Potential modeling with uncertain covariables

Helmut Schaeben[1],
Raimon Tolosana Delgado[2], K. Gerald van den Boogaart[1,2]
[1] Geophysics and Geoinformatics, TU Bergakademie Freiberg, Germany
[2] Helmholtz Institute Freiberg for Resource Technology, Germany

18th Annual Conference IAMG2017
Fremantle, Australia, Sep 2-9, 2017

Potential modeling with uncertain covariables

Contents

Potential modeling with uncertain covariables

Objective of potential modeling

The ultimate goal of “potential modeling” is to recognize locations for which the estimated conditional probability of a “target” event T - like a specific mineralization - is a relative maximum – **regression modeling**.

Potential modeling with uncertain covariables

Objective of potential modeling

The ultimate goal of "potential modeling" is to recognize locations for which the estimated conditional probability of a "target" event T - like a specific mineralization - is a relative maximum – **regression modeling**.

The event must be sufficiently well understood in terms of cause and effect to collect data corresponding to spatially referenced covariables $B_\ell, \ell = 0, \ldots, m$, in favor or against the event T to occur – **causality modeling**.

Potential modeling with uncertain covariables

Objective of potential modeling

The ultimate goal of "potential modeling" is to recognize locations for which the estimated conditional probability of a "target" event T - like a specific mineralization - is a relative maximum – **regression modeling**.

The event must be sufficiently well understood in terms of cause and effect to collect data corresponding to spatially referenced covariables $\mathrm{B}_\ell, \ell = 0, \ldots, m$, in favor or against the event T to occur – **causality modeling**.

Then spatially referenced "posterior" probabilities given the covariables can be estimated by several approaches including weights-of-evidence, logistic regression, compositional regression, artificial neural nets, and many others;
generally by **statistical learning** or **machine learning**.

Potential modeling with uncertain covariables

Deficiencies of Raster-Based Models

- **training region** must provide perfect knowledge (*no uncertain covariables, no missed occurrence*)
- **spatial resolution** a user decision: different resolutions may result in inconsistent models, cf. Baddeley et al., 2010
- **non-spatial** models, neglecting spatially induced dependences (*shuffle pixels, no change in results*)
- **Weights-of-Evidence** (WofE) require joint **conditional independence**: assumption cannot be relaxed while preserving method simplicity
- **logistic regression**: canonical generalization of WofE; a proper regression model may be **difficult to design**.
- **artificial neural nets** (ANN) focus on prediction, **do not provide insight and understanding**.

Potential modeling with uncertain covariables

Towards Poisson and Cox point processes

- Logistic regression with respect to (very) small pixels is approximately a Poisson point process.
- A **Poisson point process with a log-linear model of the intensity field is a canonical generalization of logistic regression: Cox process**

$$f(\boldsymbol{x};\lambda) = \exp\left(-\int_W (\lambda(u)-1)\,du\right)\prod_{i=1}^{n(\boldsymbol{x})}\lambda(x_i),$$

$$\ln\lambda(u) = B(u) + \sum_{\ell}\theta_\ell\, Z_\ell(u)$$

Baddeley, Rubak, Turner, 2015, Spatial Point Patterns: Methodology and Applications with R: Chapman and Hall/CRC
Addrian Baddeley, IAMG Matheron Lecturer 2008

Potential modeling with uncertain covariables

Cox point process

a Poisson process with locally-varying **intensity field** $\lambda(u)$

- total number $\mathrm{N}(\mathbf{X} \cap B)$ of random points inside a domain $B \subset \mathbb{R}^n \sim$ Poisson with parameter $\lambda(B)$

$$\mathbb{P}\Big(\mathrm{N}\left(\mathbf{X} \cap B\right) = n(\boldsymbol{x})\Big) = \exp\left(-\lambda\left(B\right)\right) \frac{\Big(\lambda\left(B\right)\Big)^{n(\boldsymbol{x})}}{n(\boldsymbol{x})!}$$

$$\mathbb{E}\Big(\mathrm{N}\left(\mathbf{X} \cap B\right)\Big) = \mathrm{Var}\Big(\mathrm{N}\left(\mathbf{X} \cap B\right)\Big) = \lambda(B) = \int_B \lambda(u) du,$$

- The Cox process with intensity field $\lambda(u) > 0$ has **density**

$$f(\boldsymbol{x}; \lambda) = \exp\left(-\int_W \left(\lambda(u) - 1\right) du\right) \prod_{i=1}^{n(\boldsymbol{x})} \lambda(x_i),$$

where $n(\boldsymbol{x})$ denotes the total number of points in $\boldsymbol{x}$, and $|W|$ is the volume of W.

Potential modeling with uncertain covariables

Point processes

A **spatial covariable** is a random function $(\mathbf{Z}(u), u \in W)$ observable at every location $u \in W$.

Spatial covariables may be used to model the intensity $\lambda(u)$ as

- deviations from a "baseline" $B(u)$

$$\lambda(u) = \rho(Z(u))\, B(u),$$

  with some function ρ to be estimated; or

- in terms of a **log-linear model**

$$\ln \lambda(u) = B(u) + \sum_{\ell} \theta_\ell\, Z_\ell(u)$$

If little is known about $\mathbf{Z}(u)$, modeling assumptions concerning its distribution or geostatistics may apply.

Potential modeling with uncertain covariables

Likelihood of a general inhomogeneous Poisson point process

For a **general inhomogeneous Poisson point process** governed by a parameter $\boldsymbol{\theta}$ given a point pattern $\boldsymbol{x}$ inside a window W the likelihood function is

$$L(\boldsymbol{\theta}|\boldsymbol{x}) = \exp\left(-\int_W (\lambda_{\boldsymbol{\theta}}(u) - 1)\, du\right) \prod_{i=1}^{n_W(\boldsymbol{x})} \lambda_{\boldsymbol{\theta}}(x_i).$$

Maximization of the log-likelihood

$$\ln L(\boldsymbol{\theta}|\boldsymbol{x}) = \sum_{i=1}^{n_W(\boldsymbol{x})} \ln \lambda_{\boldsymbol{\theta}}(x_i) - \int_W \lambda_{\boldsymbol{\theta}}(u) du$$

with classical techniques is **numerically unfeasible** in practice (too many variables, largely unbalanced datasets).

Potential modeling with uncertain covariables

Likelihood of log-linear Poisson point process

Log-linear model of λ

$$\lambda_{\boldsymbol{\theta}}(u) = \exp\left(\sum_{\ell=0}^{m} \theta_\ell Z_\ell(u)\right) = \exp\Big(g_{\boldsymbol{\theta}}(\boldsymbol{Z}, u)\Big)$$

leads to the **likelihood function**

$$L(\boldsymbol{\theta}|\boldsymbol{x}) = \exp\left(-\int_W \exp\Big(g_{\boldsymbol{\theta}}(\boldsymbol{Z}, u)\Big) du\right) \prod_{i=1}^{n_W(\boldsymbol{x})} \exp\Big(g_{\boldsymbol{\theta}}(\boldsymbol{Z}, x_i)\Big).$$

Potential modeling with uncertain covariables

Likelihood of log-linear Poisson point process

Log-linear model of λ

$$\lambda_{\boldsymbol{\theta}}(u) = \exp\left(\sum_{\ell=0}^{m} \theta_\ell Z_\ell(u)\right) = \exp\left(g_{\boldsymbol{\theta}}(\boldsymbol{Z}, u)\right)$$

leads to the **log-likelihood function**

$$\begin{aligned}\ln L(\boldsymbol{\theta}|\boldsymbol{x}) &= \sum_{i=1}^{n_W(\boldsymbol{x})} g_{\boldsymbol{\theta}}(\boldsymbol{Z}, x_i) - \int_W \exp\left(g_{\boldsymbol{\theta}}(\boldsymbol{Z}, u)\right) du \\ &= \sum_{\ell=0}^{m} \theta_\ell \sum_{i=1}^{n_W(\boldsymbol{x})} Z_\ell(x_i) - \int_W \exp\left(\sum_{\ell=0}^{m} \theta_\ell Z_\ell(u)\right) du.\end{aligned}$$

Under mild conditions, the MLE exists and is unique (Baddeley et al., 2015, p. 344).

Potential modeling with uncertain covariables

Likelihood of log-linear Poisson point process

Numerically, the MLE is the solution of the **score equations** $\boldsymbol{U}(\boldsymbol{\theta}|\boldsymbol{x}) = 0$, where the score function is

$$\boldsymbol{U}(\boldsymbol{\theta}|\boldsymbol{x}) = \sum_{i=1}^{n_W(\boldsymbol{x})} \boldsymbol{Z}(x_i) - \int_W \boldsymbol{Z}(u)\lambda_{\boldsymbol{\theta}}(u)du.$$

The score is a vector with components

$$U_\ell(\boldsymbol{\theta}|\boldsymbol{x}) = \sum_{i=1}^{n_W(\boldsymbol{x})} Z_\ell(x_i) - \int_W Z_\ell(u)\lambda_{\boldsymbol{\theta}}(u)du, \;\; \ell = 0, \ldots, m.$$

Potential modeling with uncertain covariables

Likelihood of log-linear Poisson point process

The log-linear model of λ initially presumes that $Z_\ell, \ell = 0, \ldots, m$, are known **for all** $u \in W \subset \mathbb{R}^n$.

Potential modeling with uncertain covariables

Likelihood of log-linear Poisson point process

The log-linear model of λ initially presumes that $Z_\ell, \ell = 0, \ldots, m$, are known **for all** $u \in W \subset \mathbb{R}^n$.

If they are known for a **finite set** $s_i \in \mathbb{R}^n, i = 1, \ldots, q$, of locations only, then the expectation of the likelihood function with respect to the probability of the random function $(\mathbf{Z}(s), s \in W)$ given the data $\mathbf{Z}(s_i), i = 1, \ldots, q$, applies, i.e.,

$$\mathbb{E}_{\mathbb{P}(\circ|\mathrm{data})}\Big(L(\boldsymbol{\theta}|\boldsymbol{x})\Big) =$$

$$\mathbb{E}_{\mathbb{P}(\circ|\mathrm{data})}\Big(\exp\Big(-\int_W \exp\Big(g_{\boldsymbol{\theta}}(\boldsymbol{Z}, u)\Big)du\Big) \prod_{i=1}^{n_W(\boldsymbol{x})} \exp\Big(g_{\boldsymbol{\theta}}(\boldsymbol{Z}, x_i)\Big)\Big)$$

which appears as of today **numerically intractable**.

Potential modeling with uncertain covariables

Likelihood of log-linear Poisson point process

Log-linear model of λ, i.e., log-linear random function,

$$\ln\big(\lambda_{\boldsymbol{\theta}}(u)\big) = \boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Z}(u)$$

leads to the likelihood function

$$L(\boldsymbol{\theta}|\boldsymbol{x}) = \exp\left(-\int_W \exp\left(\boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Z}(u)\right) du\right) \prod_{i=1}^{n_W(\boldsymbol{x})} \exp\left(\boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Z}(x_i)\right).$$

Potential modeling with uncertain covariables

Likelihood of log-linear Poisson point process

Log-linear model of λ, i.e., log-linear random function,

$$\ln\big(\lambda_{\boldsymbol{\theta}}(u)\big) = \boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Z}(u)$$

leads to the log-likelihood function

$$\ln L(\boldsymbol{\theta}|\boldsymbol{x}) = -\int_W \exp\big(\boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Z}(u)\big)du \; + \; \sum_{i=1}^{n_W(\boldsymbol{x})} \boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Z}(x_i).$$

Potential modeling with uncertain covariables

Likelihood of log-linear Poisson point process

Log-linear model of λ, i.e., log-linear random function,

$$\ln\big(\lambda_{\boldsymbol{\theta}}(u)\big) = \boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Z}(u)$$

leads to the log-likelihood function

$$\ln L(\boldsymbol{\theta}|\boldsymbol{x}) = -\int_W \exp\left(\boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Z}(u)\right) du \; + \sum_{i=1}^{n_W(\boldsymbol{x})} \boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Z}(x_i).$$

Assuming that $\boldsymbol{Z}$ is a **Gaussian random function** with expectation $\boldsymbol{\mu}(u)$ and variogram Γ, taking the conditional expectation, i.e., kriging, leads to

$$\ln L(\boldsymbol{\theta}|\boldsymbol{x}) = -\int_W \exp\left(\boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Z}^*(u) + \frac{1}{2}\boldsymbol{\theta}^{\mathsf{T}} \Sigma_{\text{Krige}}\, \boldsymbol{\theta}\right) du + \sum_{i=1}^{n_W(\boldsymbol{x})} \mathbb{E}\big(\boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Z}(x_i)\big).$$

Potential modeling with uncertain covariables

Non-spatial analogue:
Poisson regression with error-prone covariables

For given $\beta_0, \beta_1, \beta_2$ and simulated normal distributed Z_1, Z_2

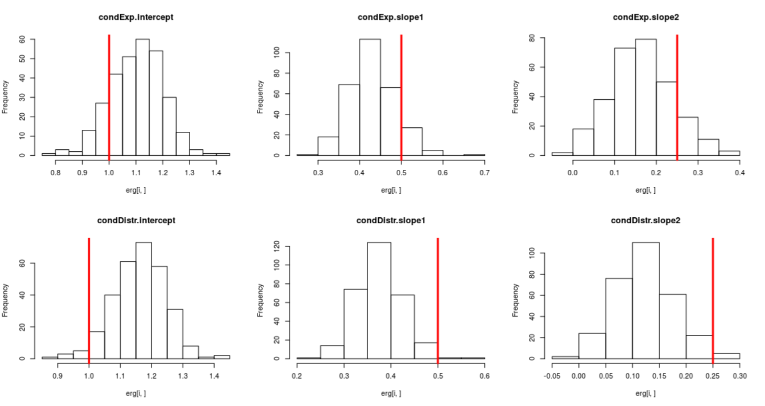
$$\ln \lambda = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2$$

a random variable $X \sim \text{Po}(\lambda)$ is simulated.
Adding some noise $\varepsilon \sim N(0, \sigma)$

1. apply ML to estimate $\beta_0, \beta_1, \beta_2$ using X, Z_1', Z_2' with $Z_1' = Z_1 + \varepsilon, Z_2' = Z_2 + \varepsilon$,
2. apply MC simulation of Z_1', Z_2' to estimate $\beta_0, \beta_1, \beta_2$, introducing additional error,

and compare.

Potential modeling with uncertain covariables



Estimation of regression parameters $\boldsymbol{\theta}$ of non-spatial Poisson regression model $\lambda_{\boldsymbol{\theta}} = \mathbb{E}(Y|\boldsymbol{Z}) = \exp(\theta_0 + \theta_1 Z_1 + \theta_2 Z_2)$ of Poisson distributed Y and Gaussian covariables Z_1, Z_2 using either "kriging" (top row) or "simulation" (bottom row).

Potential modeling with uncertain covariables

Conclusions

- Raster-based methods suffer from
  - requiring complete knowledge within training region,
  - neglecting spatially imposed dependence,
  - lacking consistency with respect to spatial resolution.
- Logistic regression with smaller resolution approximates Cox process, i.e.,
- Cox process is a canonical generalization of logistic regression.
- Including uncertainty in spatial predictors appears lethal.