

## Técnicas de sensoriamento remoto aplicadas à endemia de Leishmaniose Visceral em Teresina.

Virginia Ragoni de Moraes Correia<sup>1,2</sup>  
Wagner de Souza Tassinari<sup>2</sup>  
Marília Sá Carvalho<sup>2</sup>  
Guilherme Loureiro Werneck<sup>3,4</sup>

<sup>1</sup>Instituto Nacional de Pesquisas Espaciais - INPE  
Caixa Postal 515 - 12201-970 - São José dos Campos - SP, Brasil  
virginia@dpi.inpe.br

<sup>2</sup>Departamento de Epidemiologia e Métodos Quantitativos em Saúde,  
Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz.  
Rua Leopoldo Bulhões 1480,  
Rio de Janeiro, RJ  
21041-210, Brasil  
wagnert@ensp.fiocruz.br  
carvalho@ensp.fiocruz.br

<sup>3</sup>Instituto de Medicina Social  
Universidade do Estado do Rio de Janeiro,  
Rua São Francisco Xavier 524, 7<sup>o</sup> andar, Bloco D,  
Rio de Janeiro, RJ  
20559-900, Brasil

<sup>4</sup>Núcleo de Estudos de Saúde Coletiva  
Universidade Federal do Rio de Janeiro.  
Av. Brigadeiro Trompowsky s/n,  
Rio de Janeiro, RJ  
21941-590, Brasil  
gwerneck@nesc.ufrj.br

**Abstract.** We have used two different supervised classifications – Maxver and Bhattacharya – to describe the land use pattern related to numbers of cases of visceral leishmaniasis, a vector born disease, endemic in Teresina/PI/Brazil. Results of both classification procedures were allocated to census tracts, area unit used to geographical reference for diseases cases. Generalized additive models, with spatially structured and non-structured random effects were used to estimate association between socioeconomic and remote sensing classes and cases of leishmaniasis.

**Key-words:** remote sensing, image processing, public health

### 1. Introdução

A ocorrência de doenças endêmicas transmitidas por vetores caracteriza-se por padrões espaço-temporais distintos relacionados a aspectos ambientais – chuvas, temperaturas e vegetação – mais ou menos favoráveis à proliferação das populações de reservatórios, vetores e agentes infecciosos (Correia et al, 2004). A oferta de sensores com diversas resoluções espectrais, espaciais e temporais, aliado aos sistemas de informação geográfica e técnicas de modelagem estatística, aumenta as possibilidades de uso do Sensoriamento Remoto (SR) no estudo de endemias.

A vegetação vem sendo a variável mais freqüentemente utilizada nos estudos que buscam relacionar ambiente e doenças endêmicas, principalmente devido a sua alta correlação com

precipitação, temperatura e umidade (Hay et al., 1996). Esta relação é mediada pelo perfil sócio-econômico local, que atua sobre a suscetibilidade e as possibilidades de contato da população com vetores e reservatórios (Werneck et al, 2002 ). O interesse recente da saúde pública nas possibilidades de uso do sensoriamento remoto (SR) para o estudo ambiental tem se concentrado nas áreas rurais (Correia et al, 2004). Entretanto, devido ao processo de urbanização de diversas endemias, como a leishmaniose cutânea e visceral, a dengue, a malária, a leptospirose, a esquistossomose, a doença de Chagas e a febre amarela, é oportuno explorar o uso potencial destas imagens em área urbana e peri-urbana.

Em aplicações do SR em endemias urbanas, o satélite Landsat é o mais indicado, pois oferece possibilidades de estudos retrospectivos do ambiente capacitando formular hipóteses sobre fatores de risco a endemia. (Correia et al, 2004)

A Leishmaniose Visceral, LV, conhecida pelo nome Calazar, é uma doença sistêmica grave, com letalidade que chega a atingir 10% (FUNASA, 2002). O parasita denominado *Leishmania chagasi*, é transmitido por um flebotomíneo do gênero *Lutzomyia* . Este vetor vive ao nível do solo, próximo a vegetação em raízes e/ou troncos de árvores, podendo ser encontrados em tocas de animais. Gostam de lugares com pouca luz, úmidos, sem vento e que tenham alimento por perto. De um modo geral, para seu desenvolvimento necessitam temperaturas entre 20 e 30°C, umidade superiores a 80% e matéria orgânica. Costumam picar a partir do por do sol até a madrugada. Os reservatórios mais importantes são o cão e a raposa, que agem como mantenedores do ciclo da doença. Os canídeos apresentam intenso parasitismo cutâneo, o que permite uma fácil infecção do inseto, e, por este fato, são os mais importantes elos na manutenção da cadeia de transmissão (SUCEN, 2003).

Os dados da Fundação Nacional de Saúde (Fundação Nacional da Saúde - FUNASA), mostram que a doença ocorre em todas as regiões brasileiras, exceção na região Sul, sendo que no Nordeste a situação é mais crítica, graças à expansão e urbanização da doença. A cidade de Teresina tem sido foco de algumas epidemias de LV, a primeira registrada em 1980 (Costa et al., 1990). Entre 1993 e 1995 foram contabilizados cerca de 1200 novos casos. Um dos motivos dessa endemia rural emergir em centro urbano é atribuída aos movimentos migratórios rurais-urbanos ocorridos nas últimas décadas. A maior parte dessa população que chega ocupa as áreas da periferia da cidade sendo exposta às áreas de florestas e ao ciclo silvestre da *Leishmania chagasi*, promovendo a transmissão do parasita (Werneck, 2000).

Na área da saúde a dimensão em estudo – casos de doença – é estimada através de contagens de casos por área, usualmente setor censitário, considerada uma unidade de tamanho adequado (RIPSA, 1999), abrangendo população em torno de 800 pessoas (Censo,2000) e sendo por isso a unidade de análise de escolha.

O objetivo desse trabalho é estudar a distribuição espacial da Leishmaniose Visceral (LV) em Teresina (Piauí) em função de diversos fatores de risco, ambientais e sócio-econômicos, utilizando ferramentas de SR, SIG e técnicas de modelagem estatística. Os classificadores supervisionados, Maxver e Bhattacharya, forneceram padrões de uso do solo que são testados em conjunto com variáveis sócio-econômicas como indicadores de situações que potencializam a ocorrência da endemia.

## **2. Material e método**

### **2.1. Área de estudo**

Teresina localiza-se entre rios Parnaíba e Poti, centralizada a 5°05' de latitude sul e na longitude oeste de 42°48'. A região de estudo abrange a área que contém a malha de setores censitários de 1991. A cidade tem um clima tropical com uma temperatura média anual de 27°C e uma precipitação pluviométrica média anual de 1300mm<sup>3</sup>. A estação chuvosa ocorre

de Janeiro a Abril, e as maiores temperaturas vão de Agosto a Dezembro. A vegetação predominante consiste de árvores esparsas e arbustos, existindo na periferia da cidade áreas de pastagem e floresta tropical. O uso do solo na região é predominantemente urbano residencial, comercial e industrial.

## 2.2. Dados do estudo -

Para o estudo utilizou-se uma imagem Landsat 5, passagem em 17/08/1990, e uma Landsat 7 de 06/07/2001, órbita 219, ponto 64, projeção UTM. As bandas utilizadas foram as 3, 4 e 5, de acordo com a maior parte dos estudos urbanos, por serem as bandas com maiores diferenças espectrais (Zha et al, 2003). O período escolhido foi o de seca pela facilidade de se obter imagens sem cobertura de nuvens, o que muitas vezes impossibilita o uso das imagens óticas. Procurou-se, dentro do possível, unificar os períodos de aquisição, para que as diferenças de iluminação solar e de condições atmosféricas fossem mínimas.

A malha vetorial utilizada corresponde aos setores censitários de 1991 da cidade de Teresina (Piauí). Os dados da doença, relativos aos anos de 1993 a 1996, foram cedidos pela Fundação Nacional de Saúde (FUNASA). As variáveis sócio-econômicas utilizadas: proporção de alfabetizados, de domicílios com recolhimento de lixo, e domicílios com abastecimento de água, foram obtidas diretamente do Censo 1991 e da contagem de 1996. As variáveis que indicam os padrões de uso do solo foram geradas pelas classificações e agregadas em algumas subclasses significativas para a endemia.

O aplicativo *SPRING* (Sistema de Processamento de Informações Georeferenciadas) foi utilizado no processamento e análise das imagens (Camara et al, 1996), e o *BayesX* para a regressão Bayesiana, (Brezger et al., 2003)

## 2.3. Modelo estatístico

Para mapear a variação geográfica de doenças em pequenas áreas, quando o número de eventos esperados é pequeno e a variabilidade do estimador muito grande, deve-se utilizar modelos que permitam suavizar a flutuação utilizando a estrutura local de vizinhança, através do processo conhecido por "*borrowing strength*". Os modelos que estimam, simultaneamente, o efeito de covariáveis fixas, da superdispersão relacionada a determinantes desconhecidos e permite suavizar a flutuação aleatória relacionada a pequenas populações são os denominados modelos lineares generalizados mistos. Estes são modelos de regressão que incorporam este conjunto de aspectos e permitem introduzir diferentes estruturas de dependência. A estimação de efeitos espaciais, mais recente, tornou-se possível com a inclusão de efeitos aleatórios estimados não-parametricamente, o que permite generalizar o tipo de modelo para Modelos Aditivos Generalizados Mistos (GAMM) (Lang & Fahrmeir, 2001) (Clayton & Kaldor, 1987).

O método de estimação para estes modelos complexos baseia-se em inferência Bayesiana, utilizando técnicas de simulação Monte Carlo via Cadeias de Markov (MCMC) (Gilks et al, 1996; Gelman et al, 1995). A abordagem Bayesiana tem sido utilizada devido o desempenho eficiente do modelo e à flexibilidade na incorporação de estruturas aleatórias e de dependência.

Nesse trabalho, investiga-se a associação entre os casos de LV e as variáveis explicativas, sócio-econômicas e padrões de uso do solo urbano. A variável resposta é uma contagem (de casos), tendo como offset a população de cada setor  $i$ , portanto, uma distribuição de Poisson com valor esperado  $\lambda_i$ . O modelo adotado é um GAMM que pode ser denominado geoaditivos quando se introduz os efeitos espaciais. Segue a equação do modelo para a aplicação:

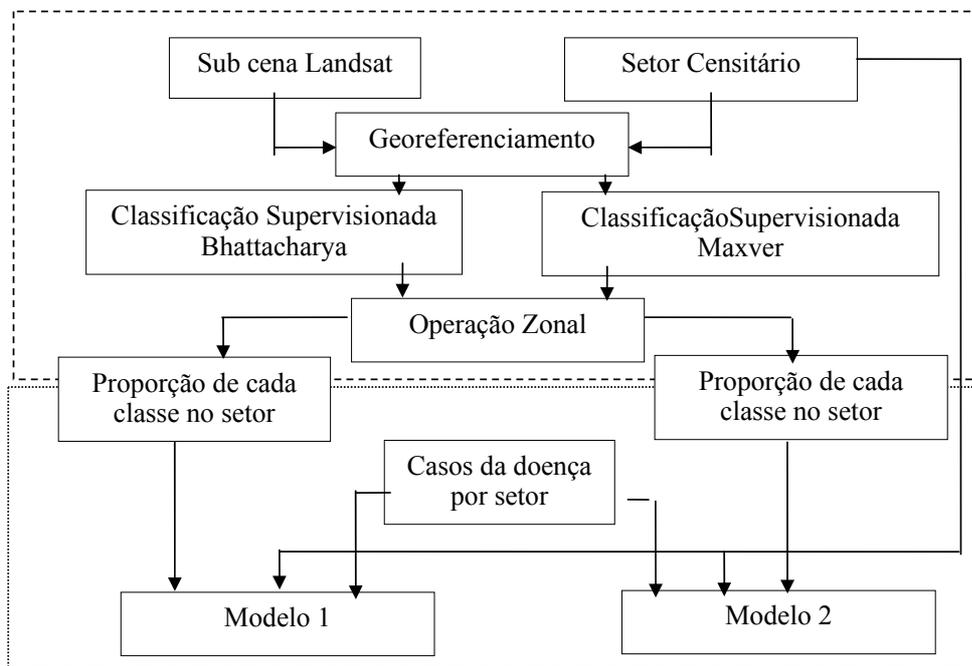
$\log(\lambda_i) = f_{\text{espaço}}^{\text{estrut}}(s_i) + f_{\text{espaço}}^{\text{Nestrut}}(s_i) + \delta_{\text{effixos}} + \text{offset}(\text{pop})$  onde  $f_{\text{espaço}}$  é efeito aleatório espacial (estruturado e não estruturado), (Lang & Fahrmeir, 2001).

Na inferência bayesiana, as funções desconhecidas  $f$  e os parâmetros de efeitos fixos  $\delta$  são considerados variáveis aleatórias com distribuição *a priori* apropriadas. Para os efeitos espaciais estruturados, a priori é uma generalização de um passeio aleatório de primeira ordem para duas dimensões e é denominado um campo aleatório Markoviano. Para os efeitos não estruturados pode-se pensar em efeitos Gaussianos. (Fahrmeier et al, 2003)

## 2.5. Fluxograma da metodologia

O fluxograma abaixo, **Figura 1**, ilustra a metodologia utilizada. Anteriormente à classificação por região, a imagem é segmentada. Para o treinamento da classificação supervisionada utiliza-se a informação de algumas amostras cuja verdade terrestre é conhecida.

**Figura 1** – Fluxograma da metodologia utilizando SPRING (-----) e BayesX(.....)



A imagem de 1990 foi utilizada para a classificação. A de 2001, única ETM-7, com a banda pancromática de 15 metros passou pela operação de restauração e fusão com a finalidade de se obter uma imagem colorida com 5m de resolução, possibilitando o registro do setor censitário. Apesar da imagem ser de 2001 e os setores de 1991, alguns cruzamentos de ruas puderam servir de referência para os ajustes. A importância de sobrepor o setor censitário de 1991 à imagem, é permitir o cálculo da proporção de classes de uso em cada setor.

A partir da cena original completa Landsat5 de 1990, uma subcena cobrindo a cidade de Teresina foi recortada. Pela necessidade de sobrepor a outras cenas e camadas de informação foi necessário georeferenciá-la. O mesmo processo foi utilizado para a Landsat7. O registro da primeira foi manual, utilizando-se uma imagem de referência fornecida pela NASA. Para o registro da Landsat7 foi utilizado o processo automático, usando a Landsat5

como referência. O erro do registro foi próximo de meio pixel, o esperado para uma área urbana.

Para a segmentação utilizou-se 12 para o parâmetro de similaridade e 50 para o de área.

No treinamento dos classificadores supervisionados foram utilizadas amostras localizadas na imagem por especialista e rotuladas como agricultura, residencial/densa, residencial/muito/verde, residencial/verde, vegetação densa, solo, entre outras.

Um programa utilizando a linguagem LEGAL (Câmara et al, 1994) foi desenvolvido para calcular a proporção de cada classe nos setores censitários, para cada um dos classificadores.

Os dados atualizados bem como a malha de setores são exportados do programa SPRING e importados no aplicativo BayesX onde é executada a regressão bayesiana. Na **Figura 1**, modelo 1 significa modelo final que utiliza variáveis obtidas pelo classificador Maxver, e o modelo 2, se refere ao resultado para o Bhattacharya.

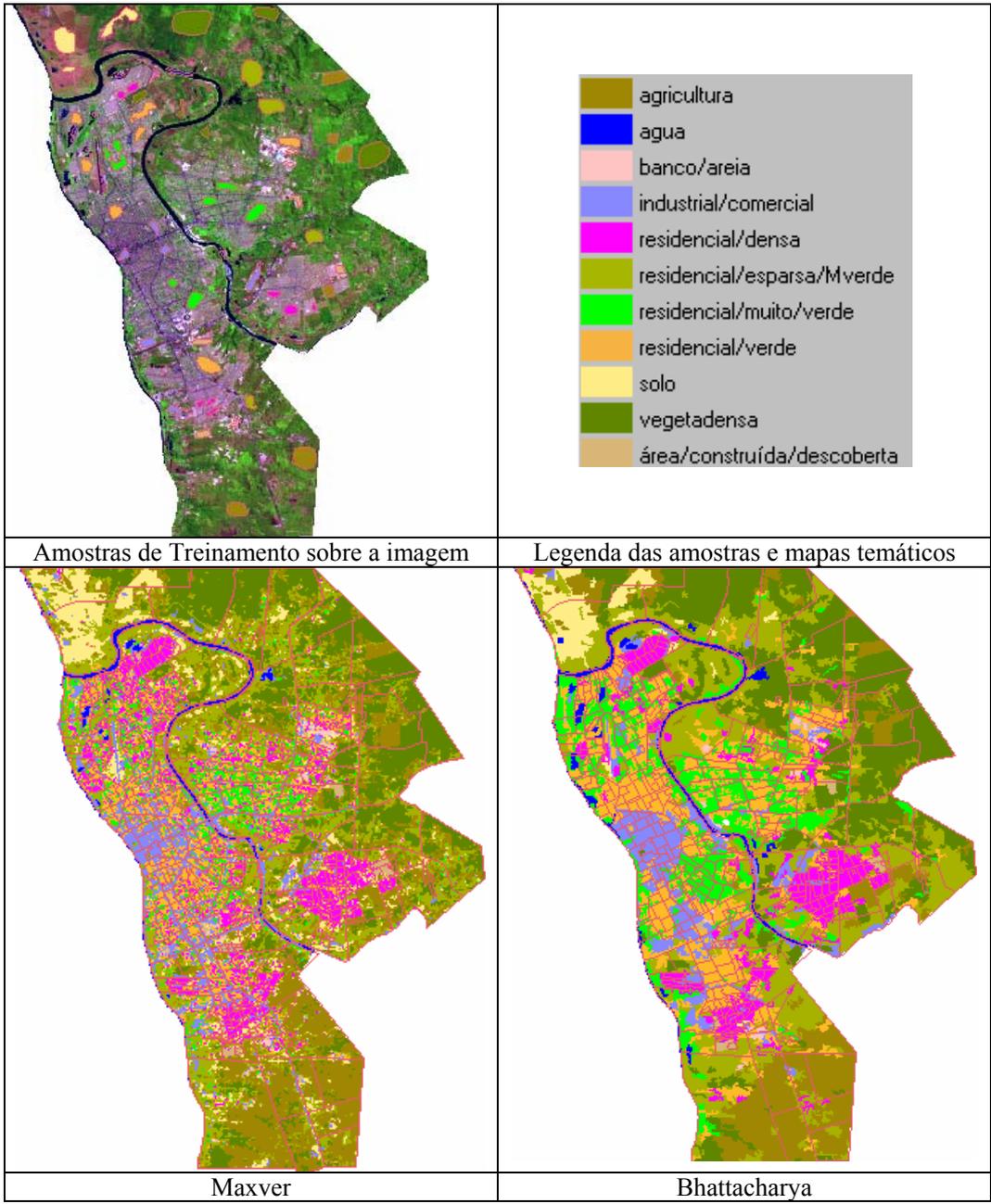
3. Resultados As amostras de treinamento e a legenda e o resultado das classificações Maxver e Bhattacharya, com os setores censitários superpostos, são apresentados na **Figura 2**. Nos setores situados na fronteira urbano-rural, mais homogêneos, o cálculo da proporção de cada classe no setor foi semelhante para os dois classificadores. Nos setores centrais o resultado não foi bom. Para cada classificador, existe uma confusão entre as classes residencial verde, residencial com muito verde, residencial densa e área industrial-comercial. Esse achado já era esperado após classificação das amostras na fase de treinamento e também pela dificuldade de classificar áreas urbanas com o sensor TM. Além disso, as duas classificações foram treinadas apenas com as amostras fornecidas. Observa-se também que, em alguns setores o resultado é diferente para cada classificador o que indicou a necessidade de agrupar algumas classes mantendo o significado epidemiológico e minimizar o efeito da confusão ocorrida na classificação.

No modelo, apresentado na **Figura 3**, as classes industrial-comercial e urbana densa, vegetação densa e residencial espaço muito verde, foram agrupadas. Observa-se que no modelo 1, utilizando as variáveis obtidas com o classificador Bhattacharya somente a classe agricultura não foi significativa enquanto no modelo 2, além da área de agricultura, as classes: proporção de vegetação densa, e urbana densa, agregada com industrial-comercial, também não foram significativas. Nos dois modelos as variáveis sócio-econômicas proporção de água e lixo não foram significativas.

Observou-se que o efeito de introduzir os efeitos espaciais estruturados foi pequeno mas relevante. Uma avaliação desse modelo indica que, apesar dos resultados semelhantes nos parâmetros do modelo de regressão, o classificador Maxver (DIC=-672,438) teve um desempenho melhor na explicação da distribuição espacial da LV em Teresina (**Figura 3**). Pode-se pensar entretanto que o classificador Bhattacharya passou pela operação de segmentação que já envolvia uma certa correlação espacial.

Introduzir variáveis ambientais aos modelos epidemiológicos é fundamental dada a importância das condições ambientais ao desenvolvimento dos parasitas e vetores das endemias. Que variáveis introduzir no modelo depende de acompanhar o comportamento das endemias no tempo e no espaço associando os casos ocorridos às alterações ambientais. Na Leishmaniose Visceral em Teresina é interessante um estudo retrospectivo, mapeando essas alterações. A temperatura e a chuva, por exemplo, que nesse estudo não seriam variáveis informativas, podem ser importantes num estudo longitudinal. Esse trabalho teve como motivação contribuir com o resgate do uso do Sensoriamento Remoto em endemias apresentando uma ligação entre técnicas utilizadas em imagens, os recursos dos sistemas de informação geográfica e os modelos estatísticos bayesianos.

**Figura 2** – Amostras de treinamento com legenda e classificações



**Figura 3:** Tabela dos modelos Bayesianos

Variáveis	Modelo Bhattacharya			Modelo Maxver		
	Média	5% quant.	95% quant.	Média	5% quant.	95% quant.
Const	<b>-2.841</b>	<b>-4.216</b>	<b>-1.468</b>	<b>-1.849</b>	<b>-2.981</b>	<b>-0.712</b>
PAGUA	-0.233	-0.695	0.236	-0.270	-0.749	0.216
PLIXO	0.302	-0.203	0.771	0.209	-0.337	0.743
PALF	<b>-2.218</b>	<b>-3.488</b>	<b>-0.881</b>	<b>-2.343</b>	<b>-3.720</b>	<b>-1.050</b>
TXCRESC	<b>0.616</b>	<b>0.062</b>	<b>1.153</b>	<b>0.571</b>	<b>0.051</b>	<b>1.094</b>
DENSPOP	<b>-0.045</b>	<b>-0.061</b>	<b>-0.029</b>	<b>-0.050</b>	<b>-0.069</b>	<b>-0.030</b>
PVD2	<b>1.508</b>	<b>0.475</b>	<b>2.680</b>	0.492	-0.412	1.418
PIC2	<b>1.340</b>	<b>0.367</b>	<b>2.385</b>	0.622	-0.119	1.377
PRV	<b>1.539</b>	<b>0.544</b>	<b>25.898</b>	<b>1.178</b>	<b>0.354</b>	<b>2.012</b>
PRMV	<b>1.659</b>	<b>0.603</b>	<b>274.581</b>	0.910	-0.388	2.329
PAG	0.270	-1.988	2.320	0.632	-0.406	1.686
Ef. Esp. ñ estrut.	0.137	0.024	0.262	0.132	0.005	0.282
Ef. Esp. estrut.	0.752	0.372	1.235	0.798	0.366	1.343
DIC	-669.462			-672.438		

## Referências

### Artigo em Revista:

- Brezger, A., Kneib, T. & Lang, S. (2003) BayesX - Software for Bayesian Inference based on Markov Chain Monte Carlo simulation Techniques. Available under <http://www.stat.uni-muenchen.de/~lang/Bayesx>.
- Camara, G.; Souza, R.C.M.; Freitas, U.M.; Garrido, J.; Spring. Integrating remote sensing and GIS by object oriented data modeling. *Computers & Graphics*, v.20,n.3, p.347-471, 1996.
- Clayton, D.E, Kaldor, J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, v 43, p.671-681, 1987.
- Correia, V. M. C; Carvalho, M. S.; Sabroza, P. C.; Vasconcelos, C.; Remote sensing as a tool to survey endemic diseases in Brazil. *Cadernos de Saúde Pública*, v. 20, no.4, p.891-904. 2004.
- Costa, C.H.N.; Prira, H.F.; Araújo, M.V. epidemia d lishmaniose visceral no stado do Piauí, Brasil, 1980-1986. *Revista de saúde Pública*, v. 24, p. 361-372. 1990
- Fahrmeier, L; Lang, S.; Bayesian Semiparametric Regression Analysis of Multicategorical Time-Space Data. *Annals of the Institute of Statistical Mathematics*, v.53, no.1.p.10-30, 2001
- Guindon, B. ;Edmonds, C. M. Large-area land cover mapping through scene-based classification composition. *Photogrammetric Engineering & Remote Sensing*. v. 68, n. 6, p.589-596, 1992.
- Hay, S.I; Tucker, C.J.; Rogers, D.J; Packer, M.J;. Remotely sensed surrogates of meteorological data for the study of the distribution and abundance of arthropod vectors of disease. *Annals of Tropical Medicine and Parasitology*. v. 90, no 1, p.1-19,1996.
- Werneck G.L; Rodrigues,L.;Santos M.V; Araújo, I.B; Moura,L.S.; Lima, S.S; Gomes, R.B; Maguire, J.H, Costa, C.H. The burden of *Leishmania chagasi* infection during an urban outbreak of visceral leishmaniasis in Brazil. *Acta Tropica* , v.83, no.1, p.13-18, 2002.
- Zha, Y; Gao, J.; Ni, S. Use of normalized difference build-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, v. 24. , no.17, p.583-594, 2003.
- SUCEN, 2000. Superintendência de Controle de endemias. Leishmaniose Visceral Americana: informações para profissionais de saúde - acesso 24-10-2203 as 16:00 < <http://www.sucen.sp.gov.br/doencas/index.htm>>.
- RIPSA. Comitê Temático Interdisciplinar: Dados Espaciais em Saúde - relatório., OPAS/MS: Brasília. p. 10, 1999
- FUNASA. Guia de Vigilância Epidemiológica. Fundação Nacional da Saúde, 5ª edição. Brasília, 2002

IBGE (Instituto Brasileiro de Geografia e Estatística), Censo demográfico de-1991.

IBGE (Instituto Brasileiro de Geografia e Estatística), Contagem populacional de-1996.

**Livro:**

Gelman, A, Carlin, J.B, Stern, H.S, Rubin, D.B., **Bayesian data analysis**. London: Chapman & Hall, 1995, 526p..

Gilks, W. R., Richardson, S., Spiegelhalter, D.J. **Markov Chain Monte Carlo in practice**. London: Chapman&Hall, 1996, 486p.

**Tese:**

Werneck, G.L., **Spatial approaches to study the epidemiology of visceral leishmaniasis in Teresina, Brasil**. Doutorado em Saúde Pública. Harvard University, HARVARD, Cambridge, Estados Unidos, 2000.

**Eventos:**

Camara, G.; Freitas, U.M.; Cordeiro, J.P; Towards an Algebra of Geographical Fields. In: VII Simpósio Brasileiro de Computação Gráfica Processamento de Imagens. **Anais...**, Curitiba, 1994, p. 205-212.