

Análise de correlação espacial de variáveis relacionadas à colonização de macrófitas submersas

Ligia Flávia Antunes Batista^{1,2}
Nilton Nobuhiro Imai¹
Luiz Henrique da Silva Rotta¹
Fernanda Sayuri Yoshino Watanabe¹

¹ Universidade Estadual Paulista – PPGCC/FCT/UNESP
19060-900 – Presidente Prudente – SP, Brasil
nnimai@fct.unesp.br, {luizhrotta,fernandasyw}@yahoo.com.br

² Universidade Tecnológica Federal do Paraná – UTFPR
86036-370 – Londrina - PR, Brasil
ligia@utfpr.edu.br

Abstract. The aim of this work is to analyze the spatial correlation among pair of maps related to submerged macrophytes growth in Taquaruçu reservoir, between São Paulo and Paraná states. The study of submerged vegetation is mainly important due to its excessive proliferation, which causes ecological unbalance and economical losses to hydroelectric power plants. Three survey fields were carried out between March and April of 2010, in order to map the submerged canopy height and depth, with the hydro-acoustic technique. With these data, other spatial variables related to macrophytes growth were derived: slope, water column height, margin distance, thalweg distance, colonized neighborhood cells count and colonized neighborhood height sum. The correlation analysis was divided in two sets of pairs of maps. One set was formed by morphometric ones and the other set had the pairs of maps whose variables were related to vegetation, with temporal variation. The spatial correlation was evaluated using Cramers index, based on chi-square statistic, and Joint Information Uncertainty, based on entropy concept. The results showed that the most correlated variables were depth and thalweg distance. Some applications, such as modeling, require that one of the highly associated variable of the map pair is removed. In this work it is suggested that depth and thalweg distance be excluded.

Keywords: reservoirs, contingency tables, entropy reservatórios, tabelas de contingência, entropia

1. Introdução

A possibilidade de aquisição de informações georreferenciadas por meio de diversos instrumentos e sensores, somada à disponibilidade de dados cartográficos sistemáticos, gerados por diversas instituições e usuários, faz com que seja necessário integrar dados espaciais multivariados e multifonte durante a investigação de problemas espaciais.

No entanto, não raro tais dados possuem alto grau de associação entre si, sendo portanto necessário tratar e eliminar essa redundância a fim de que o resultado não seja comprometido.

Em trabalhos de modelagem espacial dinâmica, por exemplo, o uso de variáveis com alto grau de associação dificulta a calibração dos coeficientes do modelo, podendo inclusive mascarar a influência de outras variáveis. Na aplicação de métodos de inferência espacial, como pesos de evidência, eliminar alta correlação entre mapas é imprescindível para obtenção de bons resultados (BONHAM-CARTER, 1994).

Para tratar tais situações, é importante analisar a associação espacial entre os mapas e com isso determinar se é necessário excluir alguns deles da análise. Dentre as métricas de associação espacial aplicadas a pares de mapas, cita-se análise de correlação cruzada, baseadas em tabelas de contingência e aquelas fundamentadas no conceito de entropia (BONHAM-CARTER, 1994).

Exemplos de aplicações em que análise de associação espacial é necessária incluem os mais diversos domínios. Cita-se problemas geológicos, em que se deseja identificar fatores indicativos de depósitos minerais ou estudar padrões de terremoto relacionados a feições geológicas para estabelecer zonas de risco (ZIAII; CARRANZA; ZIAEI, 2011; FAN et al., 2011). Na Geomorfologia, verifica-se a associação entre áreas para investigar ocorrências de deslocamento de massa em função de pedologia, uso e cobertura do solo, vegetação, entre outros. No que se refere a questões de desmatamento, também faz-se necessário avaliar a associação espacial das variáveis que atuam como condicionantes, dentre as quais distância de rios, distância de estradas, solos, vegetação, declividade (SOARES FILHO; CERQUEIRA; PENNACHIN, 2002).

Neste trabalho, o foco é a presença de vegetação aquática em reservatórios. No Brasil, há vários que apresentam crescimento excessivo de macrófitas aquáticas, o que tem gerado problemas ecológicos e econômicos. Além do desequilíbrio do ecossistema, há prejuízos para a produção de energia elétrica, pois este tipo de vegetação pode se acumular nas turbinas de usinas hidrelétricas, causando entupimento das mesmas (ROCHA; MARTINS, 2011; BATISTA et al., 2012).

Acrescenta-se que tal tipo de vegetação é difícil de ser mapeada no caso de espécies submersas. Imagens de sensoriamento remoto óptico, podem ser aplicadas (ROTTA et al., 2011), mas apresentam certas restrições, devido à atenuação da radiação eletromagnética quando penetra na coluna d'água, limitando a identificação de vegetação submersa até uma determinada profundidade. Uma das alternativas é a utilização do método hidroacústico, com equipamentos como o ecobatímetro (ROTTA et al., 2012).

Neste contexto, há diversas variáveis que influenciam a colonização de macrófitas submersas, sejam morfométricas, as quais caracterizam o reservatório, sejam limnológicas, de natureza física, química ou biológica. Neste caso, a avaliação de correlação espacial é muito importante para determinar os principais fatores a serem analisados com o intuito de reduzir a dimensionalidade do problema sem perda de informação significativa.

O objetivo deste trabalho é analisar a correlação espacial entre variáveis espaciais que influenciam o crescimento de macrófitas submersas. Este tipo de análise auxilia na definição dos fatores mais adequados para serem utilizados em modelagem espacial, que requer que os dados possuam baixa associação espacial.

2. Domínio da aplicação

Sabe-se que há influência de diversos fatores limnológicos no desenvolvimento de vegetação submersa, dentre os quais: luz, vento, macronutrientes, carbono, oxigênio, condutividade elétrica, pH, temperatura, profundidade, declividade, velocidade da corrente, vegetação pré-existente (BATISTA, 2011).

Dentre as variáveis elencadas para avaliação neste trabalho, citam-se: profundidade, declividade, velocidade da corrente, radiação e vegetação pré-existente, que poderiam ser mapeados ou inferidos a partir do levantamento ecobatimétrico.

Uma das características morfométricas relacionadas à disponibilidade de luz, essencial para a fotossíntese, é a profundidade, uma vez que em águas muito profundas, plantas enraizadas têm dificuldades para se desenvolverem, pois a luz pode não chegar até o fundo (CARRILLO; GUARÍN; GUILLOT, 2006).

Em áreas com declive acentuado, o sedimento torna-se instável, o que dificulta a fixação de plantas enraizadas. Carrillo, Guarín e Guillot (2006) em um trabalho realizado no reservatório da Colômbia, avaliaram a declividade como um dos fatores que interferem no aumento da biomassa da *Egeria densa*, uma das principais espécies submersas encontradas em reservatórios brasileiros.

A velocidade da corrente, quando moderada, promove oxigenação da água, sendo portanto um fator positivo para o crescimento. Se for muito veloz, o fluxo dificulta a fixação das macrófitas enraizadas (BIUDES; CAMARGO, 2008). Neste trabalho, o dado da corrente não estava disponível. Uma forma de considerá-lo indiretamente foi a partir da sua relação com a distância da margem e do talvegue (parte mais profunda do leito do rio). Junto às margens, o fluxo é menor, aumentando à medida em que se aproxima do talvegue.

Outro fator importante quando se trata de colonização de macrófitas submersas é a vegetação pré-existente, uma vez que a principal forma de reprodução é a vegetativa, por propagação de fragmentos de caule (COOK; URMI-KÖNIG, 1984).

Relacionada à penetração da radiação, tem-se a informação da altura da coluna d'água, determinada pela diferença entre o topo do dossel e a interface ar/água. Dependendo dos valores do coeficiente de atenuação da luz e da altura da coluna d'água, o dossel recebe ou não luz no ápice das plantas, o que pode interferir no seu desenvolvimento (BATISTA, 2011).

3. Análise de correlação

Para realizar a verificação da associação espacial entre pares de mapas, tabelas de contingência ou testes χ^2 podem ser usados, embora não possam ser usados para fazer testes formais, pois o valor máximo de χ^2 varia conforme as unidades de medida (BONHAM-CARTER, 1994).

Para usar a estatística χ^2 , a tabulação cruzada é usada como tabela de contingência. Sendo a tabela de área entre mapas A e B denotada de matriz T , com elementos T_{ij} , onde $i = 1, 2, \dots, n$ são as classes dos mapa B (linhas da tabela) e $j = 1, 2, \dots, m$ são as classes do mapa A (colunas da tabela). Os totais marginais de T são definidos como T_{i+} para a soma da i -ésima linha e T_{+j} para a soma da j -ésima coluna, sendo T_{++} a soma total de linhas e colunas (BONHAM-CARTER, 1994). Segundo o autor, se dois mapas são totalmente independentes, a área esperada em cada combinação T_{ij} de classes é dada pelo produto dos totais marginais dividido pelo total geral, como na Equação 1.

$$T_{ij}^* = \frac{T_{i+}T_{+j}}{T_{++}} \quad (1)$$

A estatística χ^2 é então definida conforme a Equação 2. Essa expressão possui limite inferior igual a zero quando os dois mapas são completamente independentes. O limite máximo é variável, dependendo da unidade de área, o que dificulta sua avaliação. Derivado de χ^2 , tem-se o coeficiente de Cramers, definido como na Equação 3, em que M é o mínimo entre $(n-1, m-1)$, sendo m e n o número de classes dos mapas A e B. A magnitude do índice de Cramers é independente de unidade de área, variando de 0 (sem correlação) ao máximo menor que 1, dependendo de χ^2 , área total e dimensões da tabela de contingência (BONHAM-CARTER, 1994).

$$\chi^2 = \sum_{i=1}^n \sum_{j=1}^m \frac{(T_{ij} - T_{ij}^*)^2}{T_{ij}^*} \quad (2)$$

$$C = \sqrt{\frac{\chi^2}{T_{++}M}} \quad (3)$$

Outra métrica para medir associação entre mapas é baseada no conceito de entropia, também conhecida como estatística de informação, a qual também usa a matriz de tabulação cruzada T , em que cada valor é transformado em proporção de área p_{ij} , dividindo-se cada T_{ij} por T_{++} . Como proporções são adimensionais, o valor da entropia também não é afetado pelas unidades de medida.

Assim, pode-se determinar a entropia (H) de A e B usando as Equações 4 e 5. A entropia conjunta da combinação de dois mapas A e B é dada pela Equação 6. A métrica de Incerteza de Informação Conjunta (IIC) pode ser usada como medida de associação, sendo definida pela Equação 7, que varia entre 0 (independente), quando $H(A, B) = H(A) + H(B)$ e 1 (dependente), quando $H(A) = H(B) = H(A, B) = 1$ (BONHAM-CARTER, 1994).

$$H(A) = - \sum_{j=1}^m p_{+j} \ln p_{+j} \quad (4)$$

$$H(B) = - \sum_{i=1}^n p_{i+} \ln p_{i+} \quad (5)$$

$$H(A, B) = - \sum_{i=1}^n \sum_{j=1}^m p_{ij} \ln p_{ij} \quad (6)$$

$$IIC(A, B) = 2 \left[\frac{H(A) + H(B) - H(A, B)}{H(A) + H(B)} \right] \quad (7)$$

Bonham-Carter (1994) considera como alta associação quando o índice de Crammer ou a IIC apresentem valores acima de 0.5.

4. Materiais e métodos

A área de estudo compreende o reservatório de Taquaruçu, no rio Paranapanema, no trecho situado no município de Santo Inácio (PR), próximo à divisa com o estado de São Paulo, delimitado aproximadamente pelas coordenadas $51^{\circ}44'W$, $22^{\circ}38'S$ e $51^{\circ}43'W$, $22^{\circ}37'S$, como ilustrado na Figura 1. A região em que foram feitos os levantamentos de dados abrange uma área de 1.83 km^2 , em que a profundidade máxima é de aproximadamente 23 m e a média em torno de 5 m.

O levantamento de dados foi dividido em três trabalhos de campo para obtenção das medidas ecobatimétricas, realizados entre março e abril de 2010, com intervalos variados, referenciados neste texto como c1, c2 e c3. No lado direito da Figura 1 está ilustrado o percurso do barco em um dos levantamentos.

A técnica hidroacústica foi utilizada para obtenção dos dados. Os instrumentos necessários para realizar as coletas em campo foram: 2 unidades portáteis GPS Garmin, sendo uma acoplada ao ecobatímetro Biosonics modelo DT-X, e outra para navegação. Os dados do ecobatímetro mostram frequência da cobertura de plantas a cada dez pulsos emitidos e altura dos dosséis, bem como a profundidade. Esses dados foram utilizados para caracterizar a distribuição espacial da vegetação submersa, a partir da grandeza da altura média do dossel (AMD), dada pela média de altura a cada dez pulsos, ponderada pela frequência de cobertura das plantas.

Os dados adquiridos foram organizados de modo a compor um banco de dados geográfico, com o software Spring (CÂMARA et al., 1996). Inicialmente foram importados os dados da medida AMD de cada levantamento, bem como os dados de profundidade. Foi feita então a geração de superfícies a partir da interpolação dos dados de vegetação e de batimetria (relevo submerso), além da geração da superfície de declividade a partir da profundidade. O método aplicado foi o da média ponderada, considerado por Valley, Drake e Anderson (2005) com desempenho bastante semelhante à krigeagem em aplicações de macrófitas submersas mapeadas por técnica hidroacústica. Foram derivadas outras variáveis espaciais, a saber: (a) altura da coluna d'água; (b) distância do talvegue, que foi definido a partir de edição vetorial, orientada pela visualização do modelo batimétrico; (c) distância das margens; (d) declividade;

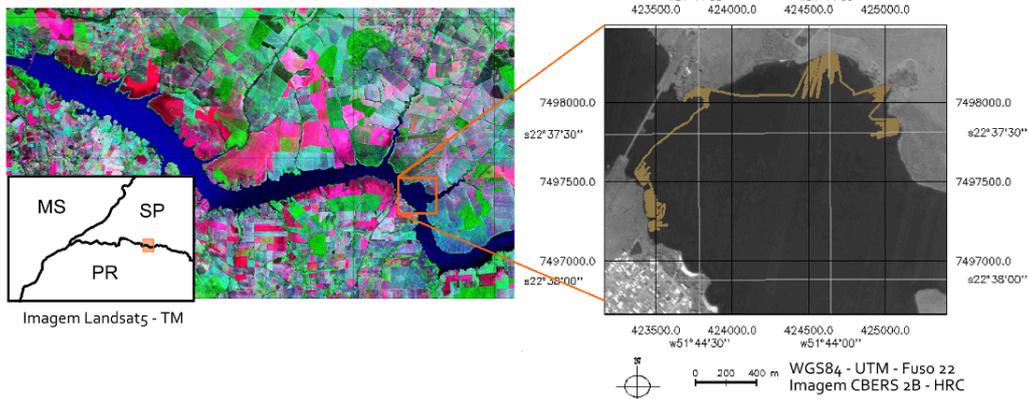


Figura 1: Área de estudo.

(e) soma da AMD da vizinhança (SAV), calculada com base nas 8 células vizinhas, somando-se seus valores de AMD; (f) contagem de vizinhos com plantas (CVP), obtido com base nas 8 células da vizinhança, contando o número de vizinhos colonizados de cada célula.

Optou-se por usar o índice de Cramers, cujo resultado independe da unidade de área e a medida de IIC, baseada no conceito de entropia. Como tais índices trabalham com tabela de contingência, foi feita a discretização dos intervalos de cada variável espacial, inicialmente representadas como Modelos Numéricos de Terreno. Foi aplicado um algoritmo que adequa as faixas de valores, de modo a preservar a estrutura dos dados, estabelecendo um delta mínimo e intervalos incrementais (SOARES FILHO; RODRIGUES; COSTA, 2009). Os cálculos do índice de Cramers e IIC foram feitos com o *software* Dinamica-EGO (SOARES FILHO; CERQUEIRA; PENNACHIN, 2002; SOARES FILHO; RODRIGUES; COSTA, 2009).

Para as variáveis morfométricas (batimetria, declividade, distância da margem e do talvegue), os índices de associação espacial foram calculados uma única vez, pois considerou-se que não houve grande variação no período considerado. Para as variáveis que tinham relação com a presença de vegetação, as métricas de associação espacial foram calculadas para cada levantamento, devido a sua variação temporal.

5. Resultados e discussão

As Figuras 2(a), 2(b), 2(c) e 2(d) apresentam as cartas de batimetria, declividade, distância das margens e distância do talvegue, cujos dados foram utilizados neste trabalho. A Figura 3(a) apresenta a representação espacial da altura da coluna d'água para o levantamento c1 e na Figura 3(b) está ilustrada uma parte do mapa que caracteriza a CVP.

A Tabela 1 apresenta o resultado do índice de Cramers e IIC para as variáveis morfométricas e a Tabela 2 apresenta os mesmos índices para as variáveis temporais.

Tabela 1: Avaliação da associação espacial entre as evidências morfométricas

Mapa 1	Mapa 2	Cramers	IIC
batimetria	declividade	0.36	0.67
batimetria	dist. margem	0.24	0.12
batimetria	dist. talvegue	0.25	0.12
declividade	dist. margem	0.16	0.08
declividade	dist. talvegue	0.21	0.1
dist. margem	dist. talvegue	0.22	0.07

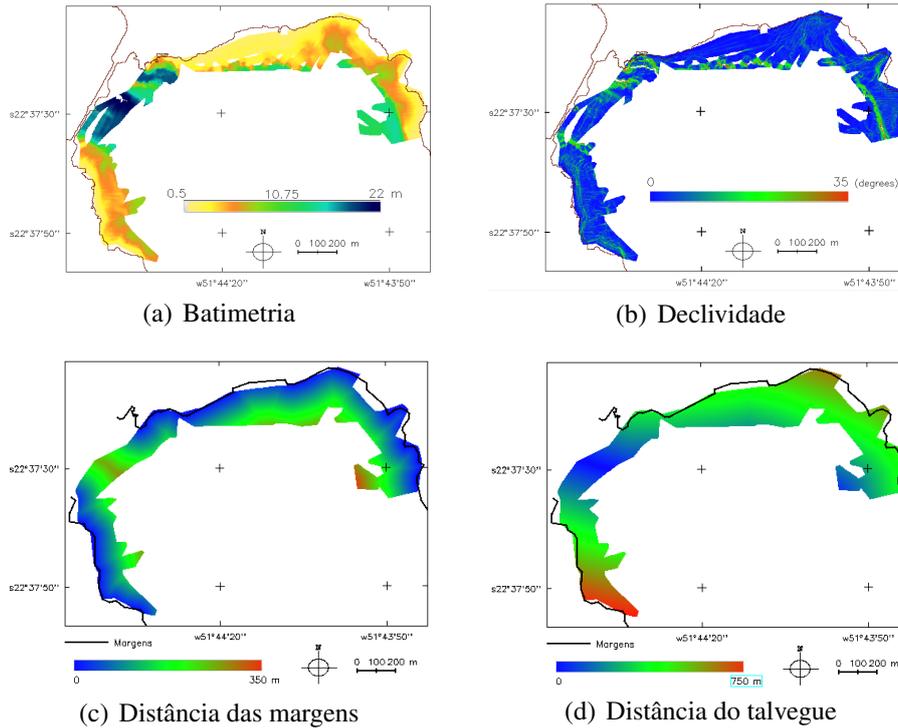


Figura 2: Evidências morfométricas. WGS84/ UTM 22S.

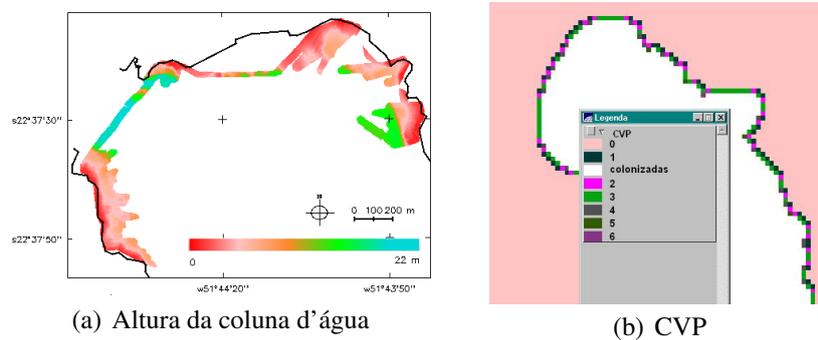


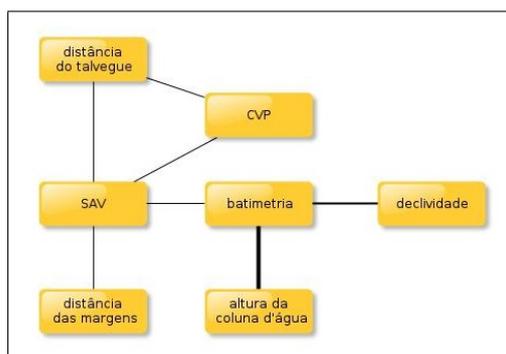
Figura 3: Evidências relacionadas a vegetação. WGS84/ UTM 22S.

A Figura 4(a) mostra as associações espaciais por meio de arestas, sendo a espessura proporcional ao grau da correlação. Em termos de valores, os maiores índices de associação são entre batimetria/coluna d'água (maior que 0.9) e entre batimetria/declividade (0.67). Os pares de mapas que apresentaram alta correlação segundo as duas métricas, Cramers e IIC, são batimetria/coluna d'água e distância do talvegue/CVP. Para avaliação temporal, apresentaram-se correlacionados em mais de um levantamento os pares batimetria/coluna d'água (3 levantamentos), distância do talvegue/SAV (2 levantamentos) e distância do talvegue/CVP (3 levantamentos).

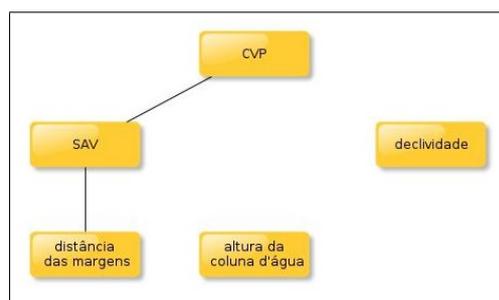
Com base nessa análise, percebeu-se que as principais correlações envolviam o mapa de batimetria e o de distância do talvegue. Nos casos de modelagem, em que recomenda-se a retirada de um dos mapas dos pares correlacionados, esses seriam os primeiros a serem excluídos (BONHAM-CARTER, 1994). Os demais mapas, cujas correlações estão pouco acima de 0.5, poderiam ser mantidos, a fim de evitar perda de informação. Dessa forma, apenas as associações mais fracas seriam mantidas, como ilustrado na Figura 4(b).

Tabela 2: Avaliação da associação espacial entre as evidências temporais

Mapa 1	Mapa 2	c1		c2		c3	
		Cramers	IIC	Cramers	IIC	Cramers	IIC
coluna d'água	batimetria	0.94	0.93	0.93	0.96	0.94	0.95
coluna d'água	declividade	0.12	0.04	0.16	0.05	0.17	0.06
coluna d'água	dist. margem	0.36	0.22	0.35	0.21	0.32	0.21
coluna d'água	dist. talvegue	0.28	0.13	0.21	0.09	0.28	0.09
coluna d'água	SAV	0.13	0.03	0.15	0.06	0.16	0.06
coluna d'água	CVP	0.05	0.01	0.09	0.01	0.05	0.01
batimetria	SAV	0.33	0.37	0.41	0.4	0.42	0.5
batimetria	CVP	0.36	0.34	0.33	0.35	0.3	0.34
declividade	SAV	0.28	0.29	0.29	0.29	0.31	0.36
declividade	CVP	0.33	0.33	0.31	0.34	0.29	0.32
dist. margem	SAV	0.43	0.4	0.45	0.4	0.43	0.51
dist. margem	CVP	0.41	0.44	0.44	0.49	0.38	0.47
dist. talvegue	SAV	0.44	0.26	0.55	0.39	0.5	0.46
dist. talvegue	CVP	0.5	0.36	0.57	0.51	0.5	0.49
SAV	CVP	0.42	0.41	0.45	0.46	0.43	0.52



(a) Associações com todas as variáveis



(b) Associações após exclusões

Figura 4: Evidências relacionadas a vegetação. WGS84/ UTM 22S.

6. Considerações Finais

Com este trabalho, percebeu-se a importância de análises de correlação entre variáveis espaciais. Quantificar o grau de associação pode ajudar na seleção dos fatores espaciais mais importantes para descrever um domínio da aplicação, além de reduzir a dimensionalidade do problema.

Neste trabalho, a principal correlação encontrada foi entre coluna d'água e batimetria. Tal constatação está relacionada com a definição da variável coluna d'água, dada pela diferença entre profundidade e altura do dossel submerso. Como nos levantamentos realizados os dados mostraram dosséis com altura baixa, a altura da coluna d'água ficou muito semelhante à profundidade, o que explica a correlação. Em situações em que a vegetação estivesse maior, tal análise poderia apresentar resultados diferentes. Dessa forma, no que refere-se a macrófitas submersas, percebe-se que a associação espacial deve ser verificada em cada caso específico.

7. Agradecimentos

Agradecemos ao Programa de Pós-graduação em Ciências Cartográficas da FCT/UNESP, pelo apoio para realização do trabalho e à Universidade Tecnológica Federal do Paraná, Campus

Londrina.

Referências

- BATISTA, L. et al. Análise da dinâmica espaço-temporal de macrófitas submersas mapeadas com técnicas hidroacústicas no reservatório de Taquaruçu, rio Paranapanema. *Planta Daninha*, scielo, v. 30, p. 241 – 252, 06 2012.
- BATISTA, L. F. A. *Modelagem espaço-temporal da colonização de macrófitas submersas no Reservatório de Taquaruçu*. Tese (Doutorado) — Universidade Estadual Paulista, Faculdade de Ciências e Tecnologia, Pós-Graduação em Ciências Cartográficas, Presidente Prudente, 2011.
- BIUDES, J. F. V.; CAMARGO, A. F. M. Estudos dos fatores limitantes à produção primária por macrófitas aquáticas no Brasil. *Oecologia Brasiliensis*, v. 12, p. 7–19, 2008.
- BONHAM-CARTER, G. F. *Geographic Information Systems for Geoscientists: Modelling with GIS*. Ontario: Pergamon, 1994. (Computer Methods in the Geosciences, v. 13).
- CÂMARA, G. et al. SPRING: Integrating remote sensing and GIS by object-oriented data modelling. *Computers & Graphics*, v. 20, n. 3, p. 395–403, May-Jun 1996.
- CARRILLO, Y.; GUARÍN, A.; GUILLOT, G. Biomass distribution, growth and decay of *Egeria densa* in a tropical high-mountain reservoir (NEUSA, Colombia). *Aquatic Botany*, v. 85, p. 7–15, 2006.
- COOK, C. D. K.; URMI-KÖNIG, K. A revision of the genus *Egeria* (Hydrocharitaceae). *Aquatic Botany*, v. 19, p. 73–96, 1984.
- FAN, D. et al. Weight of evidence method and its applications and development. *Procedia Environmental Sciences*, v. 11, Part C, n. 0, p. 1412 – 1418, 2011. ISSN 1878-0296. 2011 2nd International Conference on Challenges in Environmental Science and Computer Engineering (CESCE 2011).
- ROCHA, D.; MARTINS, D. Levantamento de plantas daninhas aquáticas no reservatório de Alagados, Ponta Grossa-PR. *Planta Daninha*, Scielo, v. 29, p. 237 – 246, 06 2011.
- ROTTA, L. et al. Sensoriamento remoto hidroacústico no mapeamento de macrófitas aquáticas submersas. *Planta Daninha*, scielo, v. 30, p. 229 – 239, 06 2012.
- ROTTA, L. H. S. et al. Detecção de macrófitas submersas por meio de imagem quickbird. estudo de caso: Usina hidrelétrica de porto colômbia, mg/sp. In: *Simpósio Brasileiro de Sensoriamento Remoto, 15 (SBSR)*. Curitiba: INPE, 2011. p. 4460–4467.
- SOARES FILHO, B.; CERQUEIRA, G.; PENNACHIN, C. DINAMICA - A stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecological modelling*, v. 154, p. 217–235, 2002.
- SOARES FILHO, B. S.; RODRIGUES, H.; COSTA, W. *Modeling Environmental Dynamics with Dinamica EGO*. Belo Horizonte, 2009. Disponível em www.csr.ufmg.br/dinamica. Acessado em 13/04/2010.
- VALLEY, R.; DRAKE, M.; ANDERSON, C. Evaluation of alternative interpolation techniques for the mapping of remotely-sensed submersed vegetation abundance. *Aquatic Botany*, v. 81, n. 1, p. 13–25, 2005.
- ZIAI, M.; CARRANZA, E. J. M.; ZIAEI, M. Application of geochemical zonality coefficients in mineral prospectivity mapping. *Computers and Geosciences*, v. 37, n. 12, p. 1935 – 1945, 2011. ISSN 0098-3004. Geocomputation of Mineral Exploration Targets – Geocomputation, Carranza.