e Serviços de Saúde, v.16(4), p.295-302, 2007.
- BONAT, W. H. *Aplicações de inferência bayesiana aproximada para modelos Gaussianos latentes espaço temporais*, 2010. 80p. Dissertação de Mestrado - UFPR - Universidade Federal do Paraná, Curitiba, 2010.
- BONAT, W. H. ; KRAINSKI, E. T. ; RIBEIRO Jr, P. J. ; ZEVIANI, W. M. *Métodos computacionais para inferência com aplicações em R*. João Pessoa: ABE - Associação Brasileira de Estatística, 2012. 260p.

- BONAT, W. H. ; DALLAZUANNA, H. S. ; RIBEIRO Jr., P. J.; REGIS, L. N. ; MONTEIRO, A. M. V. ; SILVEIRA, J. C. ; ACIOLI, R. ; SOUZA, W. V. *Investigando fatores associados a ocorrência de ovos do mosquito aedes aegypti coletados em ovitampas em Recife/PE*. Revista Brasileira de Biometria, v.27, p.519-537, 2009.
- CLAYTON, D. G. Generalized linear mixed models. In: GILKS, W. R. ; RICHARDSON, S. ; SPIEGELHALTER, D. J. (Ed.). *Markov Chain Monte Carlo in Practice*. London: Chapman & Hall, 1996. p.275-301.
- CONNOR, M. E. ; MONROE, W. M. *Stegomyia indices and their value in yellow fever control*. American Journal of tropical Medicine and Hygiene, v.3 , p.9-19, 1923.
- FAY, R. W. ; ELIASON, D. A. *Laboratory studies of ovipositional preferences of Aedes aegypti*. Mosquito News, v.25, p.270-281, 1965.
- FAY, R. W. ; ELIASON, D. A. *A preferred ovoposition site as a surveillance method for Aedes aegypti*. Mosquito News, v.26, p.531-534, 1966.
- MARTINO, S. ; RUE, H. *Implementing approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations: A manual for the inla-program*, 2009. 84p. Technical report - Department of Mathematical Science, Norwegian University of Science and Technology, Trondheim, Norway, 2009.
- MONTEIRO, A. M. ; CARVALHO, M. S. ; ASSUNÇÃO, R. W. ; RIBEIRO Jr, P. J. ; VIEIRA, W. ; DAVIS, C. ; REGIS, L. *SAUDAVAL: Bridging the Gap between Research and Service in Public Health Operational Programs by Multi - Institutional Networking Development and use o Spatial Information Technology Innovative Tools*. 2006.
- REGIS, L. ; MONTEIRO, A. M. ; MELO SANTOS, M. A. ; SILVEIRA, J. C. ; FURTADO, A. F. ; ACIOLI, R. V. ; SANTOS, G. M. ; NAKASAWA, M. ; CARVALHO, M. S. ; RIBEIRO Jr, P. J. ; SOUZA, G. M. *Developing new approaches for detecting and preventing Aedes aegypti population outbreaks: bases for surveillance, alert and control system*. Memórias do Instituto Oswaldo Cruz, v.103 , p.50-59, 2008.
- RUE, H. ; MARTINO, S. ; CHOPIN, N. *Approximate Bayesian inference for latent Gaussian models using integrated Laplace approximations*. Journal Royal Statistical Society (Series B), v.71, p.319-392, 2009.
- RUE, H. ; FOLLESTAD, T. *GMRFLib: A C-library for fast and exact simulation of Gaussian Markov random fields*, 2002. Technical report - Department of Mathematical Science, Norwegian University of Science and Technology, Trondheim, Norway, 2002.
- RUE, H. ; HELD, L. *Gaussian Markov Random Fields: Theory and Applications*. London: Chapman and Hall, 2005. 263p.
- R Development Core Team. 2009. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- SCHRODLE, B. ; HELD, L. *A primer on disease mapping and ecological regression using INLA..* Computational Statistics, v.26 p.241-258, 2011.

- SPIEGELHALTER, D. J. ; BEST, N. G. ; CARLIN, B. P. ; VAN DER LINDE, A. *Bayesian measures of model complexity and fit (with discussion)*. Journal Royal Statistical Society (Series B), v.64 p.583-639, 2001.
- STONE, C. J. *Polynomial splines and their tensor products in extended linear modelling*. Annals of Statistics, v.25, p.1371-1470, 1997.
- TAUIL, P. L. *Aspectos críticos do controle do dengue no Brasil*. Cadernos de Saúde Pública, v.18, p.867-871, 2002.