

Uso de Regras de Associação *Fuzzy* e Técnicas de Exploração Espacial na Mineração dos dados Sócio-Ambientais do Projeto PIATAM, Amazônia Central

Fábio Roque da Silva Moreira¹

Fernando Pellon de Miranda²

Alexandre Gonçalves Evsukoff¹

¹ Universidade Federal do Rio de Janeiro – UFRJ/COPPE
Caixa Postal 68.552 - 21949-900 – Rio de Janeiro - RJ, Brasil
fmoreira@labsar.coppe.ufrj.br
alexandre.evsukoff@coc.ufrj.br

² Centro de Pesquisas e Desenvolvimento da Petrobras -CENPES
Avenida Horácio Macedo, 950, Cidade Universitária, Ilha do Fundão, CEP 21941-915 – Rio de Janeiro-RJ, Brasil
fmiranda@petrobras.com.br

Abstract. The present study proposes a methodology based on fuzzy association rules and spatial analysis for knowledge discovery, expressed through linguistic rules, of the information pertaining to ichthyofauna and limnology themes of the PIATAM Project. An algorithm with the ability to deal with numerical and categorical data and to associate objects with geographical properties through spatial analysis was developed. The methodology included pre-processing (data selection, cleaning, summarizing and transformation), data mining, visualization and evaluation. In the pre-processing step, the themes were spatially associated according to the criteria of proximity. Analyses of distance of the community centers were also incorporated. In the transformation stage, the variables were partitioned in linguistic values. The data mining algorithm uses one-dimensional candidates (linguistic values) to produce others of higher dimensions. The last step consisted of the construction of association rules through the application of Boolean operators. The strong rules were then geocoded through the fuzzy membership of the linguistic values. The performance of this mining study was considered reasonable, based in specific rules. The spatial visualization of the strong rules through the geocoded fuzzy membership of the linguistic values showed to be useful, inasmuch as it allowed a better inspection of the strength of the rule in the different communities.

Palavras-chave: knowledge discovery in spatial databases, GIS, fuzzy sets, rainforest ecosystems, extração de conhecimento em banco de dados espacial, SIG, conjuntos nebulosos, ecossistemas de floresta tropical úmida.

1 – Introdução

A Base de Dados Integrada (BDI) do PIATAM constitui uma fonte de informação e de conhecimento sócio-ambiental para estudos científicos e de monitoramento, bem como de insumo para processos de tomada de decisão, relacionados com as atividades de transporte de óleo e gás no trecho do Rio Solimões entre Coari e Manaus, na Amazônia Central (Tocantins, 2006). Paradoxalmente, esse grande volume de dados de diversas áreas do conhecimento, com seus atributos espaço-temporais, também representa um desafio para que o especialista consiga fazer uma análise precisa e abrangente. Uma forma de extrair conhecimento da BDI, considerando tais características, é a aplicação de métodos de mineração de dados (*data mining*) combinados com técnicas de exploração espacial. Tal abordagem confere ao intérprete uma grande capacidade de lidar com massivos bancos de dados espaciais.

Nesse contexto, o presente estudo empregou regras de associação *fuzzy* e técnicas de exploração espacial para adquirir conhecimento de modo semi-automático da base de dados BDI PIATAM. Inicialmente, foi efetuada a seleção dos dados de interesse referentes aos temas ictiofauna e limnologia, seguida de sua limpeza, com o intuito de eliminar valores espúrios e inconsistentes. Na etapa de pré-processamento, técnicas de sumarização e de análise exploratória espacial foram aplicadas segundo critérios espaço-temporais. A seguir, os dados foram reformatados para um padrão binário, no caso dos atributos categóricos, e em conjuntos *fuzzy*, para os atributos numéricos. Na etapa de mineração de dados, o algoritmo de

regras de associação *fuzzy* adotado seleciona primeiramente os conjuntos de itens (*itensets*) frequentes, *i.e.* as variáveis linguísticas com frequência maior que o suporte mínimo especificado. Em seguida, constrói as regras de associação linguísticas através da combinação dos conjuntos de itens frequentes.

Por fim, as regras geradas foram espacializadas em um sistema de informações geográficas (SIG). A visualização espacial dos valores de pertinência das variáveis linguísticas que compõem a regra permitiu uma inspeção mais adequada de sua efetividade nas diferentes comunidades estudadas pelo projeto (Moreira, 2010).

2 – Projeto PIATAM: Localização e Natureza dos Dados

O Projeto PIATAM (Inteligência Socioambiental Estratégica da Indústria do Petróleo e Gás no Amazonas) foi idealizado para fornecer subsídios objetivando minimizar possíveis impactos ambientais nas rotas fluviais de transporte utilizadas pela indústria do petróleo na Amazônia. A principal meta do projeto é a caracterização e monitoramento socioambiental, através de séries históricas de dados no trecho do Rio Solimões onde atua a Petrobras (Rivas et al., 2007). Nessa região, foi construído o gasoduto Coari-Manaus (397 km de extensão), que liga o Terminal da Petrobras, em Coari, até a REMAN (Refinaria de Manaus) (Figura 1).

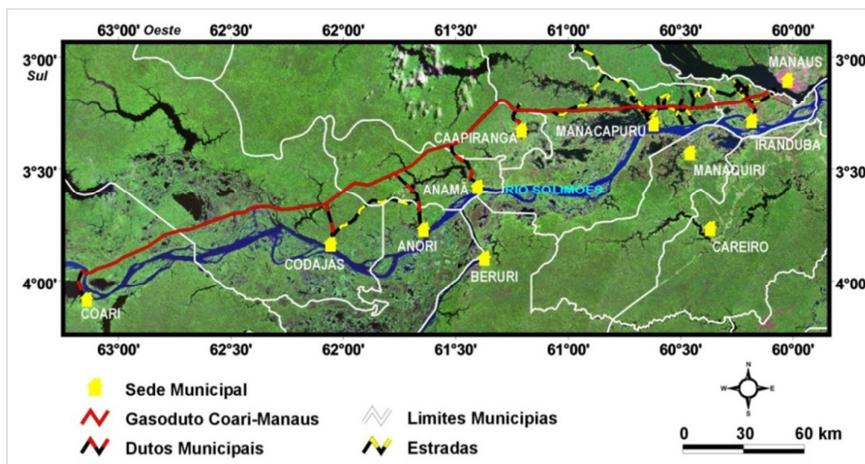


Figura 1 - Imagem Landsat-5 TM, composição 5R4G3B, com a região que abrange os municípios atravessados pelo gasoduto Coari-Manaus. Fonte: Beisl (2009).

O PIATAM é um projeto de natureza multi-institucional e interdisciplinar (www.piatam.ufam.edu.br). A equipe executora inclui pesquisadores das mais diversas áreas do conhecimento, que periodicamente (quatro vezes ao ano) vão a campo coletar informações sobre a fauna, a flora, o meio físico e a socioeconômica na área investigada. As visitas ao campo coincidem com as fases do ciclo hidrológico amazônico (seca, enchente, cheia, e vazante).

Os dados obtidos na região investigada são distribuídos nos treze temas listados a seguir: Arqueologia, Crustáceos, Entomologia, Flora, Herpetofauna, Ictiofauna, IDH-C, Limnologia, Macrófitas Aquáticas, Malária, Sócioeconomia, Solo, e Varzeometria. Esses dados, conjuntamente com arquivos de imagens individuais de satélites e aerotransportadas, mosaicos georreferenciados e dados temáticos (organizados pelo IBGE), compõem uma base de dados integrada (BDI PIATAM).

A região se caracteriza por um complexo sistema de rios, lagos, planícies inundáveis (várzeas), canais, diques e ilhas, denominado de sistema Solimões-Amazonas. As áreas de várzea são sujeitas à alternância entre períodos seca e cheia, o que resulta em modificações drásticas nas condições ambientais.

3 - Regras de Associação

Uma regra de associação (Agrawal et al., 1993), é a associação direta $A \rightarrow B$ com $A, B \subseteq R$ e $A \cap B = \emptyset$. O significado expresso por tal regra é que A normalmente ocorre associado a B . O lado esquerdo do símbolo ' \rightarrow ' é o antecedente da regra e o lado direito é a consequência. De modo a encontrar regras de associação "interessantes" ou "fortes" em um banco de dados, uma regra $A \rightarrow B$ é geralmente avaliada de acordo com algumas medidas, que não podem ocorrer abaixo de um limiar especificado. Uma das mais comumente empregadas é o suporte, que define o número de objetos que satisfazem ambos A e B , i.e.,

$$\text{sup}(A \rightarrow B) \stackrel{\text{def}}{=} |\{x \in D | A(x) \cap B(x)\}| \quad (1)$$

onde $|\cdot|$ denomina a cardinalidade. O suporte pode ser também definido pela proporção em lugar do número absoluto de objetos ($n = |D|$). O conjunto de itens que satisfaz o suporte mínimo especificado é denominado conjunto de itens frequentes. O procedimento para minerar as regras de associação consiste em encontrar todos os conjuntos de itens frequentes na base de dados, para então definir as regras de associação "fortes" entre os conjuntos de itens frequentes.

A outra medida de avaliação é a confiança, que corresponde à proporção das aplicações corretas da regra e pode ser entendida como a probabilidade condicional, $\text{Prob}(B|A)$, de que a consequência é verdadeira dado que o antecedente esta presente (Dubois et al., 2006), conforme:

$$\text{conf}(A \rightarrow B) \stackrel{\text{def}}{=} \frac{|\{x \in D | A(x) \cap B(x)\}|}{|\{x \in D | A(x)\}|} \quad (2)$$

As regras de associação são geradas a partir desses conjuntos de itens frequentes, sendo consideradas apenas as que satisfazem o suporte e a confiança mínimos especificados.

3.1 - Regras de associação fuzzy

Na abordagem *fuzzy*, uma regra de associação *fuzzy* é entendida na forma $A \rightarrow B$, onde A e B são, respectivamente, conjuntos de atributos *fuzzy* A^i e B^i . O domínio dos atributos numéricos é definido através de um conjunto *fuzzy* $F^1 \dots F^n$, onde cada um de seus elementos pode ser visto como um atributo com valor $[0, 1]$, tal que $A^i(x) = F^i(A(x))$.

Com relação às medidas de significância das regras de associações *fuzzy*, Dubois et al. (2006) lembram que várias generalizações das medidas de qualidade de (1) e (2) têm sido propostas na literatura. Modelando o produto cartesiano e a cardinalidade de um conjunto *fuzzy*, como proposto em (1) e (2), deve-se obter:

$$\text{sup}(A \rightarrow B) \stackrel{\text{def}}{=} \sum_{x \in D} A(x) \otimes B(x) \quad (3)$$

$$\text{conf}(A \rightarrow B) \stackrel{\text{def}}{=} \sum_{x \in D} A(x) \otimes B(x) / \sum_{x \in D} A(x) \quad (4)$$

A escolha mais comum para a *t-norma* \otimes é o mínimo ou o produto.

3.2 - Regras de Associação Espaciais

Nas regras de associação espaciais, $A \rightarrow B$, pelo menos um dos elementos (A ou B) é um predicado espacial (Koperski, 1995). Enquanto nas regras de associação clássicas cada linha de um conjunto de dados é uma transação e cada coluna é um item, na mineração de regras de associação espaciais cada linha corresponde a uma instância (*e.g.* Manaus) de um objeto espacial (*e.g.* cidade), denominado de feição do tipo alvo, e suas colunas são predicados.

Cada predicado tem relação com um atributo não-espacial (*e.g.* população) da feição tipo alvo ou com um predicado espacial. Um predicado espacial corresponde a uma feição do tipo relevante que possui alguma relação espacial (*e.g.* distância, ordem, topologia) com instâncias da feição tipo alvo (*e.g.* contém_fabricas) (Bogorny, 2006). As relações no terreno são

computadas através das associações espaciais entre todas as instâncias t (e.g. Manaus) das feições do tipo alvo T (e.g. cidade) e todas as instâncias o (e.g. Rio Negro) de cada feição do tipo relevante O (e.g. rio). Verifica-se, assim, o conjunto de feições do tipo relevante S (e.g. rio, domicílio) que tenha qualquer relação espacial (e.g. toca contém, perto, longe) com T .

4 - Metodologia

Tendo em vista que o presente estudo trata basicamente de um processo de aquisição de conhecimento, foram contempladas as etapas de seleção, pré-processamento, transformação e mineração dos dados, bem como de visualização e avaliação dos resultados.

4.1 - Seleção

Para esse primeiro estudo de mineração, foram usadas apenas as variáveis levantadas em campo pelos temas ictiofauna e limnologia, bem como algumas feições espaciais temáticas da base cartográfica do IBGE contidas na BDI. Os dados de campo e as bases temáticas encontram-se armazenadas em um repositório dados de acesso remoto controlado via web.

4.2 - Limpeza e Estatística

O intuito da limpeza foi identificar e eliminar erros e inconsistências eventualmente presentes. Sobre os dados editados, foi realizado o cálculo da estatística descritiva. Os dados de limnologia foram adquiridos ao longo de quinze excursões de campo nos anos de 2004 (4 excursões), 2005 (4 excursões), 2006 (4 excursões) e 2007 (3 excursões). No total, foram realizadas 227 coletas (valores após edição). O tema ictiofauna efetuou 130 coletas durante dezessete trabalhos de campo. As excursões compreenderam os anos de 2004 (4 excursões), 2005 (4 excursões), 2006 (4 excursões), 2007 (4 excursões) e 2008 (1 excursão).

4.3 – Sumarização

Na etapa de pré-processamento, técnicas de sumarização e de análise exploratória espacial foram aplicadas. A sumarização foi aplicada apenas nos dados do tema ictiofauna. A tabela de dados inicial, que continha 33 atributos e onde cada registro correspondia a um peixe capturado (total de 23.897 peixes), foi sumarizada de modo que cada novo registro corresponde as informações por coleta. Como resultado, obteve-se uma nova tabela com 130 registros (coletas) e 845 atributos.

4.4 - Análise Exploratória Espacial

Na análise exploratória espacial foi realizada a associação entre os dados de ictiofauna e limnologia, assim como as análises de distância entre os objetos (temas) e as feições espaciais da BDI (e.g., marco PIATAM). O procedimento consistiu em, primeiramente, identificar as excursões (atributo temporal) que efetuaram coletas de ambos os temas. As coletas de uma mesma excursão de ambos os temas foram espacializadas (feições do tipo ponto). Sobre esses pontos, foi aplicada a função de associação espacial que integra os registros dos diferentes objetos segundo um critério de proximidade expresso pela distância euclidiana. Assim, foram relacionados apenas os pontos mais próximos dos diferentes temas (objetos) que se encontravam dentro da distância máxima especificada ($d_{max} = 2$ km). O resultado final é um conjunto de pontos com as informações de ambos os objetos. No total, 97 pontos de ambos os temas foram associados. As informações do tema limnologia (feição do tipo relevante) foram associadas a coletas do tema ictiofauna (feição do tipo alvo).

A próxima etapa no processo de análise exploratória espacial consistiu na análise da distância das coletas (feição alvo T) em relação às comunidades (feição relevante O). Os novos atributos identificam a comunidade (instância da feição relevante o) mais próxima a coleta (feição alvo t), assim como a respectiva distância.

4.5 - Transformação dos Dados

A próxima etapa consistiu na transformação dos dados para o formato de entrada do processo de mineração. Nesse processo, os atributos são entendidos como variáveis linguísticas, as quais devem ser particionadas em valores linguísticos (Hu et al., 2003). As variáveis categóricas e numéricas foram particionadas em valores binários e em conjuntos *fuzzy*, respectivamente. As variáveis linguísticas categóricas compreendem valores finitos, a princípio não ordenados. Por exemplo, o atributo categórico “período hidrológico” (*A período hidrológico*) foi particionado em quatro valores linguísticos binários: cheia, vazante, seca e enchente.

Para os atributos numéricos, os valores linguísticos (partições), por simplicidade, foram definidos segundo funções triangulares aplicadas sobre seus domínios. Os atributos numéricos de ambos os temas foram particionados em três valores linguísticos (baixo, médio e alto).

4.6 – Mineração dos Dados

O algoritmo de regras de associação *fuzzy* adotado é semelhante ao proposto por Hu et al. (2003) e consiste basicamente em dois processos: geração dos candidatos k -dimensionais e construção das regras.

4.6.1 - Geração dos candidatos k -dimensionais do *grid fuzzy*

Nessa etapa, primeiramente, cada valor linguístico gerado com valor de suporte superior ao mínimo especificado ($\text{sup_min} = 0,2$), é visto como um candidato unidimensional do *grid fuzzy*. Posteriormente, esses candidatos unidimensionais foram empregados na geração dos outros candidatos do *grid fuzzy* de dimensões maiores. Um possível candidato k dimensional do *grid fuzzy* $A_{i_1}^{x_1} \times A_{i_2}^{x_2} \times \dots \times A_{i_{k-1}}^{x_{k-1}} \times A_{i_k}^{x_k}$ foi derivado a partir da união de dois outros membros do *grid fuzzy* com dimensão $k - 1$ (i.e. $A_{i_1}^{x_1} \times A_{i_2}^{x_2} \times \dots \times A_{i_{k-1}}^{x_{k-1}}$ e $A_{i_1}^{x_1} \times A_{i_2}^{x_2} \times \dots \times A_{i_k}^{x_k}$), e que compartilham $k - 2$ valores linguísticos $A_{i_1}^{x_1} \times A_{i_2}^{x_2}$.

O grau com o qual a tupla t_p pertence ao candidato do *grid fuzzy* $A_{i_1}^{x_1} \times A_{i_2}^{x_2} \times \dots \times A_{i_{k-1}}^{x_{k-1}} \times A_{i_k}^{x_k}$ foi definido como a seguir: $\mu_{i_1}^{x_1}(t_p) \times \mu_{i_2}^{x_2}(t_p) \times \dots \times \mu_{i_{k-1}}^{x_{k-1}}(t_p) \times \mu_{i_k}^{x_k}(t_p)$. Para verificar se esse candidato pode ser considerado ou não um candidato frequente do *grid fuzzy*, calculou-se o valor de suporte *fuzzy* $\text{sup}(A_{i_1}^{x_1} \times A_{i_2}^{x_2} \times \dots \times A_{i_{k-1}}^{x_{k-1}} \times A_{i_k}^{x_k})$, como a seguir:

$$\text{sup}(A_{i_1}^{x_1} \times A_{i_2}^{x_2} \times \dots \times A_{i_{k-1}}^{x_{k-1}} \times A_{i_k}^{x_k}) = \left[\sum_{p=1}^n \mu_{i_1}^{x_1}(t_p) \times \mu_{i_2}^{x_2}(t_p) \times \dots \times \mu_{i_{k-1}}^{x_{k-1}}(t_p) \times \mu_{i_k}^{x_k}(t_p) \right] / n \quad (5)$$

Caso o valor do $\text{sup}(A_{i_1}^{x_1} \times A_{i_2}^{x_2} \times \dots \times A_{i_{k-1}}^{x_{k-1}} \times A_{i_k}^{x_k})$ obtido seja maior ou igual ao suporte mínimo *fuzzy* especificado pelo usuário, pode-se considerar $A_{i_1}^{x_1} \times A_{i_2}^{x_2} \times \dots \times A_{i_{k-1}}^{x_{k-1}} \times A_{i_k}^{x_k}$ um membro do *grid fuzzy* com dimensão k (Hu et al., 2003).

4.6.2 - Construção das regras

As regras são obtidas através da combinação de pares de membros do *grid fuzzy*, através de um operador de disjunção não exclusiva (XOR). O algoritmo impõe três condições para que os membros *fuzzy* do *grid* sejam combinados.

A primeira é que o primeiro membro *fuzzy* a ser combinado sempre apresente uma posição superior de registro em relação ao segundo, i.e., $GF[u] \text{ XOR } GF[v]$ só é válido caso $u < v$ seja verdadeira. A segunda condição é que todos os valores linguísticos do primeiro *grid fuzzy* estejam presentes no segundo *grid fuzzy* (i.e. $GF[u] \subset GF[v]$). A última condição é que a regra apresente um valor de confiança superior ao especificado ($\text{conf_min} = 0,8$). Os valores linguísticos do primeiro membro passam a compor o antecedente da regra e os valores linguísticos resultantes da operação **XOR** compõem a consequência da regra.

Como a regra R é gerada a partir de dois membros frequentes do *grid fuzzy*, i.e., $A_{i_1}^{x_1} \times A_{i_2}^{x_2} \times \dots \times A_{i_{k-1}}^{x_{k-1}} \times A_{i_k}^{x_k}$ e $A_{i_1}^{x_1} \times A_{i_2}^{x_2}$, a confiança pode ser definida como a seguir:

$$conf(R) = \sup(A_{i_1}^{x_1} \times A_{i_2}^{x_2} \times \dots \times A_{i_{k-1}}^{x_{k-1}} \times A_{i_k}^{x_k}) / \sup(A_{i_1}^{x_1} \times A_{i_2}^{x_2}) \quad (6)$$

4.7 – Visualização das Regras

As regras foram visualizadas no espaço através do georrefenciamento dos valores linguísticos que compõem a regra. Os pontos foram definidos cartograficamente através das coordenadas geográficas capturadas por GPS na coleta dos dados. Os valores de pertinência do antecedente e do consequente à regra são associados a uma escala de cores. As cores mais fortes (vermelho) indicam os valores de pertinência próximos a ‘1’, enquanto as cores mais fracas (verde) os valores próximos a ‘0’ (Figura 2). A visualização espacial em um SIG permitiu uma inspeção mais adequada da efetividade da regra em cada comunidade.

5 – Resultados

O desempenho da mineração nesse primeiro estudo foi considerado razoável. Entretanto, as regras geradas necessitam de uma análise mais detalhada por parte dos especialistas. A avaliação inicial do estudo baseou-se na identificação de regras pontuais, as quais, embora óbvias, atestaram o desempenho coerente do estudo de mineração. Diversas regras foram classificadas nesta condição, sendo as duas regras a seguir exemplos ilustrativos:

- Se a “**Temperatura Profundidade**” é baixa, Então a “**Temperatura em Superfície**” é média;
- Se a “**Peixe Total peso**” é baixa, Então “**Peixe Total comprimento**” é baixa.

A primeira regra capturou de modo coerente a variação de temperatura nas diferentes profundidades. A segunda expressa a relação esperada entre peso e comprimento de peixes pequenos.

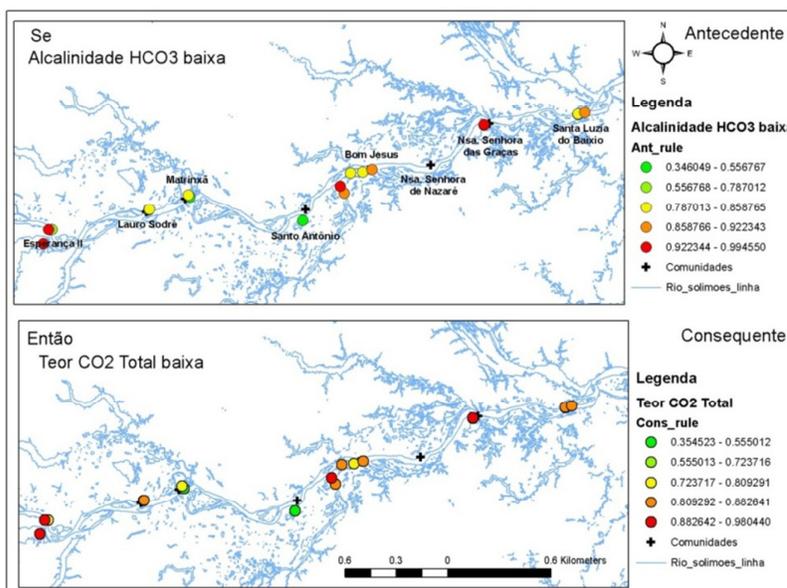


Figura 2 – Mapa com os valores de pertinência representados em cores para os valores linguísticos presentes no antecedente (Alcalinidade HCO₃ baixa) e no consequente (Teor de CO₂ Total) da regra. Fonte: Moreira (2010).

Adicionalmente, algumas regras foram identificadas como interessantes e coerentes pelos pesquisadores especialistas (coordenadores do projeto PIATAM). A seguir são apresentadas duas regras que exemplificam essa situação:

- Se a “Alcalinidade HCO_3 ” é baixa, Então “Teor de CO_2 Total” é baixa (Figura 2);
- Se a “Distância Centro Comunidade” é baixa, Então “Peixe Total comprimento” é baixa (Figura 3).

Na primeira, as relações entre a alcalinidade HCO_3 e o teor de CO_2 Total foram consideradas coerentes, uma vez que são comumente encontradas em amostragens na Bacia do Solimões. A visualização espacial da regra através da pertinência dos valores lingüísticos permitiu verificar que as comunidades Esperança II e N. Sra. das Graças foram as que apresentaram os maiores valores de pertinência (pontos de cor vermelha). Isso indica que a relação expressa pela regra é mais provável de ocorrer nessas comunidades. Por outro lado, a comunidade Santo Antônio foi a que apresentou os menores valores de pertinência, indicando que a relação expressa pela regra talvez não seja tão forte como nas demais comunidades.

Na segunda, a regra indica uma relação de proximidade entre peixes pequenos e os centros comunitários. Tal relação desconhecida merece ser investigada. A visualização espacial da regra mostra que as coletas próximas à comunidade St. Luzia do Baixio apresentam pertinências altas, indicando que, nesta localidade, a regra é mais provável. De modo inverso, as comunidades Santo Antônio e Matrinxã apresentaram os menores valores de pertinência indicando que a relação expressa pela regra talvez não seja forte como nas demais comunidades.

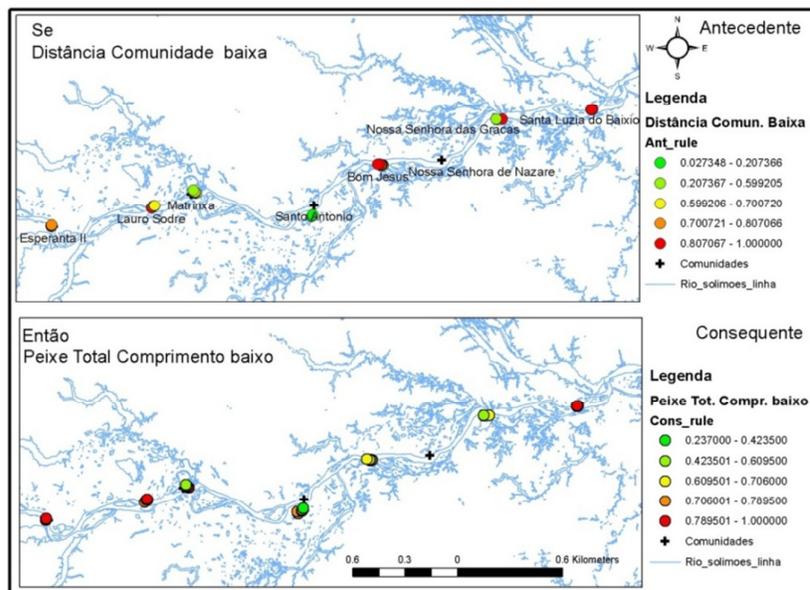


Figura 3 – Mapa com os valores de pertinência representados em cores para os valores lingüísticos presentes no antecedente (Distância Comunidade baixa) e no consequente (Peixe Total Comprimento baixo) da regra. Fonte: Moreira (2010).

6 - Conclusões

A aplicação de regras de associação *fuzzy*, em conjunto com técnicas de exploração espacial, demonstrou ser uma abordagem adequada para adquirir conhecimento, expresso por regras lingüísticas, das informações levantadas pelos temas ictiofauna e limnologia do Projeto PIATAM. O uso de conjuntos *fuzzy* na representação dos atributos numéricos em variáveis lingüísticas demonstrou ser um modo ideal de lidar com esses dados, uma vez que seus domínios apresentam uma propriedade contínua, na qual é muito difícil definir partições rígidas.

A associação espacial entre as coletas dos temas limnologia e ictiofauna, segundo critérios de proximidade expressos pela distância euclidiana, permitiu integrar suas variáveis respeitando as propriedades espaciais de dependência e a heterogeneidade. O cálculo da

distância dos pontos de coleta em relação aos centros comunitários mais próximos possibilitou também verificar a existência de alguma dependência espacial indicativa de uma possível influência da comunidade sobre as coletas de campo. A visualização espacial das regras consideradas fortes, através da espacialização das pertinências dos valores linguísticos presentes no antecedente e no conseqüente, demonstrou ser útil, uma vez que permitiu uma melhor inspeção de suas forças nas diferentes comunidades.

Agradecimentos

Agradecemos à coordenação do projeto PIATAM, em nome do prof. Alexandre A. F. Rivas, pelo apoio, como também aos professores Tereza Cristina S. de Oliveira e Carlos Edwar C. Freitas, coordenadores dos temas limnologia e ictiofauna, respectivamente. Tais profissionais contribuíram com a liberação dos dados de campo e com discussões técnicas.

Referência Bibliográfica

Agrawal, R.; Imielinski, T.; Swami, A. Mining Association Rules Between Sets of Items in Large Databases, In: **Proceedings** of the 1993 ACM SIGMOD international conference on Management of data, 1993, p. 207-216.

Beisl, C.H. Utilização de mosaicos JERS-1 SAR e de lógica *fuzzy* para elaboração de mapas de sensibilidade ambiental temporal a derrames de óleo na Amazônia Central. 2009. 196 p. Tese (Doutorado em Sistemas Computacionais) – Universidade Federal do Rio de Janeiro, Programa de Engenharia Civil - COPPE /UFRJ, Rio de Janeiro, 2009.

Bogorny, V. Enhancing Spatial Association Rule Mining in Geographic Databases. 2006. P.120. Tese (Doutorado em Computação) - Universidade Federal do Rio Grande do Sul - UFRGS, Porto Alegre. 2006.

Dubois, D.; Hüllermeir, E.; Prade, H. A systematic approach to the assessment of fuzzy association rules. **Data Min Knowl Disc**, n.13, p. 167-192, 2006.

Hu, Y.-C.; Chen, R.-S.; Tzeng, G.-H. Discovering Fuzzy Association Rules Using Fuzzy Partition Methods. **Knowledge-Based Systems**, v. 16, n. 3, p. 137.-147, 2003.

Koperski, K.; Han, J. Discovery of Spatial Association Rules in Geographic Information Databases. In: **Proceedings** of the 4th International Symposium on Advances in Spatial Databases, 1995, p. 47-66.

Moreira, F. R. S. Uso de Regras de associação *fuzzy* e técnicas de análise e exploração espacial na mineração dos dados sócio-ambientais do projeto PIATAM, Amazônia Central. 2010. 168 p. Tese (Doutorado em Sistemas Computacionais) – Universidade Federal do Rio de Janeiro, Programa de Engenharia Civil - COPPE /UFRJ, Rio de Janeiro, 2010.

Rivas, A. A. F.; Cavalcante, K. V.; Freitas, C.E.C. O Sistema de indicadores PIATAM. In: Cavalcante, K. V.; Rivas, A. A. F. e Freitas, C.E.C. (Eds.). **Indicadores Sócioambientais e atributos de referência para o trecho Urucu-Coari-Manaus, Rio Solimões, Amazônia Ocidental**, Parte I, Edua., 2007, p. 15-20.

Tocantins, C. A.; Botelho, A. A.; Cavalcante, K. V.; Costa, M. G.; Miranda, F. P.; Moreira, F. R. S.; Prado, W. F.; André Neto, M. A. Base de dados Integrada PIATAM: aplicação multidisciplinar e integração espaço temporal. **Anais...** do 1º Simpósio de Geotecnologias no Pantanal, 2006, Campo Grande, p.757-762.