I Encontro Luso-Galaico de Estatística em Ambiente e Ecologia

Sessão Convidada

Modelagem da densidade relativa de peixes numa rede fluvial

Alexandra M. Schmidt
Universidade Federal do Rio de Janeiro, Brasil, alex@im.ufrj.br
João Batista M. Pereira
Universidade Federal do Rio de Janeiro, Brasil, joao@dme.ufrj.br
Marco A. Rodríguez
Université du Québec à Trois-Rivières, Canadá, marco.rodriguez@uqtr.ca

Palavras–chave: Convolução, Inferência Bayesiana, Processo Gaussiano.

Resumo: Processos ecológicos que operam em múltiplas escalas podem induzir uma autocorrelação espacial na densidade relativa de espécies. Em redes fluviais, a autocorrelação na contagem de peixes pode ser gerada (1) por processos direcionais restritos pelo fluxo da água, como o movimento de peixes a montante e a jusante, o transporte de nutrientes e poluição; (2) ou, por processos geológicos que influenciam a estrutura do solo da bacia hidrográfica, topografia local, e a composição química da água em diferentes escalas espaciais. Utilizamos aqui uma abordagem por convolução espacial para modelar a densidade relativa de peixes numa rede fluvial. Em particular, assumimos uma mistura Poisson-lognormal que incorpora efeitos aleatórios para o tempo e o espaço. A componente espacial incorpora distância euclideana e efeitos direcionais, para capturar efeitos geológicos como, também, considera distância hidrológica ("como o peixe nada"). Adicionalmente, covariáveis são utilizadas para capturar efeitos ambientais esperados. O procedimento de inferência é realizado sob o paradigma de Bayes e ajustamos diferentes modelos que representam diferentes combinações de efeitos temporais e espaciais, incluindo um que assume independência completa no espaço. Os modelos são comparados a partir de critérios de informação e, também, a partir da comparação da variância relativa capturada por cada componente espacial. Comparamos as distribuições a priori e a posteriori dessas variâncias para investigar o ganho de informação relativo à contribuição dos diferentes processos para explicar a variação na contagem de peixes.

1 Introdução

Geralmente, ecologistas buscam interpretar a variação na distribuição espacial de populações em termos de respostas à características ambientais. Entretanto, a distribuição de populações é influenciada, simultaneamente, por uma grande quantidade de variáveis ambientais que variam no espaço e no tempo, e nem sempre podem ser observadas diretamente. Separar os efeitos individuais dessas variáveis em populações e, mais geralmente, interpretar os efeitos ambientais nas análises ecológicas pode ser desafiador se a influência de variáveis latentes, e também, das correlações temporais e espaciais, não são consideradas. Nosso principal objetivo neste estudo é entender a variação espacial na abundância relativa da truta das fontes (*Salvenilus fontinalis*). A contagem relativa de truta das fontes e as medidas de variáveis ambientais, foram quantificadas em 600 seções ditribuídas entre 120 pontos amostrais e 22 afluentes na bacia do Rio Cascapedia, no Quebec, Canadá. As localizações foram visitadas numa sequência aleatória durante a época do verão no hemisfério norte, de meados de junho ao fim de agosto nos anos de 2000, 2001 e 2002.
A modelagem da dependência espacial de observações feitas em redes fluviais é um grande desafio devido às características intrínsecas deste meio, que fazem com que a metodologia geoestatística tradicional não seja adequada. Ver Hoef *et al.* (2006) e Cressie *et al.* (2006) desenvolveram uma metodologia específica para a modelagem da dependência espacial que leva em conta aspectos dos

rios e córregos que compõem as redes fluviais. Estes autores desenvolvem modelos cujas funções de correlação dependem de distância fluvial ao invés da, tradicional, distância euclideana. A distância fluvial é a distância calculada ao longo da rede fluvial. Estes modelos estão baseados em médias móveis que possuem cauda para baixo ou, para cima, dependendo se a corrente do riacho está a jusante ou a montante, respectivamente. Este trabalho estende os modelos propostos por Ver Hoef e co-autores para o caso de contagens e apresenta um procedimento de inferência baseado no paradigma de Bayes. Este documento está organizado da seguinte forma. A próxima seção apresenta o modelo proposto e descreve brevemente o procedimento de inferência adotado. A seção 4 descreve os principais resultados a serem apresentados a partir da análise dos dados do Rio Cascapedia.

2 Modelo Proposto

Seja $Y(s)$ o número de indivíduos (contagens) observados na localização s. Assumimos que, condicional a $\alpha(s)$, $Y(s)$ segue uma distribuição de Poisson independente, de modo que

$$Y(s) \mid \alpha(s) \sim Poi(\alpha(s)), \tag{1}$$

onde $\log\,\alpha(s)$, o logaritmo da contagem média, é decomposta na soma de efeitos de covariáveis e efeitos latentes, tal que

$$\log \alpha(s) = \log e(s) + X(s)'\beta + \gamma(t_i(s)) + \nu(s), \tag{2}$$

onde $e(s)$ é um termo de *offset*, descrito pela área média da seção; $X(s)'\beta$ descreve o efeito de covariáveis que capturam efeitos do habitat local, e β é um vetor p dimensional de coeficientes a serem estimados; $\gamma(t_i(s))$ captura efeitos temporais comuns as n_t localizações observadas num particular dia, $t_i(\cdot)$ é uma função determinística que associa a cada localização a ordem do dia juliano, dentro de um particular ano i ($i = 2000, 2001, 2002$), relativo à localização s, isto é, $\{t_i(\cdot) \in 1, \cdots, T_i\}$, onde T_i é o número total de dias observados no ano i, e $\nu(s)$ captura efeitos locais remanescentes, depois de considerar os efeitos temporais e das covariáveis. Em seguida, descrevemos como modelamos cada uma destas componentes.

2.1 Modelando o efeito temporal $\gamma(t_i(\cdot))$

Observações são obtidas ao longo dos dias de verão de três anos consecutivos. Acreditamos ser razoável assumir que observações são indepentendes ao longo dos anos mas permitimos uma correlação temporal entre observações de um mesmo ano. Por esta razão, assumimos que os efeitos aleatórios, $\gamma(t_i(s))$, seguem uma distribuição normal multivariada com vetor de médias igual a zero e estrutura de covariâncias dada por

$$Cov(\gamma(t_i(s)), \gamma(t_k(s'))) = \begin{cases} 0, & i \neq k \\ \tau^2 \exp\left\{-\frac{1}{\phi_\gamma}|t_i(s) - t_k(s')|\right\}, & i = k, \end{cases}$$

onde $i = 2000, 2001, 2002$, e $|t_i(s) - t_i(s')|$ é a diferença absoluta entre os dias julianos das observações feitas nas localizações s e s' no ano i. O parâmetro $\tau^2 > 0$, que é considerado comum ao longo dos anos, captura a estrutura de variância de $\gamma(t_i(s))$; e $\phi_\gamma > 0$ captura quão rápido a correlação decai a zero, à medida que os dias dentro de um ano, se tornam mais distantes. Note que observações feitas num mesmo dia possuem um efeito $\gamma(t_i(\cdot))$ comum.

2.2 Modelando o efeito local $\nu(s)$

A componente $\nu(s)$ está presente para capturar os efeitos locais relacionados a localização s depois de considerar os efeitos temporais e de covariáveis. É razoável assumir que $\nu(s)$ e $\nu(s')$ tendem a ser mais correlacionados se s e s' estão próximas, e mais independentes se s e s' estão mais distantes.

A correlação espacial na contagem de peixes pode ocorrer por causa de (1) processos direcionais restritos pelo fluxo da água, como o movimento de peixes a montante e a jusante, o transporte de nutrientes e poluição; (2) processos geológicos que influenciam a estrutura do solo da bacia hidrográfica, topografia local, e a química da água em diferentes escalas espaciais. Assumimos aqui que $\nu(\cdot)$ segue um processo Gaussiano com média zero e exploramos diferentes estruturas de correlação. Como apontado por Peterson e Ver Hoef (2010) há situações onde seria útil considerar modelos que permitem uma correlação mais forte entre localizações conectadas pelo fluxo pela água, do que não conectadas pelo fluxo. Por exemplo, [8] estimaram que *truta cerrada* (*Oncorhynchus clarki*) move mais frequentemente a montante do que a jusante.
Seguimos aqui os modelos propostos por Ver Hoef e Peterson (2010) e Peterson e Ver Hoef (2010) e permitimos que a estrutura de covariância geral de $\nu(\cdot)$ siga uma modelo de mistura entre componentes *tail-up*, tail-down, distância euclideana e efeitos pepita. Seja $\nu = (\nu(s_1), \cdots, \nu(s_n))'$, onde n é o número total de localizações observadas. Então assumimos que $\nu \sim N(0, \Sigma)$, onde Σ, é a matriz de covariâncias de ν.
Ver Hoef e Peterson (2010) propõem diferentes estruturas de covariância para modelos *tail-up* e tail-down. A seguir nos concentraremos nos modelos baseados em funções médias móveis com caudas mais pesadas (Modelos Mariah) já que os resultados a serem apresentados estão baseados nestes modelos. Outras estruturas de covariância foram ajustadas como, por exemplo, a exponencial, mas os resultados não foram muito diferentes daquele obtidos sob os modelos Mariah.

Modelo Mariah *tail-up* Pode ser mostrado que sob a função núcleo Mariah (Ver Hoef e Peterson, 2010), a covariância do processo é dada por

$$C_c(h \mid \sigma_u^2, \phi_u) = \begin{cases} \sigma_u^2 \left(\frac{\log(h/\phi_u+1)}{h/\phi_u} \right), & \text{se } h > 0 \\ \sigma_u^2, & \text{se } h = 0, \end{cases}$$

onde $\sigma_u^2 > 0$ é a variância relacionada à componente *tail-up* e $\phi_u > 0$ é o parâmetro que mede quão rápido a correlação decai a zero à medida que a distância entre as localizações h aumenta. Seja C_c a matriz de covariâncias n dimensional resultante do alinhamento das covariâncias entre as n localizações. No modelo *tail-up* a função núcleo é partida à medida que vai a montante. E isto é obtido através da atribuição de pesos em cada seguimento de fluxo, de modo que a matriz de covariâncias resultante é dada por

$$C_u = W \odot C_c, \tag{3}$$

onde $\odot$ é o produto de Hadamard, e W é uma matriz de pesos com elementos $\pi_{i,j}$ se duas localizações i e j são conectadas pelo fluxo, e 0 se essas localizações não são conectadas pelo fluxo.
Os pesos da matriz W podem ser escolhidos baseados em diferentes características da bacia sob estudo. Por exemplo, Ver Hoef *et al.* (2006) fixaram esses pesos como a raiz quadrada da percentagem do volume de fluxo que uma localização a jusante recebe de uma localização a montante. Cressie *et al.* (2006), por outro lado, utilizou pesos baseados na ordem de Shreve (Shreve, 1967). Ver Hoef e Peterson (2010) mostram que essas abordagens são similares. Como não possuímos informação sobre as características de fluxo de volume para os nossos dados, consideramos pesos iguais a $\sqrt{1/2}$ Ver Hoef *et al.* (2006) para todos os modelos que consideram componentes *tail-up*.

Modelos Mariah *tail-down* Diferentemente dos modelos *tail-up*, nos modelos *tail-down* é preciso considerar tanto localizações conectadas, como não conectadas, pelo fluxo. A função de covariância Mariah para os modelos *tail-down* (Ver Hoef e Peterson, 2010) são dadas por

$$C_d(h,a,b \mid \sigma_d^2,\phi_d) = \begin{cases} \sigma_d^2\left(\frac{\log(h/\phi_d+1)}{h/\phi_D}\right), & \text{se são conectadas pelo fluxo e } h>0 \\ \sigma_d^2, & \text{se são conectadas pelo fluxo e} h=0 \\ \sigma_d^2\left(\frac{\log(a/\phi_d+1)-\log(b/\phi_D+1)}{a-b/\phi_d}\right), & \text{se não são conectadas pelo fluxo e } a\neq b \\ \sigma_d^2\left(\frac{1}{a/\phi_d+1}\right), & \text{se não são conectadas pelo fluxo e } a=b. \end{cases} \quad (4)$$

Note que a função de covariância acima depende da distância hidrológica h para localizações conectadas pelo fluxo, e para localizações não conectadas pelo fluxo a função depende de a e b, que são as distâncias hidrológicas entre essas localizações não conectadas e sua bifurcação mais próxima. Como apontado por Ver Hoef e Peterson (2010), a covariância entre localizações conectadas pelo fluxo é a mesma para modelos *tail-up* e *tail-down*, exceto pelos pesos. Denotaremos por C_d a função de covariância entre as n localizações monitoradas sob o modelo *tail-down* Mariah.

Modelo exponencial baseado na distância euclideana Diferentemente dos modelos anteriores, este modelo não usa informação baseada na distância hidrológica. Aqui a estrutura de covariância é apenas função da distância euclideana entre as localizações. Em particular, assumimos que $C_e(d) = \sigma_e^2\exp(-d/\phi_e)$, onde d é a distância euclideana entre duas localizações quaisquer na bacia, $\sigma_e^2 > 0$ é o parâmetro de variância, e ϕ_e é o parâmetro de decaimento que mede quão rápido a correlação decai a zero à medida que a distância aumenta. Denotamos por C_e, a função de covariância entre as n localizações observadas sob o modelo exponencial, como função da distância euclideana entre os pontos.

Modelo de mistura Poisson-lognormal Se consideramos o vetor $Y = (Y(s_1),\cdots,Y(s_n))'$ obtido do empilhamento de todas as contagens, ao longo de todos os dias, durante os diferentes anos, podemos escrever o modelo em (1) como um modelo de mistura Poisson-lognormal, de modo que

$$\begin{aligned} Y \mid \alpha &\sim Poi(\alpha) \\ \log\alpha &\sim N(\log e + X'\beta + A\gamma, \Sigma), \end{aligned} \quad (5)$$

onde X é uma matriz $n\times p$ contendo informação das p covariáveis, γ é o vetor contendo todos os efeitos temporais para todos os anos, de modo que $\gamma = (\gamma_1,\gamma_2,\gamma_3)'$, onde γ_i é um vetor de dimensão T_i com componentes correspondendo aos efeitos temporais para o ano i. Assim, γ tem dimensão $T = T_1+T_2+T_3$, o número total de dias visitados ao longo dos três anos. Assumindo que observamos n localizações espaciais durante os 3 anos de estudo, podemos associar os efeitos temporais aleatórios às respectivas localizações considerando a matriz $A = \text{blockdiag}(A_{T_1}, A_{T_2}, A_{T_3})$, onde A_{T_i} é uma matriz de dimensão T_i cujos elementos são 1's e 0's. Se integramos $\log\alpha$ com respeito a $A\gamma$, então $\log\alpha \sim N(\log e + X'\beta, \Lambda)$, onde $\Lambda = \Omega + \Sigma$, com $\Omega = \text{blockdiag}(\Omega_1,\Omega_2,\Omega_3)$, e $\Omega_i(k,l) = \tau^2\exp\left\{-\frac{1}{\phi_\gamma}|t_i(s_k) - t_i(s_l)|\right\}$, e $\Sigma = \sigma_u^2 R_u + \sigma_d^2 R_d + \sigma_e^2 R_e + \sigma^2 I_n$, onde I_n é a matriz identidade de dimensão n. Aqui, R_u, R_d e R_e são as matrizes de correlação asssociadas as matrizes de covariância C_u, C_d e C_e. Note que σ^2 joga o papel de um efeito pepita para $\log\alpha$, em outras palavras, captura tudo o que sobra depois de considerar as diferentes estruturas de correlação espacial.

Portanto, o modelo proposto pode ser visto como um modelo de mistura Poisson-lognormal (Aitchison e Ho, 1989), cuja componente de mistura captura sobredispersão e diferentes fontes de estruturas de correlação. Casos particulares do modelo geral são obtidos fazendo um dos parâmetros de escala, τ^2, σ_u^2, σ_d^2 σ_e^2, ou σ^2, igual a 0. Se todos os parâmetros de escala forem simultaneamente iguais a zero, não capturamos sobredispersão, nem correlações temporal e espacial, possivelmente presentes nos dados.

3 Procedimento de Inferência

Considere o modelo Poisson descrito em (1). Seja $y = (y(s_1), \ldots, y(s_n))'$ um vetor de contagens observadas em n localizações. A função de verossimilhança pode ser descrita como

$$L(\mathbf{y}; \boldsymbol{\Theta}) \;=\; \prod_{i=1}^{n} \frac{1}{y(s_i)!} \exp\{-\exp(\eta(s_i))\} \exp\{\eta(s_i) y(s_i)\},$$

em que $\boldsymbol{\Theta}$ representa o vetor cujas componentes são todos os parâmetros envolvidos na definição do modelo.

O procedimento de inferência é baseado no paradigma de Bayes. Para completar a especificação do modelo, atribuímos uma distribuição *a priori* para o vetor paramétrico $\boldsymbol{\Theta}$. Para o modelo geral descrito na seção 2, as componentes de $\boldsymbol{\Theta}$ são $(\beta, \tau^2, \phi_\gamma, \sigma_u^2, \sigma_d^2, \phi_u, \phi_d, \sigma^2)$.

Assumimos independência *a priori* entre as componentes de $\boldsymbol{\Theta}$, assim

$$p(\boldsymbol{\Theta}) \;=\; p(\beta)p(\tau^2)p(\phi_\gamma)p(\sigma_u^2)p(\sigma_d^2)p(\phi_u)p(\phi_d)p(\sigma^2). \tag{7}$$

Para o vetor de coeficientes de regressão β será assumida uma distribuição *a priori* normal multivariada, isto é, $\beta \sim N(m_\beta, C_\beta)$, em que o vetor de médias m_β e a matriz de covariância C_β são conhecidos. Para os parâmetros de variância temporal e espacial serão assumidas distribuições *a priori* gama invertida, isto é,

$$\tau^2 \sim IG(a_\tau, b_\tau), \quad \sigma_u^2 \sim IG(a_u, b_u) \quad \sigma_d^2 \sim IG(a_d, b_d) \text{ e } \sigma^2 \sim IG(a_n, b_n), \tag{8}$$

em que os valores de a_τ, b_τ, a_u, b_u, a_d, b_d, a_n e b_n são conhecidos. Os parâmetros ϕ_γ, ϕ_u e ϕ_d são assumidos seguirem uma distribuição gama, isto é,

$$\phi_\gamma \sim Gama(c_\gamma, d_\gamma), \quad \phi_u \sim Gama(c_u, d_u), \quad \phi_d \sim Gama(c_d, d_d), \tag{9}$$

em que c_γ, d_γ, c_u, d_u, c_d e d_d são valores conhecidos.

Uma vez atribuída uma distribuição *a priori* para o vetor paramétrico $\boldsymbol{\Theta}$ e seguindo o teorema de Bayes, é possível obter o núcleo da a distribuição *a posteriori*, isto é

$$\begin{aligned} p(\boldsymbol{\Theta} \mid \mathbf{y}) \;\propto\; & \prod_{i=1}^{n} p(y(s_i) \mid \eta(s_i))\, p(\boldsymbol{\eta} \mid \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_u^2, \sigma_d^2, \phi_u, \phi_d, \sigma^2) \\ & p(\boldsymbol{\beta})p(\boldsymbol{\gamma} \mid \tau^2, \phi_\gamma)p(\tau^2)p(\phi_\gamma)p(\sigma_u^2)p(\sigma_d^2)p(\phi_u)p(\phi_d)p(\sigma^2) \end{aligned} \tag{10}$$

A equação acima não é analiticamente tratável. Para que seja possível realizar o procedimento de inferência e obter amostras dos parâmetros a partir da sua distribuição *a posteriori*, faz-se necessária a utilização de métodos de simulação estocástica. Métodos de Monte Carlo via Cadeias de Markov (MCMC, na sigla em inglês) serão utilizados. Em particular, será utilizado o amostrador de Gibbs, introduzido por Geman e Geman (1984) e adaptado no contexto de simulações estocásticas por Gelfand e Smith (1990), com passos de Metropolis-Hastings, algoritmo introduzido por Metropolis *et al.* (1953) e estendido por Hastings (1970).

4 Discussão

A metodologia proposta por Ver Hoef *et al.* (2006), Cressie *et al.* (2006) e Ver Hoef e Peterson (2010) para dados espacialmente referenciados observados numa rede fluvial incorpora características hidrológicas importantes através da convolução de diferentes funções de médias móveis. Neste trabalho, focamos em utilizar esta metodologia para o caso de contagens e realizamos o procedimento de inferência sob o paradigma de Bayes.

A partir do modelo geral apresentado na equação (1) ajustamos diferentes modelos às contagens observadas de truta das fontes na bacia do rio Cascapedia. Essencialmente, os modelos ajustados diferem na estrutura de covariância associada aos efeitos latentes locais $\nu(\cdot)$. Todos os modelos abaixo assumem efeitos fixos, representados pela componente $X'\beta$ na equação (1), que contém as seguintes covariáveis: velocidade média da corrente, índice do tamanho médio de substrato, resíduos lenhosos grandes, incremento máximo do nível na enchente, indicadora do ano de amostragem, comprimento médio e profunidade média da seção. Abaixo listamos os cinco modelos ajustados:

M1. estrutura de covariância de $\nu(\cdot)$ é uma matriz diagonal, isto é, $\Lambda = \Omega + \sigma^2 I_n$;

M2. estrutura de covariância de $\nu(\cdot)$ é função da distância euclideana e efeito pepita, isto é, $\Lambda = \Omega + \sigma_e^2 R_e + \sigma^2 I_n$;

M3. estrutura de covariância de $\nu(\cdot)$ possui estrutura *tail-up* e efeito pepita, isto é, $\Lambda = \Omega + \sigma_u^2 R_u + \sigma^2 I_n$;

M4. estrutura de covariância de $\nu(\cdot)$ possui estrutura *tail-down* e efeito pepita, isto é, $\Lambda = \Omega + \sigma_d^2 R_d + \sigma^2 I_n$;

M5. model completo, isto é,$\Lambda = \Omega + \sigma_u^2 R_u + \sigma_d^2 R_d + \sigma_e^2 R_e + \sigma^2 I_n$.

Durante a apresentação mostrarei resultados baseados nos ajustes destes modelos. A comparação de modelos será feita através de critérios de informação como o *Deviance Information Criterion* (DIC) (Spiegelhalter et al., 2002) e o *Ranking Probability Score* (RPS) (Gneiting et al., 2007). Notaremos que a distribuição a posteriori do intercepto não é sensível à especificação da priori de $\nu(\cdot)$. Velocidade média da corrente e largura média da seção possuem um efeito negativo, enquanto que resíduos lenhosos grandes possuem um efeito positivo no logaritmo da média da contagem relativa de peixes. As outras covariáveis consideradas na análise não mostraram efeito na explicação da média da contagem relativa de truta das fontes. A análise dos resultados sugere que algum tipo de correlação espacial deve ser considerada, já que o modelo que assume independência entre os elementos de $\nu(\cdot)$ é o que resulta nos piores valores de DIC e RPS. Observamos também que cada um dos modelos que assumem diferentes estruturas de correlação espacial (M2, M3 e M4) fornecem resultados semelhantes, sugerindo que nenhum deles é melhor do que o outro na captura da estrutura remanescente nos dados após considerar o efeito de covariáveis e o efeito temporal. Também observamos que o modelo completo, M5, fornece resultados similares quando comparado aos modelos M2, M3 e M4. Discutiremos detalhadamente porque acreditamos que isso acontece.

Referências

[1] Aitchison, J., Ho, C.H. (1989). The multivariate Poisson-log Normal distribution. *Biometrika* 76, 643–653.

[2] Cressie, N., Frey, J., Harch, B., Smith, M. (2006). Spatial Prediction on a River Network. *Journal of Agricultural, Biological, and Environmental Statistics* 11, 127–150.

[3] Gelfand, A., Smith, A. (1990). Sampling-based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association* 85, 398–409.

[4] Geman, S., Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 721–741.

[5] Gneiting, T., Balabdaoui, F., Raftery, A.E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 69(2), 243–268.

[6] Hastings, H. (1970). Monte Carlo Sampling Methods Using Markov chains and their Applications. *Biometrika* 57, 97–109.

[7] Metropolis, N., Rosenbulth, A., Rosenbulth, M., Teller, A., Teller, E. (1953). Equation of State Calculations by Fast Computing Machine. *Journal of Chemical Physics* 21, 1087–1091.

[8] Peterson, D.P., Fausch, K.D. (2003). Upstream movement by nonnative brook trout (*Salvelinus fontinalis*) promotes invasion of native cutthrat trout (*Oncorhynchus clarki*) habitat. *Canadian Journal of Fisheries and Aquatic Sciences* 60, 1502–1516.

[9] Peterson, E., Theobald, D.M., Ver Hoef, J.M. (2010). Geostatistical modelling on stream networks: developing valid covariance matrices based on hydrologic distance and stream flow. *Freshwater Biology* 52, 267–279.

[10] Shreve, R. (1967). Infinite Topologically Random Channel Networks. *Journal of Geology* 75, 178–186.

[11] Spiegelhalter, D., Best, N., Carlin, B., Linde, A. (2002). Bayesian Measures of Model Complexity and fit (with discussion). *Journal of Royal Statistical Society B* 64, 583–639.

[12] Ver Hoef, J., Peterson, E. (2010). A Moving Average Approach for Spatial Statistical Models of Stream Networks. *Journal of the American Statistical Association* 105, 6–18.

[13] Ver Hoef, J., Peterson, E., Theobald, D. (2006). Spatial Statistical Models that Use Flow and Stream Distance. *Environmental and Ecologial Statistics* 13, 449–464.