

# RDengue um ambiente para o monitoramento de ovos do mosquito *Aedes aegypti*

Wagner Hugo Bonat<sup>1</sup>, Henrique Silva Dallazuana<sup>1</sup>, Paulo Justiniano Ribeiro Junior<sup>1</sup>, Antônio Miguel Vieira Monteiro<sup>2</sup>

<sup>1</sup>Laboratório de Estatística e Geoinformação – Universidade Federal do Paraná (LEG-UFPR)  
CEP 81531-990 – Curitiba – PR – Brazil

<sup>2</sup>Instituto Nacional de Pesquisa Espacial (INPE)  
São José dos Campos, SP – Brasil

{wagner, henrique, paulojus}@leg.ufpr.br, miguel@dpi.inpe.br

**Abstract.** *This work describes the development of an integrated environment providing support for monitoring amounts of eggs of the *Aedes aegypti*, the main vector of the dengue disease. The proposed environment combines computational resources and tools for spatial and/or temporal statistical analysis. The system is used on eggs count data obtained from the an experiment conducted within the context of the SAUDAVEL project. The environment integrates the statistical software R and the GIS tools Terralib and TerraView, using the package aRT and includes the development of RDengue, an R package implemented the proposed methodology and tailored for continuous an automatic surveillance systems.*

**Resumo.** *Este trabalho descreve o desenvolvimento de um ambiente integrado de apoio ao monitoramento de ovos de *Aedes aegypti*, o principal vetor da dengue. O ambiente proposto combina recursos computacionais e ferramentas para análise estatística espacial e/ou temporal. O sistema é utilizado em dados de contagens de ovos obtidos em um experimento conduzido dentro do contexto do projeto SAUDAVEL. O ambiente integra o programa estatístico R e as ferramentas de SIG Terralib e TerraView através do pacote aRT e inclui o desenvolvimento de um pacote do R, RDengue, que implementa a metodologia proposta, direcionada para o monitoramento contínuo e automático em um sistema de vigilância entomológica.*

## 1. Introdução

A degradação do meio-ambiente e os problemas sócio-culturais afetam o cenário epidemiológico brasileiro, levando-o a ser destaque na mídia nacional e internacional, decorrente de epidemias de dengue, leptospirose, a recorrência da tuberculose, entre outras. Diante disso, constatou-se a importância de utilizar sistemas de vigilância, capazes de detectar precocemente situações que caracterizam surtos epidêmicos, modelar e identificar fatores de risco e proteção nas situações endêmicas e epidêmicas. A detecção precoce de surtos por doenças transmissíveis como a dengue é importante para ativar ações de investigação e controle por partes das agências de saúde pública, o que reforça a necessidade de sistemas de vigilância. Nesta perspectiva, foi elaborado o "Projeto SAUDAVEL"[Monteiro et al. 2006].

Segundo [Tauil 2002] a dengue é hoje a principal doença re-emergente no mundo. Na ausência de uma vacina preventiva eficaz, de tratamento etiológico e quimioprofilaxia efetivos, o único elo vulnerável para reduzir a sua transmissão é o mosquito *Aedes aegypti*, seu principal vetor.

Neste contexto, o objetivo deste trabalho é descrever o desenvolvimento de um ambiente R [R Development Core Team 2007] capaz de efetuar análises exploratórias, de forma automática e semi-automática. Tais análises, constituem um sistema de vigilância entomológica, dedicado ao monitoramento da atividade reprodutiva do mosquito *Aedes aegypti*. Através de contagens de ovos coletados com ovitrampas, em um experimento de campo conduzido na cidade de Recife/PE. Tal ambiente, acopla-se a um sistema de coleta de dados regular e continuada, e provê instrumentos para geração e disponibilização de resultados de análises dos dados em ambientes de fácil acesso para os profissionais envolvidos com o sistema.

Este sistema foi desenvolvido para os dados provenientes do experimento de coleta de ovos do mosquito *Aedes aegypti*, desenvolvido no escopo do projeto SAUDAVEL na cidade de Recife/PE. Aborda-se duas resoluções espaciais na análise: bairros e armadilhas.

Com o conjunto de rotinas propostas, é possível desenvolver análises espaciais e temporais, capazes de determinar limites de controle, verificar relações espaciais, identificar sazonalidades e tendências, estimar a ocorrência do fenômeno em áreas não observadas, definir áreas prioritárias para intervenção e também projetar possíveis cenários futuros, com o intuito de fornecer informações para subsidiar a definição de ações de controle do vetor da dengue.

## **2. Material e Métodos**

Nesta seção serão descritos a área de estudo e os aspectos computacionais envolvidos na análise.

### **2.1. Área de estudo**

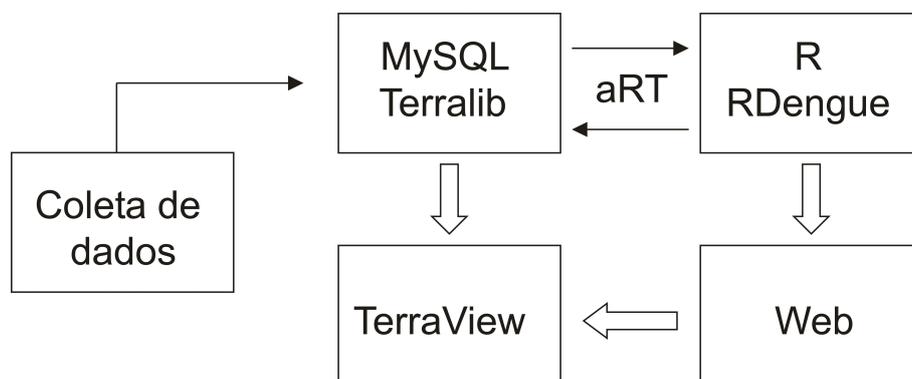
O experimento de Recife/PE consiste de 564 armadilhas, distribuídas entre 6 dos 94 bairros da cidade e instaladas de modo a cobrir toda a superfície do bairro. As armadilhas começaram a ser monitoradas a partir de 03/2004, até a data de 05/2007 foram realizadas 19.068 coletas, nas quais foram contados 14.829.557 ovos do mosquito. A cada sete dias era feita a contagem em aproximadamente um quarto das armadilhas, assim em um ciclo de vinte e oito dias todas as armadilhas são monitoradas.

Neste experimento, cada armadilha contém uma lâmina na qual a fêmea do mosquito coloca os ovos, essas lâminas são recolhidas e a contagem é feita em laboratório especializado. Os dados são inseridos em um banco de dados geográfico através de uma interface *Web*, desenvolvida pelo INPE - Instituto Nacional de Pesquisas Espaciais, esta interface foi projetada para evitar formas complexas de entrada de dados. A descrição detalhada do experimento dentro da proposta da rede SAUDAVEL é apresentada em [Regis et al. 2008].

### **2.2. Aspectos computacionais: RDengue**

Uma parte importante em projetos desta magnitude é a forma de armazenamento, tratamento e visualização dos dados. Um experimento como este gera uma grande quanti-

dade de dados que não são facilmente manipuláveis, requerendo para isto ferramentas específicas para validação e análise. O banco de dados do Recife SAUDAVEL, está implementado segundo o modelo Terralib [Câmara et al. 2000], tecnologia de código aberto e MySql Database Server como um repositório e sistema gerenciador de dados espaço-temporais baseado no modelo espaço-temporal da Terralib [Silveira et al. 2004].



**Figura 1. Formato geral de análise, as setas vazadas indicam visualização opcional, a setas sólidas indicam trocas persistentes.**

A Figura 1 ilustra o formato geral de análise, iniciando pela obtenção dos dados e armazenagem em um banco de dados geográficos. Isto feito, pode-se conectar o banco a partir do ambiente R, através do pacote aRT [Andrade Neto et al. 2005]. Todas as análises estatísticas são realizadas em um ambiente especializado em implementações de métodos estatísticos, no caso o R, através das rotinas especializadas implementadas no RDengue. Após as análises serem concluídas, tem-se através do RDengue e do aRT a opção de retornar resultados das análises para o banco geográfico. Opções de disponibilização dos resultados incluem tabelas e camadas de informação adicionais no banco, sendo ainda possível, definir no ambiente R, opções de visualização a partir do acesso via uma ferramenta de SIG, como é o caso do TerraView, ou então gerar uma página Web para a visualização pública dos resultados. Os resultados são associados às unidades espaciais do experimento, bairros ou armadilhas. Tal integração de ferramentas torna possível disponibilizar de forma regular e automática os resultados e relatórios das análises, de forma personalizada ao usuário da informação, tornando transparentes, se desejado o ambiente computacional e estatístico subjacente. Ressalta-se, que sem essa integração de ferramentas seria difícil e custoso, proceder análises como as que serão posteriormente apresentadas. Porém, a ferramenta RDengue provê opções de análise em ambiente R com opções de disponibilização de resultados para outras ferramentas de SIG como o TerraView. As rotinas de análises implementadas no RDengue são exclusivas do ambiente, usando rotinas residentes do R aplicadas e/ou adaptadas para o caso do experimento de Recife/PE.

### 3. Resultados e discussões

O monitoramento da quantidade de ovos de *Aedes aegypti* visa a detecção precoce de surtos de dengue. Para isto, é necessário que métodos estatísticos monitorem momentos de anormalidades no crescimento da população do mosquito, distinguindo entre número de ovos capturados dentro do esperado, e uma situação em que o número de ovos sugere uma

ação imediata, com vistas a evitar ou reduzir o impacto de um possível surto da doença. Para atender a estes propósitos, o RDengue utiliza ferramentas estatísticas e de SIG, com o intuito de desenvolver análises espaciais e temporais, de forma automática e/ou semi-automática, tendo em vista a estrutura de coleta regular de dados. As possibilidades de uso de tais rotinas são amplas, sendo que estas são flexíveis e podem ser analisadas de diversas formas, além disso são de fácil interpretação por não especialistas.

A seguir, serão descritas como exemplo algumas opções de análises contidas no ambiente desenvolvido. A primeira opção corresponde a uma análise descritiva por armadilha, através de gráficos de controle e ajuste de modelos para séries temporais da classe *arima*[Hyndman and Khandakar 2008]. A segunda consiste em uma abordagem de análise de superfície. O pacote implementa ainda outras estratégias mais avançadas de modelagem que se incorporam ao ambiente, opta-se por não detalhar tais métodos aqui. O RDengue pode ser expandido acoplando metodologias que vêm sendo estudadas e avaliadas.

### 3.1. Análises por armadilhas

A Figura 2 consiste na captura de uma tela de análise ilustrando a explícita integração entre o R e o TerraView através da Terralib e do SGBD. Após gráficos para cada uma das 564 armadilhas serem produzidos pelas rotinas do RDengue, os resultados são acoplados ao banco de dados do SAUDAVEL através de uma tabela de *médias*. Desta forma, uma ferramenta específica de SIG, no caso o TerraView, pode acessar os resultados da análise estatística produzida no R, com todas as facilidades e potencialidades que este ambiente tem para análise de dados. E a análise ganha o potencial de visualização própria de uma ferramenta de SIG como o Terraview. Além disso, as equipes de campo tem o acesso a tais análises muito facilitado, através do TerraView, o que não seria possível apenas através do R, pois demandaria maior familiaridade com a ferramenta estatística. O gráfico apresentado na Figura 2 utiliza as idéias das cartas de *Shewhart* [Montgomery 2000], para construir um gráfico de controle para a contagem de ovos do mosquito *Aedes aegypti*, possibilitando a comparação da situação de cada armadilha em relação as outras, ao bairro e também a cidade. O gráfico estabelece limites de controle baseados em medidas de dispersão, para o nível local (bairro) e também global (município), apresenta as médias do município, do bairro, a média mensal e também a média das armadilhas vizinhas, a fim de verificar algum tipo de efeito espacial. Neste gráfico as coletas são agrupadas por mês. É possível também através destes gráficos identificar tendências e sazonalidades. Este tipo de gráfico é uma ferramenta simples, que permite obter informações comparativas sobre a evolução do número de ovos em cada armadilha. O RDengue possui ferramentas capazes de desenvolver este tipo de gráfico para todas as armadilhas, de forma automática e facilmente atualizável conforme a disponibilidade de dados de novas coletas, formando assim o sistema de vigilância e monitoramento desejado.

Um outro tipo de análise, considera as observações de cada armadilha como uma série temporal, ajustando um modelo da classe *arima*. A rotina do RDengue procura o modelo que melhor se ajusta as observações segundo o critério de *Akaike* [Akaike 1974]. São apresentados dois gráficos clássicos na identificação da ordem do modelo: o de autocorrelação e autocorrelação parcial. Esta análise identifica sazonalidade e tendências, permitindo, através do modelo ajustado, projetar um cenário futuro para a armadilha, tal como, prever o número esperado de ovos para um ou dois meses a frente. Exemplos de tais análises, podem ser vistas no complemento online deste artigo.

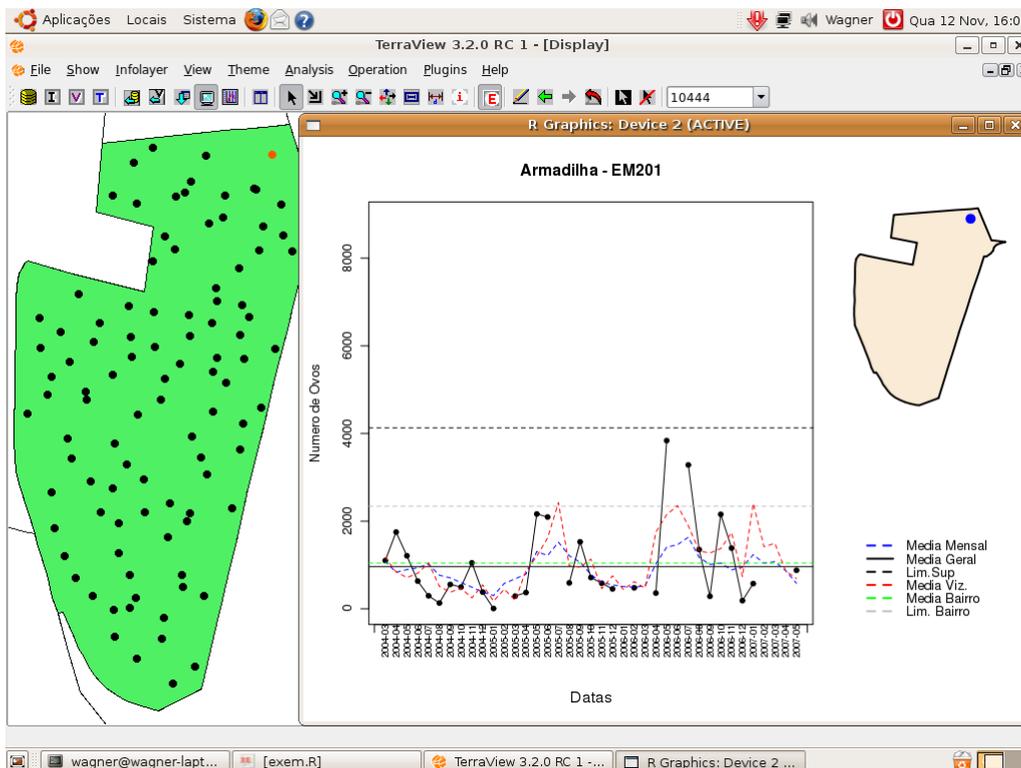
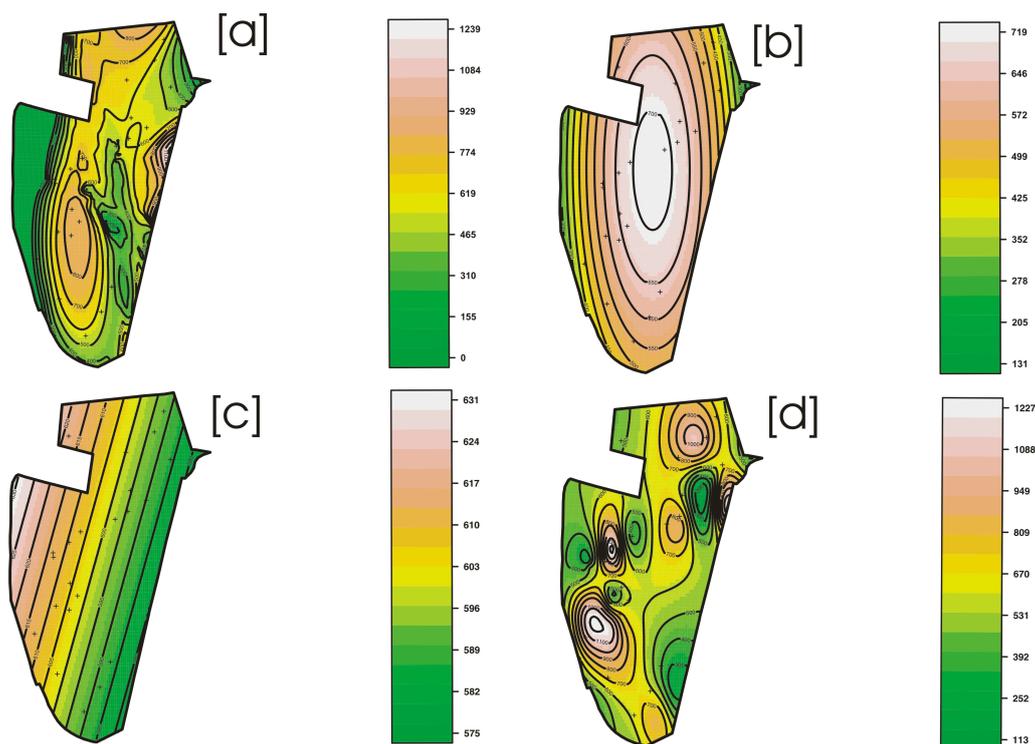


Figura 2. Análise descritiva por armadilha.

### 3.2. Análise por bairro

Um segundo tipo de análise, é um mapa da superfície de densidade de ovos, que permite a identificação das chamadas *zonas quentes*, buscando evidenciar regiões críticas do bairro, que devem receber prioridade de atendimento do setor de vigilância entomológica, considerando que maior número de ovos aumenta o risco da população desta área ser infectada. Vários métodos para estimar esta superfície estão implementados no RDengue, para citar alguns tem-se: modelos aditivos generalizados [Wood 2006], superfícies de tendência [Druck et al. 2004], regressão local [Cleveland 1978], e diferentes ajustes utilizando funções suavizadoras (*splines*) [Wahba 2000]. Busca-se com este tipo de análise, estimar uma superfície suave cobrindo todo o bairro. Conforme o método utilizado pode-se explorar diferentes pontos de análise. Por exemplo, utilizando a regressão local se evidencia os efeitos espaciais locais tais como efeitos de vizinhança entre armadilhas. Enquanto, que a superfície de tendência evidencia efeitos globais, como tendências direcionais. Cada método tem suas particularidades, e pode ser mais ou menos adequado de acordo com o objetivo da análise, e a configuração observada nos dados. Uma outra opção, é usar a transformação logarítmica, como os dados são originalmente uma contagem esta transformação é recomendada constantemente na literatura por estabilizar a variância e trazer gaussianidade, pressuposto de alguns dos métodos utilizados para estimar as superfícies. E ainda, outra opção seria gerar as superfícies por uma, duas ou todas as técnicas implementadas e compará-las, já que, este procedimento é facilitado pelas rotinas do RDengue. A Figura 3 mostra a superfície estimada por quatro métodos, para o bairro Engenho do Meio da cidade de Recife/PE. A superfície (a) traz o ajuste usando



**Figura 3. Exemplos de mapas de suavização.**

uma Regressão Local (loess). A superfície em (b) foi ajustada utilizando o método de superfície de tendência. A superfície em (c) utiliza um Modelo Aditivo Generalizado e a superfície em (d) foi ajustada usando funções splines multiníveis. A partir dos mapas de suavização é possível investigar de forma exploratória padrões espaço-temporais gerando animações. Para isto, utiliza-se uma interpolação *pixel a pixel*, o que resulta em um 'filme' que mostra a evolução do fenômeno tanto no espaço como no tempo. Os complementos online deste trabalho, mostram tal filme utilizando a escala de cores topográfica, em que valores altos da densidade de ovos são representados por cores claras sendo o valor mais alto pela cor branca, no outro extremo, valores baixos são representados por tons de verde. Considera-se, que este tipo de ferramenta exploratória é muito importante nas fases iniciais de modelagem estatística, onde se busca explorar padrões dos dados que possam sugerir estratégias adequadas de modelagem.

#### 4. Conclusões

Com a concepção de construção de ambientes para o monitoramento e vigilância entomológica, o RDengue soma-se aos objetivos do SAUDAVEL, contribuindo para o aumento da capacidade do setor de saúde no controle de doenças transmissíveis, desenvolvendo instrumentos para a prática da vigilância entomológica, incorporando aspectos ambientais, identificadores de risco e proteção, e métodos automáticos e semi-automáticos que permitem a detecção de surtos e seu acompanhamento no espaço e no tempo.

As rotinas implementadas no RDengue mostram-se satisfatórias, para gerar análises exploratórias, tendo inclusive um procedimento para a visualização espaço temporal, que evidencia picos e tendências espaciais mostrando a evolução do fenômeno em toda a área, dando a possibilidade de ver os dados, procedimento nada trivial devido ao tamanho do

experimento. Também possibilita, um monitoramento da intensidade de ovos do mosquito em toda a área com base em amostras pontuais do fenômeno, sendo que podem ser definidos sinais de aviso quando, por exemplo, a intensidade de ovos capturados for muito superior ao padrão considerado normal. Através dos modelos *arima* escolhidos de forma ótima para cada uma das 564 armadilhas, pode-se fazer a previsão para  $k$ -passos a frente. O intuito de uma análise exploratória é dar subsídios para a busca de modelos que se adequem aos dados, fornecendo a visualização de como o fenômeno se desenvolveu tanto no espaço como no tempo.

Desenvolvimentos futuros devem incorporar outras possíveis covariáveis aos modelos que estimam as superfícies tais como, temperatura, umidade e precipitação do ar. Métodos de previsão devem considerar extensões e alternativas aos modelos *arima*, gerando possíveis cenários futuros das armadilhas com maior acurácia. A avaliação de medidas de combate ao mosquito guiada por modelos estatísticos deve ser avaliada, e incorporar a capacidade de gerar cenários sob determinados tipos de intervenção.

As tecnologias utilizadas aqui e as rotinas do RDengue devem se acoplar aos ferramentais desenvolvidos pelo projeto SAUDAVEL e disponibilizadas a órgãos responsáveis pela vigilância entomológica, contribuindo desta forma para o setor de saúde na prevenção e detecção de surtos de dengue.

O pacote RDengue é disponibilizado em [www.leg.ufpr.br/RDengue](http://www.leg.ufpr.br/RDengue), junto a uma breve descrição de suas funcionalidades e suas dependências. Os códigos, resultados e telas de análises utilizados aqui são disponibilizados em [www.leg.ufpr.br/papercompanions](http://www.leg.ufpr.br/papercompanions) que corresponde ao complemento online deste artigo.

## Agradecimentos

Este trabalho é parcialmente financiado pelo projeto SAUDAVEL, CNPq 552044/2002-4 e desenvolvido no LEG/UFPR, em parceria com integrantes da rede SAUDAVEL da DPI/INPE, ENSP/Fiocruz, RJ e CPqAM/Fiocruz. Agradecemos a toda equipe da rede bem como aos técnicos de campo e laboratório da Secretaria de Saúde do Recife/CVA e também a população dos bairros de Engenho do Meio, Brasília Teimosa, Morro da Conceição, Alto José do Pinho, Sítio dos Pintos, Dois Irmãos, Casa Forte e Parnamirim pela colaboração.

## Referências

- Akaike, H. (1974). A new look statistical model identification. *IEEE Transactions on Automatic Control* AU-19.
- Andrade Neto, P. R., Ribeiro Jr., P. J., and Fook, N. D. (2005). Integration of Statistics and Geographic Information System: the R/Terralib case. *VII Simpósio Brasileiro de Geoinformática*.
- Cleveland, W. (1978). Robust locally weighted regression and smoothings scatterplots. *Journal of the American Statistical Association*.
- Câmara, G., Souza, R. C. M., Pedrosa, B. M., Vinhas, C., Monteiro, A. M. V., Paiva, J. A., Carvalho, M. T., and Gattass, M. (2000). Terralib: Technology in Support of GIS Innovation. *II Simpósio Brasileiro de Geoinformática*.
- Druck, S., Carvalho, M. S., Câmara, G., and Monteiro, A. M. (2004). *Análise Espacial de Dados Geográficos*. Embrapa, Brasília,DF.

- Hyndman, R. J. and Khandakar, Y. (2008). Automatic times series forecasting: The forecasting package for R. *Journal of Statistical Software*.
- Monteiro, A. M., Carvalho, M. S., Assunção, R., Vieira, W., Ribeiro Jr, P. J., Davis Jr, C., and Regis, L. (2006). SAUDAVEL: Bridging the Gap between Research and Services in Public Health Operational Programs by Multi - Institutional Networking Development and Use of Spatial Information Technology Innovative Tools.
- Montgomery, D. C. (2000). *Introduction to Statistical Quality Control*. John Wiley e Sons, 4th edition.
- R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Regis, L., Monteiro, A. M., Santos, M. A. M., Silveira, J. C., Furtado, A. F., Acioli, R. V., Santos, G. M., Nakazawa, M., Carvalho, M. S., Jr, P. J. R., and Souza, W. V. (2008). Developing new approaches for detecting and preventing Aedes aegypti population outbreaks: bases for surveillance, alert and control system. *Memórias do Instituto Oswaldo Cruz*.
- Silveira, J. C., Souza, W. V., Régis, L. N., Santos, M. A., Lapa, T. M., Portugal, J. L., Braga, T. S., and Monteiro, A. M. (2004). Recife em pedaços: Geotecnologias para a Detecção e Acompanhamento em Vigilância epidemiológica. *VI Congresso brasileiro de epidemiologia*.
- Tauil, P. L. (2002). Aspectos críticos do controle do dengue no brasil. *Cadernos de Saúde Pública [online]*.
- Wahba, G. (2000). *(Smoothing) Splines in Nonparametric Regression*. University of Wisconsin - Technical Report, Madison, Department of Statistics.
- Wood, S. N. (2006). *Generalized Additive Models: Introduction with R*. Chapman and Hall, Boca Raton.