BayesX: Analysing Bayesian structured additive regression models

Andreas Brezger¹, Thomas Kneib¹ and Stefan Lang²

¹ Department of Statistics, University of Munich
² Institute of Empirical Economic Research, University of Munich

What is BayesX?

- The software BayesX provides powerful regression tools for analysing structured additive regression (STAR) and survival models. STAR models cover a number of well known model classes as special cases, e.g. generalised additive models, generalised additive mixed models, geospatial models, dynamic models, varying coefficient models, and geographically weighted regression.
- BayesX supports both full Bayesian inference based on Markov chain Monte Carlo simulation techniques and empirical Bayes inference based on a mixed model representation of STAR models.
- In its current form, BayesX runs only under the various versions of the Windows operating system. A Linux version and an interface to R are work in progress.
- Contributions by Christiane Belitz, Eva-Maria Fronk, Andrea Hennerfeind, Manuela Hummel, Alexander Jerak, Petra Kragler and Leyne Oxna Echausin.

Bayesian structured additive regression

- Structured additive regression extends and unifies several additive and geospatial regression approaches.
- For exponential family models, a structured additive predictor is of the form
  \[ y = f_1(x_1) + \ldots + f_p(x_p) + \epsilon, \]
  where the \( f_j \) are nonparametric functions of \( x_j \).
- Supported model terms include:
  - Penalised splines and random walk priors for nonparametric effects \( f(x) \) of continuous covariates \( x \).
  - Random intercepts and random slopes.
  - Bayesian tensor product penalised splines for interaction surfaces \( f(x, z) \).
  - Varying coefficient terms with continuous and spatial effect modifiers.
  - State space models for time-varying seasonal patterns \( f(t) \).
  - (Intrinsic) Markov random field priors for spatial effects \( f_{\text{spatial}}(s) \) based on regional data \( s = 1, \ldots, S \).
- Stationary Gaussian random field priors for spatial effects \( f_{\text{spatial}}(s) \) based on point-referenced data \( s = (x_s, y_s) \).
- Supported univariate response distributions:
  - Gaussian responses with identity link.
  - Binary responses with logit, probit and complementary log-log link.
  - Poisson responses with log-link.
  - Gamma responses with log-link.
  - Negative Binomial responses with log-link.
- Extensions for categorical regresion models:
  - Multinomial logit (and probit) models for unordered responses.

Example: Leukemia survival data

- Survival time of adults after diagnosis of acute myeloid leukemia.
- 1,043 cases diagnosed between 1982 and 1998 in North West England.
- 16 % (right) censored.
- Continuous and categorical covariates:
  - age at diagnosis,
  - white blood cell count at diagnosis,
  - sex of the patient,
  - Townsend deprivation index.
- Spatial information in different resolution.

### Further information

- Download and further information:
  [http://www.stat.uni-muenchen.de/bayesx](http://www.stat.uni-muenchen.de/bayesx)
- Contact:
  bayesx@stat.uni-muenchen.de
- Selected references: