



O DESENVOLVIMENTO E A APLICAÇÃO DE TÉCNICAS GEOESTATÍSTICAS EM ECOLOGIA.

A.L.C. Rochelle¹

1 - Programa de Pós - Graduação em Biologia Vegetal, Instituto de Biologia, Caixa Postal 6109, Universidade Estadual de Campinas-UNICAMP, 13083 - 970 Campinas, SP, Brasil.(rochelle@unicamp.br)

INTRODUÇÃO

O avanço do conhecimento nas diversas áreas que compõem a Ecologia tem evidenciado as limitações dos métodos estatísticos tradicionais no tratamento e descrição da variabilidade espacial das variáveis que integram os ecossistemas e comunidades biológicas. Isto ocorre, principalmente porque as técnicas utilizadas pela estatística clássica assumem que todas as amostras são aleatórias e independentes entre si e sua aplicação não envolve nenhum conhecimento da posição espacial das amostras ou do relacionamento entre elas enquanto as técnicas geoestatísticas conhecem a posição geográfica de cada amostra e por causa disso podem calcular o grau de dependência entre elas.

Os métodos geoestatísticos foram desenvolvidos pelos estudos do engenheiro de minas Georges Matheron, na França, entre o final da década de 50 e início da década de 60. Estes métodos estão fundamentados na “Teoria das Variáveis Regionalizadas” (Matheron, 1963) que foi formalizada por Matheron, a partir de estudos práticos desenvolvidos por Daniel G. Krige para o cálculo de reservas de ouro nas minas de Rand na África do Sul. A geoestatística é a aplicação destes métodos matemáticos a problemas das Ciências da Terra, como Geologia, Geografia, Pedologia e Ecologia com os objetivos de estimar, mapear e correlacionar simultaneamente um conjunto de variáveis espacialmente estruturadas (variáveis corregionalizadas). A utilização de algumas destas técnicas se mostra extremamente útil e versátil proporcionando estimativas de precisão aceitável até mesmo quando uma das variáveis não foi amostrada em quantidade suficiente devido a dificuldades experimentais ou altos custos.

OBJETIVOS

Este trabalho tem três objetivos. Apresentar o relevante papel do espaço sobre as relações ecológicas. Definir o que é autocorrelação espacial e apresentar as suas influências sobre os testes estatísticos convencionais. Apresentar os novos

conceitos e técnicas estatísticas disponíveis que permitem manipulação e análise de tais tipos de dados e a inclusão explícita do espaço no processo de análise e interpretação dos dados.

MATERIAL E MÉTODOS

Foram analisadas algumas obras principais (Pielou, 1977; Legendre & Fortin, 1989; Legendre, 1993; Fortin & Dale, 2005; Rangel *et al.*, 2006) para resumir em três tópicos as principais teorias e técnicas do conjunto de análises estatísticas conhecidas como análises espaciais.

RESULTADOS

1 - O papel do espaço sobre as relações ecológicas

O mundo e principalmente a natureza são inevitavelmente espaciais. Cada organismo é uma entidade única que existe e interage com seus arredores mais próximos e o mundo real é formado por variados organismos que possuem coordenadas x, y e z específicas, além de vários outros atributos que também estão vinculados as mesmas coordenadas, como taxas de fecundidade e mortalidade que se alteram com o tempo de maneiras que dependem dos seus próprios genótipos e dos genótipos de outros organismos aos seus arredores e dos potencialmente mutáveis atributos físicos de seus habitats.

Para entender esta complexa realidade tem se tornado imprescindível a inclusão explícita do espaço tanto nas análises como nas teorias e modelos ecológicos e várias pesquisas têm elucidado os papéis potencialmente críticos do espaço sobre as relações ecológicas o que e tem levado a um crescente interesse em Ecologia e Estatística Espacial. Atualmente várias teorias e modelos ecológicos já assumem que elementos de um ecossistema que estão próximos uns dos outros no espaço e no tempo são mais prováveis de serem influenciados pelos mesmos processos. Este é o caso, por exemplo, de modelos de epidemias e parasitismo e

das teorias de competição, sucessão, evolução e adaptação, genética de populações e manutenção de biodiversidade, interações predador - presa e comportamento social (Legendre & Fortin, 1989).

Apesar de óbvias, estas realidades são muitas vezes ignoradas na maioria dos estudos em biologia e ecologia, porque elas podem complicar bastante o trabalho de campo ou a modelagem e análise das variáveis. Como resultado, temos que o desenho amostral da maioria dos estudos biológicos e ecológicos não satisfaz as premissas básicas dos testes estatísticos paramétricos e a maioria das técnicas estatísticas comumente usadas em estudos biológicos e ecológicos se mostram inadequadas em responder questões envolvendo os tipos de dados e desenhos amostrais geralmente presentes nestes tipos de estudo.

2 - Autocorrelação Espacial e sua influência sobre testes estatísticos convencionais

A autocorrelação espacial é uma propriedade estatística muito comum e presente em muitas variáveis biológicas observadas ao longo do espaço geográfico. Este fenômeno é representado pela propriedade que as variáveis possuem de assumirem valores mais similares entre pares de amostras localizadas a certa distância do que o esperado entre pares de amostras localizadas aleatoriamente (Legendre, 1993). A autocorrelação pode ser positiva, quando os valores são mais similares, ou negativa quando os valores são mais diferentes. Esta propriedade é muito comum não somente para variáveis ecológicas, mas em qualquer variável contínua observada ao longo do espaço geográfico (autocorrelação espacial) ou ao longo de uma série temporal (autocorrelação temporal). A maioria dos fenômenos ecológicos naturais apresenta agregação geográfica, pelo menos em alguma escala espacial, de micrometros a quilômetros.

Esse fenômeno tem suas origens tanto endógenas quanto exógenas. Suas causas endógenas estão relacionadas à distribuição de recursos no ambiente e a limitação de dispersão dos organismos e suas causas exógenas relacionam-se com as pressões exercidas pela força física das variáveis ambientais. De acordo com a teoria de hierarquia (Allen & Starr, 1982), nós podemos olhar para o ambiente como sendo primeiramente estruturado por processos físicos de larga escala, como processos geomorfológicos na superfície terrestre e correntes (ventos e líquidos) em ambientes fluidos que através de entradas de energias, causam a ocorrência, por um lado, de gradientes, e por outro, de estruturas agregadas (mosaicos), separadas por discontinuidades (interfaces). Estas estruturas em larga escala induzem a formação de respostas similares nos sistemas biológicos, tanto espacialmente quanto temporalmente. Dentro destas zonas relativamente homogêneas, tomam lugar processos bióticos contagiosos que causam a aparição de mais estruturação espacial através dos processos de reprodução, mortalidade, interações presa - predador, disponibilidade de recursos, parasitismo, dispersão, entre outros.

É amplamente relatado na literatura (Legendre, 1993; Legendre & Fortin, 1989; Fortin & Dale, 2005; Rangel *et al.*, 2006) que a autocorrelação espacial representa um problema estatístico, pois sua presença nos dados viola a premissa da independência assumida pela maioria das técnicas estatísticas convencionais. Esta independência espacial é

parte do problema maior das pseudoréplicas (Hurlbert, 1984) frequentemente encontrado em experimentos manipulativos e mensurativos.

A pressuposição de aleatoriedade e independência entre as amostras, requerida pelas técnicas estatísticas convencionais são situações raramente conseguidas na maioria dos estudos de populações e comunidades biológicas. Amostras aleatoriamente distribuídas significa que todas as amostras tem a mesma chance de serem alocadas em qualquer local da área de estudo, em outras palavras, significa dizer que todos os locais do universo amostral (área de estudo) tem a mesma chance de serem amostrados. Amostras independentes significa dizer que a alocação de uma amostra não interfere no posicionamento das outras. Ou seja, utilizar um transecto e posicionar amostras a cada 10 metros deste transecto não configura independência pois a partir do momento que você posicionar a primeira amostra, o local da próxima já está automaticamente determinado e os outros locais não tem mais a mesma chance de serem amostrados.

Este problema estatístico prejudica nossa habilidade de realizar testes estatísticos de hipótese convencionais. A razão pode ser melhor ilustrada através do caso do coeficiente de correlação. O problema está no fato que, quando ambas as variáveis são positivamente autocorrelacionadas, o intervalo de confiança estimado pelo procedimento clássico ao redor de um coeficiente de correlação de Pearson é menor do que ele realmente é quando calculado corretamente, então se declara muito frequentemente que o coeficiente é significativamente diferente de zero. Com todos os outros testes estatísticos convencionais a autocorrelação positiva induz ao mesmo erro e o valor calculado do teste é frequentemente declarados significativos sob a hipótese nula. Autocorrelação negativa pode produzir o efeito oposto, por exemplo, em ANOVA (Legendre, 1993).

Esta questão pode ser contemplada do ponto de vista dos graus de liberdade, em testes estatísticos clássicos, conta-se um grau de liberdade para cada nova observação independente e este procedimento permite a escolha da distribuição estatística apropriada para o teste em questão. O problema com dados autocorrelacionados é a sua falta de independência ou, em outras palavras, o fato que cada nova observação não acrescenta um grau completo de liberdade pois o conhecimento do valor da variável em alguns locais da ao observador algum conhecimento prévio do valor que a variável irá assumir em novos locais. A consequência é que estas novas observações não podem ser contadas como um grau inteiro de liberdade. Estão disponíveis correções para o número de graus de liberdade e elas serão mencionadas. Isto pode surgir como uma surpresa para ecólogos que foram treinados sob a crença que a natureza segue as prerrogativas da estatística clássica, uma delas sendo a independência das observações.

3 - Estatística espacial-a inclusão do espaço nas análises ecológicas

Procedimentos estatísticos são frequentemente utilizados para organizar e resumir grandes conjuntos de dados para que inferências significativas possam ser feitas a respeito do fenômeno de interesse. Comumente em ecologia, o fundamento de tal inferência é um teste estatístico como um teste - *t*, *F* ou qui - quadrado ou um procedimento do

tipo ANOVA. Estas ferramentas são convenientes e fáceis de empregar, mas elas geralmente assumem que os dados são independentes entre si e igualmente distribuídos, pressuposições estas raramente plausíveis para a maioria das pesquisas ecológicas.

Assumir dependência espacial é mais prático e realista uma vez que o que nós interpretamos como fenômeno ecológico envolve o reconhecimento de uma correlação. Análises ecológicas normalmente incluem investigações sobre a dispersão e os padrões em associações entre diferentes espécies em diferentes locais e em diferentes épocas ou períodos de tempo (Pielou, 1977), padrões que refletem dependência espacial e não independência. Dependência ou continuidade espacial ou temporal deveria ser facilmente aparente para a maioria dos ecólogos, pois há vários exemplos práticos de suas implicações e ecólogos de campo sabem por experiência que na natureza os seres vivos não estão distribuídos nem de maneira completamente uniforme nem aleatória. O mesmo se aplica as variáveis físicas que nós usamos para descrever os ambientes (Legendre, 1993), por exemplo, em experimentos em casas de vegetação, as plantas são rotineiramente mudadas de posição para se eliminar efeitos microclimáticos e microambientais (do espaço e do tempo), ou seja, fugir ou minimizar os efeitos da autocorrelação espacial.

As análises de padrões ecológicos espaciais compreendem duas famílias de métodos. A “análise de padrão de pontos” e a “análise de padrão de superfícies”. A “análise de padrão de pontos” relaciona - se com a distribuição de objetos individuais no espaço, por exemplo, indivíduos isolados de plantas ou animais. O propósito principal deste tipo de análise é determinar o quanto a distribuição geográfica dos dados dos pontos é aleatória ou não e descrever o tipo de padrão, na intenção de inferir sobre que tipo de processo poderia ter criado a estrutura observada.

As “análises de padrão de superfícies” relacionam - se com o estudo de fenômenos espacialmente contínuos. A distribuição espacial das variáveis é conhecida através de amostras em pontos separados. Uma ou mais variáveis são observadas em cada ponto de observação e representam sua porção circundante no espaço. As “análises de padrão de superfícies” incluem uma família de métodos desenvolvidos para responder uma variedade de questões que podem ser resumidas em três tópicos principais. (1) Detecção e descrição da autocorrelação espacial, (2) Estimação (interpolação) e mapeamento das variáveis e (3) Testes de hipótese e modelos causais que consideram o espaço.

3.1 - Detecção e descrição da autocorrelação espacial

Ecólogos estão primeiramente interessados em quantificar a autocorrelação presente nos seus dados com o objetivo de demonstrar que não há presença de nenhuma autocorrelação significativa e tornar válido o uso dos convencionais testes estatísticos de hipótese ou o investigador pode estar interessado em demonstrar que existe a presença de significativa autocorrelação espacial com a finalidade de usá - la em modelos estatísticos ou conceituais.

As estruturas espaciais são primariamente descritas através de “funções de estrutura” que nos permitem quantificar a dependência espacial e dividi - la através das varias classes de distância. As “funções de estrutura” mais comumente usadas são os correlogramas e semi - variogramas que são

gráficos onde são plotados os valores de autocorrelação na ordenada contra distâncias geográficas ou classes de distância entre as unidades amostrais na abcissa (Legendre, 1993). O semivariograma (frequentemente chamado de variograma por simplicidade) utiliza a semivariância como medida de autocorrelação espacial enquanto o correlograma utiliza um coeficiente de autocorrelação (coeficiente I de Moran ou c de Geary). A semivariância é uma medida bastante relacionada ao coeficiente de autocorrelação c de Geary. Modelos estatísticos podem ser ajustados aos variogramas (modelo linear, exponencial, esférico, Gaussiano, etc.) o que permite ao investigador relacionar a estrutura observada aos possíveis processos geradores hipotetizados.

3.2. - Estimação (interpolação) e mapeamento

Devido ao fato que as suas formas podem corresponder a mais de um tipo de estrutura espacial, as funções de estrutura requerem uma interpretação e devem ser complementadas por mapas representando a variação espacial da variável de interesse. Boas revisões estão em Legendre e Fortin (1989) e vários programas estão disponíveis para o mapeamento (Surfer, PASSaGE, SAM). Estes métodos são obtidos por diferentes processos chamados de interpolação. Os mapas mais sofisticados podem ser obtidos pelo método geoestatístico da krigagem que faz uso de variogramas para levar em consideração a autocorrelação espacial durante o processo de interpolação. Entre outras aplicações, a krigagem é interessante para ecólogos por que ela permite uma estimativa com pouco erro da quantidade total de um recurso disponível em uma determinada área (análise de estoques) mesmo com amostragens não aleatórias.

3.3. - Modelos causais e testes de hipótese que consideram o espaço.

Uma vez que um dos objetivos principais da Ecologia de comunidades é estudar as inter - relações entre um grande número de variáveis bióticas (as espécies) e outro grande número de variáveis abióticas que descrevem o ambiente e que a maioria dessas variáveis carregam uma grande carga de autocorrelação espacial foram desenvolvidas maneiras de se testar hipóteses e correlações levando em consideração o importante papel do espaço sobre essas relações, principalmente nos casos onde a estrutura espacial é vista não como ruído, mas como uma parte do processo ecológico sob estudo. Neste caso, a alternativa é modificar o método estatístico para levar em consideração a autocorrelação espacial.

Cliff & Ord (1981) propuseram métodos para a correção do erro padrão da estimativa dos parâmetros da regressão linear simples na presença de autocorrelação. O método foi estendido a correlações lineares, regressões múltiplas e testes t por e para a ANOVA de um fator por Griffith (1987). Dutilleul (1993) propôs um teste de significância do coeficiente de correlação entre dois processos espaciais através da estimativa de um número modificado de graus de liberdade, usando uma aproximação da variância do coeficiente de correlação da amostra.

Um método versátil de testar relações causais considerando o espaço é o teste de Mantel que trabalha com duas ou três matrizes (teste de Mantel parcial) de distância ou similaridade, obtidas independentemente e que descrevem as relações entre as mesmas unidades amostrais. Neste

caso, o espaço físico é expresso por uma terceira matriz de distâncias geográficas entre as unidades amostrais e a relação entre duas matrizes pode ser calculada controlando o efeito da terceira matriz, no caso, o espaço (Legendre & Fortin, 1989).

CONCLUSÃO

É imprescindível que o papel do espaço seja corretamente estudado e inserido nas teorias e análises ecológicas. É evidente que a heterogeneidade espacial é funcional nos ecossistemas e não apenas o resultado de algum processo aleatório e gerador de ruído. Especula - se até que seria improvável que ecossistemas sem estruturação espacial funcionassem. Imaginem as conseqüências, a homogeneidade em larga - escala iria diminuir a diversidade de habitats e consumidores não seriam encontrados próximos aos seus alimentos, casos seriam localizados aleatoriamente através da paisagem, condições de solo na vizinhança imediata de uma planta não seriam mais propícias para as suas plântulas do que em qualquer outra localidade, recém nascidos seriam dispersos por ai ao invés de permanecerem em ambientes favoráveis e assim por diante. Apesar da irreabilidade deste cenário, ainda é comum ele ser assumido em muitas de nossas teorias e modelos descrevendo populações e funcionamento de comunidades. Esta mudança de pontos de vistas se traduz em um novo paradigma para ecólogos, no qual a estruturação espacial é um importante componente para o ecossistema.

REFERÊNCIAS

- Allen, T.F.H. & Starr, T.B. 1982.** *Hierarchy: perspectives for ecological complexity.* University of Chicago Press, Chicago.
- Cliff, A.D. & Ord, J.K. 1981.** *Spatial processes: models and applications.* Pion Limited, London.
- Dutilleul, P.; Clifford, P.; Richardson, S.; Hemon, D. 1993.** Modifying the *t* Test for Assessing the Correlation Between Two Spatial Processes. *Biometrics* 49:305 - 314.
- Fortin, M. - J. & Dale, M.R.T. 2005.** *Spatial analysis: a guide for ecologists.* Cambridge University Press, Cambridge.
- Griffith, D.A. 1987.** *Spatial autocorrelation: a primer.* Association of American Geographers, Washington, D.C.
- Hurlbert, S.H. 1984.** Pseudoreplication and the Design of Ecological Field Experiments. *Ecological Monographs* 54(2): 187 - 211.
- Legendre, P. 1993.** Spatial autocorrelation: Trouble or new paradigm? *Ecology* 74(6): 1659 - 1673.
- Legendre, P. & Fortin, M. - J. 1989.** Spatial pattern and ecological analysis. *Vegetatio* 80: 107 - 138.
- Matheron, G. 1963.** Principles of Geostatistics. *Economic Geology* 58: 1246 - 1266.
- Pielou, E.C. 1977.** *Mathematical ecology.* J. Wiley & Sons, New York.
- Rangel, T.F.L.V.B.; Diniz - Filho, J.A.L. & Bini, L.M. 2006.** Towards an integrated computational tool for spatial analysis in macroecology and biogeography. *Global Ecology and Biogeography* 15: 321-327.