Bayesian Structured Additive Regression for Identifying Tipping Points of Tropical Forests

Nikolaus Umlauf

http://eeecon.uibk.ac.at/~umlauf/

Overview

- Introduction
- Models with Structured Additive Predictor
- Software
- Software & Results

Introduction

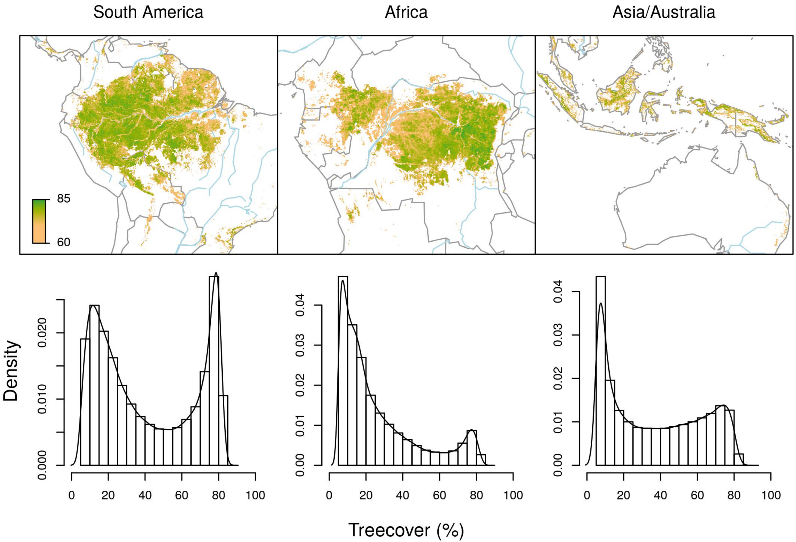
Project in collaboration with: Remote sensing, ecologist and environmental science group.

Jan Verbesselt, Milena Holmgren, Egbert H. Van Nes, Martin Herold, Marten Scheffer (University of Wageningen, Netherlands); Marina Hirota (University of Santa Catarina, Brazil); Achim Zeileis (Universität Innsbruck).

- Recent work suggests that in some regions tropical forest may be close to a tipping point.
- A relatively small perturbation might invoke a self-propagating collapse into a savanna state.
- Do tropical forests have a tipping point? How can we monitor?

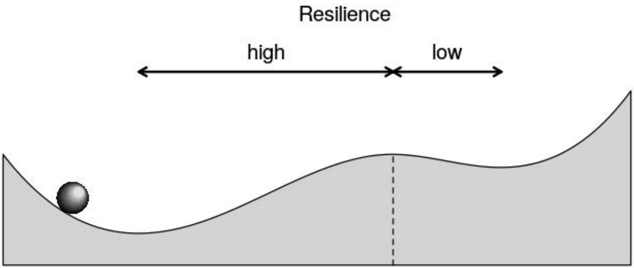
Introduction

Satellite data: 0.05° (5600 meters), $>$ 6 million observations.

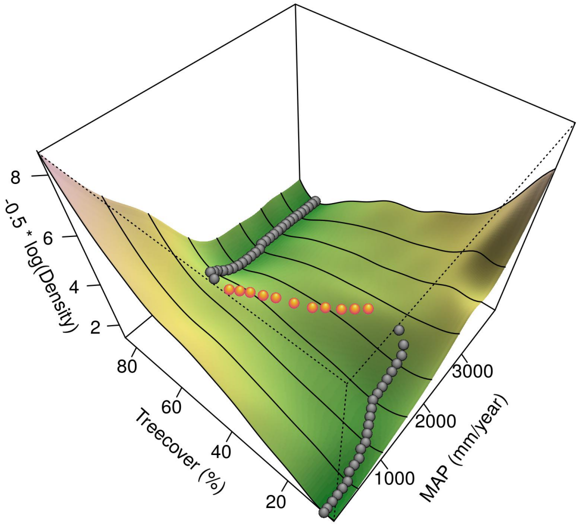


Introduction

- Loss of resilience in the vicinity of a tipping point should result in a phenomenon known as “critically slowing down”.
- Recovery rates from small perturbations decline, to eventually reach zero in a tipping point.



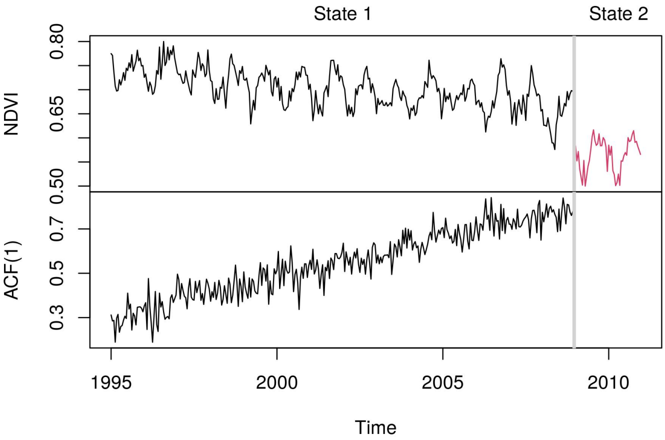
Introduction



Potential landscape derived from data.

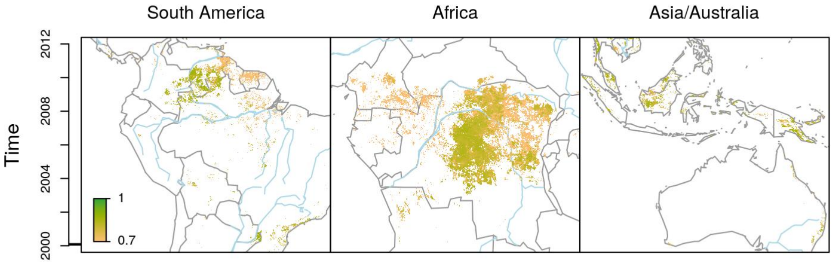
Introduction

- Detection by changes in the correlation structure of a time series.
- Critical slowing down causes an increase in the "short-term memory" of a system prior to a transition.



Introduction

- Satellite platforms: MODerate-resolution Imaging Spectroradiometer (MODIS, 2000–2011), Advanced Very High Resolution Radiometer (AVHRR, 1982–2011).
- Look at **undisturbed** forest pixels ($\geq$ 60% treecover).
- Remove time and seasonal trends of NDVI time series.
- Compute temporal correlation (`TC`): ACF(1), rolling window ACF(1), time-varying AR(1).



Introduction

We specified the following structured additive regression model (STAR)

$$\texttt{TC} = \eta + \varepsilon$$

with $\varepsilon \sim N(0, \sigma^2)$, and

$$\begin{aligned}\eta &= \beta_0 + f_1(\texttt{MAP}) + f_2(\texttt{Pre TC}) + f_3(\texttt{Temp}) + \\ &\quad f_4(\texttt{Soil}) + f_5(\texttt{\%NAs}) + f_6(\texttt{Long}, \texttt{Lat}).\end{aligned}$$

`Pre TC` is the temporal correlation of precipitation time series in the same period.

Functions $f_1, \ldots, f_6$ are (possibly) nonlinear estimated by semiparametric regression techniques.

Models where fitted for each satellite platform, continent, detrending and temporal correlation method, seperately.

STAR Models

Distributional and structural assumptions, given covariates and parameters, are based on generalized linear models with

$$E(y|\mathbf{x}, \mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\theta}) = h^{-1}(\eta)$$

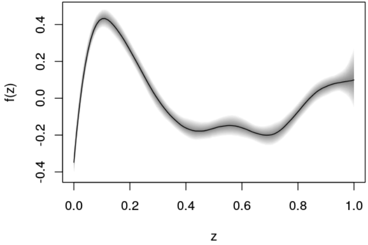
and structured additive predictor

$$\eta = f_1(\mathbf{z}) + \ldots + f_p(\mathbf{z}) + \mathbf{x}^\top \boldsymbol{\beta}$$

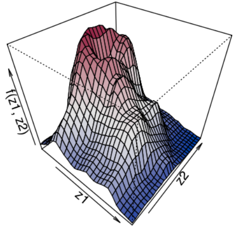
- $\mathbf{x}^\top \boldsymbol{\beta}$ parametric part of the predictor.
- $\mathbf{z}$ represents a generic vector of all nonlinear modeled covariates, e.g. may include continuous covariates, time scales, location or unit or cluster indexes.
- The vector $\boldsymbol{\theta}$ comprises all parameters of the functions $f_1, \ldots, f_p$.
- f_j one-/two-/higher-dimensional, not necessarily continuous functions.

STAR Models

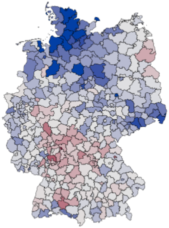
Nonlinear effects of continuous covariates



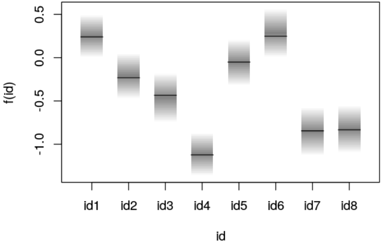
Two-dimensional surfaces



Spatially correlated effects f(z) = f(s)



Random intercepts

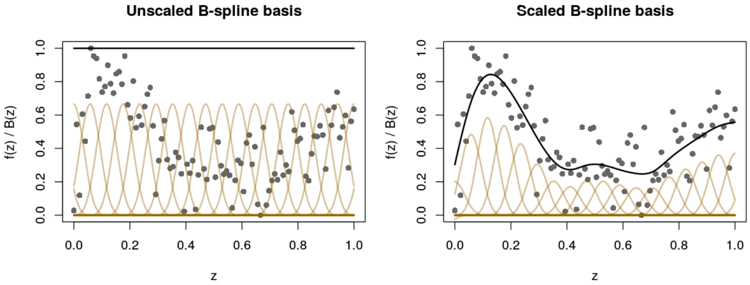


STAR Models

Within the basis function approach, the vector of function evaluations $\mathbf{f}_j = (f_j(\mathbf{z}_1), \ldots, f_j(\mathbf{z}_n))$ of the $i = 1, \ldots, n$ observations can be written in matrix notation

$$\mathbf{f}_j = \mathbf{Z}_j \gamma_j,$$

with $\mathbf{Z}_j$ as the design matrix, where γ_j are unknown regression coefficients. Form of $\mathbf{Z}_j$ only depends on the functional type chosen.



STAR Models

Penalized least squares:

$$\text{PLS}(\gamma, \boldsymbol{\lambda}) = ||\mathbf{y} - \boldsymbol{\eta}||^2 + \lambda_1 \gamma_1' \mathbf{K}_1 \gamma_1 + \ldots + \lambda_p \gamma_p' \mathbf{K}_p \gamma_p.$$

A general Prior for γ in the corresponding Bayesian approach

$$p(\boldsymbol{\gamma}_j | \tau_j^2) \propto \exp\left(-\frac{1}{2\tau_j^2} \gamma_j' \mathbf{K}_j \gamma_j\right),$$

τ_j^2 variance parameter, governs the smoothness of f_j.

Structure of $\mathbf{K}_j$ also depends on the type of covariates and on assumptions about smoothness of $\mathbf{f}_j$.

The variance parameter τ_j^2 is equivalent to the inverse smoothing parameter in a frequentist approach.

STAR Models

Distributional regression framework

$$\mathbf{y} \sim \mathcal{D}\left(h_1(\boldsymbol{\theta}_1) = \boldsymbol{\eta}_1,\ h_2(\boldsymbol{\theta}_2) = \boldsymbol{\eta}_2, \ldots,\ h_K(\boldsymbol{\theta}_K) = \boldsymbol{\eta}_K\right),$$

where $\mathcal{D}$ denotes any parametric distribution available for the response variable. Each parameter is linked to a structured additive predictor

$$h_k(\boldsymbol{\theta}_k) = \boldsymbol{\eta}_k = \mathbf{Z}_{1k}\boldsymbol{\gamma}_{1k} + \ldots + \mathbf{Z}_{pk}\boldsymbol{\gamma}_{pk} + \mathbf{X}_k\boldsymbol{\beta}_k,\ \ k = 1, \ldots, K,$$

where $h_k(\cdot)$ are known monotonic link functions.

The observations y_i are assumed to be independent and conditional on a pre-specified parametric density $f(y_i|\boldsymbol{\theta}_{i1}, \ldots, \boldsymbol{\theta}_{iK})$.

Estimation based on the idea of generalized additive models for location, scale and shape (GAMLSS).

Software

An implementation is provided in the open source package **BayesX**

`http://www.BayesX.org/`

An R implementation is provided in the **BayesR** project

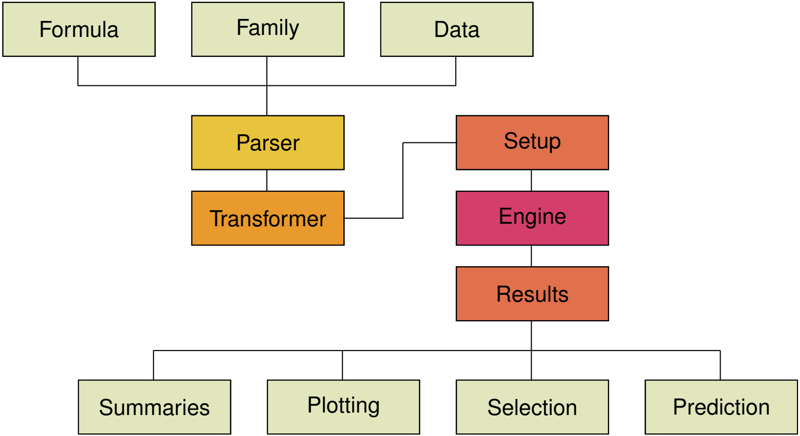
`https://R-Forge.R-project.org/projects/bayesr/`

which includes a full interactive interface to **BayesX** (package **BayesR** and **R2BayesX**).

- Complex multi-parameter models with (hierarchical) STAR predictor.
- Standard extractor and plotting functions: `summary()`, `plot()`, `fitted()`, `residuals()`, `predict()`, `coef()`, `DIC()`, `samples()`, ...
- Enhanced visualization possibilities.

Software

General architecture:



In principle, the setup does not restrict to any specific type of engine (Bayesian or frequentist).

Software & Results

Linking NDVI temporal correlation to environmental conditions:

$$\texttt{tc.vi} \sim N(\boldsymbol{\mu} = \boldsymbol{\eta}_{\mu}, log(\boldsymbol{\sigma}^2) = \boldsymbol{\eta}_{\sigma^2}).$$

```
R> load("data/ews.rda")

R> f <- tc.vi ~ s(map) + s(tc.pre) + s(mt) +
+    s(soil) + s(pNAs) + s(long, lat, k = 200)

R> f <- list("mu" = f, "sigma2" = f)

R> b <- bayesr(f, family = gaussian2, data = ews,
+    subset = continent == "South America",
+    method = c("backfitting", "MCMC"))

R> summary(b)
```

Software & Results

```
Call:
bayesr(f, family = gaussian2, data = ews,
  subset = continent == "South America",
  method = c("backfitting", "MCMC"))

Family: gaussian2
Link function: mu = identity, sigma2 = log
---

Results for mu:
---
Formula:
tc.vi ~ s(map) + s(tc.pre) + s(mt) + s(soil) + s(pNAs) + s(long,
    lat, k = 200)

Parametric coefficients:
                 Mean        Sd      2.5%       50%     97.5% alpha
(Intercept) 0.3778700 0.0003405 0.3772275 0.3778603 0.3785254     1

Smooth effects variances:
                Mean       Sd     2.5%      50%    97.5% alpha
s(map)       775.337  349.468  327.017  704.895 1707.003     1
s(tc.pre)      6.616    4.049    1.768    5.884   16.293     1
s(mt)         39.687   22.134   11.733   34.890   95.163     1
s(soil)       37.464   19.426   15.025   32.630   88.196     1
s(pNAs)      448.281  179.583  213.219  407.350  900.103     1
s(long,lat)  136.295   10.362  117.552  135.663  158.301     1
```

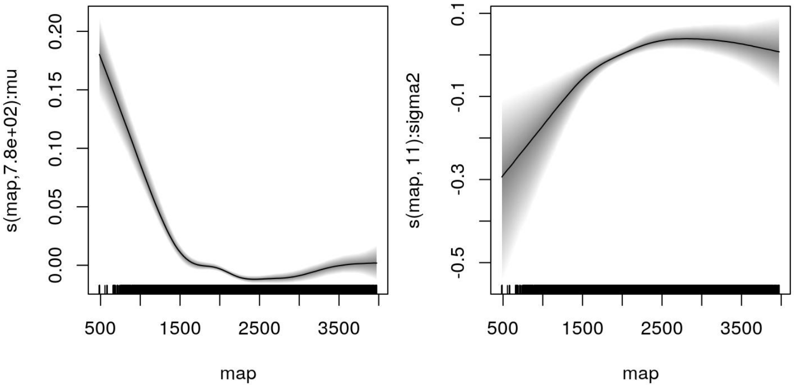
...continued on next slide...

Software & Results

```
Results for sigma2:
---
Formula:
tc.vi ~ s(map) + s(tc.pre) + s(mt) + s(soil) + s(pNAs) + s(long,
    lat, k = 200)

Parametric coefficients:
                 Mean        Sd      2.5%       50%     97.5% alpha
(Intercept) -5.171684  0.006642 -5.184547 -5.171879 -5.158544 0.998

Smooth effects variances:
               Mean      Sd    2.5%     50%   97.5% alpha
s(map)      11.4202  8.2235  2.2955  9.6416 32.2137 0.981
s(tc.pre)    1.7462  1.1993  0.3883  1.4447  4.8741 0.992
s(mt)       14.3092  7.8473  4.0933 12.7130 34.5610 0.968
s(soil)      1.7812  1.1413  0.3994  1.4663  4.8618 0.985
s(pNAs)      2.2825  1.5782  0.5035  1.9082  6.0365 0.994
s(long,lat) 12.2267  1.2149  9.9473 12.2507 14.8331 0.818

---
DIC = -1.132e+05 pd = 576.7 N = 48524
```

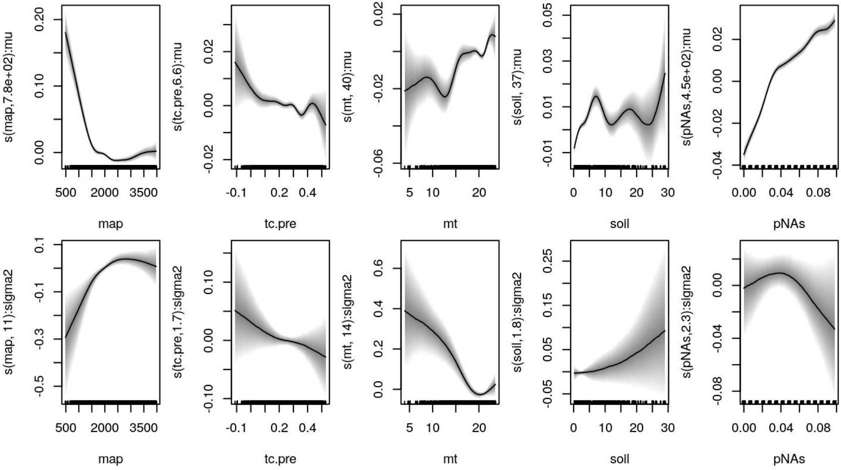
Software & Results

```
R> plot(b, term = "s(map)")
```



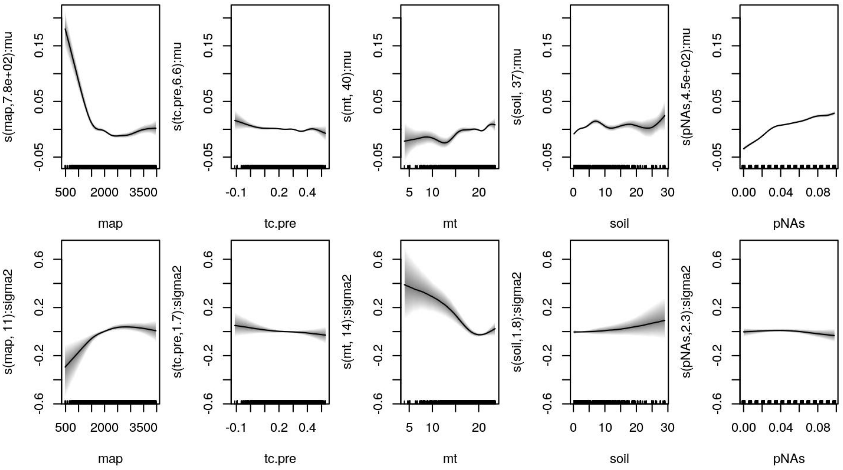
Software & Results

```
R> plot(b, scale = 0)
```



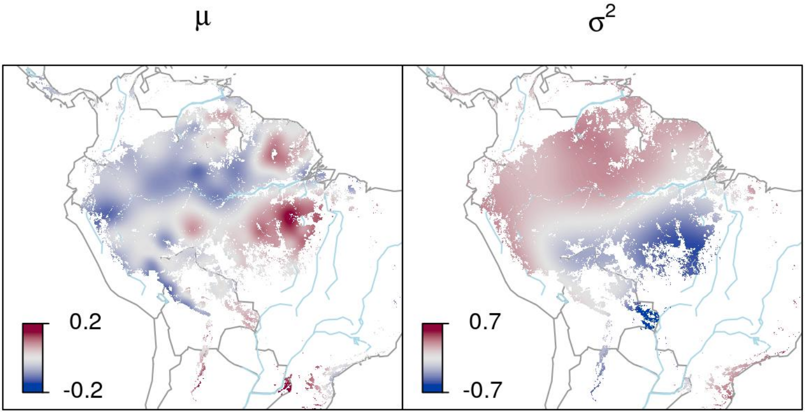
Software & Results

```
R> plot(b, scale = 1)
```

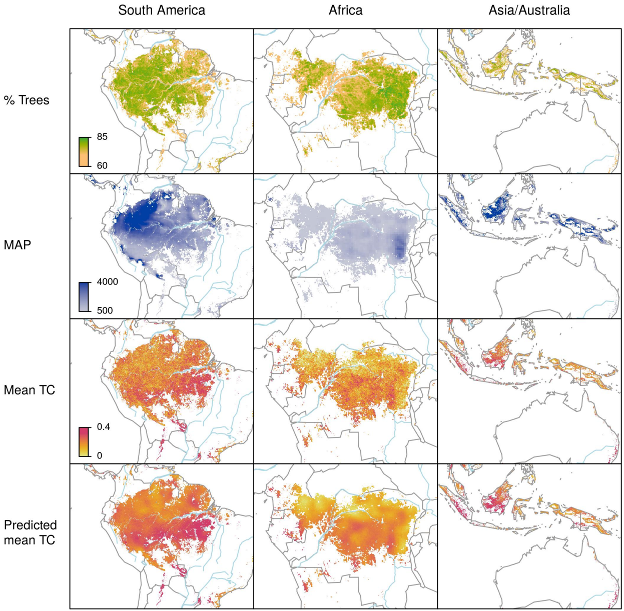


Software & Results

```
R> plot(b, term = "s(long,lat)")
```



Software & Results



Thank you!

References

Klein, N., Kneib, T., Lang, S. (2013): Bayesian structured additive distributional regression. *Working Papers in Economics and Statistics*, Research Platform Empirical and Experimental Economics, Universität Innsbruck, Oktober 2013.

Lang, S., Umlauf, N., Wechselberger, P., Harttgen, K., Kneib, T. (2013): Multilevel structured additive regression. *Statistics and Computing*, 24(2), 223–238.

Hirota, M., Holmgren, M., Van Nes, E., Scheffer, M. (2011): Global Resilience of Tropical Forest and Savanna to Critical Transitions. *Science*, 334(6053), 232–235.

Umlauf, N., Adler, D., Kneib, T., Lang, S., Zeileis, A. (2014). Structured additive regression models: An R interface to **BayesX**. *Journal of Statistical Software*. (to appear)

de Jong, R., Verbesselt, J., Zeileis, A., and Schaepman, M. E. (2013). Shifts in global vegetation-activity trends. *Remote Sensing*, 5(3):1117–1133.

Zeileis, A., Shah, A., and Patnaik, I. (2010). Testing, monitoring, and dating structural changes in exchange rate regimes. *Computational Statistics & Data Analysis*, 54(6):1696–1706.