

**INVESTIGANDO FATORES ASSOCIADOS A CONTAGENS DE
OVOS DE *Aedes Aegypti* COLETADOS EM OVITRAMPAS EM
RECIFE/PE.**

Wagner Hugo BONAT¹

Henrique Silva DALLAZUANNA²

Paulo Justiniano RIBEIRO JR ³

Lêda Narcisa REGIS ⁴

Antônio Miguel Vieira MONTEIRO ⁵

José Constantino SILVEIRA ⁶

Ridelane ACIOLI ⁷

Wayner Vieira de SOUZA ⁸

¹Laboratório de estatística e geoinformação, Universidade Federal do Paraná, CEP:81531-990, Curitiba, Paraná, Brasil , E-mail: *wagner@leg.ufpr.br*

²Laboratório de estatística e geoinformação, Universidade Federal do Paraná, CEP:81531-990, Curitiba, Paraná, Brasil , E-mail: *henrique@leg.ufpr.br*

³Laboratório de estatística e geoinformação, Universidade Federal do Paraná, CEP:81531-990, Curitiba, Paraná, Brasil , E-mail: *paulojus@leg.ufpr.br*

⁴Departamento de Entomologia, Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, CEP: 50670-420, Recife, Pernambuco, Brasil, E-mail: *leda@cpqam.fiocruz.br*

⁵Divisão de Processamento de Imagens & Programa Espaço e Sociedade, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, São Paulo, Brasil , E-mail: *miguel@dpi.inpe.br*

⁶Departamento de Entomologia, Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, CEP: 50670-420, Recife, Pernambuco, Brasil, E-mail: *constantino@cpqam.fiocruz.br*

⁷Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, CEP: 50670-420, Recife, Pernambuco, Brasil, E-mail: *ridelane@cpqam.fiocruz.br*

⁸Departamento de Saúde Coletiva, Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, CEP: 50670-420, Recife, Pernambuco, Brasil, E-mail: *wayner@cpqam.fiocruz.br*

- RESUMO: O *Aedes aegypti* é o vetor do vírus da dengue, doença para a qual tem-se observado a ocorrência de diversas epidemias. Estudos entomológicos contribuem para o entendimento da dinâmica de proliferação do mosquito. Este artigo tem como objetivo, propor um protocolo de análise combinando métodos estatísticos para identificar fatores associados à intensidade de ovos de *Aedes aegypti*, coletados com o uso de armadilhas de oviposição. A metodologia é desenvolvida a partir de dados coletados em um experimento de campo conduzido pelo projeto SAUDAVEL, que combina dados de coletas de ovos do mosquito e condições das armadilhas e ambientais em bairros selecionados para monitoramento na cidade de Recife/PE. Os principais resultados apontam para maior frequência/intensidade de oviposição em situações residências, onde o abastecimento de água é constante. Evidências sugerem a competição das armadilhas, com recipientes como tanques, caixas d'água e tonéis. Covariáveis climáticas possibilitaram identificar a preferência do vetor por alta precipitação e umidade. Características espaço-temporais do experimento foram levadas em consideração na modelagem, através de funções suaves das coordenadas geográficas das armadilhas e das datas de coleta. Utilizou-se a metodologia de modelos aditivos generalizados, que permitiu a identificação de áreas onde as contagens de ovos são elevadas. O termo temporal no modelo identificou os meses do ano de maior abundância, além de efeitos sazonais e a tendência da série.
- PALAVRAS-CHAVE: Estudos entomológicos, *Aedes aegypti*, dengue, modelos aditivos generalizados, modelos espaço-temporais.

1 Introdução

A degradação do meio ambiente e os aspectos sócio-culturais afetam o cenário epidemiológico brasileiro, com impacto na mídia nacional e internacional, em virtude de epidemias de dengue, leptospirose, a recorrência da tuberculose, entre outras.

Diante disso, constatou-se a importância de modelar e identificar fatores de risco e proteção, nas situações endêmicas e epidêmicas visando equipar sistemas de vigilância e alerta de metodologias de análise que possam ser implementadas de forma rotineira.

Neste sentido, foi elaborado o Projeto SAUDAVEL¹, 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).

Dentre as inúmeras doenças transmissíveis conhecidas nos dias de hoje, a dengue é uma das mais importantes, sendo atualmente um dos principais problemas de saúde pública no país. O agente etiológico da dengue é um vírus, e seus transmissores são mosquitos do gênero *Aedes*, popularmente conhecidos

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

como “mosquitos da dengue”. As condições do meio-ambiente favorecem o desenvolvimento e a proliferação do mosquito *Aedes aegypti*, principal transmissor da dengue nos países latinos (Tauil, 2002).

O *Aedes aegypti*, o mais comum do gênero *Aedes*, é um mosquito fortemente domiciliado e urbano. A transmissão da doença ocorre a partir da picada das fêmeas adultas do mosquito, pois somente elas são hematófagas (Yang e Thomé, 2007).

Após ter sugado o sangue de uma pessoa contaminada e passado o período de incubação de oito a dez dias, a fêmea infectada está apta a transmitir o vírus a indivíduos suscetíveis.

A partir da década de 1980 iniciou-se um processo de intensa circulação dos vírus da dengue no Brasil, com epidemias explosivas que 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 e utilizadas no país.

O monitoramento de fatores de risco biológico relacionados aos vetores de doenças tem como finalidade o mapeamento de áreas de risco em determinados territórios, combinando informações das vigilâncias entomológicas e epidemiológicas. Na primeira pode-se verificar a presença do vetor, índices de infestação, características biológicas, tais como susceptibilidade aos inseticidas e aos vírus, avaliação da eficácia dos métodos de controle, as ações de controle químico, biológico ou físico. Exploram-se relações com informações de incidência e soro-prevalência da doença provenientes da vigilância epidemiológica.

Nos programas de controle de dengue, a vigilância entomológica é feita principalmente a partir da coleta de larvas, de acordo com proposta de Connor e Monroe (1923) para medir a densidade de *Aedes aegypti* em áreas urbanas. 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) segundo FUNASA (2001) e OPS (1995).

Segundo Braga e Valle (2007) 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, embora venha sendo usada com essa finalidade (Nelson, 1995 e Focks, 2000).

Outra metodologia adotada é a coleta de mosquitos adultos, cuja operacionalização para a estimativa do risco de transmissão é custosa e demorada. Ainda segundo Braga e Valle (2007), a correlação entre o número de vetores coletados (mosquitos adultos) e o número de humanos na área de coleta, que poderia fornecer o número de vetores adultos por pessoa, não é suficiente para quantificar o risco. Contudo, essa correlação se aproxima mais da realidade que os índices larvários, ver Focks (2000).

Apesar disso, para avaliar a infestação 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 'ovitampa', é destinada à coleta de ovos. Em um recipiente de cor escura, adere-se um material áspero que permite a fixação dos ovos depositados. Em 1965, iniciou-se o uso de ovitrampas para a vigilância das 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 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.

Diversos métodos para medir a infestação por *Aedes aegypti* em uma determinada área já foram propostas na literatura, a maioria tendo como base o uso de armadilhas para capturar ovos do mosquito. Na mensuração da densidade vetorial, cada método tem sua particularidade mas todos lançam mão de amostras, nem sempre representativas, da população estudada. Um estudo bastante completo sobre medidas dos níveis de infestação urbana para *Aedes* é encontrado em Gomes (2002). Apesar do uso intenso de ovitrampas em estudos entomológicos, poucos foram os trabalhos encontrados que buscam relacionar a ocorrência das formas imaturas, com fatores relacionados a própria condição da armadilha, como presença de recipientes potenciais para o armazenamento de água, tais como tanques, caixa d'água, cisternas, além de fatores amplamente conhecidos, como vasos de plantas, presença de pneus entre outros.

O trabalho de Santos (1999) estudou fatores associados à ocorrência de formas imaturas de *Aedes aegypti* na Ilha do Governador, no estado do Rio de Janeiro. Contudo, em um contexto diferente, sem o uso de ovitrampas e um delineamento amostral propriamente dito.

O estudo foi realizado a partir dos dados da Fundação Nacional de Saúde (FNS). Os principais resultados obtidos pelo autor mostram que 58,04% dos criadores inspecionados foram constituídos por suportes para vasos com plantas, vasilhames de plástico ou vidro abandonados no peridomicílio. Maiores percentuais de criadouros positivos foram observados para pneus (1,41%), tanques, poços e cisternas (0,93%), e barris, toneis e tinas (0,64%). Maiores percentuais de criadouros positivos durante o verão foram encontrados em grandes reservatórios de água, e a dos criadouros provenientes do lixo doméstico. No inverno foi verificado maior ocorrência em pequenos reservatórios de água para o uso doméstico. O autor também utilizou um modelo fatorial que mostrou como fator principal para determinar a ocorrência de fases imaturas de *Aedes aegypti*, como sendo aquele que leva em consideração os fatores climáticos, as medidas utilizadas foram a média da pressão atmosférica, a temperatura máxima, temperatura mínima e a umidade relativa do ar.

O trabalho de Ferreira e Chiaravalloti (2007) buscou associar os índices de infestação larvária por *Aedes aegypti* com fatores socioeconômicos, em São José do Rio Preto, São Paulo. Os autores agruparam os setores censitários urbanos

da cidade em 4 grupos segundo variáveis socioeconômicas, utilizando análise de componentes principais, e um quinto agrupamento com nível socioeconômico inferior aos demais não pertencente aos setores censitários urbanos.

Calcularam, para cada área, os índices de Breteau (IB), predial (IP), de recipientes (IR), e as médias de recipientes existentes e pesquisados por casa. Os valores dos índices de infestação não apresentaram diferenças significativas entre as áreas socioeconômicas 1 a 4, mas foram menores que para a quinta área. Os índices larvários não mostraram associação com os diferentes níveis socioeconômicos, da área correspondente aos setores censitários urbanos, porém os loteamentos da área 5 apresentaram os maiores valores desses indicadores.

Em estudo sobre a variação sazonal de *Aedes aegypti* Serpa *et al.* (2007), em Potin, São Paulo, verificaram que os meses de julho, novembro e dezembro foram os de menor abundância do mosquito. Também foi detectado a existência de relação significativa da ocorrência de formas imaturas com a temperatura máxima, porém não encontraram relação com as variáveis climáticas, temperatura mínima, pluviosidade e umidade relativa do ar.

O cenário epidemiológico brasileiro atual reforça que estudos entomológicos devem ser feitos para buscar um melhor entendimento, sobre os fatores que influenciam no comportamento reprodutivo do vetor. Desta forma, pode-se indicar fatores de risco e proteção para evitar a proliferação do vetor e consequentemente diminuir os indicadores de incidência da dengue.

Neste contexto definem-se os objetivos deste trabalho em combinar metodologias de análise que permitam determinar fatores de risco e proteção associados à ocorrência de ovos do mosquito *Aedes aegypti*, com base em dados de um experimento conduzido pelo "Projeto SAUDAVEL" na cidade de Recife/PE. Entende-se aqui, como fatores de risco/proteção tanto covariáveis associadas à armadilha, como presença de recipientes grandes ou pequenos que possam conter água em suas proximidades, como também aspectos abióticos (climáticos) como temperatura, precipitação e umidade. Possíveis relações espaciais entre as armadilhas e também a possibilidade de uma relação temporal entre as coletas são investigadas.

Busca-se um modelo implementável em procedimentos de rotina em um sistema de vigilância que leve em consideração a complexidade do ambiente em que o experimento foi desenvolvido, e que possa ser usado tanto para entender a dinâmica de proliferação do mosquito, mostrando suas preferências para a reprodução, bem como, um modelo preditivo buscando evitar surtos entomológicos nas áreas em estudo.

Padrões espaciais eventualmente detectados pelo modelo podem destacar zonas de persistência de alta infestação. Os padrões temporais, podem auxiliar na compreensão da dinâmica temporal da população de mosquitos. Efeitos dos fatores do ambiente local dos pontos de coleta são avaliados, buscando-se aqueles que apresentam relação com as contagens de ovos. Buscou-se, detectar condições climáticas refletidas em dados de temperatura e umidade em períodos que antecedem as coletas e que possam criar condições ideais para o aumento do processo

de oviposição, bem como, a determinação de períodos relevantes de defasagem, entre a medida climática e o impacto nesse processo.

Pretende-se que a metodologia de análise proposta neste artigo, seja incorporada em um pacote para a vigilância de populações de *Aedes aegypti*. O desenvolvimento do pacote **RDengue** (Bonat, Dallazuana e Ribeiro Jr, 2008), do software **R** (R Development Core Team, 2009), faz parte das atividades de uma rede interinstitucional voltada para a criação de um sistema de vigilância entomológica, baseada em uma estrutura de software livre para o controle participativo da Dengue, (Regis, *et al.* 2009).

Nesta seção enfatizou-se a condição epidemiológica e entomológica dando uma visão ampla da situação brasileira, citando vários artigos que tratam sobre o mesmo assunto, e também apresentou os objetivos gerais do trabalho. A Seção 2 visa descrever em detalhes o delineamento e forma de condução do experimento, descreve as variáveis utilizadas para a construção dos modelos, e sua forma de mensuração pela equipe de campo.

A Seção 3 traz um breve resumo sobre as metodologias estatísticas utilizadas, dando ênfase aos modelos aditivos generalizados. A Seção 4 traz os principais resultados da aplicação do modelo estatístico. A última seção traz uma breve discussão sobre os resultados, as principais conclusões e limitações do estudo e aponta algumas possibilidades para investigações futuras.

2 Contexto e material

2.1 Área de estudo, Instrumentos e Técnicas de Campo

No experimento conduzido na cidade de Recife/PE, foram instaladas 464 armadilhas (ovitrampas) para capturar ovos do mosquito *Aedes aegypti*, cuja a fêmea é o principal vetor da dengue.

Estas armadilhas começaram a ser monitoradas em março de 2004. O experimento está sendo realizado em 7 dos 94 bairros da cidade de Recife. Os resultados apresentado aqui referem-se 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, caracterizando bem o tipo de delineamento utilizado para a coleta de dados. A cada sete dias é feita a contagem de ovos encontrados em cerca de um quarto das armadilhas.

Assim, em um ciclo de 28 dias todas as armadilhas são monitoradas. Cada armadilha contém três lâminas de material áspero na qual a fêmea do mosquito coloca os ovos, quando recolhidas são levadas para um laboratório especializado onde a contagem dos ovos é realizada. O departamento de entomologia da FIOCRUZ/PE e os serviços de saúde locais foram os coordenadores operacionais e logísticos e responsáveis pela realização do experimento (Monteiro *et al.* 2006). Estes serviços

também coletaram informações sobre as características dos locais das armadilhas que serão utilizadas nas análises.

Regis *et al.*, 2008 descrevem de forma ampla o experimento SAU-DAVEL/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.

2.2 Construção das covariáveis

O banco de dados do SAU-DAVEL Recife, possibilitou a construção de diversas covariáveis a serem utilizadas nos modelos para a contagem de ovos que foram construídas de forma adequada para a inclusão no modelo estatístico e possibilitar interpretações práticas.

O conjunto de covariáveis pode ser dividido em dois grupos, o primeiro das covariáveis “locais” referentes à armadilha e o segundo de covariáveis “ambientais” referentes a fatores abióticos, representados por variáveis climáticas.

A tabela 1 resume o conjunto de covariáveis locais e a codificação adotada para as análises.

Tabela - 1: Relação de covariáveis locais.

Covariáveis	Níveis	Descrição
Tipo de imóvel	0	Residencial
	1	Não residencial
Quintal	0	Apresenta quintal
	1	Não apresenta quintal
Água ligada a rede geral	0	Ligada
	1	Não ligada
Abastecimento de Água	0	Diário
	1	Não diário
Água canalizada no cômodo	0	Canalizada
	1	Não canalizada
Fatores de risco	0	Apresenta
	1	Não apresenta
Recipientes grandes sem tampa	0	Apresenta
	1	Não apresenta
Recipientes grandes com tampa	0	Apresenta
	1	Não apresenta
Recipientes pequenos sem tampa	0	Apresenta
	1	Não apresenta
Recipientes pequenos com tampa	0	Apresenta
	1	Não apresenta

São considerados fatores de risco, plantas em vasos, charco/poça, garrafas, fossa externa, piscina, poço elevador, laje sem telhado ou calhas. Na categoria de

recipientes grandes são considerados, tanques, caixa d'água ou toneis. Na categoria de recipientes pequenos são considerados, jarros de barro ou baldes.

Uma outra covariável não elencada na tabela, mas que também foi levada em consideração na análise é denominada de **Grupos**. Esta covariável corresponde a semana em que a armadilha é monitorada, como as coletas são realizadas semanalmente é feita em grupos de armadilhas, cada grupo tem aproximadamente um quarto das armadilhas colocadas no bairro.

As covariáveis climáticas disponíveis no banco de dados e utilizadas nas análises foram precipitação, umidade relativa do ar, temperatura máxima e mínima. A mensuração destas covariáveis é feita diariamente por uma estação de monitoramento. Para fins de modelagem estatística são necessárias decisões e manipulações destes dados, para permitir investigar relações com os de coleta de ovos, obtidas semanalmente. As observações foram agrupadas mensalmente, tomando uma média das observações diárias.

Em análises preliminares e não reportadas aqui foram considerados agrupamentos semanais, além de outras medidas resumo, como máximos, mínimos, amplitudes e quartis.

Porém, a medida que mostrou melhor desempenho nos modelos ajustados foi a média, e portanto optou-se por reportar as análises com esta medida.

Foram contabilizadas estas covariáveis com uma defasagem de até três meses da observação.

3 Modelos para dados de contagens

O modelo Poisson log-linear é o modelo canônico para dados de contagem. Porém, em alguns casos pode ocorrer o fenômeno da sobredispersão, refletindo o fato de que a variância nas observações não segue as propriedades do modelo de Poisson, sendo em geral maior do que a predita pelo modelo.

Nesses casos, a suposição de distribuição de Poisson para a resposta é inadequada, sendo necessário o uso de modelos alternativos. Neste trabalho foi utilizada uma abordagem assumindo a variável resposta como tendo distribuição Binomial Negativa. Pode ser demonstrado que a Binomial Negativa é o modelo Poisson acrescido de um efeito aleatório com distribuição Gama.

Uma causa provável da sobredispersão é a heterogeneidade das unidades amostrais, que pode ser devido a variabilidades inter-unidades experimentais. Este fato, foi detectado no caso dos dados do experimento SAUDAVEL/Recife onde as unidades experimentais, dadas pelas armadilhas, são notadamente diferentes umas das outras.

Desta forma, supondo que Y_1, \dots, Y_n são variáveis independentes de modo que $Y_i \sim BN(\mu_i, \phi)$, em que $E(Y_i) = \mu_i$ e $Var(Y_i) = \mu_i + \frac{\mu_i^2}{\phi}$ e parte sistemática dada por $g(\mu_i) = x_i^T \beta$ em que $g(\cdot)$ é uma função de ligação.

Definindo $\theta = (\beta, \phi)^T$ o logaritmo da função de verossimilhança fica dado por

$$L(\theta) = \sum_{i=1}^n \left[\log \left\{ \frac{\Gamma(\phi + y_i)}{\Gamma(y_i + 1)\Gamma(\phi)} \right\} + \phi \log \phi + y_i \log y_i - (\phi + y_i) \log(\mu_i + \phi) \right]$$

em que $\mu_i = \exp(x_i^T \beta)$. As estimativas de máxima verossimilhança podem ser obtidas através de algoritmos iterativos, como o método de Newton-Raphson (Paula, 2004).

O modelo descrito, apesar de muito utilizado, tem uma restrição quanto ao relacionamento da variável resposta com as covariáveis, assumindo que tal relacionamento seja linear na escala da função de ligação, o que pode ser restritivo não possibilitando um ajuste satisfatório. Em particular, no caso do presente estudo, deseja-se incluir efeitos espaciais através das coordenadas geográficas das armadilhas, sem assumir a priori que o relacionamento com a resposta seja linear na escala da função de ligação utilizada.

Uma das possíveis abordagens alternativas é a adoção do modelo aditivo generalizado, (Hastie e Tibishirani, 1990) que pode ser descrito como uma extensão do modelo linear generalizado (Nelder e Wedderburn, 1972), porém com um ou mais preditores lineares envolvendo a soma de funções suaves das covariáveis.

O modelo torna-se semi-paramétrico e pode ser escrito da seguinte forma:

$$g(\mu_i) = X_i^* \theta + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + \dots \quad (1)$$

em que X_i^* é uma linha da matriz do modelo do termo linear para a parte estritamente paramétrica, θ é o vetor de parâmetros correspondentes e as f_j são funções suaves das covariáveis x_k .

As funções suaves podem ter mais de uma covariável como argumento, conforme ilustrado por f_3 na expressão (1). O modelo definido desta forma, proporciona flexibilidade na especificação da forma da relação entre a variável resposta e as covariáveis.

Casos particulares também são possíveis, como a especificação do modelo apenas em termos de funções suaves.

Entretanto, a flexibilidade de tais modelos vem acompanhada de dois novos problemas teóricos: como representar as funções suaves e como estimar os parâmetros envolvidos neste modelo.

O problema de como representar a função suave univariada, pode ser resolvido usando *splines*. Um *spline* cúbico, é uma curva composta por seções de polinomiais cúbicas, juntas de modo que componham uma função contínua que permita primeira e segunda derivada.

Uma segunda opção de função *spline* para o caso de mais de uma covariável é a “*thin plate*”. Segundo Wood (2006) a *thin plate* é uma solução elegante e geral para o problema de estimar uma função suave de variáveis preditoras múltiplas.

Wahba (2000) mostra que a *thin plate splines* é uma generalização natural da *spline* polinomial univariada, para duas ou mais dimensões.

A dificuldade com *thin plate splines* é o custo computacional, dado que estes suavizadores têm tantos parâmetros desconhecidos quanto dados (estritamente, número de combinações únicas do preditor).

Exceto no caso do preditor simples, o custo computacional da estimação do modelo é proporcional ao cubo do número de parâmetros.

Apesar disto, foi adotado e amplamente usado aqui, porque uma de suas principais características é a isotropia da penalidade das ondulações, onde tais ondulações são em todas as direções igualmente tratadas, com o ajuste inteiramente invariante para a rotação do sistema de coordenadas das covariáveis preditoras. Tais funções suavizadoras são adequadas para variáveis medidas na mesma unidade tais como coordenadas geográficas, onde a isotropia é assumida como adequada. Uma vez escolhida a base de *spline* o ajuste do modelo segue por uma variação do Método de Mínimos Quadrados, denominado de PIRLS - Mínimos quadrados ponderados iterativos. Intervalos de confiança para as funções suaves podem ser derivados usando argumentos Bayesianos (Wood, 2006).

Os procedimentos necessários para o ajuste desta classe de modelos, estão implementados no pacote *mgcv* (Wood, 2008) do ambiente estatístico R (R Development Core Team, 2009), o qual foi utilizado neste trabalho.

4 Resultados

4.1 Análise exploratória

As condições ligadas às armadilhas são de papel fundamental, pois são elas que orientam as políticas de prevenção da propagação do vetor através de campanhas educacionais promovidas a fim de evitar criadouros do mosquito. Assim, a identificação das características associadas é importante para orientar as ações de tais campanhas.

O conjunto de gráficos da figura 1 faz uma comparação das contagens de ovos (em escala logarítmica) entre as categorias de cada uma das doze covariáveis locais. A análise destes gráficos permite identificar de forma exploratória e inicial, os fatores que mais afetam as contagens de ovos, orientando a seleção e a escolha de modelos, antecipando e explicando possíveis resultados da modelagem.

Pelo conjunto de gráficos da figura 1 é possível observar que as diferenças são pequenas e apenas em algumas das covariáveis.

Pode-se observar que o tipo de imóvel residencial apresenta contagens médias maiores que os imóveis não residenciais. A presença de fatores de risco também apresenta um leve aumento na contagem média de ovos. As condições ligadas a água, água canalizada, frequência de abastecimento e água ligada a rede geral, apresentam uma pequena diferença entre as categorias, sendo sempre maior nas categorias que indicam presença de água abundante. A frequência de coleta de lixo, mostra uma maior contagem média de ovos nas localizações que apresentam coleta de lixo diária. Para as outras covariáveis não é possível identificar nenhuma relação evidente.

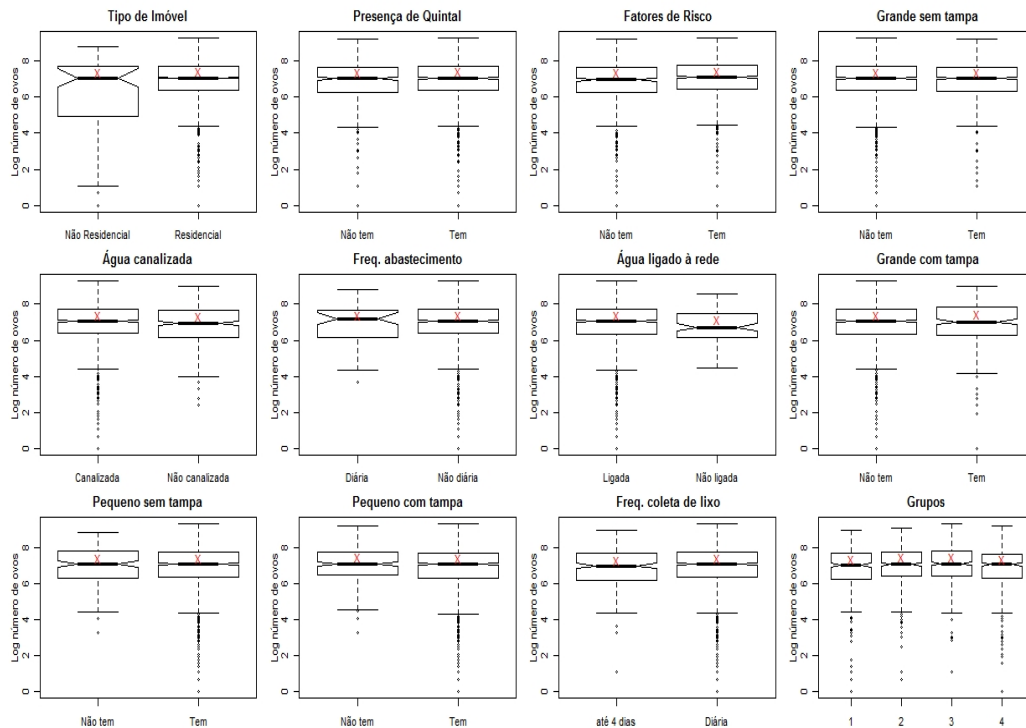


Figura - 1: Boxplots do log da contagem de ovos por categorias para cada covariável (X representa a média amostral), bairro Brasília Teimosa Recife/PE.

Uma análise exploratória para as covariáveis ambientais é dificultada pela existência de efeitos temporais e espaciais que devem influenciar na relação das contagens de ovos com essas covariáveis.

Para contornar este fato, foi inicialmente ajustado um modelo para cada uma das covariáveis ambientais em análise.

Os efeitos espaciais e temporais foram controlados por funções suaves das coordenadas geográficas e das datas de coleta evitando assim assumir a priori que o relacionamento dessas covariáveis com a resposta é linear na escala da função de ligação.

Para o ajuste do modelo a variável resposta foi assumida com Binomial Negativa com função de ligação logarítmica e funções suaves são *Thin Plate Splines*.

O conjunto de gráficos apresentados na figura 2 resume os resultados.

Pela figura 2 pode-se observar para a covariável precipitação um comportamento oscilante para o mês da observação e dois meses anteriores a observação.

Para o mês anterior o comportamento é persistente e fica claro que com o aumento da precipitação espera-se aumento nas contagens de ovos.

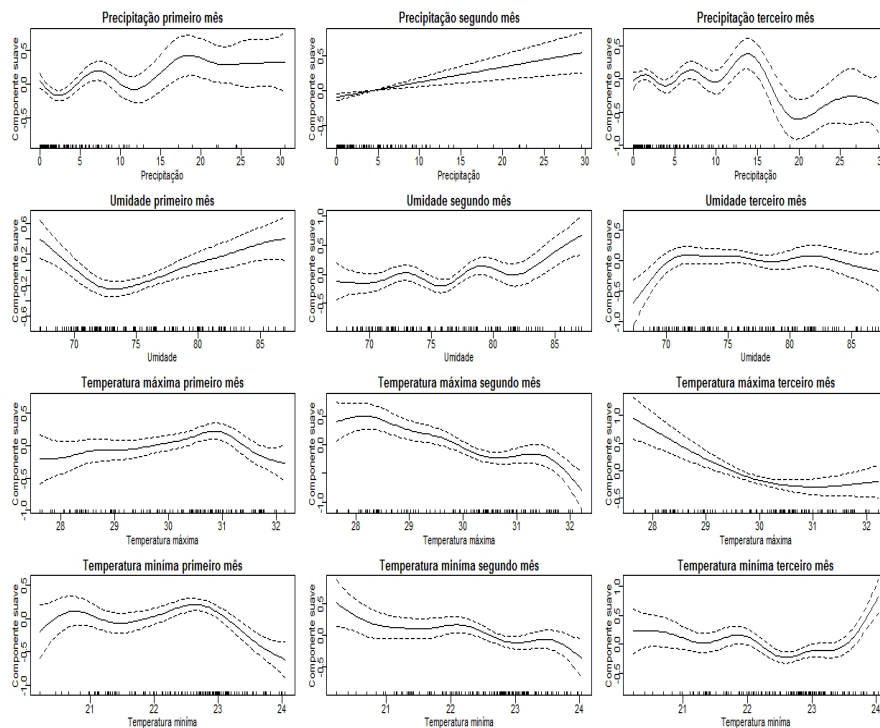


Figura - 2: Diagramas de dispersão.

Para a covariável umidade os três meses apresentam comportamentos variados. Para o primeiro mês verifica-se que umidade de 65 até aproximadamente 73 provoca uma queda nas contagens de ovos, enquanto que umidades maiores que 73 provocam aumento nas contagens de ovos. Para o mês anterior a observação o comportamento oscila mas a tendência geral parece de aumento nas contagens de ovos, apesar de alguns momentos a *spline* ter decaimentos. Para dois meses antes da observação aparece um efeito forte para umidade entre 65 e 70 aumentando as contagens, porém a partir de 70 não parece existir nenhum efeito.

Para a temperatura máxima no mês da observação verifica-se, que temperaturas de 28 a 31 graus aumentam as contagens, enquanto que temperaturas maiores que 31 provocam uma queda nas contagens de ovos. Para o mês anterior a observação a temperatura máxima é responsável por uma queda nas contagens, porém com diferentes graus de intensidade. A temperatura máxima de dois meses antes da observação provoca um queda bastante acentuada nas contagens, para temperaturas até 30,5 graus, depois disso parece não ter mais efeito.

A temperatura mínima no mês da observação de 18 até 22,7 graus é responsável por um aumento nas contagens de ovos, enquanto que temperaturas

mínimas maiores que 22,7 provocam uma queda acentuada nas contagens de ovos. A temperatura mínima no mês anterior da observação provoca uma queda nas contagens de ovos, com diferentes graus de inclinação, mas sempre decrescente. Para dois meses antes da observação a temperatura mínima de 18 a aproximadamente 23 graus provoca uma queda pouco acentuada nas coletas de ovos, porém temperatura mínima a partir de 23 graus provocam um aumento acentuado nas coletas de ovos.

Como pode ser visto de forma geral na figura 2, o relacionamento destas covariáveis com as contagens de ovos não aparece de forma imediata. Porém deve ser notado também, que as covariáveis ambientais são altamente correlacionadas, e grande parte da não linearidade destas relações pode ser causada por outras covariáveis, já que, todas estão atuando sobre a variável resposta ao mesmo tempo.

Nesse ponto espera-se que o ajuste de um modelo contemplando todas essas covariáveis conjuntamente possibilite uma visão mais clara sobre estes relacionamentos.

4.2 Aplicação do modelo estatístico

Para determinar quais são os fatores de risco e proteção para a ocorrência de ovos do mosquito *Aedes aegypti* foi utilizado o modelo aditivo generalizado considerando-se variáveis ambientais e locais.

A distribuição de probabilidade assumida para a variável resposta foi a binomial negativa com função de ligação logarítmica, as funções suaves foram representadas usando *Thin Plate Splines*. Para a seleção das covariáveis componentes do modelo foi usado o método *stepwise*.

Como mostrou a análise exploratória para as covariáveis ambientais, não pode-se a priori impôr um relacionamento linear na escala da função de ligação, destas covariáveis com a resposta. Sendo assim, foram ajustados dois modelos, o primeiro supondo relacionamento linear destas covariáveis com a resposta, e outro atribuindo uma função suave. Após o ajuste algumas medidas para comparação dos modelos foram tomadas, Logaritmo da Verossimilhança e Critério de Informação de *Akaike*.

Para o modelo assumindo relacionamento linear o logaritmo da verossimilhança foi de $-16564,68$ e o AIC foi de $33208,41$ um total estimado de $39,52$ graus de liberdade. Para o modelo com termos suaves as medidas foram $-16546,6$ e $33196,42$ respectivamente com $51,60$ graus de liberdade. Como mostram as medidas do ajuste a diferença entre os dois modelos foi muito pequena, sendo assim, optou-se por adotar o modelo com termos lineares por sua simplicidade e interpretações diretas.

O resultado do ajuste do Modelo Aditivo Generalizado, para o bairro Brasília Teimosa é apresentado na tabela 2.

Na tabela 2 PREC.MES1 e PREC.MES2 refere-se respectivamente a precipitação média no mês da observação e no mês anterior a observação. A entrada UMID.MES3 refere-se a umidade média para dois meses antes da observação.

De acordo com o modelo ajustado, tem-se que a covariável precipitação (1 e 2) é considerado um fator de risco, já que, ao se aumentar a precipitação no mês da observação em uma unidade espera-se em média um aumento nas contagens de ovos de $2,83\%$ ($1,83\% - 3,85\%$), o mesmo aumento se espera para a precipitação no mês

Tabela - 2: Ajuste do Modelo Aditivo Generalizado para o bairro Brasília Teimosa.

Covariáveis	Estimativas	Erro Padrão	Valor t	p-valor
Intercepto	5,125	0,908	5,642	< 0,001
PREC.MES1	0,028	0,005	4,697	< 0,001
PREC.MES2	0,028	0,005	4,71	< 0,001
UMID.MES3	0,026	0,010	2,207	0,027
Canalizada	-0,242	0,068	-3,539	< 0,001
Grande sem tampa	-0,112	0,0532	-2,102	0,035
Suavização	Edf	Rank	F	p-valor
Coordenadas	24,805	24	10,38	< 0,001
Tempo	8,718	9	24,39	< 0,001

anterior ao da observação. Para a covariável umidade, também considerada como fator de risco, ao se aumentar uma unidade na umidade média mensal de dois meses antes da observação, espera-se um aumento médio de 2,63%(0,64% – 4,66%).

A preferência do mosquito por residências onde a canalização de água está presente em pelo menos um cômodo, fica evidente pela significância desta covariável no modelo. Pode-se observar, que onde a localização da armadilha apresenta canalização de água em pelo menos um cômodo, tem em média 27,37%(11,48% – 45,53%) mais ovos, quando comparada as localizações que não apresentam esta canalização.

Um resultado intrigante é o fato da presença de recipientes grandes sem tampa ser considerado fator de proteção. O modelo mostra que nas localizações de armadilhas onde se encontra este tipo de recipiente as contagens são em média 11,85%(0,81% – 24,09%) menores que nas localizações onde não se encontra. Tal resultado é corroborado, por exemplo, quando se tira uma média simples do número de ovos, em cada categoria desta covariável.

Para as armadilhas onde se encontram recipientes grandes sem tampa a contagem média de ovos é de 1595,26, já para onde não tem tais recipientes a contagem média é de 1639,92.

Este resultado sugere que pode existir um efeito de competição das armadilhas com estes recipientes, ou seja, na presença destes recipientes os mosquitos apresentam uma certa preferência por eles.

Os efeitos espaciais e temporais estão sendo controlados por funções suaves, que mostram-se significativos. Para verificar as localizações do espaço onde a contagem é mais elevada a figura 3 traz um mapa apenas do efeito espacial, para o bairro Brasília Teimosa, considerando o tempo um passo a frente da última observação, ou seja, é uma previsão um passo a frente apenas do efeito espacial.

Como este efeito não tem uma escala de valores de fácil interpretação, foi calculado um risco relativo, sob a hipótese de homogeneidade espacial. Isto, consiste em tirar a média de todas as predições, considerando que a área é homogênea, depois dividir cada uma das predições pela média geral, dando uma idéia de risco relativo a média geral, assim se uma observação está próxima da média ela apresentará neste

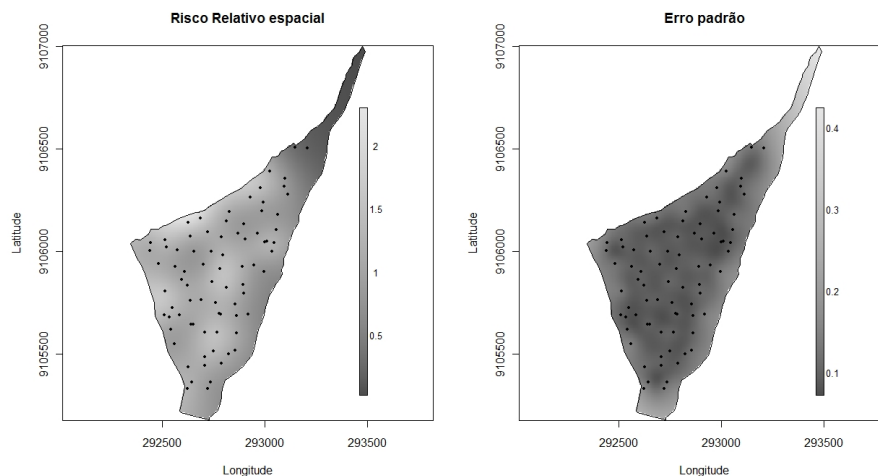


Figura - 3: Efeito espacial e incerteza associada.

mapa, um valor próximo da unidade. Se ela for maior que a média apresentará valores maior que a unidade, e se for menor que a média apresentará valores menores do que a unidade. Também traz um mapa da incerteza associada a cada predição, através de um mapa do erro padrão de cada medida predita.

De acordo com o mapa de risco relativo espacial, é possível identificar algumas áreas com risco relativo próximo de dois, mostrando que essas áreas tem o dobro de contagens do que a média geral da área, sendo, portanto áreas de alta infestação pelo mosquito, e que devem receber maior atenção pelas agências de saúde locais. Com relação ao mapa de erro padrão, o que fica evidente é que em localizações onde existem poucas observações, a incerteza aumenta consideravelmente, por exemplo, a borda superior do mapa de erro padrão, apresenta um erro padrão de aproximadamente 0,35, enquanto que onde foram feitas as observações o erro padrão varia entre 0,1 e 0,2.

Para explorar o efeito temporal, foi feita a previsão diária desde a primeira observação em 04/01/2005 até a última 15/05/2007, totalizando 862 dias. A predição é mostrada na figura 4, isolando o efeito temporal.

Pela figura 4 verifica-se quatro grandes picos, durante o período analisado. O primeiro ocorre entre os meses de março e abril de 2005, o segundo novamente em 2005 entre os meses de setembro e outubro. O terceiro ocorre entre abril e maio de 2006, e o último e maior pico ocorre entre os meses de novembro de 2006 e fevereiro de 2007. Os picos ocorrem em geral em meses mais quentes, aparentemente no início e término do verão. Os meses de frio aparentemente são os meses de menor ocorrência de ovos.

O conjunto de quatro gráficos apresentados na figura 5 apresenta diagnósticos usuais para resíduos do modelo ajustado. De forma geral os resultados são

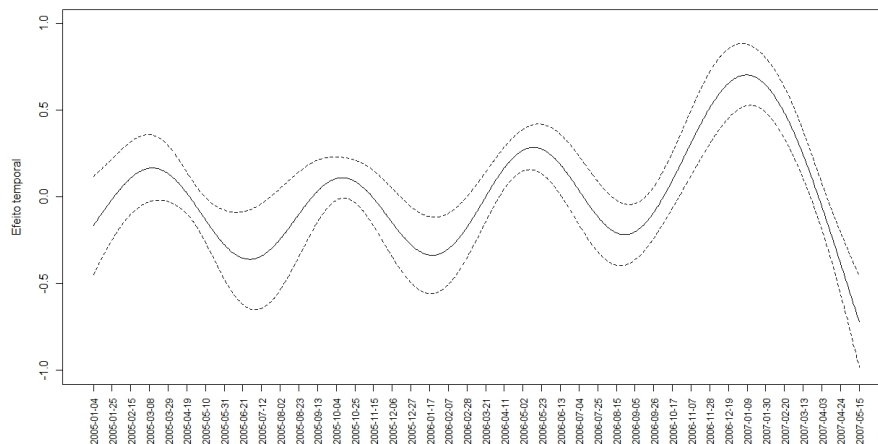


Figura - 4: Efeito temporal.

considerados satisfatórios sem apontar desvios ou problemas críticos no ajuste.

Discussão e recomendações

A determinação de fatores associados à ocorrência de ovos de *Aedes aegypti*, é um processo delicado, que envolve conhecimentos em modelagem estatística e também, aspectos ligados à biologia do indivíduo e da população do *Aedes aegypti*.

O uso de modelos estatísticos permite uma visão mais clara dos fatores relevantes em condições de campo por combinar as diversas covariáveis auxiliando na determinação precisa sobre as condições ambientais favoráveis ao aumento de contagens de ovos.

A análise da associação com fatores climáticos é uma tarefa difícil já que não se sabe a priori quais seriam as defasagens no tempo entre a data da coleta dos ovos e as condições climáticas que apresentam maior associação com tais contagens.

Com relação aos fatores climáticos os resultados mostraram que as condições de precipitação e umidade são importantes para a reprodução do mosquito, sendo que, este tem preferência por alta umidade e precipitação.

A avaliação das covariáveis ligadas às armadilhas é igualmente importante pois elas podem ajudar na elaboração de melhores programas de controle do vetor. Os resultados mostraram uma maior quantidade de ovos em residências, onde o abastecimento de água é diário, ou seja, que existe disponibilidade de água abundante sugerindo que tais locais devem ser receber atenção dos programas de controle.

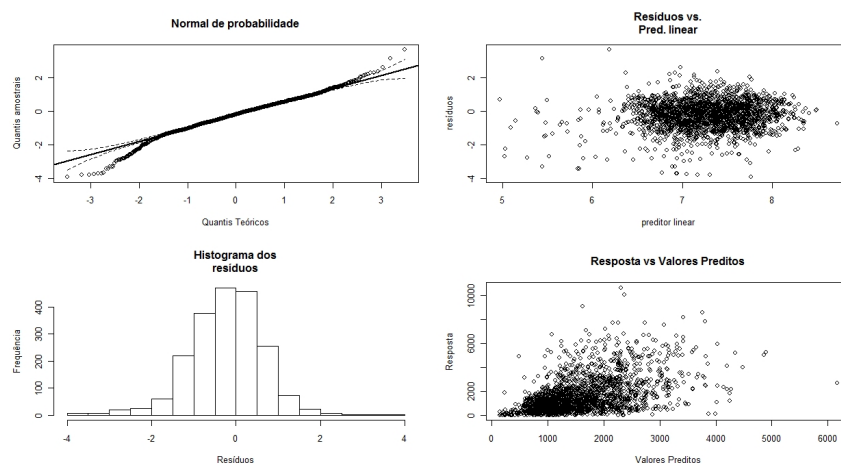


Figura - 5: Análise de resíduos.

Alguns resultados como a presença de recipientes grandes sem tampa, indicado como um fator protetor, não eram esperados. Este resultado pode indicar que existe algum tipo de competição entre estes recipientes e as ovitrampas, fazendo com que o mosquito prefira colocar os ovos em tanques, caixas d'água, ou toneis sem tampa, diminuindo assim a ocorrência de ovos nas ovitrampas.

A significância do efeito espacial é relevante, pois além de melhorar o ajuste do modelo estatístico permite verificar através de uma superfície onde estão as localizações de maior ocorrência de ovos. Isto é importante na prática, pois através de tais mapas as agência locais de saúde podem melhor direcionar suas ações de combate ao vetor.

O fato deste efeito ser significativo também indica que podem haver outras covariáveis não levadas em consideração neste trabalho que devem ser importantes para explicar a resposta. Tais covariáveis teriam padrão espacial de distribuição na área e o efeito destas covariáveis omissas são então refletidos pela significância do efeito espacial.

O ajuste do efeito temporal permite identificar os meses do ano de maior ocorrência de ovos, além de indicar possíveis efeitos sazonais e verificar a tendência de crescimento ou decrescimento da série. Os resultados mostram que durante o período em estudo, ocorreram quatro picos, que ocorrem em geral em meses quentes, aparentemente no início e término do verão, aparece claramente que os meses mais frios, são também os de menor ocorrência de mosquitos.

De forma geral pode-se concluir que o modelo escolhido apresenta um ajuste satisfatório, nenhum desvio severo dos pressupostos foi detectado o que valida as conclusões.

O fato dos dados serem provenientes de um experimento real, onde o nível de controle sobre os diversos fatores que podem afetar a variável resposta

é relativamente baixo. Sugere que este estudo deve ser encarado como um experimento *screening* ou exploratório, que busca evidências de quais seriam os possíveis fatores que estão influenciando na resposta.

Agradecimentos Os autores agradecem as equipes de campo e laboratório da Secretaria de Saúde do Recife/Programa de Saúde Ambiental e CPqAM/Fiocruz pelo intenso trabalho que produziu os dados aqui utilizados. Agradecemos também à toda equipe do projeto SAUDAVEL (<http://saudavel.dpi.inpe.br>) pela troca de informações e por disponibilizar os dados e recursos utilizados neste trabalho. Este trabalho foi realizado com o apoio e recursos do LEG/UFPR.

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 ovitrampas em Recife/PE. *Rev. Mat. Estat.*, São Paulo, v.xx, n.x, p.xx-xx, 2000. *Rev. Mat. Estat.* (São Paulo), v. 20, n.1, p. 1-10, 2000.

■ **ABSTRACT:** *Aedes aegypti* is the vector of dengue, an epidemic disease. Entomological studies aims to understand the proliferation dynamics of the mosquito population. The present work proposes a protocol for routine analysis combining statistical methods for the identification of factors related with *Aedes aegypti* egg counts. The methodology is developed based on data collected in a field experiment part of the SAUDAVEL project, combining egg count data with traps local conditions and environmental measurements taken at neighbourhoods in the city of Recife/PE. Main results show mosquito preference for residences with a regular water supply. Evidences on site competition between traps and large recipients were detected. Environmental covariates detect vector preference for elevated temperature and humidity. Spatial temporal aspected were taken into account in the modelling by smooth functions on the geographic coordinates of the traps and dates of data collection by the usage of generalised additive models and identifying hot spots for egg counts. The temporal term in the model have detected the months of higher abundance of eggs, sazonal effects and trends.

■ **KEYWORDS:** Entomological studies, dengue, *Aedes aegypti*, generalised additive models, space-time models.

Referências

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. ; DALLAZUANNA, H. S. ; RIBEIRO Jr, P. J. ; RÉGIS, L. ; MONTEIRO, A. M. V. ; SILVEIRA, J. C. *RDengue um ambiente para monitoramento de ovos do mosquito Aedes aegypti*. In: X Simpósio Brasileiro de Geoinformática, 2008.
- 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.
- FERREIRA, A. C. ; CHIARAVALLLOTI, F. *Infestação de área urbana por Aedes aegypti e relação com níveis socioeconômicos*. Revista de Saúde Pública, v.41(6), p.15-22, 2007.
- FOCKS, D. A. *A review of entomological sampling methods and indicators for dengue vectors*. Monografia na internet, 2000.
- FUNASA *Fundação Nacional de Saúde*. Dengue - instruções para o pessoal de combate ao vetor, manual de normas técnicas, 2001.
- GOMES, A. C. *Medidas dos níveis de infestação urbana para aedes (stegomyia) aegypti e aedes (stegomyia) albopictus em programa de vigilância entomológica*. Informe epidemiológico do SUS, v.11(2), p. 79-90, 2002.
- HASTIE, T. J. ; TIBISHIRANI, R. J. *Generalized additive models*. London: Chapman and Hall, 1990.
- MONTEIRO, A. M. ; CARVALHO, M. S. ; ASSUNÇÃO, R. W. ; RIBEIRO Jr, P. J. ; VIEIRA, W. ; DAVIS, C. ; REGIS, L. *SAUDAVEL: Bridging the Gap between Research and Service in Public Health Operational Programs by Multi - Institutional Networking Development and use of Spatial Information Technology Innovative Tools*. 2006.
- NELDER, J. A. ; WEDDERBURN, R. M. *Generalized linear models*. Journal of the Royal Statistical Association - Series A, 135, p. 370-384, 1972.
- NELSON, M. *The significance of indicators and indices*. Ctd/fil(den)/ic, 1995.
- OPS *Organizacion Panamericana de la Salud* Dengue y dengue hemorrágico en las Américas: guías para su prevención y control, 1995.
- PAULA, G. A. *Modelos de regressão com apoio computacional*. São Paulo: Editora da Universidade de São Paulo, 2004.
- R Development Core Team. 2009. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- 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.

SANTOS, R. S. *Fatores associados à ocorrência de formas imaturas de Aedes aegypti na Ilha do Governador, Rio de Janeiro, Brasil.* Revista da Sociedade Brasileira de medicina Tropical, v.32(4), p.373-382, 1999.

SERPA, L. L. N. ; COSTA, K. V. ; VOLTOLINI, J. C. ; KAKITANI, I. *Variação sazonal de Aedes aegypti e Aedes albopictus no município de Potim, São Paulo.* Revista de Saúde Pública, v.40(6) , p.1101-1105, 2007.

TAUIL, P. L. *Aspectos críticos do controle do dengue no Brasil.* Cadernos de Saúde Pública, v.18, p.867-871, 2002.

WAHBA, G. *Splines in nonparametric regression.* Madison, Department of Statistics: University of Wisconsin - Technical Report, 2000.

WOOD, S. N. *Generalized additive models: Introduction with R.* Boca Raton: Chapman and Hall, 2006.

WOOD, S. N. *GAMs with GCV smoothness estimation and GAMMs by REML/PQL.* R package version 1.3-31, 2008.

YANG, M. H. ; THOMÉ, R. C. *Controle ótimo do mosquito Aedes aegypti via técnica de insetos estéreis.* Congresso Nacional de Matemática Aplicada e Computacional, 2007.

Recebido em XX.XX.XXXX.

Aprovado após revisão em XX.XX.XXXX.