

GEOMEDICINA - Delimitação de Áreas de Risco à Saúde, Relações entre Incidência de Doenças e Fatores Ambientais

Humberto C. Ibanez¹, Hélio Pedrini¹, Paulo J. Ribeiro Jr.², Bonald C. Figueiredo³

¹Departamento de Informática – Universidade Federal do Paraná (UFPR)
Curitiba, PR – Brasil

²Departamento de Estatística – Universidade Federal do Paraná (UFPR)
Curitiba, PR – Brasil

³Instituto de Ensino Superior Pequeno Príncipe (IESPP)
Curitiba, PR – Brasil

humberto.ibanez@gmail.com, helio@inf.ufpr.br, paulojus@ufpr.br,
bonaldf@yahoo.com.br

Abstract. *This paper describes the development of a statistical system for visualization and navigation of maps in Internet, named Geomedicina. Spatial distribution of available chemical elements, collected from 736 locations of hydrographic basins in Parana State, as well the disease incidence rate of interest by government, are characterized in the system. In addition to such characterizations, spatial associations are investigated and described also considering the etiology between the diseases and the chemical elements, allowing the determination of risk areas to human health and aiding the decision making process by personnel who are responsible for public health care in Parana. Several open source components are employed in the construction of the system, reducing costs of software acquisition and allowing more flexibility in software component selection. The system integrates DBSM PostgreSQL and its extensions PostGIS and PL/R; the R language for statistical computing; UMN-MapServer; and HTTP Apache server. Scripts developed in PHP, Javascript and XML allow such integration and generate a Web user interface known as AJAX (Assynchronous Javascript and XML). In the proposed approach, the communication between the client and the server reduces the network traffic by only transferring essential data to the interaction.*

Resumo. *Este trabalho descreve o desenvolvimento de um sistema estatístico de visualização e navegação de mapas na Internet, denominado Geomedicina. Neste sistema são caracterizadas as distribuições espaciais de elementos químicos biodisponíveis, coletados em 736 pontos de bacias hidrográficas do Estado do Paraná, e da incidência de doenças de interesse nos municípios. Além dessas caracterizações, são investigadas e descritas associações espaciais, considerando também a etiologia, entre as doenças e os elementos químicos, através das quais são delimitadas as áreas de risco à saúde*

humana, auxiliando o processo de tomada de decisão dos responsáveis pela manutenção da saúde pública no Estado. No desenvolvimento do sistema serão empregados componentes de software de código aberto, o que consiste em economia na aquisição e flexibilidade na escolha destes componentes. O sistema é construído com a integração dos componentes SGBD PostgreSQL e suas extensões PostGIS e PL/R; linguagem R de computação estatística; servidor de mapas UMN-MapServer; e servidor HTTP Apache. Scripts desenvolvidos em PHP, JavaScript e XML fazem esta integração e produzem uma interface Web com o usuário, conhecida como AJAX (Asynchronous JavaScript and XML). Na abordagem proposta, a comunicação do cliente com o servidor reduz o tráfego na rede pela transferência de apenas dados essenciais para a interação.

1. Introdução

O uso da Internet como meio para acessar sistemas georreferenciados provém uma maior disponibilidade da informação. O acesso à informação através de um computador, em qualquer local onde haja uma conexão com a rede, sem a necessidade de instalação de softwares aplicativos, é um aspecto facilitador para o usuário. Com o uso de um simples navegador, consegue-se acessar e interagir com o sistema.

Ao visualizar e navegar em mapas temáticos o usuário tem a indicação dos locais ou regiões onde, para o tema em questão, há alguma deficiência. Esta indicação, visualizada em padrões de destaque, auxilia-o no processo de tomada de decisão.

O objetivo geral do Geomedicina é auxiliar o processo de tomada de decisão dos responsáveis pela manutenção da saúde pública, através do acesso facilitado à informação, apresentada numa interface Web, com a visualização de áreas de risco à saúde, delimitadas com o uso de tecnologias estatísticas. Para atingir este objetivo geral, são organizados alguns objetivos específicos, descritos a seguir:

- i) Caracterizar a distribuição espacial de mortalidade para doenças da faixa pediátrica e adolescente, principalmente neoplasias, por município no Estado do Paraná;
- ii) Caracterizar a distribuição espacial de elementos químicos de interesse, coletados em 736 pontos nas bacias hidrográficas do Estado do Paraná;
- iii) Identificar as correlações entre duas variáveis (química e/ou epidemiológica) através de metodologia e tecnologia estatística; e
- iv) Delimitar as áreas de risco à saúde humana.

Pretende-se também que o Geomedicina integre conhecimentos sobre a exposição aos produtos de origem natural (geológica) e manipulados pelo homem como mercúrio, chumbo, manganês, bário, cádmio, resíduos agrotóxicos, benzeno, tolueno e outros produtos tóxicos. Doenças de evolução aguda, e principalmente crônica, como câncer, doenças de pele, neurológicas, entre outras, poderão ser mais bem avaliadas num sistema integrado com um diagnóstico mais preciso do ambiente químico. O sistema terá também condições de apresentar um perfil sobre: a qualidade dos

mananciais; as variações naturais no quimismo da rede hidrográfica superficial e; os focos de poluição industrial, agrícola e doméstica, em ambientes rurais e urbanos.

2. Metodologias e Tecnologias Estatísticas

Metodologias e tecnologias estatísticas são usadas para identificar associações entre os dados do Sistema de Informações de Saúde e dados geoquímicos, numa área de estudo recente, conhecida como Geologia Médica [1].

Os dados geoquímicos são obtidos pela Minerais do Paraná (MINEROPAR) e consistem numa amostragem planejada de modo a proporcionar uma completa cobertura do território paranaense, que compreende cerca de 200.000 km², correspondendo a 83% do Estado do Paraná. As amostras são coletadas nos canais dos rios, no local de maior fluxo d'água, onde os sedimentos e a água sofrem constante homogeneização. Para a representação das superfícies geoquímicas complexas, que definem as tendências geoquímicas dos elementos aqui considerados, é necessário realizar a regularização das malhas de amostragem. Essa regularização diz respeito ao espaçamento e a distribuição espacial dos dados originais. O ponto de amostragem foi deslocado de sua posição real para o baricentro da microbacia. Esse procedimento é adotado considerando que o ponto de coleta da amostra não representa apenas um ponto mais sim uma área (a microbacia) e que os pontos originais estão concentrados ao longo das linhas das drenagens principais [2].

Os dados do Sistema de Informações de Saúde são obtidos através da Secretaria de Saúde do Estado do Paraná, Seção de Vigilância Epidemiológica. Estes dados correspondem ao número de mortalidade e morbidade por doença, ocorrida nos municípios do Estado. Dados censitários do IBGE são usados para se obter a taxa de incidência destas doenças, sobre as quais se executa um tratamento estatístico descrito nas próximas seções.

2.1. Incidência de Doenças nos Municípios

A produção de mapas temáticos requer uma padronização das contagens para que se possam comparar as diferentes populações no espaço. O estimador taxa bruta apresenta instabilidade para expressar o risco de determinada doença quando ela é rara ou quando a população de risco é pequena. Outros estimadores incluem regressão de Poisson, estimativas de riscos relativos e mapas de probabilidade. Porém estes ainda não consideram a dependência espacial do fenômeno. Para se considerar a dependência espacial das incidências de doenças nos municípios pode-se usar medidas de associação espaciais tais como o I de Moran ou G de Geary, ou métodos de suavização tais como os estimadores de Bayes empírico global ou local, que leva em consideração, além da informação da área do município, a de sua vizinhança [3].

2.2. Elementos Químicos Presentes na Água

Sobre os dados dos atributos químicos referenciados nos baricentros das microbacias são aplicadas técnicas geoestatísticas [4][5] a fim de estimar e descrever o comportamento dos atributos em todo o estado. Com o uso de tais técnicas é possível caracterizar os padrões de dependência espacial e, a partir destes, obter predições de interesse na região ou subregiões, tais como mapas de teores do atributo, zonas críticas, probabilidade de exceder valores críticos, entre outros.

2.3. Associações entre Doenças e Elementos Químicos

As informações e predições sobre atributos químicos no Estado, quanto integradas sob a área de cada município, fornecem dados a serem relacionadas com os números de casos de doenças. Neste contexto, modelos de regressão, lineares ou lineares generalizados [6], cujas respostas são variáveis de doenças, investigam a existência de associações com concentrações dos elementos químicos. Tais modelos podem ser expandidos para incluir associações espaciais e incorporar a incerteza associada aos erros de predição dos atributos químicos. Nesta investigação, a etiologia conhecida da doença é também considerada.

2.4. Computação Estatística

Os métodos estatísticos descritos acima são aplicados usando-se a linguagem de computação estatística R [7], que implementa uma diversidade de métodos estatísticos e tem se tornado o ambiente onde frequentemente novas metodologias são disponibilizadas. A linguagem é interpretada, embora diversos procedimentos utilizam código compilado para eficiência no processamento. O sistema é de código aberto disponível na internet e distribuído também em formato compilado para várias plataformas. Um sistema de pacotes adicionais permite a incorporação rápida de novas metodologias, bem como diversas opções para interfaces com outras linguagens e aplicativos. Pela flexibilidade, número de usuários e qualidade do projeto o R tem se tornado a "língua franca" para computação e disponibilização de metodologias estatísticas. Além de rotinas convencionais do sistema R, pacotes adicionais a este sistema tais como spdep [8] e geoR [9] são utilizados para procedimentos de estatística espacial descritos nas seções 2.1, 2.2 e 2.3.

3. Componentes Empregados na Construção do Sistema

O sistema, formado por componentes de software de código aberto, é instalado sobre a plataforma Linux Debian [10]. Todos os pacotes são obtidos gratuitamente pela Internet, sendo a sua maioria instalada a partir do repositório da versão estável Debian. Atualmente, apenas os pacotes Mapserver e PostGIS necessitam ser configurados apropriadamente, compilados e instalados. O conjunto dos pacotes instalados compõem um servidor de mapas. Além da linguagem de computação estatística R, citada na seção 2.4, nas próximas seções são descritos os principais pacotes que compõem o sistema.

3.1. PostgreSQL, PostGIS e PL/R

O PostgreSQL [11] é um SGBD objeto-relacional de código aberto. Uma característica importante deste SGBD para o projeto Geomedicina é a escalabilidade. O PostgreSQL pode ser estendido de várias maneiras. É possível adicionar ao PostgreSQL tipos de dados, funções, operadores, funções de agregação, métodos de índice e linguagens procedurais.

Linguagens procedurais permitem adicionar funções ao SGBD, que podem ser escritas em várias linguagens de programação, além do SQL e C, como, por exemplo: PL/PgSQL, Perl, Python, Ruby e PL/R. Uma vantagem de embutir o processamento no SGBD é a redução do tráfego na rede, entre o cliente e o servidor. Uma desvantagem desta abordagem é a dependência que se cria com o SGBD, com perda de portabilidade do sistema.

O PostGIS [12] é uma extensão do PostgreSQL que adiciona suporte a objetos geográficos ao PostgreSQL. Este módulo do PostgreSQL implementa diversas funcionalidades topológicas, possibilitando o desenvolvimento de um SIG.

O PL/R [13] é uma linguagem de procedimentos carregável que habilita a escrita de funções e triggers PostgreSQL na linguagem R. Com este recurso pode-se programar os tratamentos estatísticos dentro do SGBD.

3.2. UMN-MapServer e Mapfile

O UMN-Mapserver [14] é um ambiente de desenvolvimento para a construção de aplicações espaciais para Internet. O UMN-Mapserver produz mapas, imagens e dados vetoriais e os disponibiliza para o servidor HTTP. Suas características mais importantes para este projeto são: a automação dos elementos de mapas, como barra de escala, mapa de referência, e legenda; a produção de mapas temáticos com o uso de expressões lógicas ou expressões regulares baseadas em classes; a projeção de mapas dinamicamente. Este componente é usado no Geomedicina para desenhar os mapas em formato vetoriais, armazenados no PostGIS. As consultas SQL são feitas através do arquivo de configuração do UMN-Mapserver, o Mapfile.

O Mapfile contém a estrutura principal da aplicação. Nele são declarados o formato dos dados de entrada e os formatos das imagens produzidas: o estilo de apresentação das camadas e suas respectivas classes; o tipo de imagem (jpg, png, imagemap); o sistema de projeção; o mapa de referência; a legenda; a barra de escala; e a apresentação das consultas espaciais.

3.3. PHP e PhpMapscript

O PHP [15] é uma linguagem de script própria para desenvolvimento de interfaces Web. Através do PHP é possível fazer com que a interface com o usuário seja dinâmica. O PHP também provém acesso à base de dados PostgreSQL. Esta característica é usada para o acesso e a recuperação de dados em bases de dados não georreferenciadas.

Assim, é possível fazer uma ligação entre os dados georreferenciados e os não georreferenciados. Como exemplo, neste sistema, os dados fornecidos pela Vigilância Epidemiológica não são georreferenciados. Para ligá-los ao sistema georreferenciado, desenvolve-se um script PHP que, através de comparações sobre algum identificador do município, constante na base georreferenciada, possibilita a representação deste atributo no mapa.

Uma extensão do PHP, o PHPMapscript [16], permite que, de forma dinâmica, obedecendo a seleções feitas pelo usuário na Interface Web, seja modificado o arquivo de configuração Mapfile. Com este recurso pode-se, por exemplo, modificar o esquema de cores do mapa temático, aplicar um outro método estatístico aos dados atuais, ler um novo conjunto de dados. Este recurso aumenta a interatividade com o sistema, que adquire um comportamento dinâmico.

3.4. Servidor HTTP Apache

O sistema baseia-se na arquitetura cliente-servidor, onde o cliente é o navegador (*browser*) do usuário e o servidor HTTP é o Apache [17]. O cliente envia para o servidor requisições de conteúdos, onde estão embutidos parâmetros, como: nome do arquivo de configuração Mapfile, tamanho da imagem de saída, entre outros. O servidor trata esta requisição e retorna ao cliente o conteúdo solicitado.

4. AJAX

AJAX [18] de Assynchronous JavaScript and XML, é um conceito de desenvolvimento de conteúdo para Internet onde economiza-se o tráfego na rede, pois, com este conceito, não é necessário que a atualização de determinadas informações implique na modificação da página inteira. O AJAX proporciona um mecanismo para separar o dado da aplicação, fazendo com que somente o dado trafegue pela rede e somente quando for necessário para a aplicação.

O JavaScript monitora os eventos usados para a interação do usuário com o sistema. Assim, um evento como clicar e arrastar o mouse para deslocar o mapa é tratado pela aplicação. Nesta ação, através de controles pode-se evitar uma requisição ao servidor, recuperando-se as partes do mapa que não apareciam na tela do próprio cache do cliente.

4.2. Protótipo do Sistema

No protótipo do projeto Geomedicina empregou-se a tecnologia AJAX [19]. Neste protótipo foi usado o framework Ka-Map [20]. Neste framework scripts JavaScript e PHP controlam a interação do usuário com o servidor de mapas, com as seguintes características: deslocamento do mapa interativo e contínuo, sem o recarregamento da página; ampliação em escalas pré-definidas e; suporte para barra de escala, legenda e mapa de referência. A interface com o usuário deste protótipo é apresentada na figura 1.

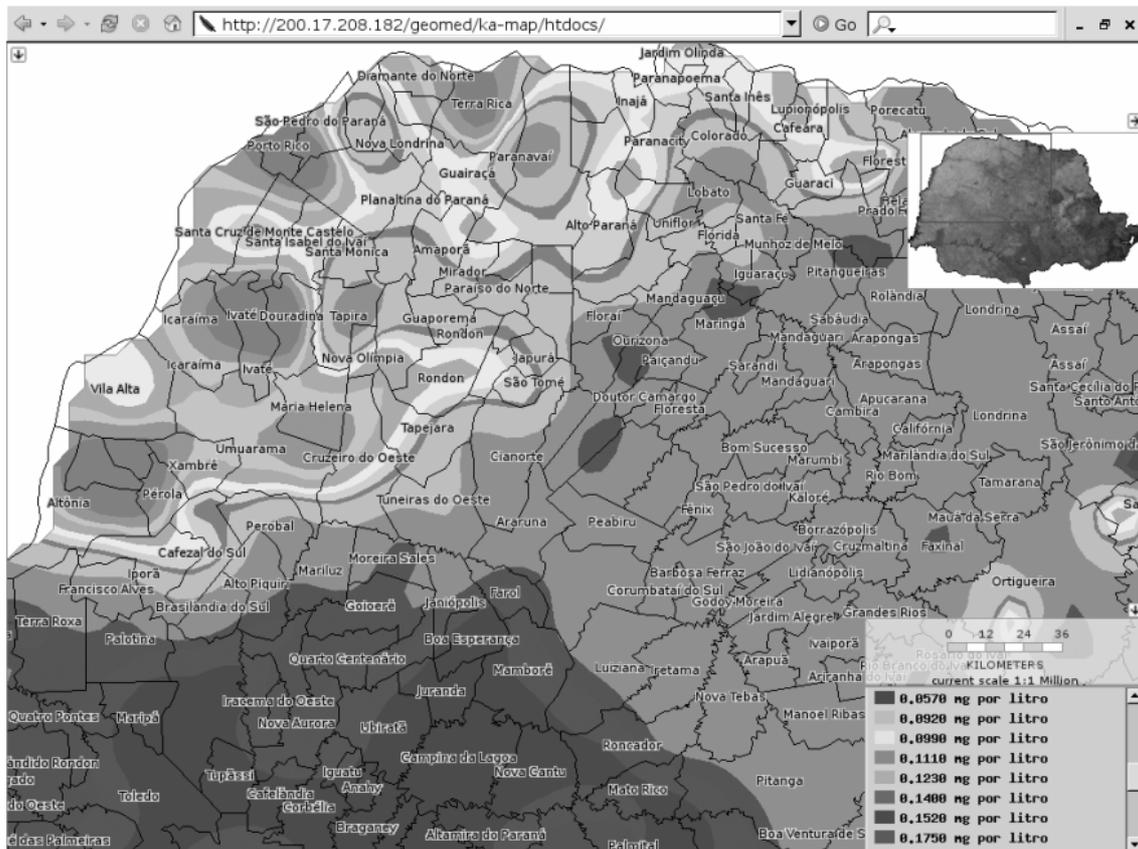


Figura 1. Interface com o Usuário do Protótipo Geomedicina

5. Conclusões

Este trabalho descreveu o desenvolvimento de um sistema estatístico de visualização e navegação de mapas na Internet, denominado Geomedicina. Distribuições e associações espaciais de elementos químicos biodisponíveis são caracterizadas, levando-se em conta a etiologia entre as doenças e o relacionamento destas com os elementos químicos, através das quais são delimitadas as áreas de risco à saúde humana.

O sistema visa auxiliar o processo de tomada de decisão dos responsáveis pela manutenção da saúde pública no Estado do Paraná. Vários componentes de software de código aberto foram empregados no desenvolvimento do sistema, com o objetivo de reduzir o custo de aquisição e manutenção de ferramentas e melhorar a comunicação do cliente com o servidor.

Referências

- [1] Medical Geology (2006), **The science dealing with the relationship between natural geological factors and health in man and animals**, URL <http://www.medicalgeology.org>.

- [2] Minerais do Paraná (2001), **Atlas Geoquímico do Estado do Paraná**, Curitiba.
- [3] Bailey, T. C. e Gatrell, A. C. (1996), **Interactive Spatial Data Analysis**, Longman, Harlow.
- [4] Diggle, P. J. e Ribeiro Jr, P.J. (2006), **Model Based Geostatistics**. Springer, New York (in press).
- [5] Chilès, J. P. e Delfiner, P. (1999), **Geostatistics: Modeling Spatial Uncertainty**, Wiley, New York.
- [6] McCullagh, P. e Nelder, J. A. (1989), **Generalised Linear Models**, Chapman and Hall, London, second edition.
- [7] R Development Core Team (2006). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- [8] Ribeiro Jr, P. e Diggle, P. **geoR: a package for geostatistical analysis** R-NEWS, 1 (2):15-18. June, 2001
- [9] Bivand, R., Anselin, L., Berke, O., Bernat, A., Carvalho, M., Chun, Y, Dormann C., Dray, S., Halbersma, R., Lewin-Koh, N., Ono, H., Peres-Neto, P., Tiefelsdorf M., Yu, D., **spdep: Spatial dependence: weighting schemes, statistics and models**. R package version 0.3-22.
- [10] Debian GNU Linux. **A free operating system (OS) for your computer**. URL <http://www.debian.org>
- [11] PostgreSQL. **A powerful, open source relational database system**. URL <http://www.postgresql.org>
- [12] PostGIS. **An extension that adds support for geographic objects to the PostgreSQL**. Refrations Research, Victoria, British Columbia, Canada, URL <http://www.postgis.org>
- [13] PL/R. **A loadable procedural language that enables you to write PostgreSQL functions and triggers in the R programming language**. Joseph E. Conway, URL <http://www.joeconway.com/plr>
- [14] UMN-Mapserver. **An Open Source development environment for building spatially-enabled Internet applications**. URL <http://ms.gis.umn.edu>
- [15] PHP. **PHP: Hypertext Preprocessor" a scripting language that is especially suited for Web development and can be embedded into HTML**. URL <http://www.php.net>
- [16] PHPMapscript. **A PHP dynamically loadable module that makes MapServer's MapScript functions and classes available in a PHP environment**. DM Solutions Group Inc, Ottawa. URL http://www.maptools.org/php_mapscript
- [17] Apache. **The Apache HTTP Server Project**. Apache Software Foundation. URL <http://httpd.apache.org>

- [18] Soares, W. (2006), **AJAX (Asynchronous JavaScript and XML) Guia Prático Para Windows**. Editora Érica, São Paulo.
- [19] **Build AJAX-Based Web Maps Using ka-Map**. O'Reilly XML.COM. URL <http://www.xml.com/pub/a/2005/08/10/ka-map.html>
- [20] Ka-Map. **An framework that is aimed at providing a javascript API for developing highly interactive web-mapping interfaces**. DM Solutions Group Inc, Ottawa. URL <http://ka-map.maptools.org>