

UNIVERSIDADE ESTADUAL PAULISTA

“Júlio de Mesquita Filho”

Pós-Graduação em Ciência da Computação

Carlos Henrique El Hetti Laurenti

**Predição de dados espaciais e temporais:
uma ferramenta *on-line* baseada em
recursos georreferenciados**

UNESP

2012

Carlos Henrique El Hetti Laurenti

Predição de dados espaciais e temporais: uma ferramenta *on-line* baseada em recursos georreferenciados

Orientador: Carlos Roberto Valêncio

Dissertação apresentada para obtenção do título de Mestre em Ciência da Computação, Área de Concentração em Sistemas de Computação, junto ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Carlos Henrique El Hetti Laurenti

Predição de dados espaciais e temporais: uma ferramenta *on-line* baseada em recursos georreferenciados

Dissertação apresentada para obtenção do título de Mestre em Ciência da Computação, Área de Concentração em Sistemas de Computação, junto ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

BANCA EXAMINADORA

Prof. Dr. Carlos Roberto Valêncio (Orientador)

UNESP – São José do Rio Preto

Prof. Dr. José Márcio Machado
UNESP – São José do Rio Preto

Prof. Dr. Pedro Luiz Pizzigatti Corrêa
USP – São Paulo

São José do Rio Preto, 13 de julho de 2012.

“A mente que se abre a uma nova ideia jamais voltará ao seu tamanho original.”

Albert Einstein

Dedico este trabalho a minha família, professores e amigos.

Agradecimentos

Ao meu orientador, Prof. Dr. Carlos Roberto Valêncio, pela orientação, sugestões, ensinamentos e, sobretudo, pela oportunidade e compreensão das dificuldades.

Ao professor Prof. Dr. José Márcio Machado por ter contribuído imensamente com sua experiência em todas as bancas.

Ao programa de pós-graduação em Ciência da Computação, em especial, aos professores e funcionários.

Aos amigos do GBD, especialmente ao Toni, que sempre me ajudou em tudo que precisei.

Aos meus amigos Tico e Ichiba, muito obrigado pela verdadeira amizade, pela ajuda nos estudos, nas discussões do trabalho e pelo companheirismo.

Agradeço imensamente a toda minha família, especialmente aos meus pais e meus irmãos que sempre me deram todo apoio e motivação.

Aos outros amigos, que mesmo distante ou perto, contribuíram para esta conquista.

Aos professores presentes nesta banca, que farão orientações para a melhora do meu trabalho.

E à Deus que sempre esteve presente na minha vida.

E, finalmente a Juliana, minha namorada, pelos momentos em que me ajudou a corrigir minha dissertação, pela amizade, pela paciência, pelos carinhos e pelo amor.

RESUMO

A união entre informações e sua posição geográfica, durante um período de tempo, é importante e necessário quando se deseja obter resultados específicos e determinísticos de uma região observada e, também, para prever estatisticamente acontecimentos no futuro. Essas informações possibilitam uma análise para tomada de decisões em metodologias de prevenção que são aplicadas em diversas áreas como em meio ambiente, sustentabilidade, segurança no trabalho, dentre outras que visam melhorar a qualidade de vida. Com o avanço de novas tecnologias e diversos outros recursos *open-source*, torna-se possível construir uma aplicação para análise e predição capaz de trabalhar com esses dados de maneira eficiente e satisfatória. Com o objetivo de integrar conceitos estatísticos espaço-temporais e conceitos computacionais este trabalho propõe uma metodologia de análise futurística de dados georreferenciados, contemplado por uma aplicação *web* desenvolvida. De maneira dinâmica qualquer base georreferenciada pode ser analisada, a metodologia permite realizar cálculos de predição em regiões selecionadas, obter gráficos estatísticos espaço-temporais e visualizar uma região predita.

Abstract

The link between the data and its position is necessary when you want specific information about a given region, and also to predict the future events statistically. This information enables an analysis for decision making in prevention methodologies that are applied in several fields such as environment, sustainability, safety, and others to improve the quality of life. With the advancement of new technologies and several other open-source resources, it becomes possible to build an application for analysis and prediction witch able to work with this data in an efficient and satisfactory way. In order to integrate spatio-temporal statistical concepts and computational concepts this paper proposes a methodology for future analysis of georeferenced data, covered by a web developed application. Dynamically any georeferenced database can be analyzed, the methodology allows performing calculations of prediction in selected regions, constructing statistical graphs and visualizing spatio-temporal region predicted.

Sumário

LISTA DE FIGURAS	XI
LISTA DE TABELAS	XIII
LISTA DE SIGLAS	XIV
CAPÍTULO 1 – INTRODUÇÃO	1
1.1. CONSIDERAÇÕES INICIAIS.....	1
1.2. MOTIVAÇÃO E OBJETIVOS.....	2
1.3. ORGANIZAÇÃO DO TRABALHO.....	2
CAPÍTULO 2 – CONCEITOS FUNDAMENTAIS	4
2.1. CONSIDERAÇÕES INICIAIS.....	4
2.2. SISTEMAS DE INFORMAÇÃO GEOGRÁFICA.....	4
2.3. FRAMEWORKS.....	6
2.4. APPLICATION PROGRAMMING INTERFACE (API) GOOGLE MAPS.....	7
2.5. ESTATÍSTICAS EM BANCO DE DADOS GEOGRÁFICOS.....	9
2.5.1. DISTRIBUIÇÃO ESPACIAL DOS PONTOS	9
2.5.2. CARACTERIZAÇÕES DAS DISTRIBUIÇÕES DOS PONTOS.....	10
2.5.3. DENSIDADE ESPACIAL	10
2.6. SÉRIES TEMPORAIS.....	10
2.6.1. CONCEITOS DE RELAÇÕES.....	11
2.6.2. MÉTODOS DE PREVISÃO	11
2.6.3. CÁLCULO DE ERROS.....	13
2.6.4. PROCESSO E ESTUDO.....	13
2.7. FERRAMENTAS E APLICAÇÕES EM ÁREAS DE INTERESSE.....	13
2.7.1. APROXIMAÇÃO MULTI-RELACIONAL PARA CLASSIFICAÇÃO ESPACIAL	14
2.7.2. PREVISÃO REGIONAL DE DESLIZAMENTO DE TERRA.....	14
2.7.3. ANÁLISE DE MUDANÇAS EM CONJUNTOS DE DADOS ESPACIAIS POR COMPARAÇÃO DE INTERESSES	14
2.7.4. REGRESSÃO REGIONAL PARA CONJUNTOS DE DADOS GEORREFERENCIADOS	15
2.7.5. ABORDAGEM BASEADA EM REGIÕES PARA DETECÇÃO DE PONTOS OUTLIERS.....	15
2.7.6. VISUALIZAÇÃO UNIFICADA DE ANÁLISE ESPACIAL E TEMPORAL EM 4D GIS	15
2.7.7. ANÁLISE ESPAÇO-TEMPORAL EM BANCOS DE DADOS GENÉRICOS EM GIS	16
2.7.8. O FRAMEWORK CUENCAS.....	16
2.7.9. O SURFMAP: UMA APLICAÇÃO WEB DE MONITORAMENTO BASEADO NA API GOOGLE	16
2.8. CONSIDERAÇÕES FINAIS.....	17
CAPÍTULO 3 – SISTEMA DE INFORMAÇÕES GEOGRÁFICAS PARA ANÁLISE E PREDIÇÃO DE DADOS ESPACIAIS E TEMPORAIS	18
3.1. CONSIDERAÇÕES INICIAIS.....	18
3.2. IDENTIFICAÇÃO DO PROBLEMA E SOLUÇÃO.....	18
3.3. ESTRUTURA DA APLICAÇÃO	19
3.3.1. ARQUITETURA.....	20
3.4. MÓDULO DE VISUALIZAÇÃO.....	22
3.4.1. PROGRAMAÇÃO E VISUALIZAÇÃO	22
3.5. MÓDULO DE PROJEÇÃO	24
3.5.1. SELEÇÃO E PREDIÇÃO.....	24
3.6. IMPLEMENTAÇÃO DO MODELO DE ANÁLISE ESTATÍSTICA	28
3.6.1. ALGORITMOS E CÁLCULO DE ERROS.....	28
3.6.2. METODOLOGIA.....	32
3.7. CONSIDERAÇÕES FINAIS.....	33
CAPÍTULO 4 – VALIDAÇÃO, ANÁLISE E RESULTADOS	34
4.1. TESTES DE BASES DIFERENTES NO MÓDULO DE VISUALIZAÇÃO	34

4.2.	TESTES DE PREDIÇÃO E ANÁLISE.....	38
4.2.1.	TESTE DE PREDIÇÃO DE GRANDES REGIÕES AGRUPADO POR DIAS	38
4.2.2.	TESTE DE SELEÇÃO PARCIAL AGRUPADO POR DIA.....	46
4.2.3.	TESTES DE COMPARAÇÃO DE PREDIÇÕES DE DIFERENTES REGIÕES AGRUPADOS POR MÊS.....	50
4.2.4.	TESTES EM GRANDES BASES DE DADOS AGRUPADOS POR ANOS E SEMANAS	60
4.3.	CONSIDERAÇÕES FINAIS.....	62
CAPÍTULO 5 – CONCLUSÕES E TRABALHOS FUTUROS.....		63
REFERÊNCIAS BIBLIOGRÁFICAS		65

Lista de Figuras

Figura 1 Uma arquitetura dos sistemas de informação geográfica (CÂMARA ET al, 2005)	5
Figura 2 Representação de cada universo (Adaptado de Câmara et. al., 2005)	5
Figura 3 Exemplo de código básico da API Google Maps JavaScript v3 (Google, 2011a)..	7
Figura 4 Exemplo de execução do código básico da API Google Maps JavaScript v3 (Google, 2011a)	8
Figura 5 Código fonte de adição simples de marco no mapa (Google, 2011b)	8
Figura 6 Visualização do marco simples (Google, 2011b)	9
Figura 7 Ferramenta SURF Map (Hofstede; Fioreze, 2009)	17
Figura 8 Arquitetura da FoPAGe	21
Figura 9 Interface de conexão com qualquer base georreferenciada	22
Figura 10 Visualização de dados espaciais no município de São José do Rio Preto	23
Figura 11 Seleção de polígono no módulo de visualização	24
Figura 12 Seleção do período para notificações de acidentes de trabalho no centro da cidade de São José do Rio Preto	25
Figura 13 Somatório de notificações de acidentes reais por dia no ano de 2009 na região central de São José do Rio Preto	26
Figura 14 Visualização da predição do Distrito Industrial de São José do Rio Preto	27
Figura 15 Resultado do método da média móvel comparado com os valores reais para o ano de 2009 na região central de São José do Rio Preto	29
Figura 16 Resultado do método do alisamento exponencial comparado com os valores reais para o ano de 2009 na região central de São José do Rio Preto	30
Figura 17 Resultado do método dos mínimos quadrados comparado com os valores reais para o ano de 2009 na região central de São José do Rio Preto	31
Figura 18 Resultados do RMSE referente ao somatório de acidentes de todos os dias do ano de 2009 para a região central de São José do Rio Preto	32
Figura 19 Resultados do MAPE referente ao somatório de acidentes de todos os dias do ano de 2009 para a região central de São José do Rio Preto	32
Figura 20 Visualização das coordenadas da tabela posicoes150 na base de dados RioPreto150	35
Figura 21 Visualização das coordenadas da tabela posicoes1000 na base de dados RioPreto1000	36
Figura 22 Deslocamento para os pontos de maior concentração na base georreferenciada chamada Distritoindustrial	37
Figura 23 Deslocamento para os pontos de maior concentração na base georreferenciada chamada Engenheiroschmidt	37
Figura 24 Notificações de acidentes de trabalho na cidade de São José do Rio Preto utilizados para cálculo de predição em janeiro de 2010	39
Figura 25 Reta calculada pelo método dos mínimos quadrados comparada com os valores reais para os dias úteis do mês de janeiro de 2010 no município de São José do Rio Preto	40
Figura 26 Curva do método do alisamento exponencial e predições para os primeiros dias de fevereiro de 2010 no município de São José do Rio Preto	40
Figura 27 Curva do método das médias móveis e predições para os primeiros dias de fevereiro de 2010 no município de São José do Rio Preto	41
Figura 28 Valores do erro RMSE de todos os métodos para os dias úteis de janeiro de 2010 no município de São José do Rio Preto	42

Figura 29 Valores do erro MAPE de todos os métodos para os dias úteis de janeiro de 2010 no município de São José do Rio Preto	42
Figura 30 Resultado dos métodos de predição nos dias úteis de janeiro de 2010 no município de São José do Rio Preto	43
Figura 31 Visualização dos mapas das notificações de acidentes de trabalho, dados reais e preditos, para o primeiro dia de fevereiro de 2010	45
Figura 32 Seleção das notificações do Centro e Parque Industrial de São José do Rio Preto para o mês de agosto de 2010	46
Figura 33 Resultados do método dos mínimos quadrados comparado com os valores reais do mês de agosto de 2010 no Centro e Parque Industrial de São José do Rio Preto ...	47
Figura 34 Curva do método do alisamento exponencial e predições para os primeiros dias de setembro de 2010 no Centro e Parque Industrial de São José do Rio Preto	47
Figura 35 Curva do método das médias móveis e predições para os primeiros dias de outubro de 2010 no Centro e Parque Industrial de São José do Rio Preto	48
Figura 36 Valores do erro RMSE de todos os métodos para o mês de agosto de 2010 no Centro e Parque Industrial de São José do Rio Preto	48
Figura 37 Valores do erro MAPE de todos os métodos para o mês de agosto de 2010 no Centro e Parque Industrial de São José do Rio Preto	49
Figura 38 Notificações de acidentes de trabalho em 2009 no Distrito Industrial selecionado e no município de São José de Rio Preto	50
Figura 39 Predição para o primeiro quadrimestre de 2010 no método dos mínimos quadrados no município de São José do Rio Preto	51
Figura 40 Predição para o primeiro quadrimestre de 2010 no método dos mínimos quadrados no distrito industrial	51
Figura 41 Método exponencial para predição do primeiro quadrimestre de 2010 de São José do Rio Preto	52
Figura 42 Método exponencial para o Distrito Industrial na predição do primeiro quadrimestre de 2010	52
Figura 43 Predição da média móvel para São José do Rio Preto em 2010	53
Figura 44 Predição da média móvel para o distrito industrial em 2010	53
Figura 45 Visualização dos mapas das notificações de acidentes de trabalho, dados reais e preditos pelo método do alisamento exponencial, para janeiro de 2010	56
Figura 46 Visualização dos mapas das notificações de acidentes de trabalho, dados reais e preditos pelo método dos mínimos quadrados, para fevereiro de 2010	57
Figura 47 Visualização dos mapas das notificações de acidentes de trabalho, dados reais e preditos pelo da média móvel, para março de 2010	58
Figura 48 Visualização dos mapas das notificações de acidentes de trabalho, dados reais e preditos pelo da média móvel, para abril de 2010	59
Figura 49 Visualização das coordenadas das árvores derrubadas na região de Manaus	60
Figura 50 Resultado de todos os métodos de predição para o desmatamento da região de Manaus durante os anos de 2001 a 2009	61
Figura 51 Resultado de todos os métodos para as semanas do primeiro semestre de 2003, referentes ao desmatamento na região de Manaus	62

Lista de Tabelas

Tabela 1 Comparação entre alguns trabalhos relacionados na área	17
Tabela 2 Comparação entre alguns trabalhos relacionados e o trabalho proposto	33
Tabela 3 Comparação das predições dos métodos com o valor real observado para os quatro primeiros dias úteis de fevereiro de 2010	43
Tabela 4 Comparação das predições dos métodos com o valor real observado para os quatro primeiros dias de setembro de 2010	49
Tabela 5 Comparação dos erros MAPE e RMSE para o município de São José do Rio Preto e o Distrito Industrial para as predições de 2010	54
Tabela 6 Comparação das predições de todos os métodos com o valor real observado para os quatro primeiros meses de 2010 em São José do Rio Preto e Distrito Industrial ...	54

Lista de Siglas

API – *Application Programming Interface*

CSR – *Complete Spatial Randomness*

GBD - Grupo de Banco de Dados

GIS – *Geographic information systems*

GML – *Geography Markup Language*

HTML – *HyperText Markup Language*

JSP – *Java Server Pages*

MAPE - Erro Percentual Médio Absoluto

OGC – *Open Geospatial Consortium*

RMSE - Raiz do Erro Médio Quadrático

SGBD – Sistema de gerenciamento de banco de dados

SIG – Sistema de informação geográfica

SIVAT – Sistema de Vigilância de Acidentes de Trabalho

SOD – *Spatial Outlier Detection*

UML – *Unified Modeling Language*

VDBSCAN - *Varied Density Based Spatial Clustering of Applications with Noise*

VWM – *Virtual World Mapping*

XML – *Extensible Markup Language*

Capítulo 1 – Introdução

1.1. Considerações Iniciais

No planejamento de um sistema de informação, deseja-se extrair de forma coerente o conhecimento e os padrões dos dados nele armazenados, porém isso não se faz de forma trivial e não se prende somente a teorias computacionais. Nos sistemas de informações espaciais, além das dificuldades convencionais, tem-se a ligação entre uma característica física de um dado e sua posição no espaço, o que torna os cálculos de extração de conhecimento ainda mais complexos. Com o uso da estatística espacial essa dificuldade pode ser tratada e assim se obter dados concretos e realísticos. Esse tipo de cálculo se mostra satisfatório em tomada de decisões e medidas de prevenção (SANTOS, 2009) (ZEILHOFER, 2007).

Em algumas bases de dados temporais são encontrados grandes conjuntos de informações espaciais. Os modos de visualização para análise dessas informações são fundamentais (Wood et al, 2007). Um dos recursos que podem ser usados para a visualização espacial é o *Google Maps Application Programming Interface* (Dunnivant, 2010). Com a visualização dos dados e algoritmos estatísticos, diversos tipos de resultados podem ser obtidos, como previsões futuras, probabilidade de acontecimentos, detecção de erros de medidas, dentre outras. Alguns projetos utilizam a *API do Google* como parte do sistema (Pejic; Pletl; Pejic, 2009) para desenvolver aplicações de maneira rápida e de grande usabilidade.

Assim é possível desenvolver uma metodologia de análise futura de dados espaço-temporais georreferenciados implementada em uma aplicação *web* que possa auxiliar nos resultados, desta forma é possível contemplar diferentes bases espaciais e obter uma qualidade diferenciada no conhecimento extraído dos dados armazenados.

Na construção desse sistema foi utilizada uma técnica de desenvolvimento rápido de aplicação, ou *Rapid Web Development*, processo que reúne diversas tecnologias existentes para tratar uma tarefa de aplicação real (PEJIC; PLETL; PEJIC, 2009).

A tecnologia da informação tem evoluído significativamente nos últimos anos, porém as necessidades dos usuários são cada vez mais abrangentes, assim utilizar sistemas *Rapid Web Propotype*, que é uma característica do *Rapid Web Development* (Pejic; Plet; Pejic, 2009), pode ser uma alternativa muito viável em termos de funcionalidades e interface (ROUSSEAU; LHOSTE, 2009).

1.2. Motivação e Objetivos

O uso de novas tecnologias permitiu um avanço importante no contexto espacial em banco de dados, o que aumentou as pesquisas e interesses nessa área. Dados espaciais armazenam informações relevantes e podem ser de grande importância em áreas como meio ambiente, sustentabilidade, segurança no trabalho, dentre outras que visam melhorar a qualidade de vida.

A pesquisa realizada na área permite observar diversas metodologias, modelos e arquiteturas que visam à análise espacial como principal identificador de soluções factíveis para serem aplicadas em regiões de interesse. Dessa forma, é necessária a proposta de uma metodologia capaz de tratar esses dados de forma especial, que analise estatisticamente as futuras probabilidades das informações e que auxilie a tomada de decisão e medidas de prevenção.

Desenvolver uma aplicação para auxiliar essa metodologia de análise supre a necessidade de se obter estatísticas sobre futuros acontecimentos em determinadas regiões de interesse. Os conceitos da estatística em tal desenvolvimento garantem um resultado viável e realístico, dentro de uma margem pequena de erro.

1.3. Organização do Trabalho

O trabalho é dividido nos capítulos descritos a seguir:

- *Capítulo 2 – Conceitos fundamentais* – relata os conceitos básicos que envolvem sistemas de informações geográficas, conceitos estatísticos, dentre outros conceitos aplicados em bases espaciais.
- *Capítulo 3 – Sistema de informações geográficas para análise e predição de dados espaciais e temporais* – apresentação da arquitetura do trabalho proposto, assim como seu desenvolvimento e funcionalidades.
- *Capítulo 4 – Validação, Análise e Resultados* – execução de testes para validação dos módulos de visualização e predição da aplicação Web, assim como a análise e tratamento dos resultados obtidos.

- *Capítulo 5 – Conclusão e trabalhos futuros* – finalização da ideia da nova metodologia de análise futura de dados contemplados por uma aplicação, apresentação dos resultados e sugestão de trabalhos futuros.

Capítulo 2 – Conceitos Fundamentais

2.1. Considerações Iniciais

Neste capítulo são tratados os conceitos e tecnologias pertinentes aos itens – sistemas de informações geográficas; *framework*; sistemas *Web-GIS*; algumas definições sobre a *API Google MAP* e conceitos fundamentais sobre estatística espacial e espaço-temporal.

2.2. Sistemas de Informação Geográfica

A principal diferença entre sistemas de informação geográfica e os sistemas de informações convencionais está no armazenamento da posição espacial do objeto no banco de dados. Assim, é possível obter a forma geométrica, as dimensões e a localização que o objeto pode assumir. As principais características deste tipo de sistema são (CÂMARA ET al, 2005):

- Integrar e inserir dados do tipo geográfico em uma única base, provenientes de diversas fontes de informação como – imagens espaciais, dados GPS, sistemas cartográficos etc.
- Suporte para tratar esses tipos de dados de forma a facilitar a consulta, visualização e recuperação dos mesmos.

Ao armazenar dados espaciais são necessários constantes estudos para o tratamento de prospecções e projeções desses dados, pois em alguns casos o fluxo de informações armazenadas pode sofrer constantes atualizações. Assim, armazenar padrões e alterações pode se tornar inviável (VADEERAT; CHRISTOPH, 2009).

Uma arquitetura para a confecção de um sistema SIG pode ser vista na Figura 1, na qual se tem, desde a interface, nível mais próximo do usuário, até o nível de armazenamento de dados.

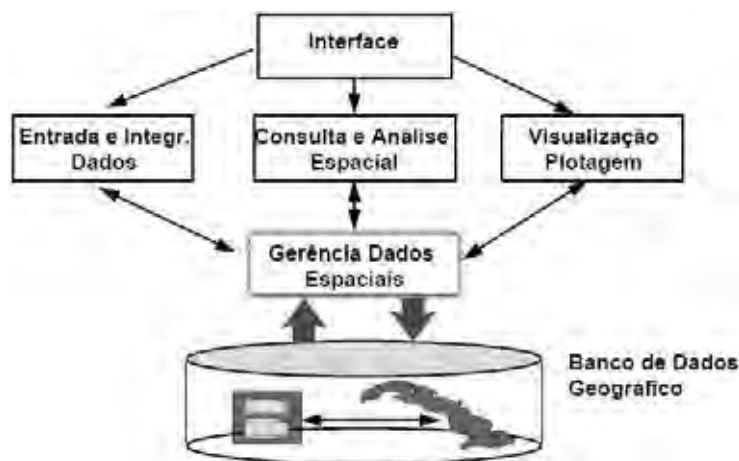


Figura 1 Uma arquitetura dos sistemas de informação geográfica (CÂMARA ET al., 2005)

As aplicações em SIG consistem em extrair a semântica mais completa dos sistemas computacionais desejados (Câmara et. al., 2005). Dessa forma, as informações obtidas podem ser mais intuitivas e utilizadas com maior frequência em medidas de prevenção ou tomada de decisões. Diversas áreas importantes como meio ambiente e saúde utilizam esses recursos, por exemplo, para controlar regiões com altos índices de homicídios (Camargo et al, 2008); para modelagem de tráfego em rodovias (Demiryurek et al, 2009); ou para identificação de volume de tráfego (GIANNOTTI ET al, 2009).

Para exemplificar o conceito de geoinformação será utilizado o paradigma dos quatro universos, adaptado de Câmara et al. (2005), e ilustrado na Figura 2.



Figura 2 Representação de cada universo (Adaptado de Câmara et. al., 2005)

O paradigma dos quatro universos consiste em traduzir as abstrações do mundo real para o computador por meio de um conjunto complexo de ferramentas, algoritmos, modelos matemáticos, dentre outros métodos durante este processo. A seguir será detalhado cada um dos universos:

- **Semântico** – Define os conceitos das abstrações do mundo real, para que sejam entendidos pelo computador, esses conceitos devem ser definidos de forma a contemplar uma diversidade semântica entre atributos semelhantes de diferentes fontes de dados. O problema inicial, em um projeto de sistema de informação, é a escolha das entidades a serem representadas em forma de conceitos. Assim, ao se tratar de dados geográficos, os conceitos são divididos em correspondências aos fenômenos físicos do mundo real,

montanhas, desfiladeiros, vales, rios, temperaturas, poluição entre outros, e conceitos correspondentes a entidades sociais e institucionais, países, municípios, lotes, quadras, desenvolvimento humano etc. O GML (*Geography Markup Language*) é um exemplo de descrição de semântica geográfica, proposto pelo consórcio OGC (*Open Geospatial Consortium*).

- **Formal** – Todos os computadores trabalham com estruturas matemáticas e se faz necessário criar entidades lógicas que possam compreender as entidades semânticas da forma mais abrangente possível. Ao se definir a representação de dados geográficos para o entendimento do computador, são adotadas escalas de medidas, que podem ser escalas nominais, ordinais, por intervalo e razão. Também se definem noções de espaço absoluto, que é isotrópico e homogêneo e, espaço relativo, onde suas características sofrem influência do ambiente. Assim a transição do universo semântico para o formal tem como objetivo classificar todos os atributos que caracterizam determinada aplicação.

- **Estrutural** – Consiste nas estruturas utilizadas na confecção dos bancos de dados geográficos, que podem ser – estruturas vetoriais, que representam as coordenadas das entidades espaciais através das formas básicas, pontos, linhas e polígonos; e, estruturas matriciais, que aplicam um sistema de grade regular para representar um objeto, onde cada célula adquire um valor ligado aos atributos de determinado objeto. Na passagem do universo formal para o estrutural, existem diversas possibilidades de estruturas que podem ser desenvolvidas a partir dos diferentes tipos de entidades, porém não é necessário tal nível de detalhamento para o presente trabalho.

- **Universo de implementação** – Onde é tratada a linguagem de programação utilizada, os tipos de bases de dados, a linguagem de consulta utilizada, dentre outros tipos de tecnologias de implementação que fornecem suporte aos sistemas de informações geográficas.

Os conceitos vistos anteriormente informam o básico para a construção de um sistema de informação geográfica que são suficientes para o total entendimento do presente trabalho.

2.3. Frameworks

Frameworks são amplamente utilizados pelos resultados oferecidos ao usuário de um sistema, para resolver problemas com alto grau de dificuldade de implementação. O projeto de um *Framework* pode ser reutilizado em diferentes aplicações e não é somente um apanhado de códigos e bibliotecas. É também um projeto proposto para resolver um

problema específico, sem ter a necessidade de se preocupar com a construção do sistema que vai utilizá-lo, basta apenas construí-lo de maneira a se adaptar a esse sistema.

Frameworks assumem diversas características especiais como o domínio de aplicação, que deve compreender todos os conceitos e relações dos requisitos do sistema a ser implantado (Ahamed et al, 2004). Existe uma grande variedade de *frameworks* voltados a diversas áreas, como exemplo, uma ferramenta de reconhecimento de gestos por meio de câmeras (Athitsos et al, 2010) e um aplicativo para teste em gerenciadores de banco de dados (Lo et al, 2010) podem ser citados. Flexibilidade, adaptabilidade, segurança e acoplamento são características importantes destes tipos de projetos.

2.4. Application Programming Interface (API) Google Maps

Application Programming Interface significa aplicação de interface de programação, ou seja, uma aplicação que contém alguns padrões e funções específicas para facilitar a programação do desenvolvedor, a fim de criar objetos de forma mais dinâmica em seu projeto. As diversas *APIs Google Maps* estão disponíveis gratuitamente e facilitam a adição de mapas em uma aplicação Web que utiliza a linguagem *JavaScript* (Luo; Shen, 2009). Um exemplo de implementação pode ser estudado:

```

1 <!DOCTYPE html>
2 <html>
3 <head>
4 <meta name="viewport" content="width=device-width, initial-scale=1.0, user-scalable=no" />
5 <meta http-equiv="content-type" content="text/html; charset=UTF-8"/>
6 <link href="http://code.google.com/apis/maps/documentation/javascript/examples/default.css" rel="stylesheet" type="text/css" />
7 <title>Google Maps JavaScript API v3 Example: Map Simple</title>
8 <script type="text/javascript" src="http://maps.googleapis.com/maps/api/js?sensor=false"></script>
9 <script type="text/javascript">
10   function initialize() {
11     var myLatLng = new google.maps.LatLng(-34.397, 150.644);
12     var myOptions = {
13       zoom: 8,
14       center: myLatLng,
15       mapTypeId: google.maps.MapTypeId.ROADMAP
16     }
17     var map = new google.maps.Map(document.getElementById("map_canvas"), myOptions);
18   }
19 </script>
20 </head>
21 <body onload="initialize()">
22   <div id="map_canvas"></div>
23 </body>
24 </html>

```

Figura 3 Exemplo de código básico da API Google Maps JavaScript v3 (Google, 2011a)

Na Figura 3, pode-se observar que somente uma função é necessária para exibir o ponto que é ilustrado em vermelho na Figura 4. O mapa é centralizado no ponto definido como *myLatLng* no código e algumas características são adicionadas, como a definição do

tamanho do zoom e o tipo do mapa. A integração da *API Google Maps* com outras tecnologias, como a .NET, torna possível a construção de um portal que permite carregar a página principal uma única vez e fazer as atualizações posteriores parcialmente (KONARSKI; ZABIEROWSKI, 2010).



Figura 4 Exemplo de execução do código básico da API Google Maps JavaScript v3 (Google, 2011a)

Para se adicionar um ponto, ou marco, no mapa, deve ser implementada a função `google.maps.Marker()`, como é mostrado na Figura 5, e assim obter de forma rápida a localização, visualizada na Figura 6.

```

1 <!DOCTYPE html>
2 <html>
3 <head>
4 <meta name="viewport" content="initial-scale=1.0, user-scalable=no" />
5 <meta http-equiv="content-type" content="text/html; charset=UTF-8"/>
6 <title>Google Maps JavaScript API v3 Example: Marker Simple</title>
7 <link href="http://code.google.com/apis/maps/documentation/javascript/examples/default.css" rel="stylesheet" type="text/css" />
8 <script type="text/javascript" src="http://maps.googleapis.com/maps/api/js?sensor=false"></script>
9 <script type="text/javascript">
10 function initialize() {
11   var myLatLng = new google.maps.LatLng(-25.363882, 131.044922);
12   var myOptions = {
13     zoom: 4,
14     center: myLatLng,
15     mapTypeId: google.maps.MapTypeId.ROADMAP
16   }
17   var map = new google.maps.Map(document.getElementById("map_canvas"), myOptions);
18
19   var marker = new google.maps.Marker({
20     position: myLatLng,
21     map: map,
22     title: "Hello World!"
23   });
24 }
25 </script>
26 </head>
27 <body onload="initialize()">
28   <div id="map_canvas"></div>
29 </body>
30 </html>

```

Figura 5 Código fonte de adição simples de marco no mapa (Google, 2011b)



Figura 6 Visualização do marco simples (Google, 2011b)

2.5. Estatísticas em Banco de Dados Geográficos

Estatísticas voltadas aos sistemas de informação geográfica são aplicadas em diversos trabalhos recentes, como em construção de modelos espaço-temporais que utilizam geoprocessamento e regressão logística (Santos, 2009), mapeamento de riscos de contaminação por pesticidas que utilizam também a regressão logística (Zeilhofer, 2007), dentre outros.

O propósito deste tópico resume-se em introduzir alguns conceitos sobre a aplicação de estatística em SIG.

2.5.1. Distribuição Espacial dos Pontos

Os tipos de dados tratados neste tópico são uma série de coordenadas de pontos de uma determinada ocorrência, o estudo deste tipo de distribuição tem por objetivo observar se os aparecimentos destes dados acontecem de forma aleatória, se aparecem em aglomerações, ou se estão distribuídos de forma regular. Resume-se em encontrar um número acima do esperado para determinados pontos próximos entre si (DRUCK ET al, 2004).

2.5.2. Caracterizações das Distribuições dos Pontos

Na estatística, os processos pontuais são pontos entendidos como irregulares, que estão distribuídos em uma região, onde sua localidade é determinada de forma aleatória. Essa determinação de localização pode ser descrita em efeitos de primeira ordem e efeitos de segunda ordem.

No processo de primeira ordem, o objetivo está na intensidade do processo que leva em conta variações no seu valor médio no espaço, pode ser global ou de larga escala. Nos efeitos de segunda ordem, locais ou de pequena escala, o objetivo é a correlação espacial, que representa sua dependência no espaço.

No processo do cálculo da correlação no espaço estima-se o relacionamento entre pares de acontecimentos esperados. Em sua maior parte, essa técnica deixa de forma clara ou oculta um comportamento estável, cuja representação pontual possui características semelhantes na análise das ocorrências (DRUCK ET al, 2004).

2.5.3. Densidade Espacial

Diversas abordagens são encontradas na literatura para identificação de densidades de objetos distribuídos espacialmente. Uma delas é a implementada pelo VDBSCAN, um algoritmo para agrupamento de dados espaciais com base na densidade do conjunto analisado. Essa abordagem consiste na análise da distribuição dos objetos georreferenciados para determinar variáveis que indicam níveis de densidade encontrados. Essas variáveis são nomeadas de Eps, obtidas a partir do cálculo da variação das distâncias entre os objetos (LIU ET al, 2007) (ZHENG ET al, 2010).

2.6. Séries Temporais

Uma série temporal trata-se de um conjunto de dados obtidos durante certo período de tempo, como exemplo, taxa de juros semanais, preços de ações, temperaturas anuais, acidentes diários, dentre diversos outros (Cryer et al, 2008). O objetivo deste tópico será explicar os conceitos fundamentais para se realizar uma predição de valores.

Ao se fazer uma análise de séries temporais, normalmente, deseja-se entender ou modelar um mecanismo estocástico. Desta forma, é possível obter informações futuras baseadas na história da série.

Normalmente, existe uma dependência entre as observações encontradas, onde se pode obter uma estrutura comum entre as informações, essa é a chave principal para qualquer tipo de análise na estatística temporal.

Em GIS a análise temporal trata duas aproximações estendidas (CRYER ET al, 2008).

Entidade discreta – quando a base da representação são lagos, estradas, montanhas, dentro outros exemplos similares.

Campo contínuo – quando a base da representação dos dados é o espaço e/ou tempo. Objetos individuais neste caso são denotados por atributos ligados a localização dada no espaço-tempo.

2.6.1. Conceitos de Relações

A sequência de variáveis aleatórias Y_t , com $\{t = \pm 1, \pm 2, \pm 3, \dots\}$ é considerada um processo estocástico na estatística e serve como modelo de séries temporais. Este modelo pode ser analisado em termos de sua correlação, covariância, variância e esperança. Uma breve definição desses conceitos é feita em seguida (CRYER ET al, 2008):

Esperança – É o valor esperado de uma variável aleatória, que é calculada pela equação, $\mu_t = E(Y_t)$, onde μ_t é o valor de ocorrência esperado no tempo t , em geral, esse valor pode ser diferente para cada tempo t observado.

Variância – É a medida do intervalo de variação do valor esperado, calculado por $Var(Y) = E((Y - \mu)^2)$.

Covariância – Mede a variação entre duas variáveis de forma conjunta, obtida pela equação, $\gamma_{t,s} = Cov(Y_t, Y_s)$, onde $\{t, s = \pm 1, \pm 2, \pm 3, \dots\}$ e $Cov(Y_t, Y_s) = E[(Y_t - \mu_t)(Y_s - \mu_s)] = E(Y_t Y_s) - \mu_t \mu_s$.

Correlação – É a medida da correlação entre duas variáveis, calculada pela equação, $\rho_{t,s} = Corr(Y_t, Y_s)$, onde $Corr(Y_t, Y_s) = \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)Var(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}$.

2.6.2. Métodos de Previsão

Existem vários motivos para o estudo de séries temporais, dentre eles são destacados, a predição estatística de valores e o estudo do comportamento da série. Essa predição pode ser realizada por meio de cálculos estatísticos que serão definidos a seguir (BROCKWELL; DAVIS, 1991) (WONNACOTT T.; WONNACOTT R., 1990):

Método do alisamento exponencial simples – Consiste em extrair dos dados seu comportamento aleatório, onde todas as observações recebem pesos para o cálculo da predição.

Seja x_t um conjunto de n valores observados $\{x_1, x_2, x_3, \dots, x_n\}$, onde:

- Não se observa sazonalidade.
- São observações aleatórias.
- As médias das observações se alteram ao longo do tempo.

Pode-se definir $a_t = \alpha x_t + (1 - \alpha)a_{t-1}$, como a equação do alisamento exponencial, onde t representa o número de observações $\{t = 1, 2, 3, \dots\}$, assim o tempo $t + 1$, quando encontrado, pode ser considerado uma observação futura.

A variável α representa o peso da observação e seu valor encontra-se entre zero e um.

Método de Médias Móveis Simples – O método de médias móveis pode ser considerado um dos métodos mais simples para a projeção de valores, porém não menos importante quanto outros métodos, devido a seu baixo custo computacional e alto grau de qualidade das previsões obtidas.

Para aplicar este método deve-se obter um valor inicial para a variável $\bar{X}_t$, onde $\bar{X}_t = \frac{x_{t-1} + x_{t-2} + \dots + x_{t-n}}{n}$, n é o número de observações iniciais definidas e t é o período escolhido para o cálculo.

O valor n é definido como a variável a ser ajustada no modelo, assim para grandes valores de n têm-se uma grande absorção do histórico da série nas próximas predições.

Método dos Mínimos Quadrados – Esse método tem por objetivo encontrar uma reta que se adapta melhor ao gráfico gerado pelas observações durante o tempo, minimizando a distância entre o ponto observado e a reta que será calculada.

A equação é dada por $y_i = \alpha + \beta x_i$, onde $\beta = \frac{n \sum y_i x_i - \sum y_i \sum x_i}{n \sum x_i^2 - (\sum x_i)^2}$ e $\alpha = \bar{Y} - \beta \bar{X}$.

A qualidade do resultado de todos os métodos será calculada pela raiz do erro médio quadrático e pelo erro percentual médio absoluto, que serão abordados em sequência.

2.6.3. Cálculo de Erros

O cálculo do erro final dos métodos é de extrema importância para validar a predição estatística como confiável. Dentre os diversos tipos de erros encontrados na literatura, a raiz do erro médio quadrático - RMSE e erro percentual médio absoluto - MAPE, são suficientes para realizar essa validação (Brockwell; Davis, 1991) (Wonnacott T.; Wonnacott R., 1990). Ambos usados como medida de variação dos resultados, onde são

dados pelas equações, $\sqrt{\frac{\sum_{j=1}^N \left(\frac{a_j - y_j}{a_j}\right)^2}{N}}$ e $\frac{\sum_{j=1}^N \left|\frac{a_j - y_j}{a_j}\right|}{N} 100$, respectivamente.

2.6.4. Processo e Estudo

Na aplicação de uma predição estatística, a base de dados deve conter informações espaciais temporais, para manter o histórico da série durante o tempo. Assim é possível saber a transformação do objeto ao longo de sua existência e então projetá-la de forma objetiva e concreta.

O estudo dos procedimentos e modelos espaciais pode ser significativamente importante quando comparados entre si, pois o conhecimento retirado de cada processo específico pode ser comparado com outros e, dessa forma, analisar padrões que possam ser encontrados. Em estudos de caso de controle considera-se que existem dois tipos diferentes de padrões de ocorrências. Na construção de um modelo para este tipo de estudo é possível encontrar várias formas diferentes de conceitos (DRUCK ET al, 2004).

2.7. Ferramentas e Aplicações em Áreas de Interesse

Os trabalhos relacionados podem ter diversos objetivos finais como a detecção, comparação e identificação de objetos espaciais (Deyan, Z; Guoqing, Z, 2007), simulação temporal e cálculo de mudanças espaciais de salinidade no solo de determinado rio (Zhou; Liu, 2010), técnicas de desempenho em grandes volumes de dados espaciais (Liu et al, 2010), cálculos de predições de crescimento urbano (Liu et al, 2009), dentre outros. Na maioria desses projetos são apresentados resultados por meio do cálculo direto em uma base espacial específica, assim como nos trabalhos relacionados com predição.

Em casos de medidas de prevenção existem novas metodologias que propõem a divisão do sistema de informação espacial em três categorias (Kakumoto et al, 2002): a primeira sendo a análise dos dados e a tomada de decisão, baseado nas informações geográficas durante o tempo dos acontecimentos; o segundo corresponde ao uso do sistema

geográfico para o público, como mapas cartográficos; e o terceiro, o uso do GIS para casos especiais, como os sistemas de navegação dos carros.

Nos próximos sub-tópicos serão apresentadas técnicas e metodologias de análises utilizadas em trabalhos relacionados e de alto impacto, que podem ou não realizar previsões.

2.7.1. Aproximação Multi-relacional para Classificação Espacial

Essa ferramenta busca padrões com o objetivo de uma classificação espacial dos objetos. Com uma abordagem multirrelacional espacial, propõe-se um método de classificação chamado de *UnMASC (Unified Multi-relational Aggregationbased Spatial Classifier)* baseado nos métodos de Voronoi. Os métodos de Voronoi são aplicados para o cálculo de vizinhanças e por meio de um diagrama mostram as relações espaciais. Os cálculos exigem um alto conhecimento de dados multirrelacionais, pois os métodos de *data mining* convencionais não tratam de forma completa os dados quando são adicionadas as informações geográficas (FRANK, ESTER, KNOBBE, 2009).

2.7.2. Previsão Regional de Deslizamento de Terra

Este trabalho apresenta um método de geoestatística para analisar, em escala regional, a relação espacial entre deslizamentos de terra e seus de fatores de influência. O objetivo é prever regiões suscetíveis a deslizamento de terra e obter informações para medidas de prevenções e tomada de decisões. A abordagem consiste no uso de bancos de dados digitais, sistema de informação geográfica e mapeamento interativo com análise estatística espacial (CAIYAN ET al, 2011).

2.7.3. Análise de Mudanças em Conjuntos de Dados Espaciais por Comparação de Interesses

Na maioria dos bancos de dados que contêm informações espaciais existe um enorme fluxo de informações, que dificulta as condições de monitoramento quando se quer comparar alterações de uma mesma região durante certo espaço de tempo. As regiões de interesse são definidas pelo usuário. O trabalho propõe um *framework* que, por meio de funções e algoritmos, detecta regiões de interesse para o conjunto de dados selecionados (VADEERAT, CHRISTOPH, 2009).

2.7.4. Regressão Regional para Conjuntos de Dados Georreferenciados

Uma proposta de *framework* que utiliza regressão estatística para descoberta de regiões de interesse com alto índice de relacionamento entre variáveis, para auxiliar no suporte às informações espaciais, pois os algoritmos de regressão tradicionais não analisam as relações espaciais das variáveis existentes em uma base de dados. O *framework* também constrói modelos de regressão para cada região específica, de uma forma geral, os métodos não são aplicados globalmente, ou seja, atuam diretamente em sub-regiões. Na confecção da ferramenta são utilizadas funções do tipo *fitness*, para cálculo de adequação de variáveis e busca de regiões onde possam surgir funções de regressão com alta precisão nos resultados. Essas funções são selecionadas por meio de uma metodologia genérica baseada na correlação entre os dados selecionados (CELEPCIKAY, EICK, 2009).

2.7.5. Abordagem Baseada em Regiões para Detecção de Pontos *Outliers*

Um *framework* criado para detecção espacial de pontos *outlier*, ou seja, pontos que estão fora de uma determinada região de interesse, otimiza a densidade utilizada para obter informações sobre os objetos espaciais. Neste trabalho é ressaltada a falta de conceitos estatísticos em sua construção, a falta de conexões teóricas para o *framework* apresentado e as comparações entre os métodos locais e globais para detecção de *outlier* (SOD). O trabalho identifica os problemas fundamentais na generalização estatística local do *framework*. O grande volume de dados espaciais é destacado como uma dificuldade na extração de conhecimento e detecção de *outliers*, devido à comparação entre atributos de dados não espaciais com os espaciais. Assim, os conceitos e algoritmos construídos têm como foco principal tratar as relações existentes entre esses tipos de relações (CHEN, BOEDIHARDJO, LU, 2010).

2.7.6. Visualização Unificada de Análise Espacial e Temporal em 4D GIS

Sistemas de informação geográfica, sensoriamento remoto, modelagem de ambiente e engenharia de sistemas de informação são áreas onde bases de dados espaço-temporais se tornam cada vez mais importantes. Neste trabalho é proposto um *framework* de visualização unificada para análise espacial em 4D (3D + tempo) GIS, que consiste em componentes como, provedor de dados, gerenciador de dados, gerenciador de componentes, *GeoVideoServer* e *GeoVideoClient*. Um conceito de serviço de dados chamados *Virtual World Mapping* (VWM) é abordado para mapear o mundo-real 3D em vídeo. O método proposto pode facilmente recuperar as informações geográficas e

atributos para reconstruir o espaço virtual 3D de acordo com a sequência de quadros dos vídeos (SUNG-SOO ET AL, 2003).

2.7.7. Análise Espaço-temporal em Bancos de Dados Genéricos em GIS

Neste trabalho é proposto o projeto e o desenvolvimento de um sistema para controle de dados georreferenciados, baseado em um conjunto genérico de ferramentas de processamento em diferentes formatos de dados espaciais que modelam os dados de forma padrão para serem usados por qualquer sistema. O trabalho enfatiza a pesquisa nos últimos anos em conceitos genéricos para bases espaciais. A modelagem genérica é feita em bases espaço temporais e as linguagens mais usadas são GML e UML (HAMDI; FARAH; AHMED, 2005).

2.7.8. O *Framework* Cuencas

Cuencas é um *framework* numérico para análise automática dos fluxos em redes hidrográficas. Na sua construção são utilizados modelos de elevação digital para extração de redes e uma abordagem geomorfológica para a divisão de uma bacia hidrográfica em vertentes e canais de ligação, que são necessários para representar as características em um terreno real. O trabalho aborda diversos conceitos estatísticos em testes de auto-semelhanças nessa representação (MANTILLA; GUPTA, 2005).

2.7.9. O *SurfMap*: uma Aplicação Web de Monitoramento Baseado na API Google

SurfMap é uma ferramenta para monitoramento de rede que fornece informações de fluxo de rede em dimensão geográfica baseado na *Google Maps API*. Com o uso da API é possível obter diferentes níveis de *zoom* ao exibir essas informações, o que resulta na criação de diferentes níveis de abstração de visualização de dados de rede. Neste trabalho é ressaltada a importância do *Google Maps* nos últimos anos, o qual referencia trabalhos de impacto significativo em diversas áreas, por meio da utilização de suas APIs. A localização geográfica é obtida por meio do *IP2Location* e a arquitetura é simples e dividida em partes, onde a principal parte é a *SurfMap Web*, uma página que é o núcleo da ferramenta e controla todas as outras funcionalidades (HOFSTEDE; FIOREZE, 2009).



Figura 7 Ferramenta SURF Map (Hofstede; Fioreze, 2009)

Na Tabela 1 é possível observar uma comparação entre os trabalhos relacionados na área, as metodologias adotadas e carências observadas.

Tabela 1

Comparação entre alguns trabalhos relacionados na área

	Wu, Wang and Qiao, 2011	Liu et al, 2009	Mantilla, Gupta, 2005	Hamdi; Farah; Ahmed, 2005	Chen, Boedihardjo, Lu, 2010
Análise de dados espaciais	√	√	√	X	√
Métodos estatísticos para cálculos	√	√	√	X	√
Cálculo de predição	√	√	X	X	X
Visualização da predição	√	X	X	X	X
Modelagem genérica	X	X	X	√	X
Permuta entre qualquer Base de dados	X	X	X	X	X

2.8. Considerações Finais

Este capítulo teve como objetivo demonstrar os principais tópicos abordados nos métodos de análise de dados espaciais e temporais, algumas das proposições da estatística de probabilidade também foram conceituadas. Encontrou-se um *déficit* na generalização de resultados de metodologia de análise nos trabalhos relacionados, assim como na predição de informações.

Capítulo 3 – Sistema de Informações Geográficas para Análise e Predição de Dados Espaciais e Temporais

3.1. Considerações Iniciais

A confecção da Ferramenta *on-line* de Predição e Análise Georreferenciada – FoPAGe, torna-se necessária para auxiliar a visualização dos resultados de predição e a análise das informações georreferenciadas obtidas. Sendo assim, foram implementadas tecnologias factíveis com a estruturação e modelagem do projeto para visualização e predição das bases de dados georreferenciadas. Ainda na predição, os cálculos dos valores futuros são baseados nos métodos estatísticos e a visualização desse resultado é realizada por meio de funções de probabilidade e cálculos de densidade por distância.

Com base nas informações propostas, os próximos tópicos tratam da confecção da ferramenta e da metodologia de generalização utilizada.

3.2. Identificação do Problema e Solução

As séries temporais da estatística são notadas em bases geográficas com informações espaço-temporal e são utilizadas de maneira satisfatória em diversas frentes de pesquisa (DEYAN, Z; GUOQING, Z, 2007) (ZHOU; LIU, 2010).

Diferentes tipos de técnicas e metodologias surgem para tratar bases no espaço tempo, tanto para analisar grandes volumes de dados espaciais temporais (Yan; Liqin; Pusheng, 2008) quanto evidenciar pesquisas sobre mudanças espaço-temporais em bacias hidrográficas (Zhao et al, 2009). Também se podem encontrar ideias de sustentabilidade, onde se observam análises espaciais de contaminação de metais pesados em rios (Cuihua et al, 2007), partículas inaláveis e seu fator de influência no espaço em função do tempo

(Wen-Hui et al, 2009) e estudos sobre imagens de satélite que evidenciam de forma temporal a degradação da natureza e algumas tendências (JUNMEI, 2009).

Diversas ferramentas *Web*, como algumas observadas nos tópicos anteriores, fazem uso de novas tecnologias em sua confecção. Arquiteturas de grande impacto utilizam a *Google Maps API* e o *PostGIS* em sua construção (Kiwon, 2009) com objetivo de desenvolver portais de dados geográficos, também são encontrados arquiteturas que envolvem logística e descrevem a importância dos sistemas *WebGIS* e de novas tecnologias em implementações deste tipo (CONG ET al, 2010).

Em alguns portais podem ser observados o uso de outras bases (Yang et al, 2009), contudo as informações geográficas nesses tipos de sistemas são obtidas diretamente pela *API Google* e não fazem uso de um gerenciador de banco de dados com suporte espacial.

Ferramentas de visualização garantem um suporte analítico em pesquisas espaciais, como a construção de uma linguagem de consulta espacial visual para dados espaço temporais (Bonhomme; Aufaure, 2002), aplicações que exibem informações sobre o uso da terra (Hao et AL, 2010), ferramentas pedagógicas (Teresco, 2011) e as usadas para prever o impacto em solo por produtos químicos (NAN; XIAOHONG; JING, 2009).

Na maioria desses trabalhos existe uma infinidade de dados temporais que poderiam ser aproveitados e utilizados em uma predição periódica. Com base nessa ideia e na carência em visualizações e análises genéricas de dados georreferenciados, este trabalho propõe projetar informações no futuro, baseado no histórico dos dados e analisá-los com uma interface em que uma base de dados espaciais pode ser exibida e projetada.

Esta proposta de metodologia pode suportar ações de gestão e oferecer benefícios em diversos segmentos tais como medidas de qualidade de vida e meio ambiente. De maneira mais abrangente pode-se encontrar problemas regionais relacionados a altos níveis de assaltos, de mortalidade infantil, de acidentes, dentre várias outras possibilidades.

Assim como já exibido por meio de trabalhos correlatos encontrados na literatura é de alto interesse a unificação de conceitos de georreferenciamento, computação e estatística no intuito de calcular dados futuros para análise e implantação de medidas de sustentabilidade.

3.3. Estrutura da Aplicação

A FoPAGe possui uma interface de conexão para qualquer base de dados espaciais provenientes do *PostGIS*, que é uma extensão espacial do gerenciador *PostgreSQL*, escolhido por ser extremamente utilizado na atualidade, possuir licença gratuita, suporte a

objetos espaciais, suporte a funções e operadores, integração com serviços *web*, seguir os padrões *OpenGIS* (Lijing; Jing, 2010), dentre outros fatores.

Após inserir o nome da base, é possível realizar a visualização dos dados apenas pelo nome da tabela que contém as coordenadas espaciais. Em seguida, pode-se realizar a projeção futurística dos dados ou selecionar uma determinada região para tal. Os resultados são exibidos após a projeção e então se analisa a região selecionada. Essa análise é um fator de grande importância e sustentada por pesquisas paralelas na área, onde são envolvidos sensoriamento remoto e sistemas GIS (LEI; DA-FANG; DONG-SHENG, 2008).

3.3.1. Arquitetura

A arquitetura da FoPAGe pode ser observada na Figura 8, onde se verifica em um primeiro nível uma separação em dois módulos: o módulo de visualização e o módulo de projeção. Novos modelos de arquitetura, que utilizam tecnologias recentes, são evidenciados em propostas de análise de informações geográficas em grandes bases espaço-temporal, como em modelos que utilizam computação paralela (Dingju, 2009) e modelos espaço-temporal distribuídos (Schulze; Wytzisk; Simonis, 2002). Os módulos da FoPAGe serão detalhados nos próximos tópicos.

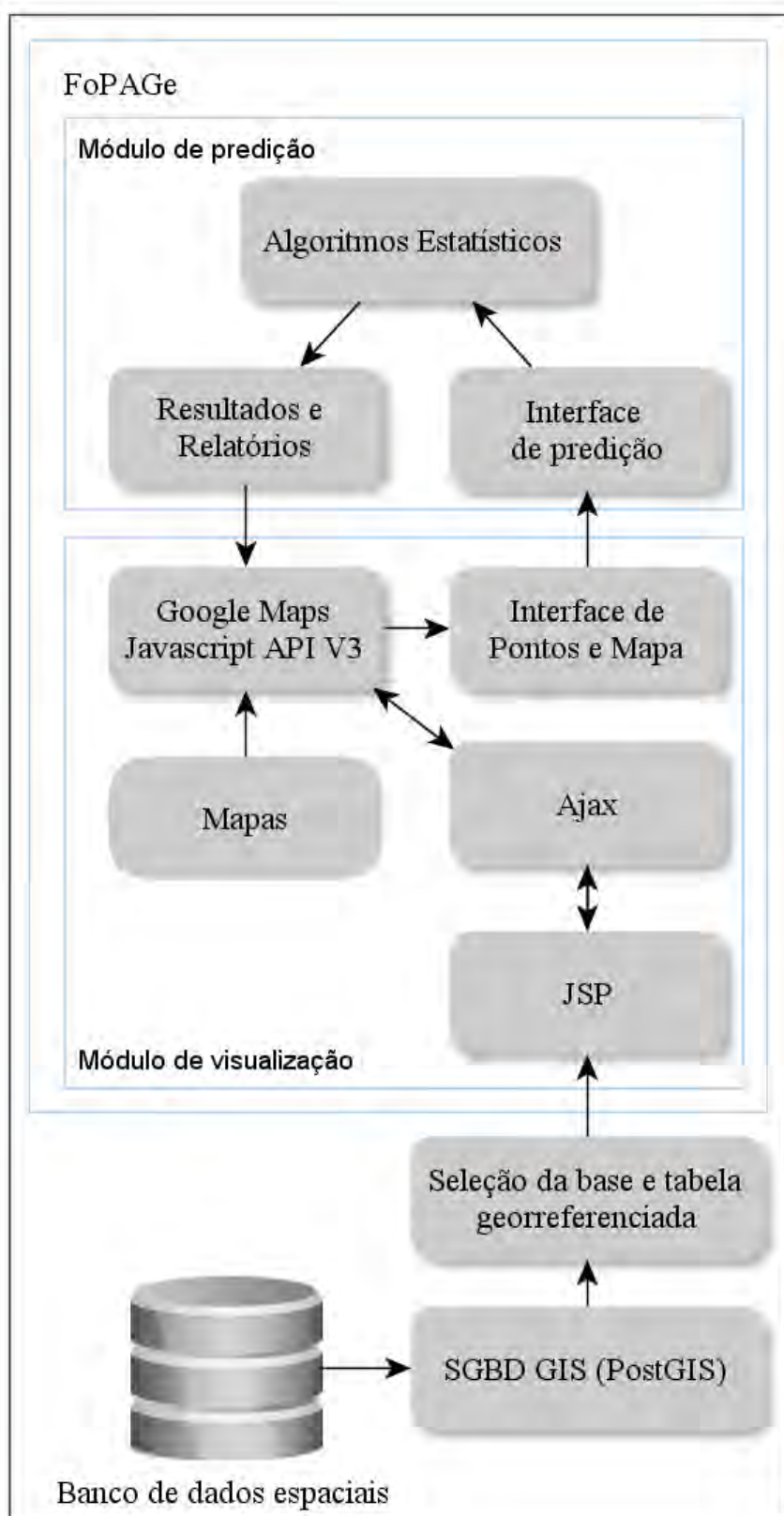


Figura 8 Arquitetura da FoPAGe

3.4. Módulo de Visualização

Este módulo é a parte inicial e essencial FoPAGe, responsável pelo controle e visualização dos dados espaciais, por meio de gráficos convencionais. A FoPAGe possui um interface de fácil utilização, com alto nível de abstração, onde o usuário pode conectar sua base de dados georreferenciada ao inserir o nome e a tabela que contém as coordenadas geográficas.

3.4.1. Programação e Visualização

A programação da FoPAGe é feita com o *Java Server Pages* que tem sido muito utilizado na confecção de diversas ferramentas de impacto considerável. O JSP permite trabalhar com a linguagem Java e *HTML*. Assim uma classe de conexão é criada para trabalhar dinamicamente com a *API Google*, por meio do *AJAX* e se conectar a qualquer base espacial proveniente do *PostGIS*, esse método generaliza um modelo de visualização espacial. A *API* permite programar as funções de seleção de polígono, de adicionar os pontos espaciais e de exibir informações.



Figura 9 Interface de conexão com qualquer base georreferenciada

Na Figura 9 é demonstrada a interface para a conexão e visualização de uma base de dados espacial. Para os testes e aplicações da FoPAGe foi utilizada uma base de dados que contém informações de notificações de acidentes de trabalho pertinentes à sede regional situada em São José do Rio Preto. Atualmente, esta base de dados que suporta o

Sistema de Informação de Vigilância e Acidentes do Trabalho – SIVAT contém por volta de 70 mil registro de notificações de acidentes do trabalho, sendo que aproximadamente 16 mil registros contêm informações para georreferenciamento e tem servido de base de experimentação para aplicação de diversas tecnologias na área de descoberta de conhecimento em banco de dados (VALÊNCIO ET al, 2012) (RODRIGUES ET al, 2011).

Na Figura 10 são exibidas as notificações referentes à região de São José do Rio, onde cada ponto vermelho representa uma coordenada geográfica.

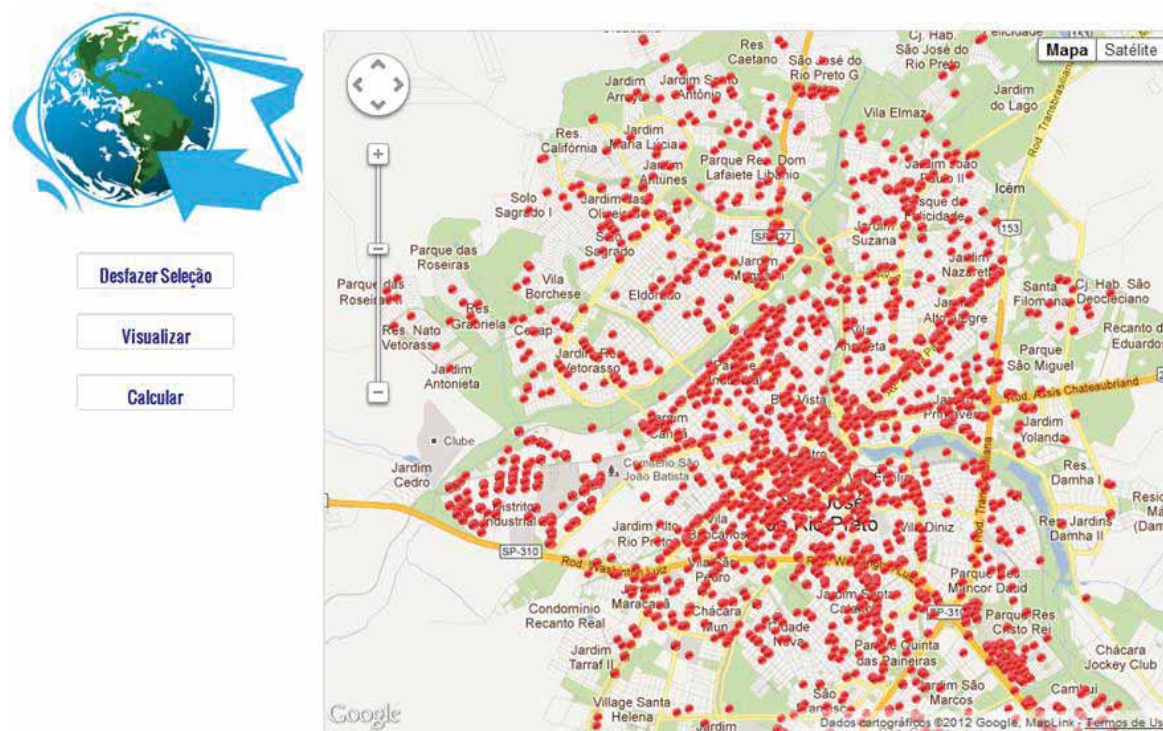


Figura 10 Visualização de dados espaciais no município de São José do Rio Preto

As coordenadas são transformadas em um formato padrão para que a *Google Maps JavaScript API* possa identificar a posição corretamente e exibi-la no mapa. O *JSP* e o *AJAX* executam as consultas dos dados na base que utilizam a linguagem de manipulação de dados espaciais do *PostGIS*.

A FoPAGe oferece o recurso de seleção de regiões de interesse, por meio da delimitação de polígonos, onde um conjunto específico de pontos podem ser adicionados para a aplicação das etapas posteriores. Na Figura 11 é possível observar um exemplo junto a esta interface.

do Rio Preto, onde os pontos em vermelho correspondem às ocorrências registradas de acidentes de trabalho de janeiro a dezembro de 2009.

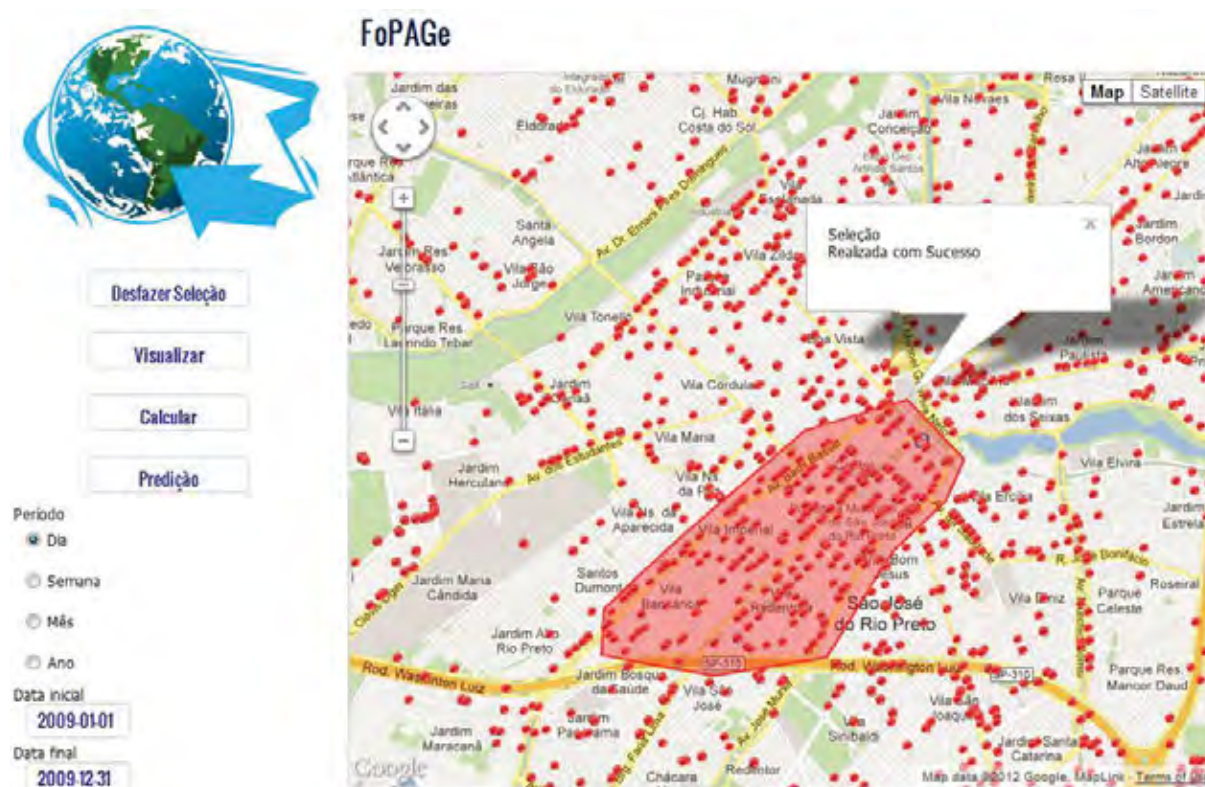


Figura 12 Seleção do período para notificações de acidentes de trabalho no centro da cidade de São José do Rio Preto

Os valores reais dos dados contidos na base, durante o período selecionado, são exibidos em um gráfico convencional, que de forma dinâmica também mostra o resultado dos cálculos estatísticos. A Figura 13 ilustra somente os dados reais, onde os nomes em cinza claro, na legenda da direita, denotam que os resultados dos métodos estão ocultos no gráfico. O eixo horizontal exibe o somatório das ocorrências de todos os dias do ano de 2009, os demais valores são referentes às predições calculadas pelos métodos, que serão visualizados posteriormente.

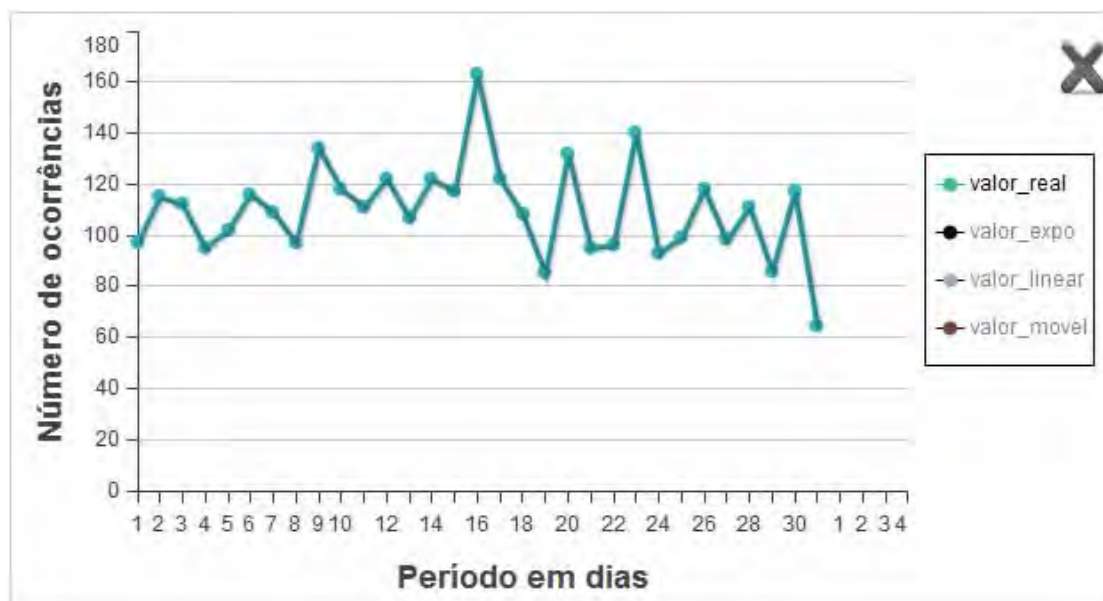


Figura 13 Somatório de notificações de acidentes reais por dia no ano de 2009 na região central de São José do Rio Preto

A predição é calculada baseada nos conceitos de séries temporais da estatística, onde os dados são não estacionários, ou seja, possuem médias e variâncias diferentes ao longo do tempo. Dessa forma, o resultado não será um gráfico constante ou um gráfico que converge para um determinado valor, o que indicaria uma predição tendenciosa.

A predição é exibida baseada nos cálculos de densidade e funções probabilísticas. As distâncias entre os objetos espaciais são calculadas e armazenadas em um vetor ordenado. Os Eps são obtidos a partir desse vetor e determinam os níveis de densidade da base espacial. Os valores obtidos como resultados das predições são divididos por proporção nos níveis de densidades observados e os pontos são escolhidos como locais de predição no mapa por meio da esperança e da variância estatística.

Na Figura 14 é possível observar o resultado de um teste de predição, onde o primeiro mapa foi estimado baseado nos dados que estão armazenados na base de dados do segundo mapa, esse teste será detalhado posteriormente.



Visualização dos pontos estimados



Visualização dos dados reais

Figura 14 Visualização da predição do Distrito Industrial de São José do Rio Preto

Nas próximas sessões serão descritos os métodos e seus respectivos resultados obtidos pela FoPAGE.

3.6. Implementação do Modelo de Análise Estatística

Não é possível encontrar metodologias genéricas para análise de uma base de dados espaço-temporal georreferenciada, mas podem ser criados novos modelos analíticos, como exemplo a análise de disparidade (Tao; Jin-Ping; He, 2010) e análise de superfícies (ZHAO ET al, 2011).

Um novo modelo analítico é proposto baseado em séries temporais da estatística, onde são utilizadas algumas funções de cálculo de probabilidade, métodos de predição de valores, agrupamento de regiões de interesse por distância e análise de gráficos e resultados.

Os modelos e métodos estatísticos são implementados em algoritmos que, de forma dinâmica, calculam as predições e seus pontos futuros. Todos esses cálculos são realizados neste módulo, posteriormente, um vetor com os resultados é enviado para o módulo de visualização, onde ele exibe os pontos preditos e os gráficos para análise.

3.6.1. Algoritmos e Cálculo de Erros

Foram criados três algoritmos com base nos métodos de séries temporais da estatística, onde são realizados cálculos para prever futuros valores. O vetor de entrada e saída dos dados contêm as coordenadas geográficas dos pontos, o período e a quantidade. O vetor de saída ainda retorna os valores dos erros obtidos que serão usados para validação.

Os algoritmos e os cálculos dos erros são descritos e exemplificados em gráficos separadamente:

- Algoritmo do método das médias móveis

Este algoritmo realiza o cálculo de uma curva média referente aos dados reais. O valor inicial para os cálculos desse método é obtido a partir de um intervalo que contém algumas posições iniciais do vetor de entrada. Para cada posição seguinte uma nova média é calculada, inserindo o próximo ponto real e removendo o último, sem exceder o intervalo de entrada. Para a predição deste método, estimam-se os quatro próximos valores, baseados no histórico da série.

Na Figura 15 é apresentado o gráfico do cálculo do método das médias móveis durante o período de janeiro a dezembro de 2009, onde os dados são agrupados por dias. A curva de cor marrom representa o resultado desses cálculos e a predição é exibida ao final, onde os números de 1 a 4, no eixo horizontal, são as predições do primeiro, segundo,

terceiro e quarto dia de 2010, respectivamente. Os demais métodos são ocultados do gráfico.

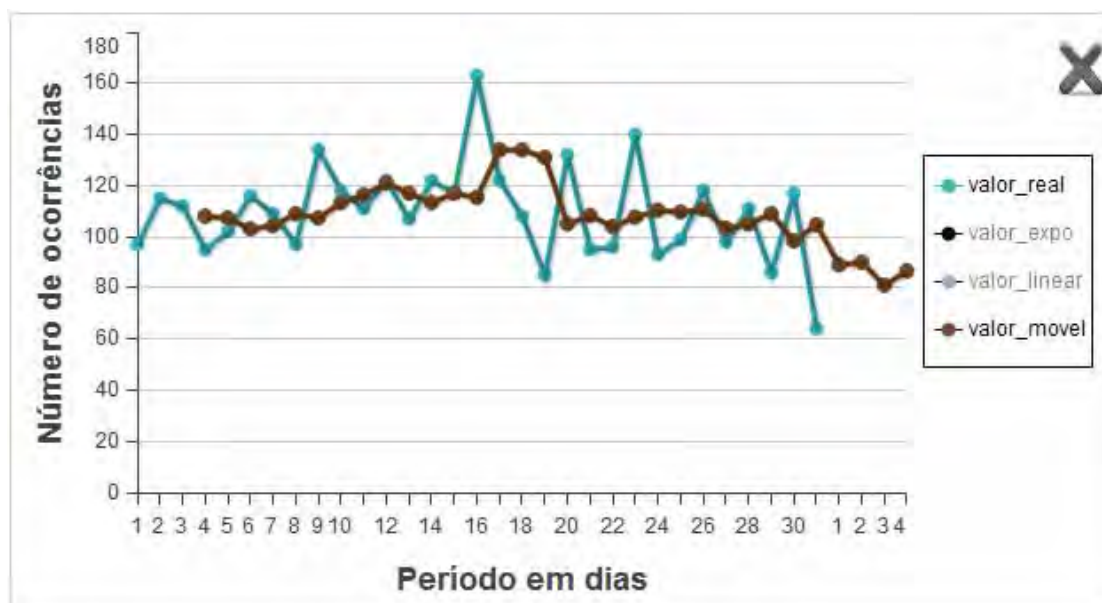


Figura 15 Resultado do método da média móvel comparado com os valores reais para o ano de 2009 na região central de São José do Rio Preto

- Algoritmo do método do alisamento exponencial simples

Para o cálculo deste algoritmo são usados cerca de 30% do histórico da série temporal, esse valor garante uma boa qualidade dos erros baseados na conceituação estatística. Este método permite construir uma curva que acompanha o gráfico, assim sua primeira predição estará mais próxima do real do que dos demais métodos propostos. As próximas predições são calculadas por valores que foram estimados, consequentemente, os erros aumentam de forma significativa, o que prioriza o uso dos outros métodos nesses casos.

Na Figura 16 é ilustrado o gráfico do cálculo do método do alisamento exponencial simples para o ano de 2009, onde os dados são agrupados por dias. A curva de cor preta corresponde ao resultado do algoritmo e a predição é exibida ao final, onde os números de 1 a 4 no eixo horizontal correspondem às predições para os primeiros quatro dias do ano de 2010. Os demais métodos são ocultados do gráfico.

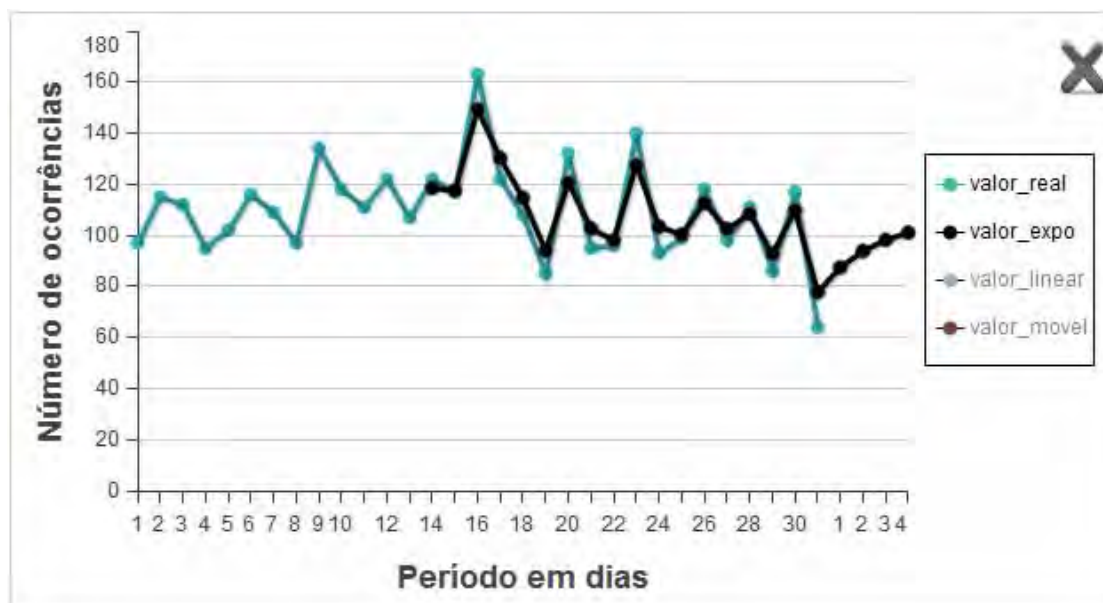


Figura 16 Resultado do método do alisamento exponencial comparado com os valores reais para o ano de 2009 na região central de São José do Rio Preto

- Algoritmo do método dos mínimos quadrados

Este algoritmo tem por objetivo encontrar uma reta que minimiza a soma dos quadrados das diferenças entre as distâncias de todos os pontos aos pontos estimados. Esse método permite concluir se os resultados são crescentes ou decrescentes, ou estão constantes ao longo do tempo. Essa orientação dos resultados obtém uma predição com valores pertencentes a uma sequência monótona, assim a predição dos demais métodos pode ser comparada. Com isso, o algoritmo pode ser usado para reforçar a validação dos resultados.

Na Figura 17 é exibido o gráfico do cálculo do método dos mínimos quadrados para o ano de 2009 agrupados por dias, os demais métodos foram ocultados do gráfico. A reta de cor azul corresponde ao resultado do algoritmo, que neste caso exibe uma tendência de decréscimo no número de notificações de acidentes. Os números de 1 a 4 no eixo horizontal correspondem às predições para os primeiros quatro dias do ano de 2010.

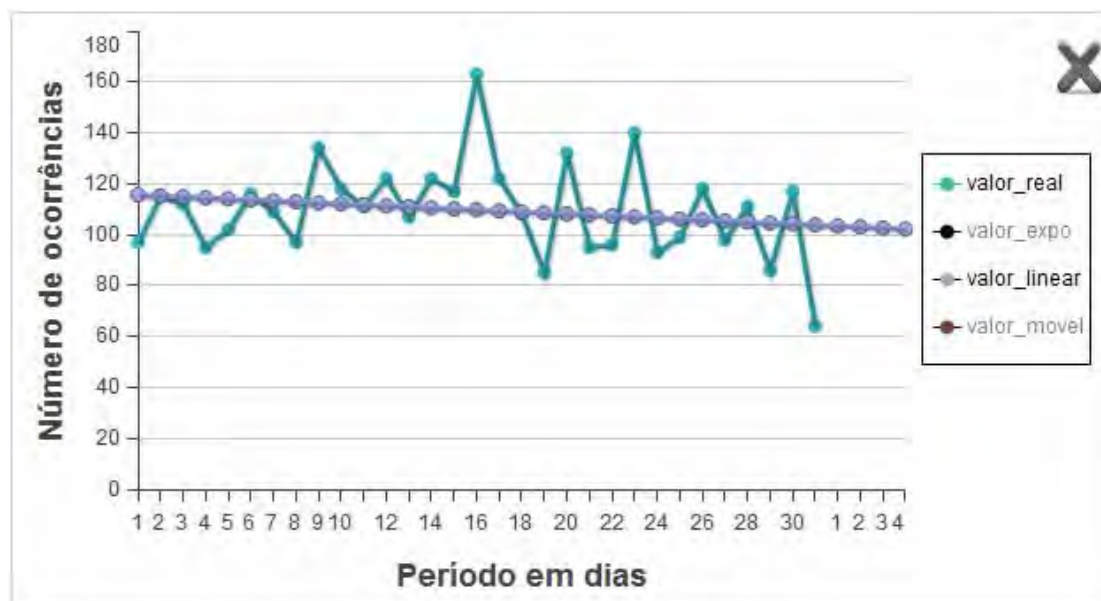


Figura 17 Resultado do método dos mínimos quadrados comparado com os valores reais para o ano de 2009 na região central de São José do Rio Preto

- Cálculo dos erros dos algoritmos

Os erros são calculados separadamente para cada método implementado. A raiz do erro médio quadrático absoluto é usada para comparar os valores obtidos. O menor resultado obtido nessa comparação pode ser escolhido como a predição mais confiável do período. Na Figura 18 são exibidos os resultados referentes ao RMSE, onde as curvas de cores marrom, preta e azul correspondem aos erros dos métodos das médias móveis, do alisamento exponencial e dos mínimos quadrados, respectivamente.

O erro percentual médio absoluto calcula uma variação confiável para o resultado das predições. O gráfico do MAPE pode ser visualizado na Figura 19, onde a representação dos métodos segue o mesmo padrão descrito anteriormente.

Os gráficos dos erros são exibidos para verificar a alteração de valores sofrida no decorrer do processo, uma vez que o valor do erro é usado somente ao final dos cálculos da predição. Essa precisão está relacionada com a quantidade temporal das coordenadas, pois os algoritmos de projeção não obtêm resultados satisfatórios em um período muito curto.

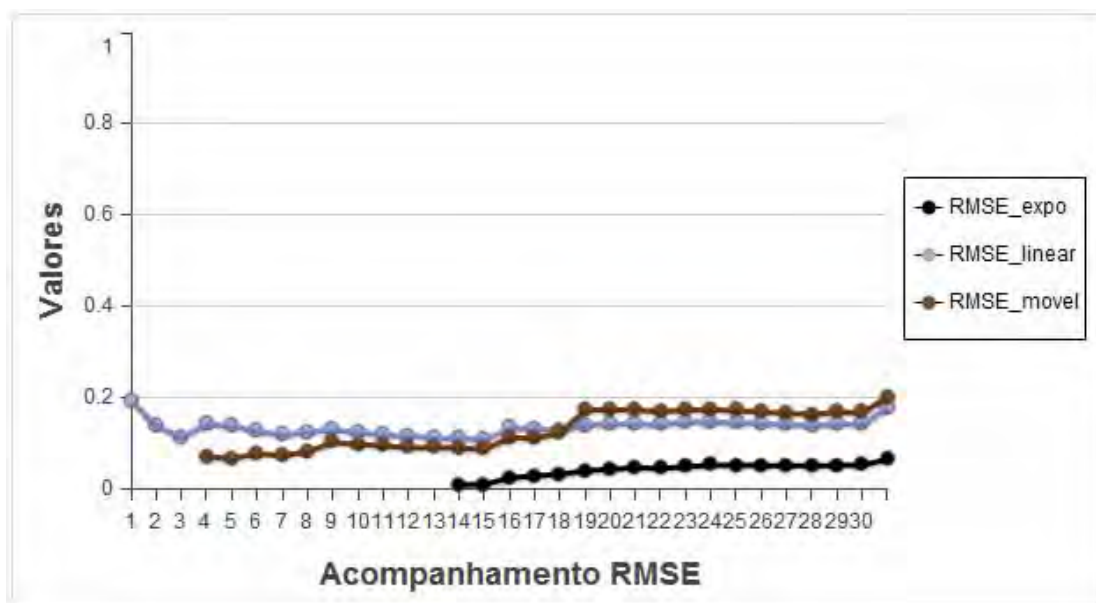


Figura 18 Resultados do RMSE referente ao somatório de acidentes de todos os dias do ano de 2009 para a região central de São José do Rio Preto

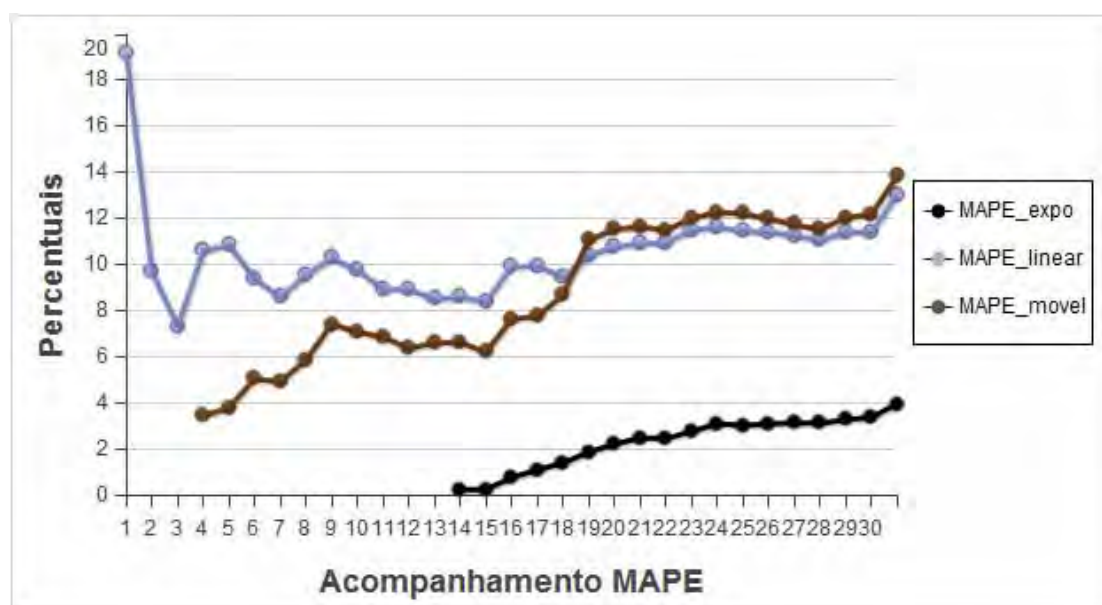


Figura 19 Resultados do MAPE referente ao somatório de acidentes de todos os dias do ano de 2009 para a região central de São José do Rio Preto

3.6.2. Metodologia

A metodologia proposta para a análise da predição estatística é baseada nos métodos de séries temporais implementados e pode ser aplicada a qualquer base georreferenciada que contém o histórico temporal dos dados. O objetivo desse modelo é observar e validar as tendências dos resultados. A estrutura da proposta segue os seguintes padrões:

- Validação dos resultados obtidos, por meio dos erros MAPE e RMSE calculados.

- Análise da orientação das projeções, fundamentada nos cálculo da reta do método dos mínimos quadrados.

-Análise da predição, onde o módulo calcula os próximos quatro valores futuros, nos três métodos estatísticos implementados. Assim se podem obter resultados determinísticos para a tomada de decisões.

- Escolha do primeiro valor futuro pelo método do alisamento exponencial, pois fornece erros menores por acompanhar o gráfico.

- Interpretação dos resultados das demais predições.

- Definição e aplicação dos resultados como metodologia de prevenção, caso os mesmo apontem o surgimento de problemas nas regiões analisadas.

Tabela 2

Comparação entre alguns trabalhos relacionados e o trabalho proposto

	Wu, Wang and Qiao, 2011	Liu et al, 2009	Mantilla, Gupta, 2005	Hamdi; Farah; Ahmed, 2005	Chen, Boedihardjo, Lu, 2010	Trabalho Proposto
Análise de dados espaciais	√	√	√	X	√	√
Métodos estatísticos para cálculos	√	√	√	X	√	√
Cálculo de predição	√	√	X	X	X	√
Visualização da predição	√	X	X	X	X	√
Modelagem genérica	X	X	X	√	X	√
Permuta entre qualquer Base de dados georreferenciada	X	X	X	X	X	√

Na Tabela 2 é exibida a comparação entre o trabalho proposto, alguns trabalhos relacionados na área e os métodos adotados para relacionamento.

3.7. Considerações Finais

A arquitetura e os algoritmos implementados que foram discutidos neste capítulo demonstraram comparações relevantes, como exemplo, o cálculo dos erros RMSE e MAPE, que são utilizados para validação dos resultados. O módulo de predição exibe diversos gráficos para análise, assim tanto os usuários com pouco ou muito conhecimento em estatística podem obter os resultados desejados. Os resultados dos métodos e a metodologia de análise proposta possibilitam a tomada de ações em medidas de prevenções em diversas áreas. No próximo tópico são detalhados os testes para validação da aplicação.

Capítulo 4 – Validação, Análise e Resultados

Para a realização dos testes propostos neste capítulo é necessário a conexão com a *internet*, pois este projeto utiliza uma aplicação *Web-GIS*, então se assume uma boa velocidade de navegação, a qual não interfira no tempo de execução e cálculo dos dados.

Em termos da velocidade do processador, memória ou periféricos, é necessário somente um computador capaz de navegar de forma eficiente na *internet*, para os experimentos foi usado um Computador com Core I3 530 PC, 4GB de memória, HD 500GB. Nos demais tópicos são detalhados os métodos de análise e testes.

4.1. Testes de Bases Diferentes no Módulo de Visualização

Para validar o módulo de visualização foram construídas bases espaciais com pontos geográficos escolhidos aleatoriamente, sendo que cada base e tabela foram criadas de formas diferentes.

As bases e tabelas foram inseridas no gerenciador *PostgreSQL* com o suporte *PostGIS* para dados espaciais e utilizadas para a execução dos testes iniciais de conexão.

Na Figura 20, pode ser observada a conexão com a primeira base de teste chamada de RioPreto150, com cerca de 150 pontos geográficos inseridos na tabela *posicoes150*. Os pontos em vermelho no mapa representam as coordenadas georreferenciada.

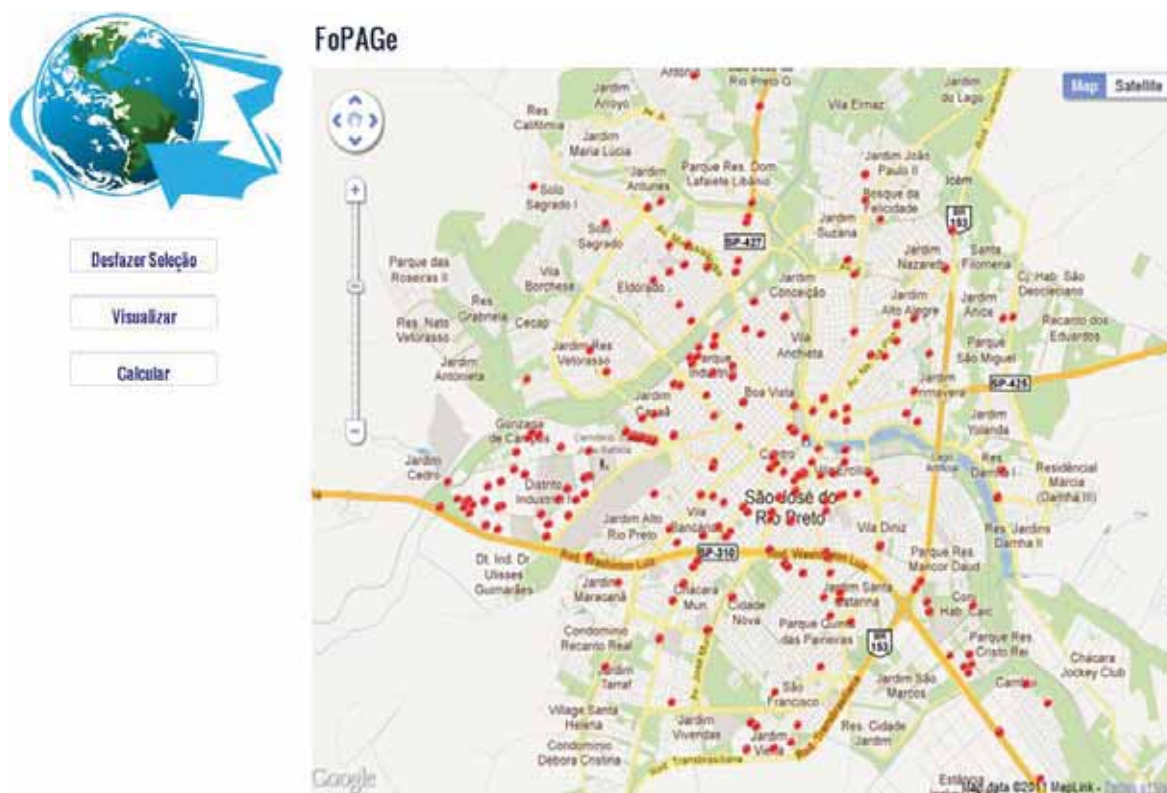


Figura 20 Visualização das coordenadas da tabela posicoes150 na base de dados RioPreto150

Na Figura 21 são exibidos os pontos georreferenciados da segunda base de testes chamada RioPreto1000, onde cerca de 1000 pontos geográficos estão armazenados na tabela posicoes1000. O tempo de visualização foi inferior a um segundo, precisamente variou entre 200ms e 398ms, nos navegadores, *Internet Explorer*, *Firefox* e *Google Chrome*, todos em versões atualizadas.

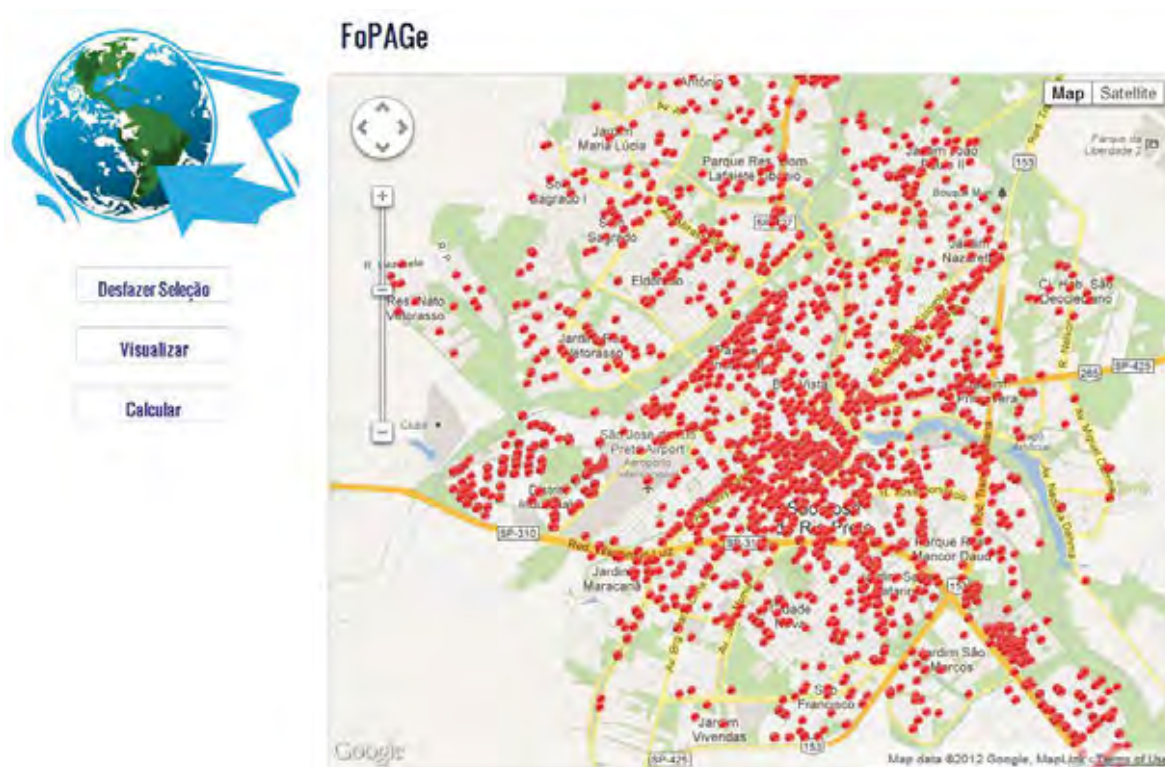


Figura 21 Visualização das coordenadas da tabela posicoes1000 na base de dados RioPreto1000

Para os testes seguintes foram criadas duas bases fictícias chamadas Distritoindustrial e Engenheiroschmidt, que contém pequenas quantidades de dados georreferenciados.

Nas Figuras 22 e 23 são exibidos mapas referentes a testes de deslocamento, onde a FoPAGe centraliza a imagem nos pontos geográficos de maior concentração. Desta forma, mesmo que se tenham pontos *outliers*, a maioria dos pontos será exibida.



Figura 22 Deslocamento para os pontos de maior concentração na base georreferenciada chamada Distritoindustrial

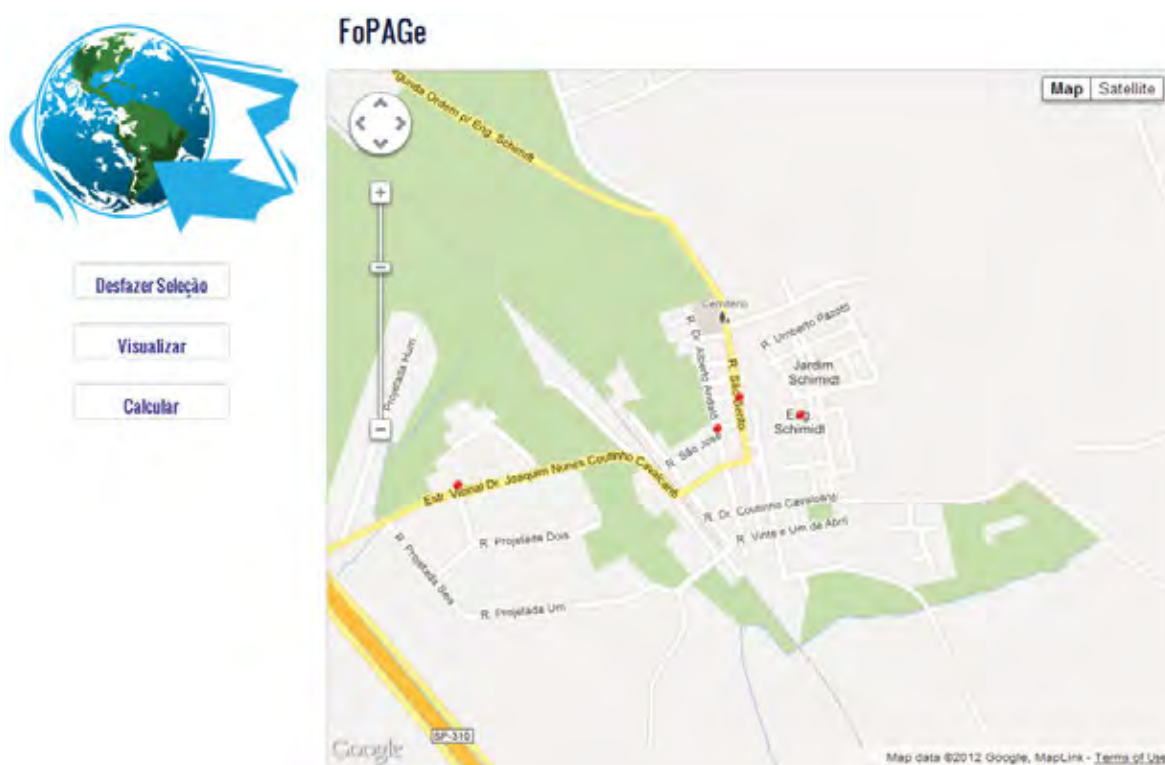


Figura 23 Deslocamento para os pontos de maior concentração na base georreferenciada chamada Engenheiroschmidt

Nestes experimentos e testes observa-se pela forma genérica de visualização, que essa aplicação Web-GIS pode ser utilizada em uma etapa inicial e pré-programada na confecção de sistemas espaciais mais robustos.

4.2. Testes de Predição e Análise

Os testes realizados com a predição de valores e análise estatística utilizam duas bases de dados espaciais, a primeira chamada SIVAT que contém por volta de 70 mil registro de notificações de acidentes do trabalho, sendo que aproximadamente 16 mil registros possuem informações georreferenciadas (Valêncio et al, 2012) (Rodrigues et al, 2011). A segunda base de testes é chamada de Grandedamazonas que contém dados criados aleatoriamente na região de Manaus, com cerca de 100.000 mil registros e, representa um desmatamento fictício. Todos os testes tiveram uma variação entre 170ms e 480ms na execução dos cálculos, nos navegadores, *Internet Explorer*, *Firefox* e *Google Chrome*. As imagens contidas nos testes são provenientes do navegador *Google Chrome*.

4.2.1. Teste de predição de grandes regiões agrupado por dias

Para o início deste teste a base de dados selecionada foi a do SIVAT, a região selecionada contempla todo o município de São José do Rio Preto. No agrupamento temporal os dias foram selecionados como variável de período, assim a predição será encontrada para os quatro dias seguintes.

Como a base de testes armazena registros de acidentes de trabalho, existe menor número de ocorrências em feriados e finais de semana e, para não resultar em pontos *outliers*, neste teste tais períodos foram descartados.

O período selecionado para os cálculos são os dias do mês de janeiro de 2010, com exceção dos feriados e finais de semana. Na Figura 24 pode ser visualizada a região de São José do Rio Preto, onde serão preditos as notificações de acidentes.

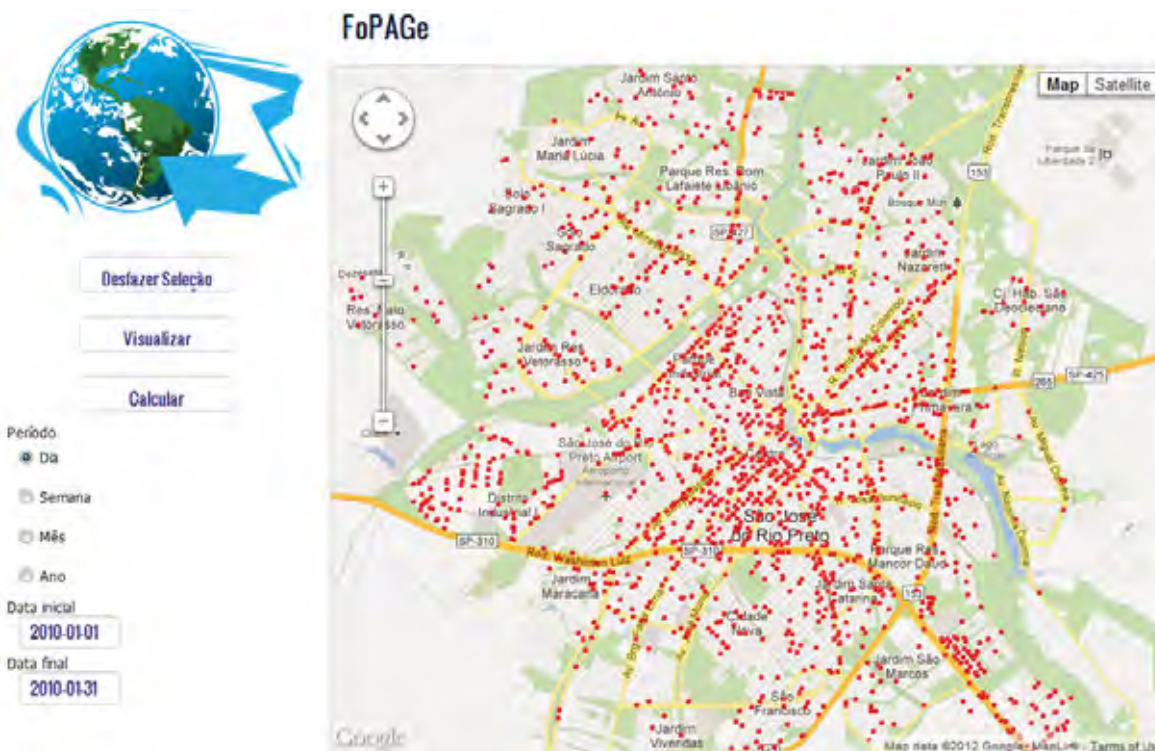


Figura 24 Notificações de acidentes de trabalho na cidade de São José do Rio Preto utilizados para cálculo de predição em janeiro de 2010

Na Figura 25 é exibida a reta do cálculo do método dos mínimos quadrados e os valores reais para essa região, os demais métodos estão ocultos. Ao comparar os resultados das duas séries obtidas, é possível concluir que a tendência dos acidentes é crescente, apesar dos valores reais não representar uma função monótona. A predição para os primeiros dias de janeiro está representada nos últimos quatro valores do eixo horizontal.

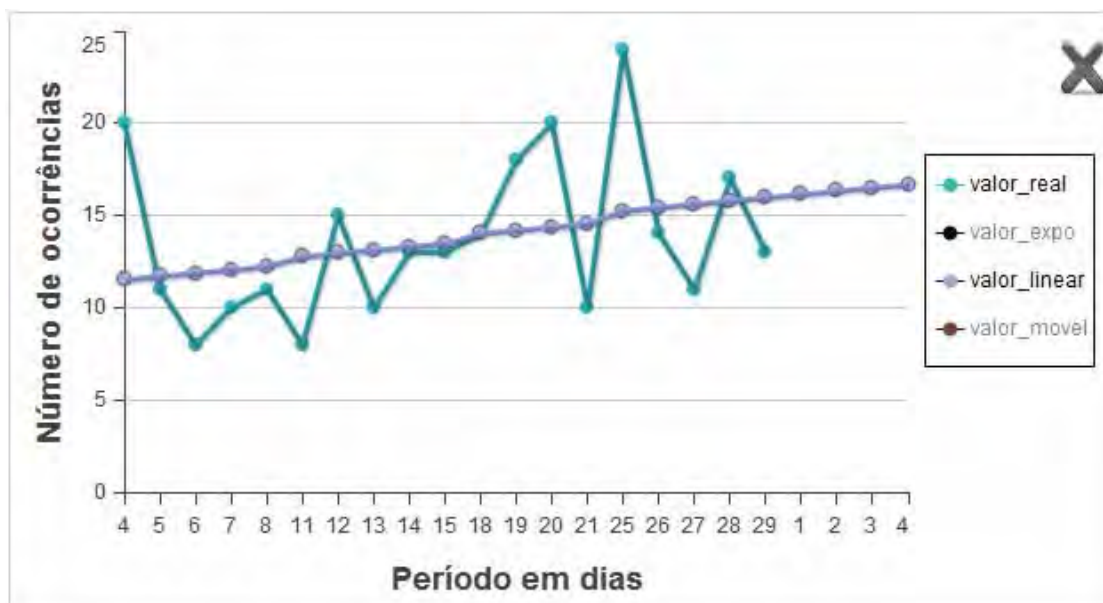


Figura 25 Reta calculada pelo método dos mínimos quadrados comparada com os valores reais para os dias úteis do mês de janeiro de 2010 no município de São José do Rio Preto

O valor da primeira previsão obtido pelo cálculo do alisamento exponencial simples é 14.09 acidentes no primeiro dia útil do mês de fevereiro. Os demais valores utilizam este último como parâmetro e obtêm baixa variação entre eles, como pode ser visualizado na Figura 26, nos quatro últimos valores do período em dia. A curva do alisamento exponencial simples acompanha os valores reais obtidos para calcular essa previsão. Os demais métodos não são mostrados no gráfico.

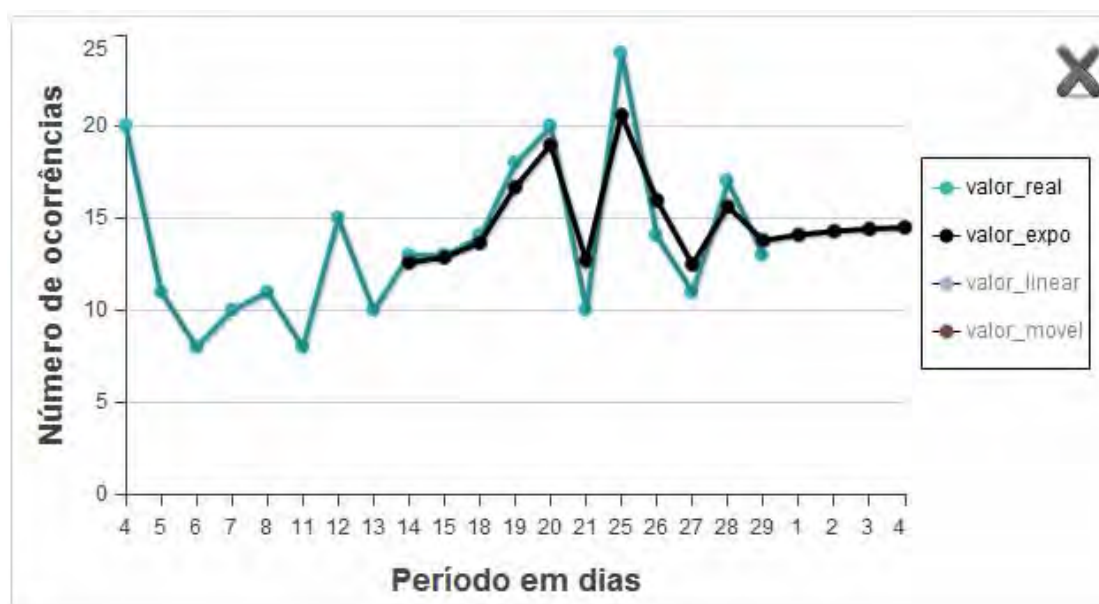


Figura 26 Curva do método do alisamento exponencial e previsões para os primeiros dias de fevereiro de 2010 no município de São José do Rio Preto

O cálculo das médias móveis configura uma boa previsão para todos os quatro dias de fevereiro de 2010, com variação de uma ocorrência, como podem ser observados nos

últimos valores dos períodos exibidos na Figura 27. A maioria dos pontos reais está acima da curva móvel, o que indica que as notificações de acidentes do mês de janeiro de 2010 para São José do Rio Preto estão acima da média esperada. Os outros métodos estão ocultos no gráfico.

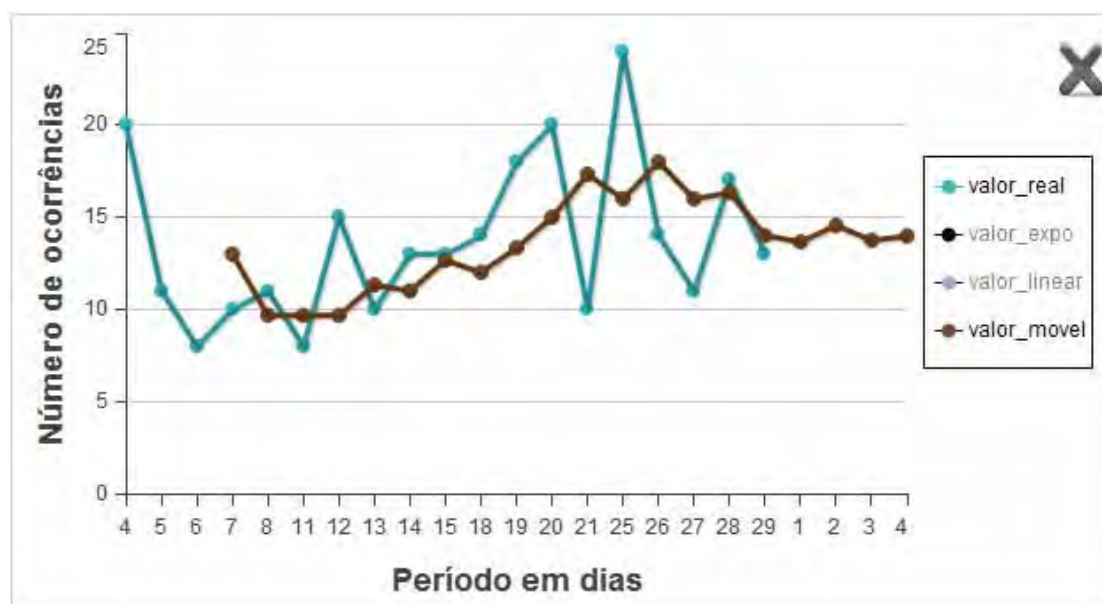


Figura 27 Curva do método das médias móveis e previsões para os primeiros dias de fevereiro de 2010 no município de São José do Rio Preto

As ocorrências observadas durante o período de um mês variaram entre 8 e 25, sendo assim os erros MAPE e RMSE podem ficar com uma porcentagem maior ao final dos cálculos, já que uma única ocorrência representa 5.8% do total.

A previsão pode ser considerada qualificada, mesmo para um intervalo pequeno de dias, pois os valores futuros de todos os métodos se aproximam entre si consideravelmente. O acompanhamento por dia dos erros dos métodos pode ser observado nas Figuras 28 e 29, onde os resultados do alisamento exponencial, dos mínimos quadrados e da média móvel correspondem às cores preta, azul e marrom, respectivamente. A curva dos valores da média móvel aproxima-se dos valores dos mínimos quadrados para ambos os erros, o que significa que as duas previsões estão variando com a mesma intensidade. Como o método do alisamento exponencial acompanha a curva dos valores reais, seu erro é menor, porém só é válido para a primeira previsão.

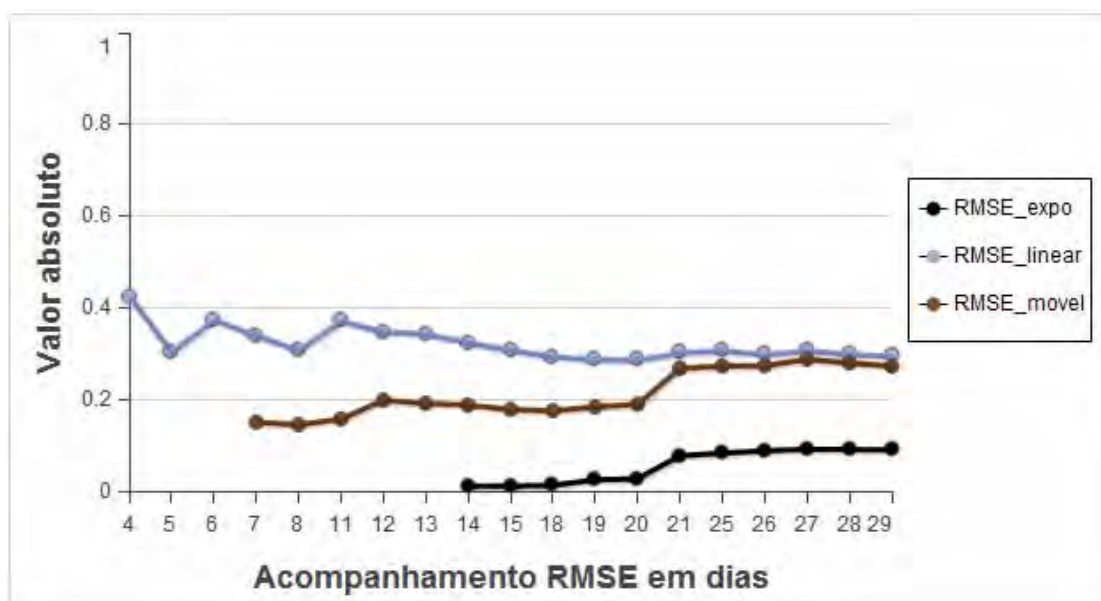


Figura 28 Valores do erro RMSE de todos os métodos para os dias úteis de janeiro de 2010 no município de São José do Rio Preto

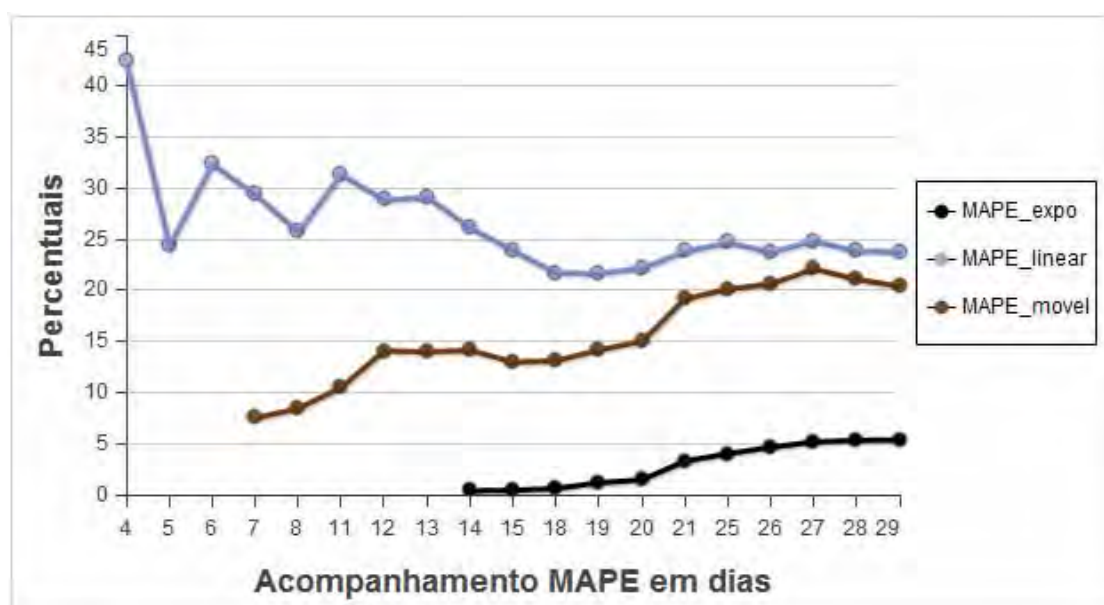


Figura 29 Valores do erro MAPE de todos os métodos para os dias úteis de janeiro de 2010 no município de São José do Rio Preto

Todos os resultados obtidos foram calculados nos dados referentes aos dias úteis do mês de janeiro de 2010, conseqüentemente, as predições indicam futuras ocorrências de acidentes nos primeiros quatro dias úteis de fevereiro de 2010.

Tabela 3

Comparação das previsões dos métodos com o valor real observado para os quatro primeiros dias úteis de fevereiro de 2010

Notificações de acidentes no mês de fevereiro de 2010	Quantidade real observada	Predições do alisamento exponencial	Predições da média móvel	Predições dos mínimos quadrados
1º Dia útil de fevereiro	12	14.09	13.66	16.13
2º Dia útil de fevereiro	12	14.29	14.55	16.31
3º Dia útil de fevereiro	10	14.42	13.74	16.49
4º Dia útil de fevereiro	14	14.50	13.98	16.66

Os resultados obtidos pelas previsões dos métodos podem ser observados na Tabela 3. Em relação ao primeiro dia útil de fevereiro de 2010, o valor da média móvel e do alisamento exponencial obtido está próximo da quantidade real observada, com uma variação de aproximadamente 13% e 17%, respectivamente. Um resultado semelhante foi observado no segundo dia útil, porém com variações maiores. Esses erros, apesar de serem altos, estão dentro do intervalo permitido pelo MAPE. Somente as previsões do terceiro dia útil de fevereiro estão discrepantes do real o suficiente para descartar este resultado. No entanto, no quarto dia útil, as variações das previsões desses métodos possuem erro inferior a 4%.

Na Figura 30 é observado o resultado obtido pelos cálculos de todos os métodos, que são realizados por fórmulas e equações estatísticas diferentes. Apesar de metodologias diferentes nos cálculos, as previsões apresentam estimativas próximas, o que garante maior robustez nos resultados.

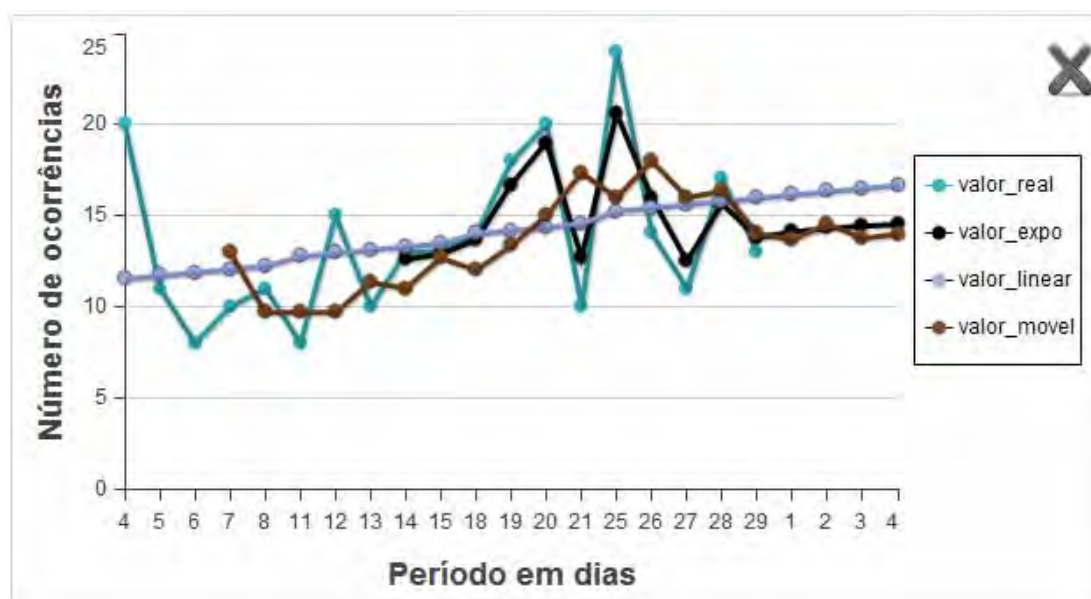


Figura 30 Resultado dos métodos de previsão nos dias úteis de janeiro de 2010 no município de São José do Rio Preto

Esse método e modelo de análise indicam que, no caso da cidade de São José do Rio Preto, poderia ocorrer uma média de 14 acidentes para os primeiros dias de fevereiro de 2010, com tendência de crescimento. Os resultados reais comparados com as previsões demonstraram uma aproximação significativa entre si. Baseado nas previsões de notificações de acidentes pode-se subsidiar ações de gestão política direcionada a fim de minimizar seus efeitos ao longo do tempo.

Devido à pequena quantidade de dados, o resultado da visualização das previsões por dia não é satisfatório.

A Figura 31 exibe os mapas dos valores reais e preditos, por meio do método do alisamento exponencial, para o primeiro dia de fevereiro de 2010.



Notificações reais do primeiro dia útil de fevereiro de 2010



Predição das notificações para o primeiro dia útil de fevereiro de 2010

Figura 31 Visualização dos mapas das notificações de acidentes de trabalho, dados reais e preditos, para o primeiro dia de fevereiro de 2010

4.2.2. Teste de Seleção Parcial Agrupado por Dia

Inicialmente é feita uma seleção da região central de São José do Rio Preto junto com o bairro Parque Industrial. A seleção neste teste é ainda mais específica, pois é feita a seleção dos acidentes diários dessa região, assim os resultados obtidos serão ainda menores que os do teste anterior. Um dos objetivos deste teste é validar resultados ínfimos em grandes quantidades temporais.

São escolhidos todos os dias do mês de agosto de 2010 para os cálculos, na Figura 32 pode ser observada a seleção da região.

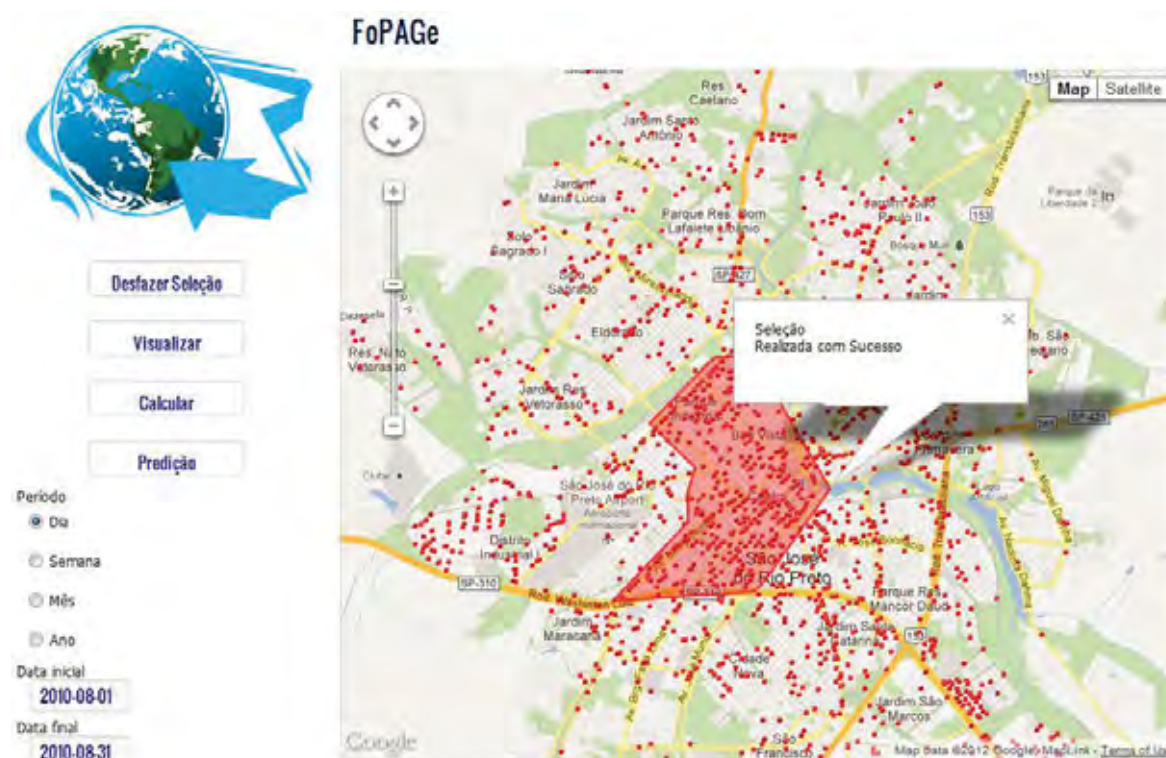


Figura 32 Seleção das notificações do Centro e Parque Industrial de São José do Rio Preto para o mês de agosto de 2010

Na Figura 33 é exibida a reta do cálculo do método dos mínimos quadrados e os valores reais para essa região. Ao comparar os resultados das duas séries obtidas, pode-se concluir que a tendência dos acidentes é decrescente. A predição para os primeiros dias de setembro está representada nos últimos quatro valores do eixo horizontal.

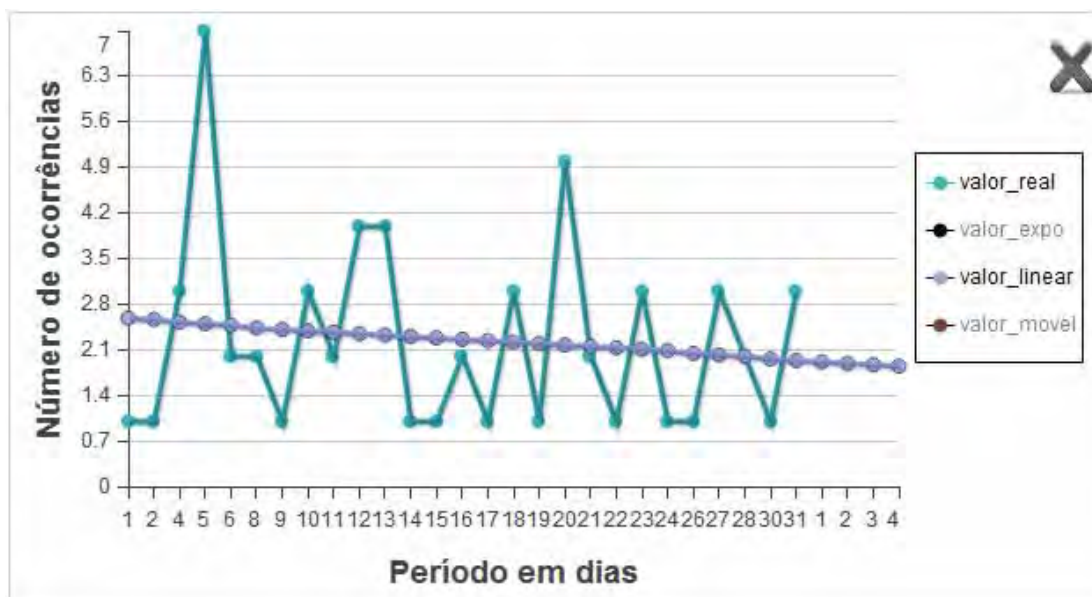


Figura 33 Resultados do método dos mínimos quadrados comparado com os valores reais do mês de agosto de 2010 no Centro e Parque Industrial de São José do Rio Preto

Na Figura 34 é visualizado o gráfico da predição do alisamento exponencial, onde os últimos valores calculados estão próximos do último valor real contido na base de notificações de acidentes. Em virtude de poucas ocorrências, a variação média é pequena para este caso.

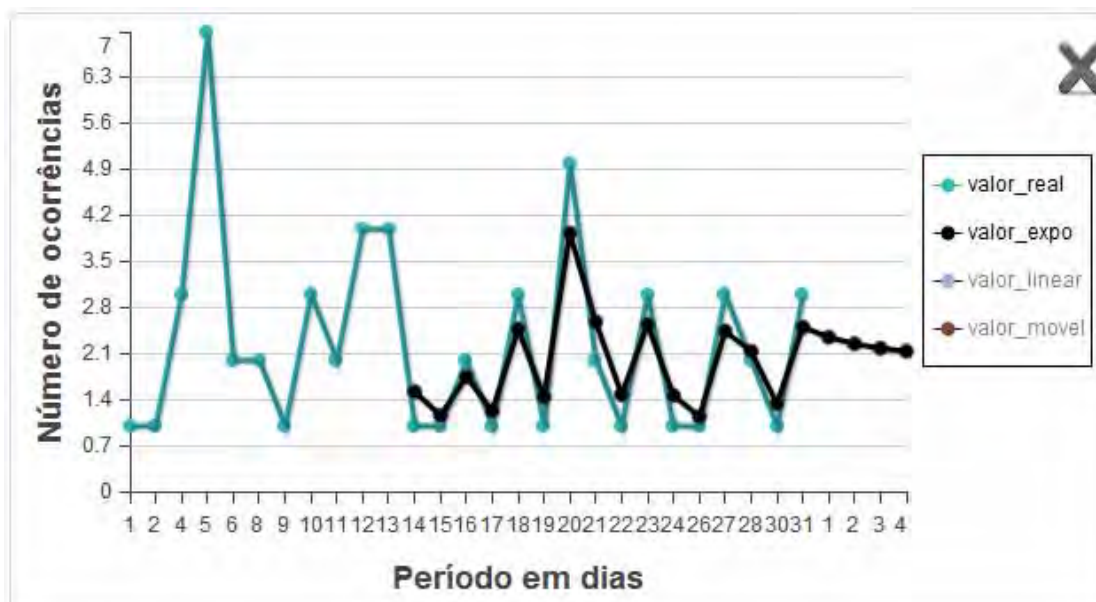


Figura 34 Curva do método do alisamento exponencial e previsões para os primeiros dias de setembro de 2010 no Centro e Parque Industrial de São José do Rio Preto

O cálculo das médias móveis exibe uma predição confiável para os primeiros dias de setembro de 2010, a variação das previsões é inferior a uma ocorrência diária, como observado na Figura 35. A maioria dos pontos reais está abaixo da curva móvel, o que indica que a média de acidentes para o mês de agosto encontra-se abaixo da média.

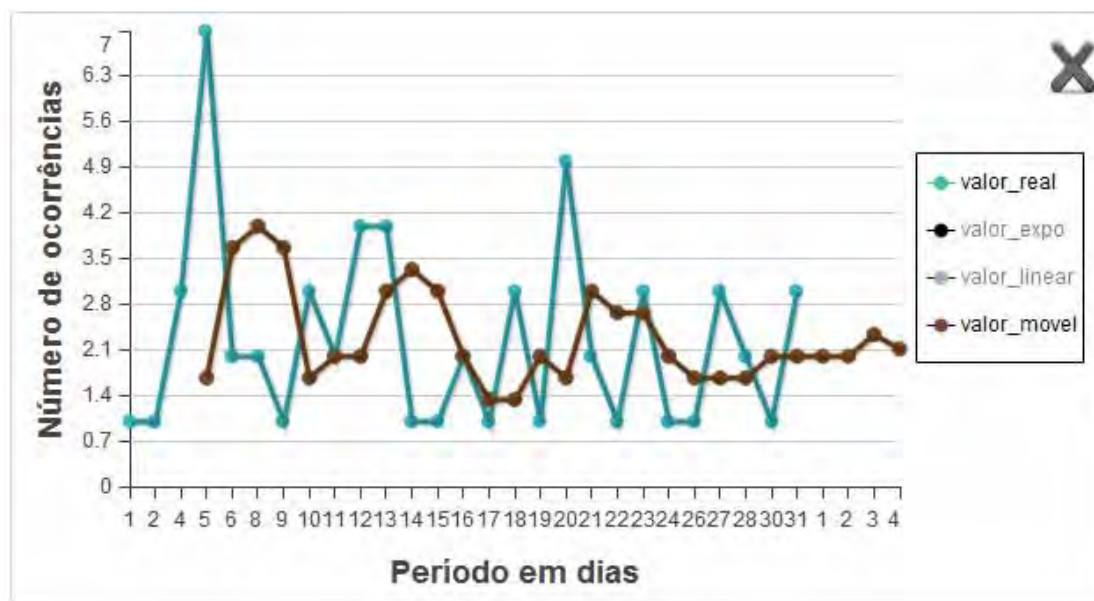


Figura 35 Curva do método das médias móveis e previsões para os primeiros dias de outubro de 2010 no Centro e Parque Industrial de São José do Rio Preto

O valor do erro final MAPE e RMSE ficou elevado para as projeções do alisamento exponencial e da média móvel, isso porque uma única ocorrência pode chegar a representar mais de 50% do total. Nas Figuras 36 e 37 pode-se visualizar o acompanhamento dos erros de todos os métodos para o mês de agosto de 2010.

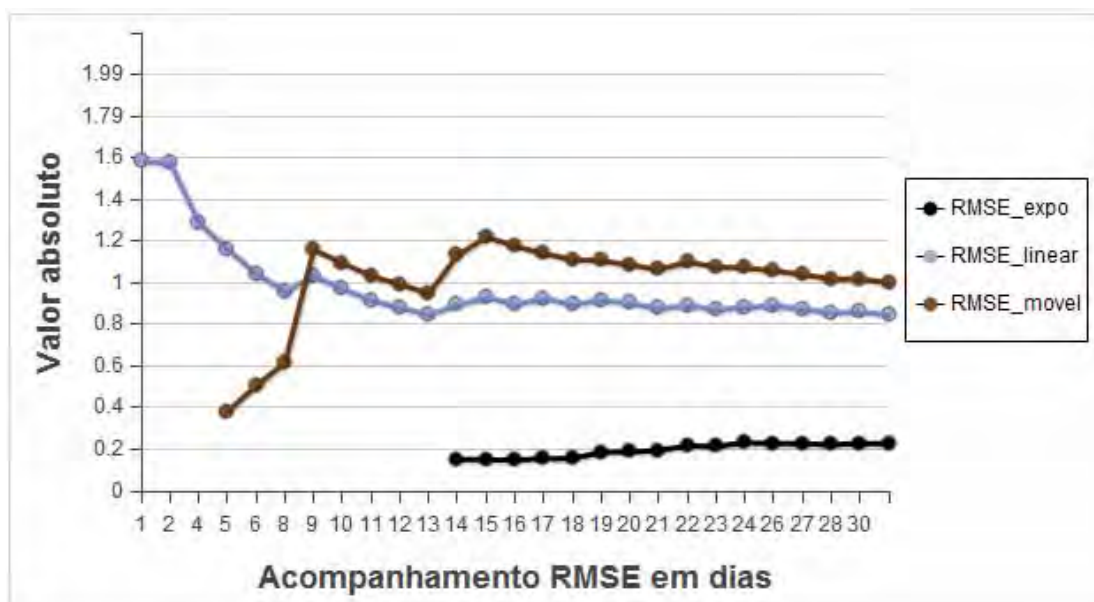


Figura 36 Valores do erro RMSE de todos os métodos para o mês de agosto de 2010 no Centro e Parque Industrial de São José do Rio Preto

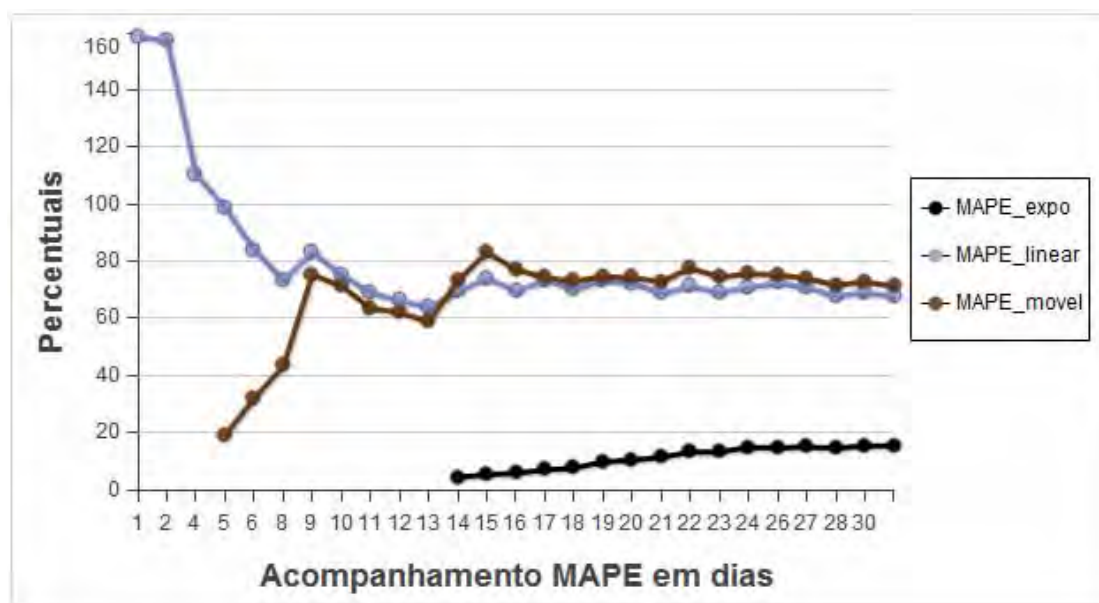


Figura 37 Valores do erro MAPE de todos os métodos para o mês de agosto de 2010 no Centro e Parque Industrial de São José do Rio Preto

Tabela 4

Comparação das previsões dos métodos com o valor real observado para os quatro primeiros dias de setembro de 2010

Notificações de acidentes no mês de setembro de 2010	Quantidade real observada	Predições do alisamento exponencial	Predições da média móvel	Predições dos mínimos quadrados
1º Dia útil de setembro	2	2.34	2	1.91
2º Dia útil de setembro	4	2.24	2	1.89
3º Dia útil de setembro	1	2.18	2.33	1.86
4º Dia útil de setembro	3	2.13	2.11	1.84

Os resultados obtidos pelas previsões dos métodos podem ser observados na Tabela 4. Somente o primeiro dia de setembro de 2010 está validado pelos erros MAPE e RMSE para todos os métodos, pois os valores de todas as previsões comparados com o valor real tiveram erros inferiores a 15%. A previsão do segundo e quarto dia de setembro são validados por meio dos erros dos métodos da média móvel e dos mínimos quadrados, apesar dos altos resultados. No terceiro dia de setembro as previsões ultrapassaram mais de 100% do valor real observado, dessa forma não foram validadas. Foram utilizados todos os dias do mês de agosto de 2010 para a execução deste teste.

O resultado desta metodologia de análise mostra que, nesta região de São José do Rio Preto, existe uma tendência de decréscimo dos acidentes e também que três dos quatro resultados da previsão estão próximos dos reais.

Neste teste não foi possível realizar a visualização das previsões, pois a quantidade de dados preditos não foi significativa.

4.2.3. Testes de Comparação de Predições de Diferentes Regiões Agrupados por Mês

Para este teste foram calculados os resultados da predição sobre a região total de São José do Rio Preto e de seu Distrito Industrial. O período selecionado para os cálculos foi o ano de 2009 e os dados foram agrupados por mês. Assim a predição foi feita para os quatro primeiros meses de 2010. A região de São José do Rio Preto e o Distrito Industrial selecionado para a representação dos cálculos são exibidos na Figura 38.

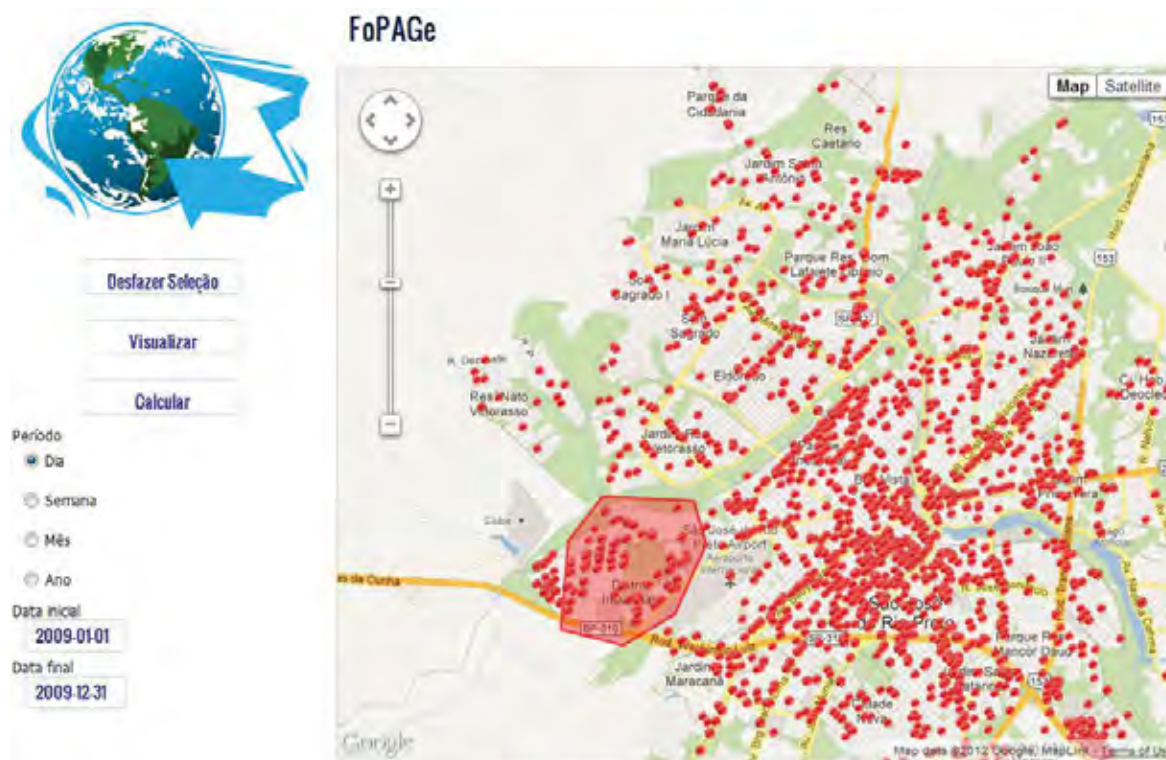


Figura 38 Notificações de acidentes de trabalho em 2009 no Distrito Industrial selecionado e no município de São José de Rio Preto

Na Figura 39 é exibido o gráfico dos valores do cálculo do método dos mínimos quadrados onde a tendência dos acidentes é crescente e entre os meses de maio e dezembro os valores reais representaram uma função monótona. No gráfico ilustrado na Figura 40 o resultado do método apresenta uma reta decrescente, dessa forma, mesmo que o município de São José do Rio Preto apresente um crescimento de notificações de acidentes, a região do Distrito Industrial não influencia esse resultado.

Esse resultado deixa evidente que diferentes áreas podem resultar em diferentes tipos de análises. Ao selecionar várias regiões para os cálculos, é possível encontrar algumas com maior nível de criticidade, onde deveriam ser concentrados os esforços para

uma melhoria na qualidade dos dados. Nas outras regiões, com baixo ou nenhum nível crítico, seria necessário apenas manter esses resultados.

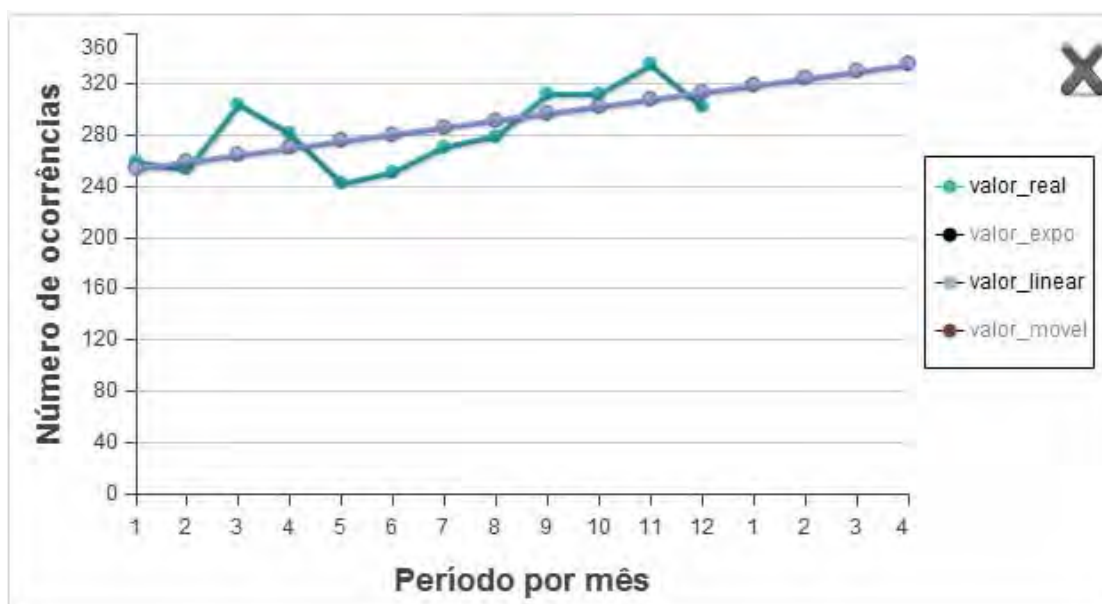


Figura 39 Predição para o primeiro quadrimestre de 2010 no método dos mínimos quadrados no município de São José do Rio Preto

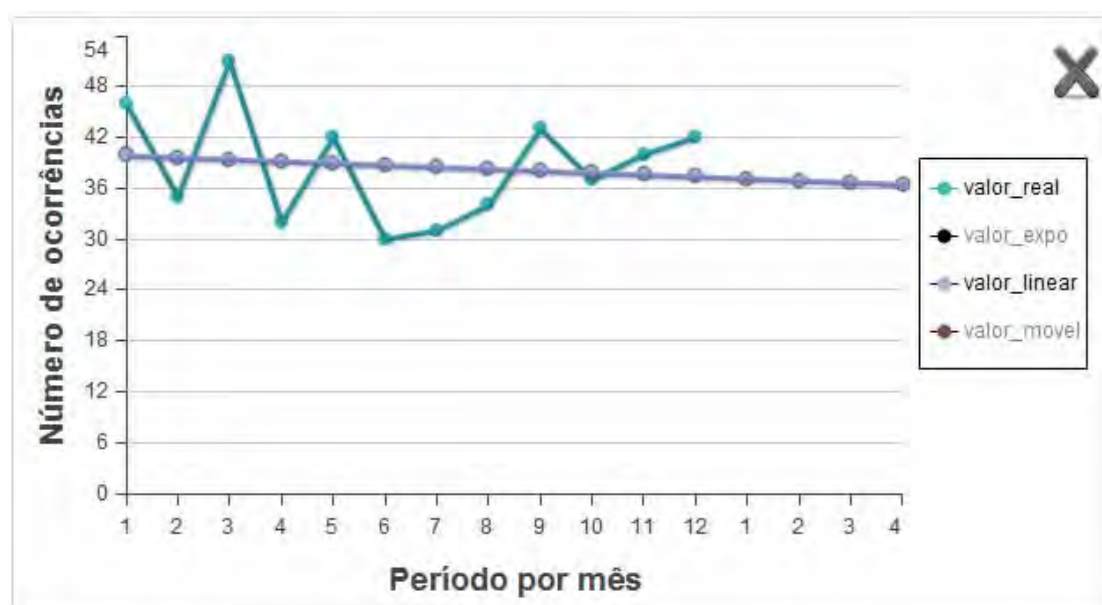


Figura 40 Predição para o primeiro quadrimestre de 2010 no método dos mínimos quadrados no distrito industrial

Na Figura 41 pode ser visualizado o método do alisamento exponencial para o município de São José do Rio Preto onde os quatro últimos pontos dessa curva são referentes às predições do primeiro quadrimestre de 2010. Esses valores representam as estimações das notificações de acidentes do trabalho.

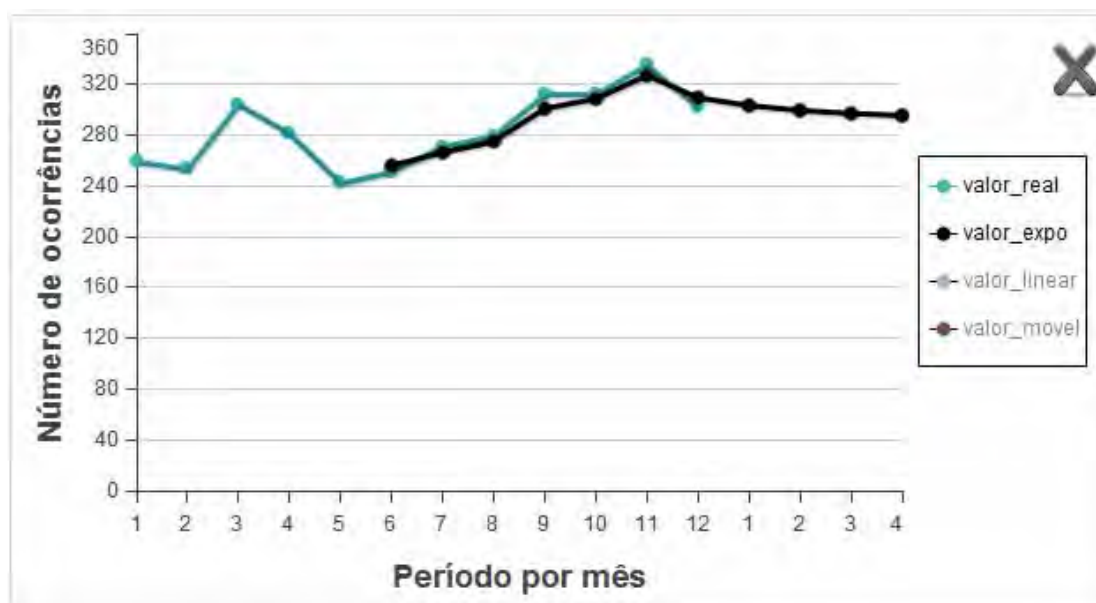


Figura 41 Método exponencial para previsão do primeiro quadrimestre de 2010 de São José do Rio Preto

Na Figura 42 é ilustrado o gráfico do método do alisamento exponencial para o Distrito Industrial, onde os quatro últimos pontos dessa curva são referentes às previsões do primeiro quadrimestre de 2010. Esses valores representam as estimações das notificações de acidentes do trabalho.

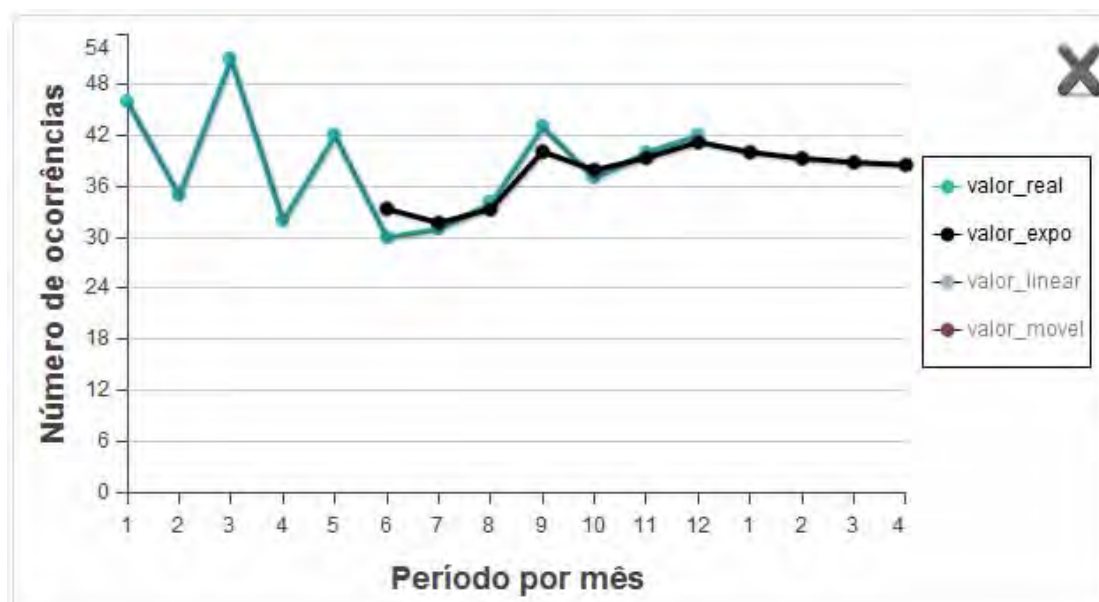


Figura 42 Método exponencial para o Distrito Industrial na previsão do primeiro quadrimestre de 2010

Os resultados obtidos no cálculo da projeção da média móvel para ambas as seleções indicaram uma continuidade nos resultados, os erros MAPE e RSME para esse método tiveram um resultado bastante confiável. Os resultados deste método para o

município de São José do Rio Preto são ilustrados na Figura 43 onde a maioria dos valores reais está acima da curva móvel, consequentemente, o ano de 2009 registrou ocorrências acima da média. Na Figura 44 é exibido o método para o Distrito Industrial onde as ocorrências reais dos dados também ficaram acima da média de ocorrências estimadas, assim nestes períodos deveriam ter sido adotados métodos de prevenção.

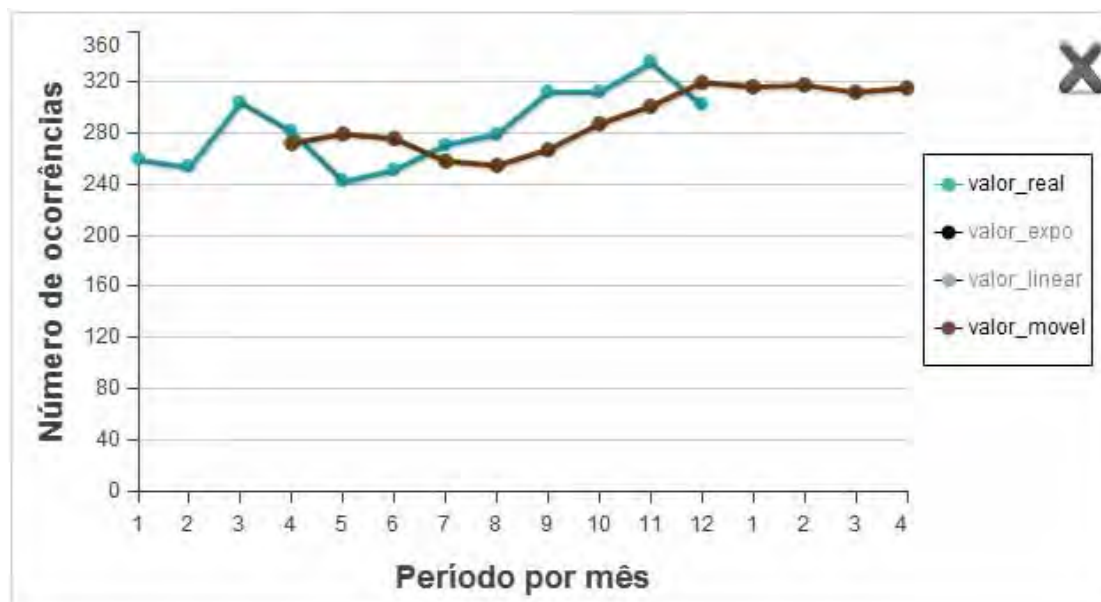


Figura 43 Predição da média móvel para São José do Rio Preto em 2010

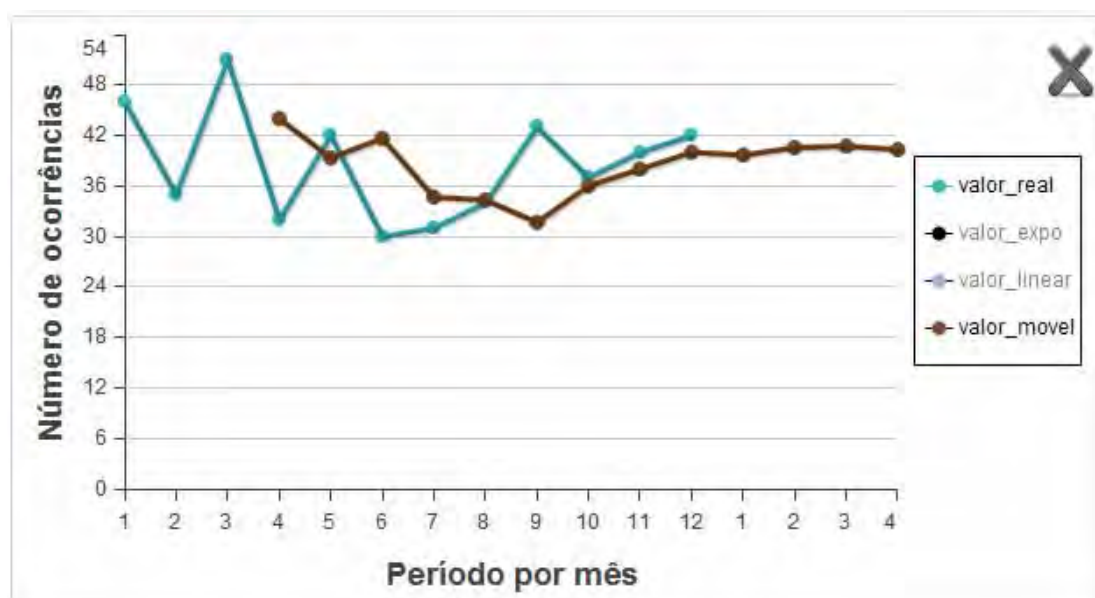


Figura 44 Predição da média móvel para o distrito industrial em 2010

Tabela 5

Comparação dos erros MAPE e RMSE para o município de São José do Rio Preto e o Distrito Industrial para as predições de 2010

Resultados dos erros para os métodos nas regiões	Método do alisamento exponencial	Método das médias móvel	Método dos mínimos quadrados
MAPE – São José do Rio Preto	0.016	0.084	0.075
RMSE – São José do Rio Preto	1.20%	6.69%	6.40%
MAPE – Distrito Industrial	0.040	0.180	0.165
RMSE – Distrito Industrial	2.35%	11.20%	14.62%

Os resultados dos erros dos métodos podem ser observados na Tabela 5. No município de São José do Rio Preto os erros das médias móveis e dos mínimos quadrados aproximaram-se do valor 0.08 para o RMSE e 6.5% em média para o MAPE. O valor do erro MAPE pode significar estatisticamente que, caso não ocorra ação de variáveis externas, os valores reais terão uma variação máxima de 6.5% dos valores preditos. Todos os erros na região do Distrito Industrial tem, aproximadamente, o dobro dos valores obtidos no município.

Tabela 6

Comparação das predições de todos os métodos com o valor real observado para os quatro primeiros meses de 2010 em São José do Rio Preto e Distrito Industrial

Predições das notificações de acidentes por regiões em 2010	Quantidade real observada	Predições do alisamento exponencial	Predições da média móvel	Predições dos mínimos quadrados
Janeiro – São José do Rio preto	298	303.36	316.33	318.71
Fevereiro - São José do Rio preto	266	299.53	317.77	324.14
Março – São José do Rio preto	354	297.11	312.03	329.57
Abril – São José do Rio preto	320	295.58	315.38	335.00
Janeiro – Distrito Industrial	41	40.03	39.66	37.10
Fevereiro - Distrito Industrial	35	39.29	40.55	36.87
Março - Distrito Industrial	43	38.83	40.74	36.65
Abril - Distrito Industrial	47	38.53	40.32	36.42

Na Tabela 6 é mostrada a comparação dos resultados reais e preditos para o município de São José do Rio Preto e para o Distrito Industrial, no primeiro quadrimestre de 2010. Na comparação dos resultados reais e preditos, com exceção do mês de fevereiro, os valores obtidos pela metodologia utilizada na aplicação da FoPAGe aproximam-se dos reais com pouca variação. Em média, estes valores são satisfatórios para a decisão de medidas de prevenção em ambos os casos. Uma justificativa para o mês de fevereiro estar fora dos resultados dá-se ao fato deste mês possuir uma quantidade menor de dias em relação aos outros.

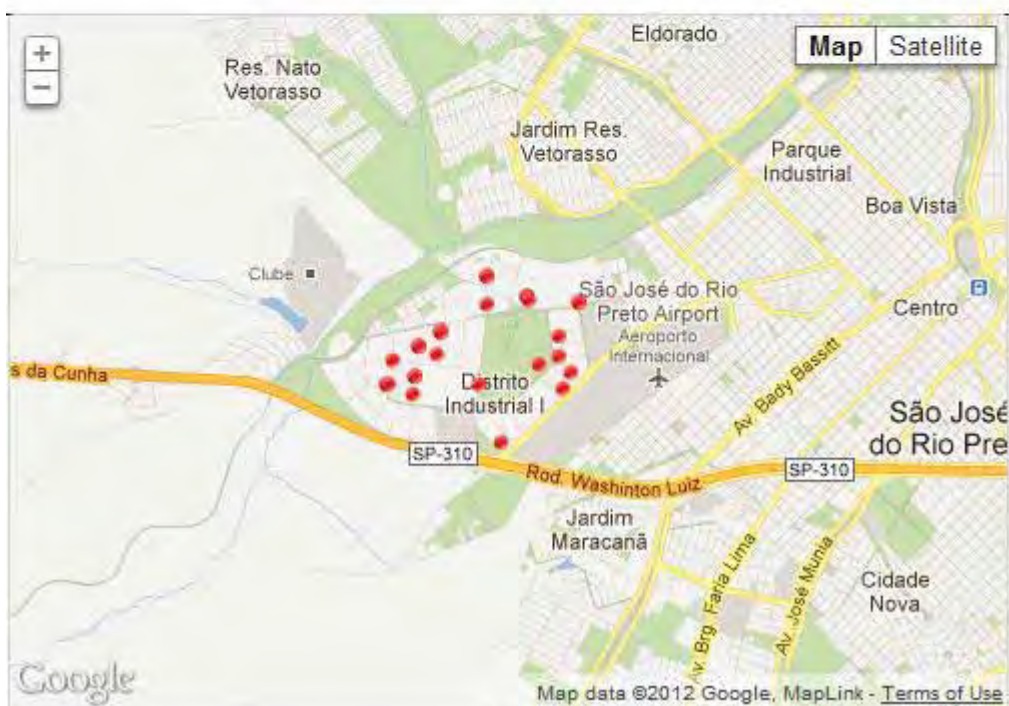
Sobre uma análise mais abrangente poderia ser concluído que a região do Distrito Industrial pode não influenciar tanto no aumento dos acidentes na região total, sendo assim

poderiam ser selecionadas outras regiões para análise, até encontrar uma na qual o crescimento seja constatado e, assim, aplicar medidas de prevenção nesta região. Com isso, certamente, ocorreria uma melhoria nos resultados futuros.

A predição para o mês de janeiro de 2010 pode ser visualizada na Figura 45, onde são ilustrados no primeiro mapa os valores reais e no segundo o resultado do método do alisamento exponencial, no Distrito Industrial.



Dados reais para janeiro de 2010



Dados estimados para janeiro de 2010

Figura 45 Visualização dos mapas das notificações de acidentes de trabalho, dados reais e preditos pelo método do alisamento exponencial, para janeiro de 2010

A predição para o mês de fevereiro de 2010, pode ser visualizada na Figura 46, onde são ilustrados no primeiro mapa os valores reais e no segundo o resultado do método dos mínimos quadrados no Distrito Industrial.



Dados reais para fevereiro de 2010



Dados preditos para fevereiro de 2010

Figura 46 Visualização dos mapas das notificações de acidentes de trabalho, dados reais e preditos pelo método dos mínimos quadrados, para fevereiro de 2010

A predição para o mês de março de 2010 pode ser visualizada na Figura 47 onde são ilustrados no primeiro mapa os valores reais e, no segundo, o resultado do método da média móvel, no Distrito Industrial.



Dados reais para março de 2010



Dados preditos para março de 2010

Figura 47 Visualização dos mapas das notificações de acidentes de trabalho, dados reais e preditos pelo da média móvel, para março de 2010

A predição para o mês de abril de 2010 pode ser visualizada na Figura 48, onde são ilustrados no primeiro mapa os valores reais e no segundo o resultado do método da média móvel, no Distrito Industrial.



Dados reais para abril de 2010



Dados preditos para abril de 2010

Figura 48 Visualização dos mapas das notificações de acidentes de trabalho, dados reais e preditos pelo da média móvel, para abril de 2010

Os resultados dos métodos para a seleção mensal apresentaram uma quantidade significativa de notificações, com isso, foi possível exibir as predições, para todos os meses, nos mapas para comparação.

4.2.4. Testes em Grandes Bases de Dados Agrupados por Anos e Semanas

Nos próximos testes, as informações utilizadas para predição de valores correspondem à base Grandedamazonas. A representação dos dados é fictícia e simboliza uma possível predição em determinada área de desmatamento na região de Manaus. Esse teste tem por objetivo a execução de grandes quantidades de dados e também mostrar que as funções estatísticas estão bem definidas e implementadas. Os dados foram obtidos de forma aleatória, porém, para cada período as distribuições de probabilidade foram iguais. Dessa forma, o caráter da série temporal será constante e o resultado da predição, baseado nos conceitos estatísticos, também deverá ser constante.

Cada ponto que é exibido na Figura 49 identifica uma árvore derrubada, assim se assume que exista uma grande concentração de árvores por toda essa região. Nem todos os pontos podem ser mostrados, pois foi selecionada uma pequena região para visualização.

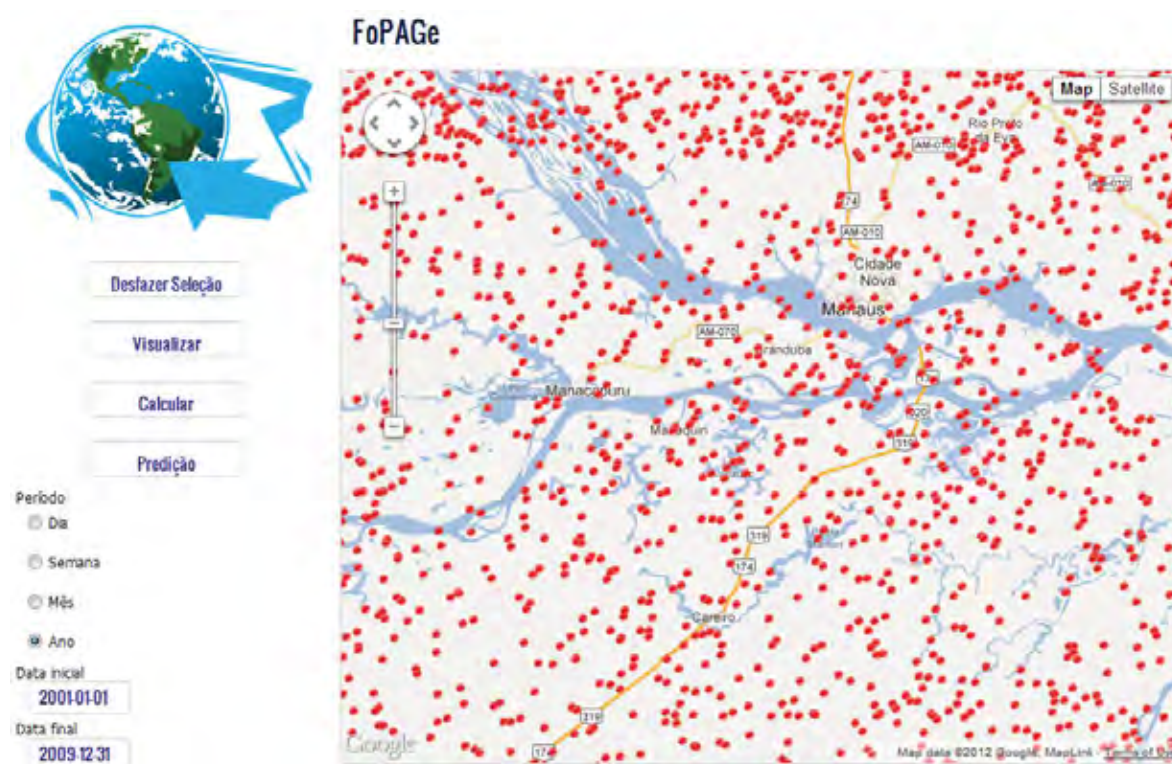


Figura 49 Visualização das coordenadas das árvores derrubadas na região de Manaus

O comportamento dos algoritmos foi testado para as seleções anuais e semanais. Na seleção anual, a variação entre todas as previsões dos métodos e a série dos valores contidos na base foram inferiores a 1%, o que garante a continuidade da série. Na Figura 50 são exibidos os resultados de todos os métodos de previsão e os valores da base Grandeamazonas, onde o período selecionado foi de 2001 a 2009, assim os valores futuros são calculados para os anos de 2010, 2011, 2012 e 2013. Esses resultados sugerem a continuidade da série, o que valida os métodos com base nos conceitos de probabilidade e estatística. Os algoritmos estatísticos implementados apesar de possuírem formas diferentes de cálculo obtêm resultados próximos, sendo condizente com os fatos fictícios gerados.

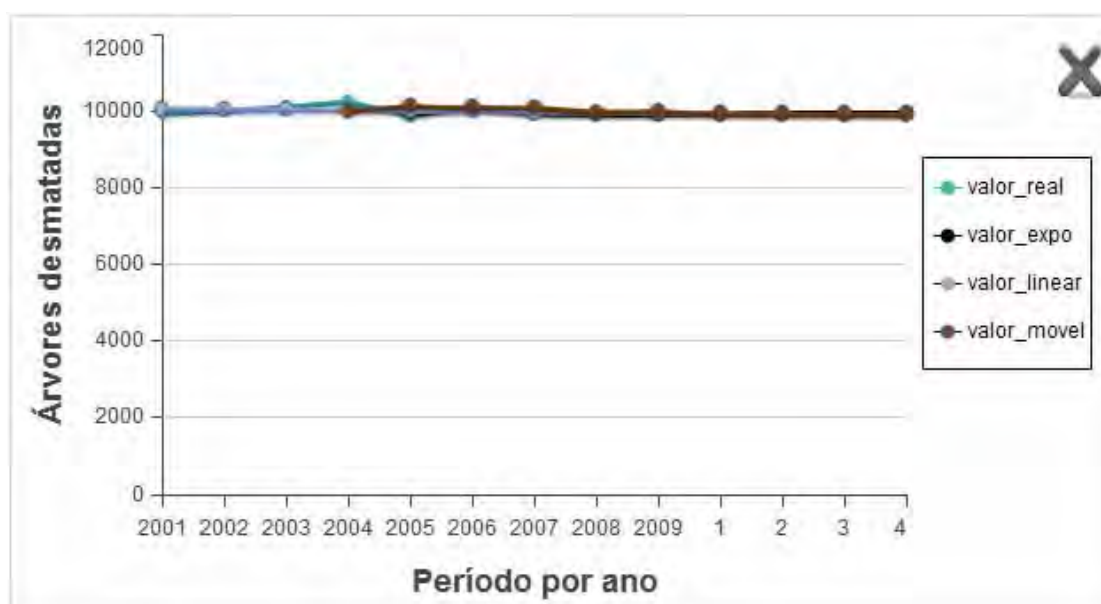


Figura 50 Resultado de todos os métodos de previsão para o desmatamento da região de Manaus durante os anos de 2001 a 2009

O desempenho na visualização dos dados foi, em média, quatro segundos maior do que os realizados anteriormente. Em todos os navegadores ocorreu uma variação entre 4.20s e 5.36s. Na seleção semanal, o período escolhido foi o primeiro semestre do ano de 2003, ocorreu uma variação relativamente maior nas previsões dos métodos em relação aos dados da base, pois alguns meses não contêm semanas completas. Na Figura 51 os métodos podem ser visualizados e, para esse período os resultados apresentam valores constantes. Todos os métodos resultaram em previsões parecidas, para as primeiras semanas do segundo semestre de 2003, o que confirma a precisão dos métodos estatísticos.

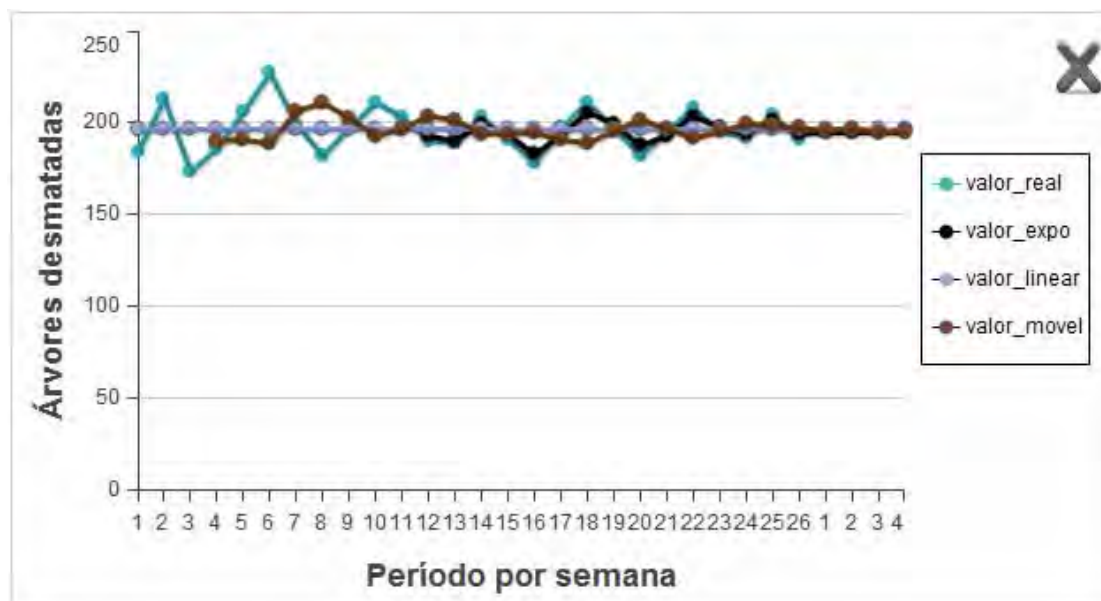


Figura 51 Resultado de todos os métodos para as semanas do primeiro semestre de 2003, referentes ao desmatamento na região de Manaus

Os resultados da predição final exibidos por todos os métodos demonstraram predições factíveis. Assim a metodologia de análise mostrou-se capaz de calcular dados temporais georreferenciados de forma a encontrar critérios nas aplicações de esforços em medidas de prevenção.

4.3. Considerações Finais

Os resultados observados demonstraram a qualidade da análise que se pode obter com a FoPAGe. Os testes realizados contemplaram as predições em pequenas regiões, que possuem quantidades ínfimas de dados e, grandes regiões, com mais de 100.000 dados georreferenciados. A FoPAGe obteve um tempo de execução satisfatório para ambos os casos.

Com uma base de dados concreta e bem definida, o SIVAT, visualizou-se estatísticas reais, que possibilitam a tomada de decisões qualitativas para uma possível melhoria dos resultados futuros.

Capítulo 5 – Conclusões e Trabalhos Futuros

Os estudos voltados à predição espacial têm alta relevância como pode ser constatado nos assuntos tratados neste trabalho. Os artigos e publicações apresentados demonstram uma área recente de pesquisa onde existe um propósito de contribuir na melhoria da qualidade de vida, meio ambiente, saúde, dentre outras áreas relacionadas.

A confecção da FoPAGe inteirou conceitos estatísticos espaço-temporais, computacionais e de probabilidade para auxiliar uma metodologia de análise futurística de dados georreferenciados. De maneira dinâmica, qualquer base georreferenciada pode ser analisada, a metodologia permitiu realizar cálculos de predição em regiões selecionadas, obter gráficos estatísticos espaço-temporais e visualizar uma região predita.

A metodologia utilizada dividiu a construção em módulos, o que possibilitou testes e resultados separados. Essa abordagem obteve resultados estatísticos e visuais satisfatórios. Para validar o contexto genérico do trabalho proposto, diversas bases de dados foram visualizadas nos testes, assim como os pontos preditos por georreferenciamento. A estimativa dos métodos estatísticos de predição aproximou-se dos valores reais na maioria dos casos.

A metodologia para cálculo da localização, dos dados preditos nos mapa, foi capaz de encontrar regiões de criticidade próximas das reais, quando comparada com os dados da base georreferenciada chamada SIVAT. Nas concentrações de ocorrências mais densas, preditas na FoPAGe, poderiam ser adotados métodos de prevenção de acidentes do trabalho, o que garantiria uma melhoria na saúde do trabalhador.

5.1 Trabalhos Futuros

Com base na metodologia usada, na confecção da FoPAGe, nas predições e todos os conceitos utilizados, alguns trabalhos futuros são propostos. Como a implementação de algoritmos de agrupamento espacial para projeção de dados previamente relacionados, pois

assim uma região de interesse poderia ser encontrada de forma automática e a predição poderia resultar em dados correlacionados.

A confecção de interfaces de visualização e predição de diversas bases permitiria metodologias de análises múltiplas, com possibilidade de comparações entre dados geográficos relacionados em diferentes bases de dados.

O estudo de novas funções estatísticas para predição de séries temporais, aplicadas nos dados espaciais, poderia aperfeiçoar o modelo proposto e, possivelmente, a análise dos gráficos e também propor novas técnicas de visualização de predição, para comparar resultados no mapa que contém os pontos georreferenciados, obtidos com a metodologia proposta.

Referências Bibliográficas

Ahamed, S.I; Pezewski, Ale; Pezewski, Al. **Towards framework selection criteria and suitability for an application framework.** Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004. International Conference on, pp 424 – 428, 2004.

Athitsos, V; Wang, H; Stefan, A. **A database-based framework for gesture recognition.** Personal and Ubiquitous Computing, Springer-Verlag London, UK, pp 511 – 526, 2010.

Bonhomme, C.; Aufaure, M.-A. **Mixing icons, geometric shapes and temporal axis to propose a visual tool for querying spatio-temporal databases.** Proceedings of the Working Conference on Advanced Visual Interfaces. Trento, Italy: ACM: 282-289 p. 2002.

Brockwell, P. J., Davis, R. A. **Time Series: Theory and Methods.** Springer Series in Statistics. 2nd ed. 1991. Corr. 6th printing, 1991, XVI, 584 p.

Caiyan, W.; Qing, W.; Jianping, Q. GIS-based regional prediction of landslides in Wanzhou County, the three Gorges Reservoir area. **Geoinformatics, 2011 19th International Conference.** 24-26 June 2011. p.1-5.

Câmara, G; Casanova, M; Davis, C; Vinhas, L; Queiroz, G; **Bancos de Dados Geográficos.** MundoGEO, Curitiba, 2005.

Camargo, E; Druck, S; Monteiro, A, M; Freitas, C, C; Câmara G. **“Mapeamento do risco de homicídio com base na co-krigeagem binomial e simulação: um estudo de caso para São Paulo, Brasil”.** Cadernos de Saúde Pública, 2008.

Celepcikay, O, U; Eick, C, F. **REG²: a regional regression framework for geo-referenced datasets.** In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS '09). ACM, New York, NY, USA, 326-335, 2009.

Chen, F; Chang-Tien Lu, Boedihardjo, A, P. **GLS-SOD: a generalized local statistical approach for spatial outlier detection.** In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '10). ACM, New York, NY, USA, 1069-1078, 2010.

Cong, F. et al. **The logistics network system based on the Google Maps API**. Logistics Systems and Intelligent Management, 2010 International Conference on, 2010. 9-10 Jan. 2010. p.1486-1489.

Cryer, Jonathan D., Chan, Kung-Sik. **Time Series Analysis With Applications in R**. Series: Springer Texts in Statistics. 2nd ed. 2008. Corr. 3rd printing, 2008, XIV, 491 p.

Cuihua, C. et al. **Assessing spatial-temporal variation of heavy metals contamination of sediments using GIS 3D spatial analysis methods in Dexing mines, Jiangxi province, China**. Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International, 2007. 23-28 July 2007. p.1650-1653.

Demiryurek, U; Pan, B; Banaei-Kashani, F; Shahabi, C. **Towards Modeling the Traffic Data on Road Networks**. ACM - Proceedings of the Second International Workshop on Computational Transportation Science, New York, NY, USA, pp 13- 18, 2009.

Deyan, Z.; Guoqing, Z. **Area spatial object co-registration between imagery and GIS data for spatial-temporal change analysis**. Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International, 2007. 23-28 July 2007. p.2597-2600.

Dingju, Z. **Parallel Spatial-Temporal Knowledge Discovery Model**. Fuzzy Systems and Knowledge Discovery, 2009. FSKD '09. Sixth International Conference on, 2009. 14-16 Aug. 2009. p.468-472.

Druck, S; Carvalho, M; Câmara, G; Monteiro, A, M; **Análise Espacial de Dados Geográficos**. Brasília, EMBRAPA, pp 208, 2004.

Dunnavant, S. **Create interactive web illustrations with google maps**. Proceedings of the 38th annual fall conference on SIGUCCS. Norfolk, Virginia, USA: ACM: 267-268 p. 2010.

Frank, R; Ester, M; Knobbe, A. **A multi-relational approach to spatial classification**. In Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '09). ACM, New York, NY, USA, 309-318, 2009.

Giannotti, F; Nanni, M; Pedreschi, D; Pinelli, F. **Trajectory Pattern Analysis for Urban Traffic**. ACM - Proceedings of the Second International Workshop on Computational Transportation Science, New York, NY, USA, pp 43- 47, 2009.

Google. **Exemplo Básico**. Disponível em: <<http://code.google.com/intl/pt-BR/apis/maps/documentation/javascript/examples/map-simple.html>>. Acesso em: 02 agosto 2011a.

Google. **Exemplo Marco Simples**. Disponível em: <<http://code.google.com/intl/pt-BR/apis/maps/documentation/javascript/examples/marker-simple.html>>. Acesso em: 02 ago. 2011b.

Hamdi, I.; Farah, I. R.; Ahmed, M. B. **Analysing spatial-temporal data based on generic GIS**. Computer Systems and Applications, 2005. The 3rd ACS/IEEE International Conference on, 2005. 6-6 Jan. 2005. p.102.

Hao, Z. et al. **Land use information release system based on Google Maps API and XML**. Geoinformatics, 2010 18th International Conference on, 2010. 18-20 June 2010. p.1-4.

Hofstede, R.; Fioreze, T. **SURFmap: A network monitoring tool based on the Google Maps API**. Integrated Network Management, 2009. IM '09. IFIP/IEEE International Symposium on, 2009. 1-5 June 2009. p.676-690.

Junmei, T. **Spatial Analysis and Temporal Simulation of Landscape Patterns in Two Petroleum-Oriented Cities**. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, v. 2, n. 2, p. 54-60, 2009. ISSN 1939-1404.

Kakumoto, S. et al. **Development of spatial temporal geographic information system and risk adaptive regional management information system - toward development of GIS based on Asian culture for disaster prevention**. SICE 2002. Proceedings of the 41st SICE Annual Conference, 2002. 5-7 Aug. 2002. p.352-357 vol.1.

Kiwon, L. **Technical architecture for land monitoring portal using google maps API and open source GIS**. Geoinformatics, 2009 17th International Conference on, 2009. 12-14 Aug. 2009. p.1-5.

Konarski, M.; Zabierowski, W. **Using Google Maps API along with technology .NET**. Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), 2010 International Conference on, 2010. 23-27 Feb. 2010. p.180-182.

Lei, L.; Da-Fang, Z.; Dong-Sheng, Q. **Spatial-Temporal Characteristics of Siping Urban Expansion Based on Remote Sensing and GIS**. Education Technology and Training, 2008. and 2008 International Workshop on Geoscience and Remote Sensing, ETT and GRS 2008. International Workshop on, 2008. 21-22 Dec. 2008. p.241-243.

Liu, F. et al. Urban Expansion and Evolution Prediction of Jinan City Based on Remote Sensing and GIS Technology. **2009 Joint Urban Remote Sensing Event, Vols 1-3**, p. 979-984, 2009.

Liu, P. et al. VDBSCAN: Varied density based spatial clustering of applications with noise. **4th International Conference on Service Systems and Service Management, 2007**. Chengdu, PEOPLES R CHINA. Jun 09-11. p.528-531.

Liu, Y. et al. **A MapReduce Approach to Gi*(d) Spatial Statistic**. Proceedings of the ACM SIGSPATIAL International Workshop on High Performance and Distributed Geographic Information Systems. San Jose, California: ACM: 11-18 p. 2010.

Lijing Zhang; Jing Yi, Management **methods of spatial data based on PostGIS**. Circuits, Communications and System (PACCS), 2010 Second Pacific-Asia Conference on , vol.1, no., pp.410-413, 1-2 Aug. 2010

Lo, E; Binnig, C; Kossmann, D; Tamer, M, Ö; Hon, W. **A framework for testing DBMS features**. The VLDB Journal - The International Journal on Very Large Data Bases, v. 19, i. 2, 2010.

Luo, R.; Shen, Y. **The Design and Implementation of Public Bike Information System Based on Google Maps**. Proceedings of the 2009 International Conference on Environmental Science and Information Application Technology - Volume 02: IEEE Computer Society: 156-159 p. 2009.

Mantilla, R.; Gupta, V. K. **A GIS numerical framework to study the process basis of scaling statistics in river networks**. Geoscience and Remote Sensing Letters, IEEE, v. 2, n. 4, p. 404-408, 2005. ISSN 1545-598X.

Nan, S.; Xiaohong, R.; Jing, A. **RS and GIS based temporal-spatial variation and multi-factor spatial analysis on nonpoint source pollution**. Geoinformatics, 2010 18th International Conference on, 2010. 18-20 June 2010. p.1-4.

Pejic, A.; Pletl, S.; Pejic, B. **An expert system for tourists using Google Maps API**. Intelligent Systems and Informatics, 2009. SISY '09. 7th International Symposium on, 2009. 25-26 Sept. 2009. p.317-322.

Rodrigues, H. P.; et al. **Sistema Computacional para Análise de Notificações de Acidentes de Trabalho por Meio de Recursos Georreferenciados**. In: 6º Congresso de Extensão Universitária da Unesp, 2011, Águas de Lindóia.

Rousseaux, F.; Lhoste, K. **Rapid Software Prototyping Using Ajax and Google Map API**. Proceedings of the 2009 Second International Conferences on Advances in Computer-Human Interactions: IEEE Computer Society: 317-323 p. 2009.

Santos, E. S; Zeilhofer, P; Ribeiro, A. L. M; Santos, M. A; Miyazaki, R. D. **Geoprocessamento e Regressão Logística para construção de modelos espaço-temporais de habitats de Anopheles darlingi na área de influência da APM-Manso - MT**. 14º Simposio Brasileiro de Sensoriamento Remoto, Natal, 2009.

Schulze, T.; Wytzisk, A.; Simonis, I. **Distributed spatio-temporal modeling and simulation**. Simulation Conference, 2002. Proceedings of the Winter, 2002. 8-11 Dec. 2002. p.695-703 vol.1.

Sung-Soo, K. et al. **A unified visualization framework for spatial and temporal analysis in 4D GIS**. Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International, 2003. 21-25 July 2003. p.3715-3717 vol.6.

Tao, W.; Jin-Ping, L.; He, Y. **Analyzing Economic Spatial-Temporal Disparities at County Level in Jiangsu Based on ESDA-GIS**. Computational and Information Sciences (ICCIS), 2010 International Conference on, 2010. 17-19 Dec. 2010. p.1224-1227.

Teresco, J. D. **Using the Google Maps API with highway mapping data as a pedagogical tool: demonstration**. J. Comput. Sci. Coll., v. 26, n. 6, p. 58-60, 2011. ISSN 1937-4771.

Vadeerat, R; Christoph, F. **Change analysis in spatial datasets by interestingness comparison**. SIGSPATIAL Special, pp 33 – 38, 2009.

Valencio C.R., et al. **MR-Radix: a multi-relational data mining algorithm**. Human-centric Computing and Information Sciences 2012, p 2-4.

Wen-Hui, Z. et al. **Temporal and Spatial Variation of Urban Airborne Inhalable Particle and It's Influence Factor Analysis using GIS & RS**. Geoinformatics, 2009 17th International Conference on, 2009. 12-14 Aug. 2009. p.1-5.

Wonnacott T. H., Wonnacott R. J.. **Introductory Statistics for Business and Economics**. New York: John Wiley and Sons, Inc.; 5th Edition. 1990, 736 p.

Wood, J. et al. Interactive Visual Exploration of a Large Spatio-temporal Dataset: Reflections on a Geovisualization Mashup. **IEEE Transactions on Visualization and Computer Graphics**, v. 13, n. 6, p. 1176-1183, 2007. ISSN 1077-2626.

Yan, H.; Liqin, Z.; Pusheng, Z. A Framework for Mining Sequential Patterns from Spatio-Temporal Event Data Sets. **Knowledge and Data Engineering, IEEE Transactions on**, v. 20, n. 4, p. 433-448, 2008. ISSN 1041-4347.

Yang, Y. et al. **Design and Implementation of Campus Spatial Information Service Based on Google Maps**. Management and Service Science, 2009. MASS '09. International Conference on, 2009. 20-22 Sept. 2009. p.1-4.

Zeilhofer, P; Oliveira, I. M; Klemp, S. M; Santos, E. S; Dores, E. F. G. C. **Sig e regressão logística para mapeamento de risco de contaminação por pesticidas nos mananciais superficiais da bacia do alto rio das mortes-MT**. xiii simpósio brasileiro de sensoriamento remoto, florianópolis, pp 3623-3630, 2007.

Zhao, X. et al. **GIS-Based Spatial and Temporal Analysis of Regional Water-Logging Confluence**. Computational Sciences and Optimization (CSO), 2011 Fourth International Joint Conference on, 2011. 15-19 April 2011. p.886-889.

Zhao, Y.-W. et al. **GIS-Based Analysis on Temporal-Spatial Changes of Land Use in Liao River Basin**. Bioinformatics and Biomedical Engineering , 2009. ICBBE 2009. 3rd International Conference on, 2009. 11-13 June 2009. p.1-4.

Zheng, H. et al. Clustering Algorithm Based on Characteristics of Density Distribution. **2nd Ieee International Conference on Advanced Computer Control (Icacc 2010), Vol. 2**, p. 431-435, 2010.

Zhou, W. Z.; Liu, G. H. **Simulating Temporal and Spatial Changes of Soil Salinity in Yellow River Delta, China Using GIS**. Multimedia Technology (ICMT), 2010 International Conference on, 2010. 29-31 Oct. 2010. p.1-4.