Space-time modelling for trend estimation of natural resources

Nicole Augustin, University of Bath, Bath, UK

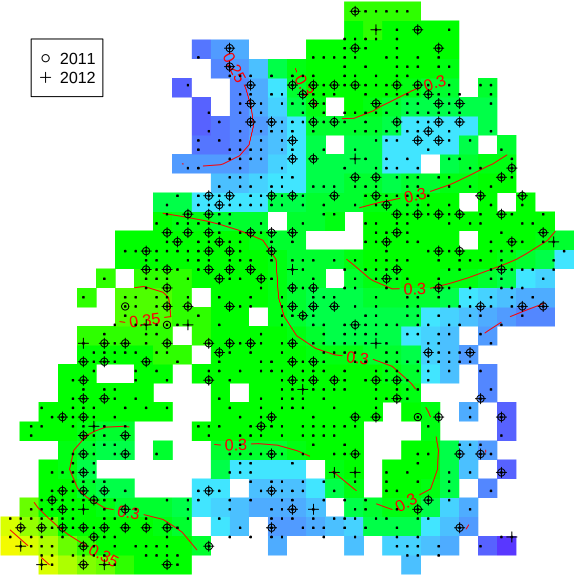
1. Forest health monitoring: estimation spatio-temporal trends of beech defoliation, with K. von Wilpert (FVA Freiburg), A. Griffiths (Bath) and S. Meining (Freiburg)

2. Estimating relative abundance trends of blue ling, with V. Trenkel, P. Lorance (IFREMER, Nantes) and S. Wood (Bath)

1. Forest health monitoring - Background

- Tree damage in Europe first noticed in 1970s
- Environmental Monitoring set up in 1983
- main influencing factors:
  - the deposition of pollutants through the air and rainfall;
  - weather (extreme heat and droughts - climate change);
  - biotic influences (e.g. biological pests).

Sampling locations - beech



Monitoring data

- response: defoliation in crown (%);
- covariates: tree (age, species) and site specific (geology, soil type, ...)
- yearly approx. 24 trees at each sampling grid point are sampled
- 1983 - 2012 with different resolutions: 4x4km, 8x8km and 16x16km.
- main species: spruce, beech, fir, pine, oak
- survey is in alignment with European level I monitoring programme of the International Cooperative Programme on Forests and Integrated Monitoring of Ecosystems

Objectives: provide a statistical tool for

- detecting trends in defoliation/tree health
- identifying possible causes
- optimising sampling scheme

Statistical tool: A spatio-temporal model for forest health

requirements

- space-time trend and interaction;
- temporal correlation;
- spatial correlation;
- account for covariates with possible non-linear effects, such as age;
- functional covariates, e.g. function of weather history at specific location.

A spatio-temporal model for forest health - a generalized additive mixed model (GAMM)

$$logitE(y_{it}) = f_1(northing_i, easting_i, year_t) + f_2(age_{it}, fruct_{it})$$

- mean defoliation $y_{it} = E(y_{it}) + \epsilon_{it}$ at location $i = 1, ..., 1500$ and year $t = 1985, ..., 2012$ ($n = 7023$).
- $\boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \boldsymbol{\Lambda}_i)$ where $\boldsymbol{\Lambda}_i$ within location i autoregressive (AR3) process;
- $logitE(y_{it})$ - for bounding fitted values in (0,1);
- $f_1(.)$ - 3-d tensor product smooth with different penalties for space and time (separate smoothing parameters, scale invariant);
- $f_2(.)$ - two dimensional smooth of mean age and mean fructification.

3-d tensor product smoother for space and time

$$f_1(northing_i, easting_i, year_t)$$

- mixing of bases: cubic regression spline basis for time and thin plate regression spline basis for space;
- obtain separate smoothing parameter for space and time;
- scale invariant due to different penalty matrices for each dimension.

Set up via marginal bases for time and space.
Allow the parameters of the temporal smooth to vary in space.

3-d tensor product smoother for space and time ...

Start from marginal smooths for time f_{year} and space $f_{n,e}$ with associated quadratic roughness penalties.

Let $a_i(year)$ and $b_l(n,e)$ be two low rank bases for f_{year} and $f_{n,e}$, then

$$f_{year}(year) = \sum_{i=1}^{P} \alpha_i a_i(year) \;\text{ and }\; f_{n,e}(n,e) = \sum_{l=1}^{L} \beta_l b_l(n,e),$$

where α_i and β_l are parameters.

To convert f_{year} into a smooth function of n and e:

$$\alpha_i(n,e) = \sum_{l=1}^{L} \beta_{il} b_l(n,e)$$

and this gives

$$f_{year,n,e}(year,n,e) = \sum_{i=1}^{P} \sum_{l=1}^{L} \beta_{il} b_l(n,e) a_i(year).$$

Parameter estimation via penalized quasi likelihood (Breslow & Clayton, 1993) using mixed model technology (Ruppert et al, 2003):

$$\begin{aligned} logitE(y_{it}) &= f_1(northing_i, easting_i, year_t) + f_2(age_{it}, fruct_{it}) \\ &= \mathbf{X}_{it}^f \theta^f + \boldsymbol{Z}_{it}^f \boldsymbol{b}^f \end{aligned}$$

reparameterise components of f relating to

smooth parts: $\boldsymbol{X}_{it}^f \theta^f$ - fixed effects, unpenalised coefficients;

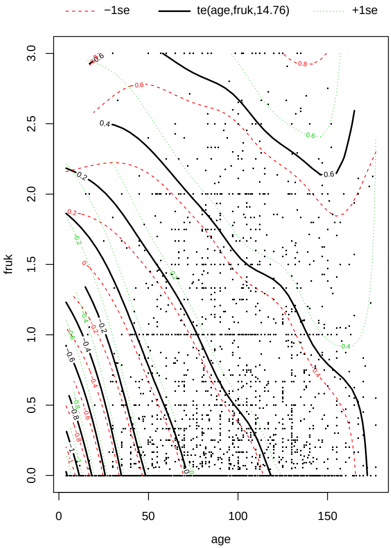
rough parts: $\boldsymbol{Z}_{it}^f \boldsymbol{b}^f$ - random effects, penalised coefficients.

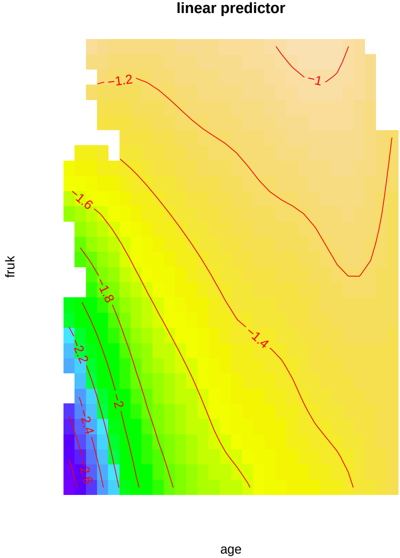
Then $\boldsymbol{b}^f \sim N(0, \psi)$ where ψ contains smoothing parameters.

Trend and variance estimation: Represent mixed model as a Bayesian model with priors for coefficients

- Use the posterior distribution of coefficients, θ^f and $\boldsymbol{b}^f$, to obtain predictive distribution of $E(y_{it}) = logit^{-1}\left(\mathbf{X}_{it}^f\theta^f + \boldsymbol{Z}_{it}^f\boldsymbol{b}^f\right)$;
- estimate spatial and temporal trends with credible intervals by appropriate averaging;
- **Prediction**: evaluate $\boldsymbol{X}_{it}^f$ and $\boldsymbol{Z}_{it}^f$ for prediction data.

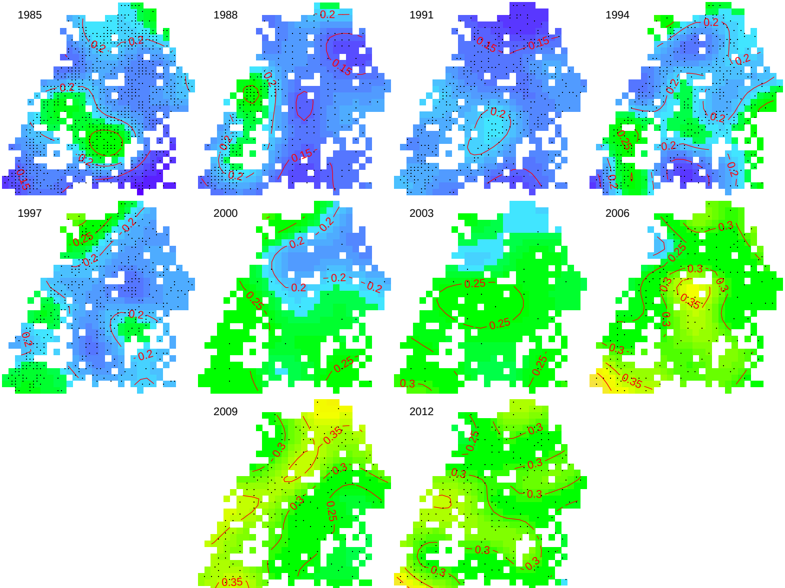
Results: age and fructification effect



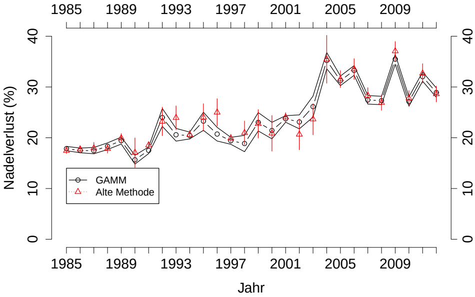


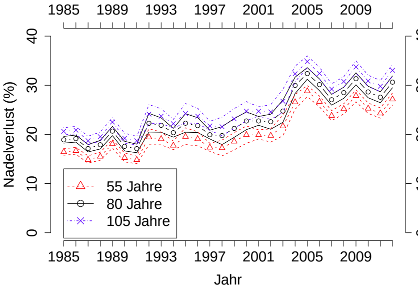
age effect is weaker with fructification

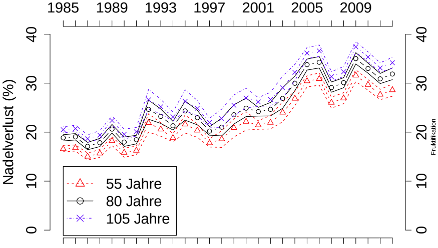
Average defoliation predictions of beech at mean age and mean fructification

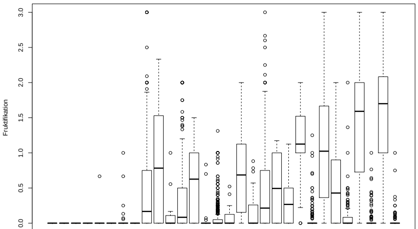


Results: Beech timetrend - adjusted and unadjusted for fructification





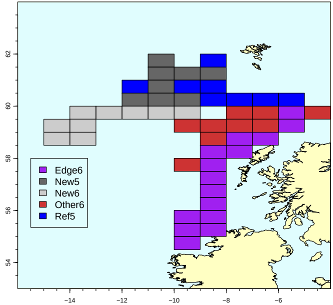




Summary on beech modelling

- Forest eco-system is damaged; since about 2003 there is evidence for an increased trend in defoliation;
  Possible causes: pollution and climate change.
- adjusting for fructification? fructification indicator for tree health?

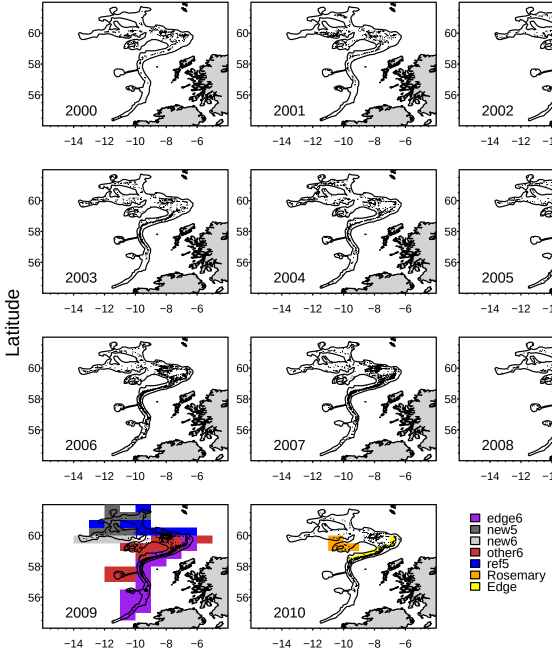
2. Spatio-temporal model for stock management of blue ling



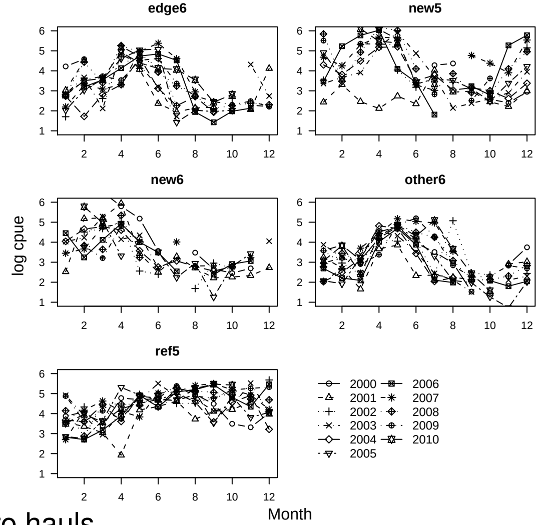
The Data

- 19 French deep-sea trawlers operating in the Northeast Atlantic during the period 2000-2010.
- Variables: landings (biomass in kg) by species, latitude and longitude, mean fishing depth, haul duration a proxy for effort .
- Use subset: haul duration between 30 and 600 mins, haul depth between 200 and 1100m
- Zero landings indicate no abundance or very low abundance of blue ling in the specific area and time.
- By-catch of blue ling is always possible (not affected by differences in fishing techniques due to targetting).

Positions of hauls per year



Observed median monthly catch (log kg) per hour by year and fishing areas.



19% of zero hauls.

The questions

- What is the relative overall trend of blue ling abundance?
- Is there any evidence for a space-time interaction, supporting the hypotheses of:
  - local depletion in areas with longer exploitation histories?

Approach

- Use a generalised additive mixed model (GAMM) incorporating a smooth function of space and time.
- The spatial boundary of blue ling is complicated, depending on the topography of the sea bed.
- Will need to test whether space-time interactions are present.
- Inappropriately imposing smoothness across boundary features might induce model mis-specification.

Blue Ling model

- Reponse 'kg blue Ling in haul *i*', y_i, n = 9031 observations
- Full model is

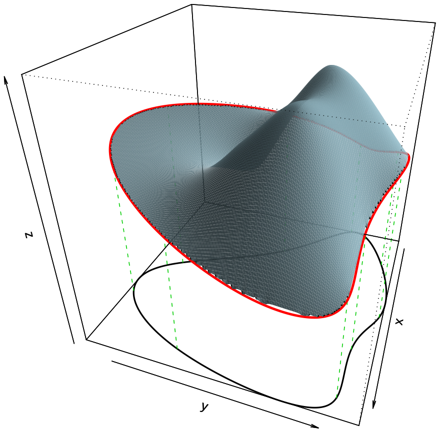
$$\begin{aligned}\log(\mu_i) = f_1(\texttt{duration}_i) + f_2(\texttt{depth}_i, \texttt{year}_i) + v_{k(i)} \\ f_3(\texttt{depth}_i) + f_4(\texttt{month}_i) + f_5(\texttt{depth}_i, \texttt{month}_i) \\ + f_6(\texttt{north}_i, \texttt{east}_i, \texttt{year}_i),\end{aligned} \tag{1}$$

- $\mu_i = E(y_i)$ and $y_i \sim \text{Tweedie}(\mu_i, \phi\mu_i^p)$, $p = 1.5$;
- $v_{k(i)}$ is a random vessel effect, assumed i.i.d. $N(0, \sigma_v^2)$;
- f_{1-6} are smooth functions of the covariates available with each haul.
- RED: effects of fisheries management and targetting;
- BLACK: biological effects.

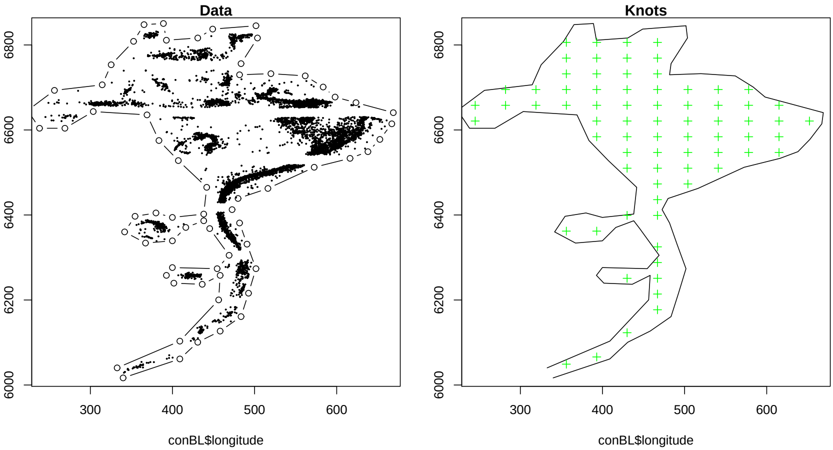
$f_6(\texttt{north}_i, \texttt{east}_i, \texttt{year}_i)$

- 3D tensor product of two *marginal* bases and penalties for time and space:
- $f_{n,e}(\texttt{north}, \texttt{east})$ and $f_y(\texttt{year})$
  2-d isotropic smoother for space and a 1-d CRS for year
- allows spatial smooth to be isotropic while being invariant to relative scaling of space and time.
- $f_{n,e}(\texttt{north}, \texttt{east})$ **soap film smooth**
  respects the biological boundary,
  but requires manual knot selection.

Soap film smoother (Wood et al, 2008)

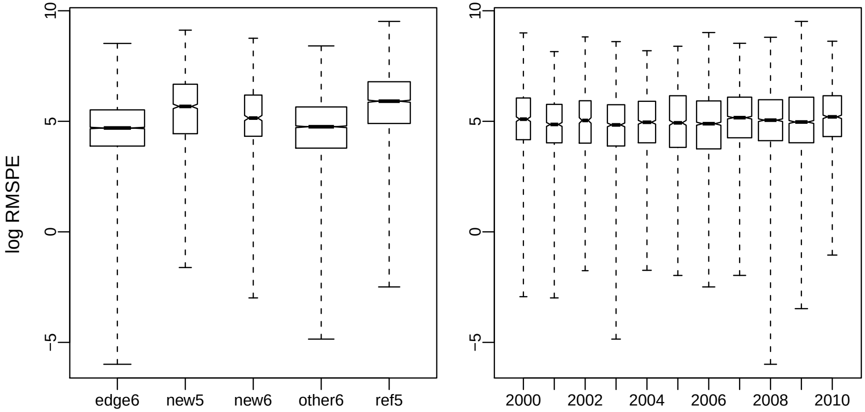


Boundary and knots

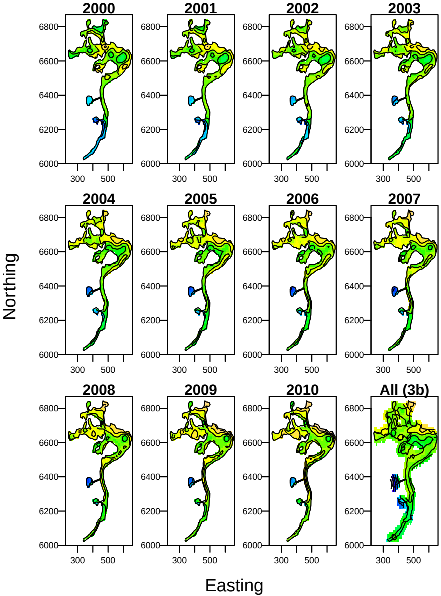


Selected model

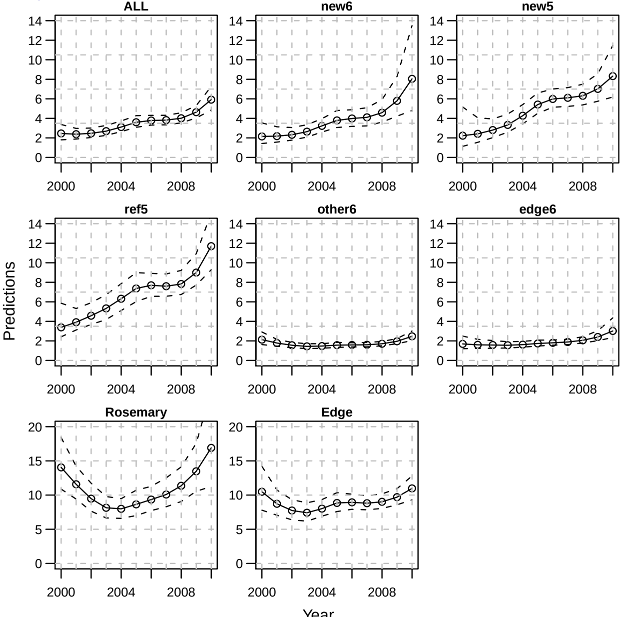
$$\log(\mu_i) = f_1(\texttt{duration}_i) + f_2(\texttt{depth}_i, \texttt{year}_i) + \beta power_k$$
$$f_3(\texttt{depth}_i) + f_4(\texttt{month}_i) + f_5(\texttt{depth}_i, \texttt{month}_i)$$
$$+ f_6(\texttt{north}_i, \texttt{east}_i, \texttt{year}_i), \qquad (2)$$



Fitted spatial model smooths by year for model 2a (soap smoother) and model 3b (TPRS smoother)



Time trends by area (January) and for two spawning areas (April)



Some conclusions about blue ling

- 3d tensor product of soap film smooth for space and CRS smooth for year allowed us to test for the presence of space-time interaction.
- Soap film smoother ensures that we are not smoothing accross the complicated boundary.
- What is the time trend of blue ling abundance? It appears to be constant/increasing.
- Is there any evidence for localised depletions in areas with a longer exploitation history? yes - maybe. Space-time interaction term is required.
- Cannot assess how biased results are since data are based on preferential sampling.
  But model allows to control for effects of fisheries management and targetting.

References

Augustin, N.H., Trenkel, V., Wood, S.N. and Lorance, P. Space-time modelling of blue ling for fisheries stock management. *Environmetrics* (accepted).

Augustin, N.H. and Musio, M. and von Wilpert, K. and Kublin, E. and Wood, S.N. and Schumacher, M. 2009. Modelling spatio-temporal forest health monitoring data. *Journal of the American Statistical Association*, 104 (487), 899-911.

Tweedie, M. C. K., 1984. An index which distinguishes between some important exponential families. Statistics: Applications and New Directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference, 579-604,editors J. K. Ghosh and J. Roy, Calcutta: Indian Statistical Institute.

Wood, S.N. and Bravington, M.V. and Hedley, S.L., 2008. Soap film smoothing. *Journal of the Royal Statistical Society*,70 (5), 931-955.