

Métodos estatísticos aplicados em saúde pública

Wagner Hugo Bonat

LEG - Laboratório de Estatística e Geoinformação
Orientador: Ricardo S. Ehlers
Universidade Federal do Paraná

October 23, 2007

Introdução

- Degradação do meio ambiente e os problemas sócio-culturais afetam o cenário epidemiológico.
- Epidemias de dengue, leptospirose, recorrência de tuberculose entre outras.
- É de fundamental importância criar métodos capazes de detectar precocemente surtos epidêmicos.
- Modelar e identificar fatores de risco e proteção nestas situações.

Introdução

- Projeto SAUDAVEL (Sistema de Apoio Unificado para Detecção e Acompanhamento em Vigilância).
- Pretende contribuir para aumentar a capacidade do setor de saúde no controle de doenças transmissíveis.
- Criação de sistemas automatizados e atualizáveis de vigilância e monitoração.

Objetivos

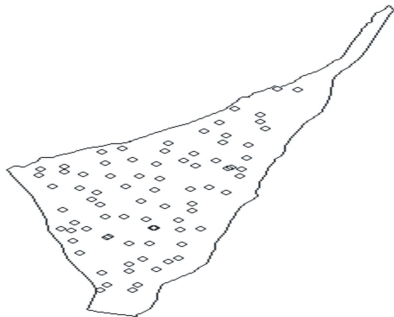
- Desenvolver um procedimento para a visualização espaço-temporal dos dados do experimento de coleta de ovos do mosquito *Aedes aegypti*.
- Demonstrar a utilização de softwares desenvolvidos pelos grupos integrantes do projeto SAUDAVEL.
- Ênfase ao **aRT**.

Área de estudo, Instrumentos e Técnicas de Campo

- O experimento foi desenvolvido em Recife/PE.
- Foram instaladas 564 armadilhas para o mosquito *Aedes Aegypti*.
- Nas quais foram realizadas 17.668 coletas.
- Contados ao todo 13.668.909.
- O período do experimento foi de 03/2004 a 12/2006.
- Foi realizado em 5 dos 94 bairros de Recife/PE.

Delineamento utilizado

No bairro Brasília Teimosa - Recife/PE foram instaladas 80 armadilhas.



Aspectos computacionais

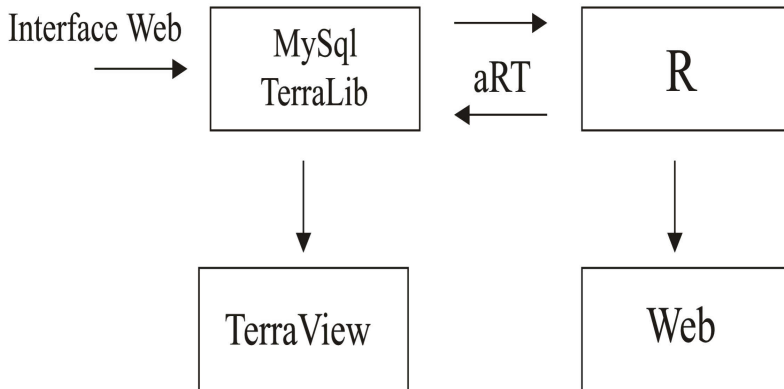


Figure: Formato geral de análise

Procedimento proposto

- 1 Ler o banco geográfico através do *aRT*.
- 2 Estimar uma superfície para cada semana do experimento através dos métodos descritos na seção 3.2.
- 3 Interpoliar 7 imagens entre cada superfície estimada.
- 4 Exportar as imagens geradas em formato **jpeg**, para um diretório do Linux.
- 5 Gerar um arquivo **.avi** para a visualização das imagens no formato de um filme mostrando a evolução dos ovos na área em estudo, através do software *mencoder* contido no *Mplayer* para Linux.
- 6 Disponibilizar o arquivo **.avi** em uma página na *Web* para visualização.

Metodologia Estatística

- Para utilizar os dados de forma eficiente, necessita-se de um procedimento de interpolação, para gerar uma superfície que represente o fenômeno em toda a área.
- Para este propósito segundo Druck *et al* 2004 pode-se considerar três tipos de abordagem.
 - 1 Modelos determinísticos de efeitos locais
 - 2 Modelos determinísticos de efeitos globais.
 - 3 Modelos estatísticos de efeitos globais e locais (Krigagem).
- Os dois primeiros serão tratados aqui, para maiores informações sobre o terceiro ver Digle and Ribeiro 2007.

Regressão Local (LOESS)

- Método não paramétrico que estima curvas e superfícies através de suavização (*smoothing*).
- As idéias básicas do modelo podem ser observadas considerando o mais simples dos modelos de regressão.

$$y_i = g(x_i) + e_i$$

- Regressão local estima "g" na vizinhança de cada ponto de interesse.
- Para utilização deste método é preciso fazer três escolhas:
 - 1 Parâmetro de suavização
 - 2 Grau do polinômio local
 - 3 Função de Kernel
- A suposição implícita é de que predominam as variações em pequena escala.

Superfície de Tendência

- A superfície é estimada por um ajuste polinomial aos dados, por um processo de regressão múltipla.
- A variável resposta é o número de ovos e as regressoras são as coordenadas geográficas das armadilhas.
- Exemplos incluem equações quadráticas do tipo:

$$w = \alpha_1 + \alpha_2x + \alpha_3y + \alpha_4x^2 + \alpha_5y^2$$

- Suposição implícita é de que predominam as variações em larga escala.

Resultados

- As superfícies foram geradas semanalmente.
- Essas foram animadas utilizando uma interpolação pixel a pixel.
- Entre 2 superfícies estimadas foram interpoladas 7.
- As superfícies foram suaves e aproximam o fenômeno em estudo.
- Reexpressão dos dados pelo seu logaritmo.

Conclusão

- A visualização de dados espaço-temporal é algo recente na estatística.
- Pouco se tem de ferramentas exploratórias para dados desta natureza.
- Para o caso do experimento em estudo o procedimento mostrou-se satisfatório.
- Permite evidenciar picos e tendências espaciais, mostrando a evolução do fenômeno em toda a área.
- Permite a visualização dos dados, tarefa nada fácil em experimentos desta magnitude.
- Verificar as potencialidades da integração entre os Sistemas de Informação Geográficas (SIG's) e ambientes estatísticos como o **R**.

Limitações e Recomendações

- A principal limitação do procedimento proposto é a escala dos dados.
- A geração de um única animação com escalas variadas é muito custosa.
- A presença de valores discrepantes tende a deixar a escala de visualização distorcida escondendo variações em pequenas escalas.
- Recomenda-se como futuras agendas de pesquisa para o projeto SAUDAVEL.
 - 1 Incorporar nas superfícies estimadas possíveis covariáveis como, condições climáticas e defasagens da variável resposta.
 - 2 Métodos mais flexíveis que não assumam formas lineares como os estimadores de superfície de tendência, e não sejam dependentes de um raio (*span*) como a Regressão Local.

Referências

- SUZANA DRUCK ; MARILIA SÁ CARVALHO ; GILBERTO CÂMARA ; ANTÔNIO MIGUEL VIEIRA MONTEIRO. *Análise Espacial de Dados Geográficos*. Embrapa, Brasília, DF, 2004.
- R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2007. ISBN 3-900051-07-0.