Understanding the Stochastic Partial Differential Equation Approach to Smoothing

David L. Miller Richard Glennie Andrew Seaton

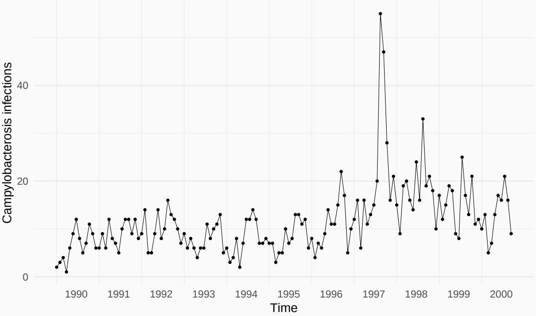
17 June 2021

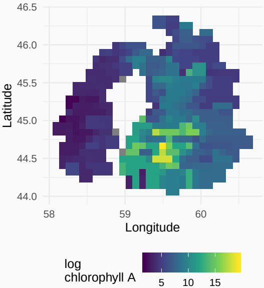
International Biometric Society Journal Club

Andy and Richard

Introduction

The kinds of data we’ll talk about

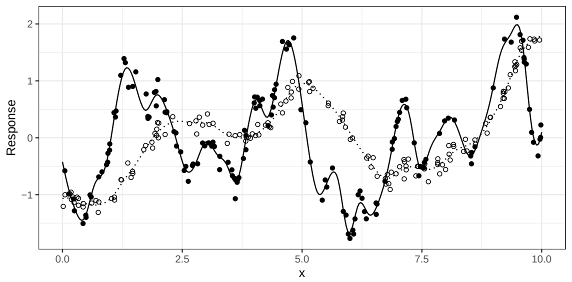




How to model dependence?

Tobler's law: "everything is related to everything else, but near things are more related than distant things."

Two ways to think about this: *Smoothness* and *Correlation*.



How to model dependence?

Two ways to think about this: Smoothness and Correlation.

Two popular approaches to modelling it: GAMs and SPDEs.

- GAMs ≈ "smoothing using spline-type things"
- SPDEs ≈ "Matérn covariance approximation using SPDEs/MRFs"

Jenny Bryan
@JennyBryan

All models are wrong, so why not start with one you actually understand?

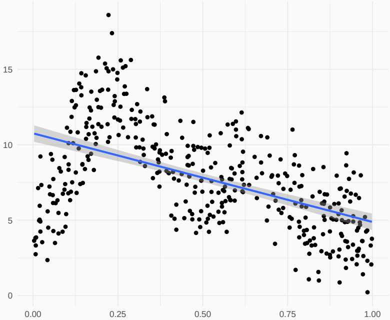
Refresh: what are GAMs made of?

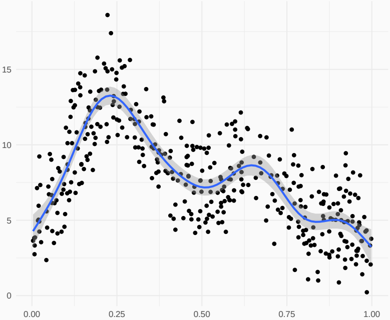
Explain: how does the SPDE method work?
Apply: how can we use these results in practice?

The GAM approach

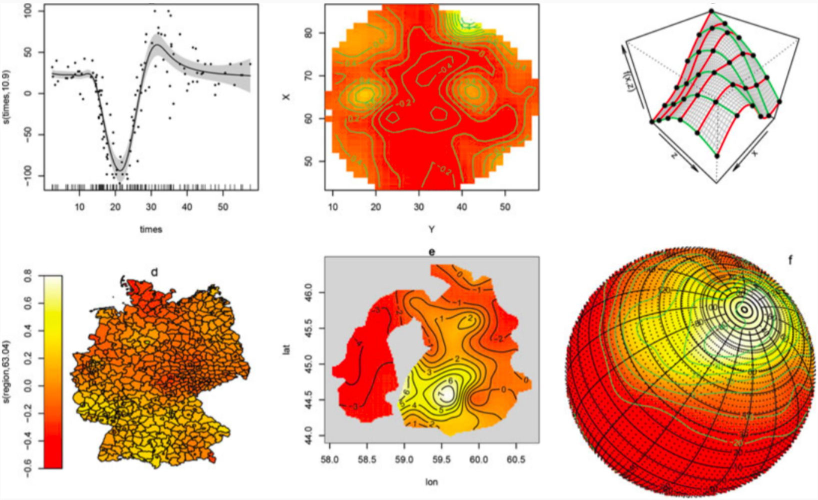
From GLMs to GAMs

- Want flexible, structured models
- GAM provides this via "smooths"
- $y \sim s(x)$, where s is a smooth function of x





Examples of smooths



Smooths come in 2 bits...

basis functions

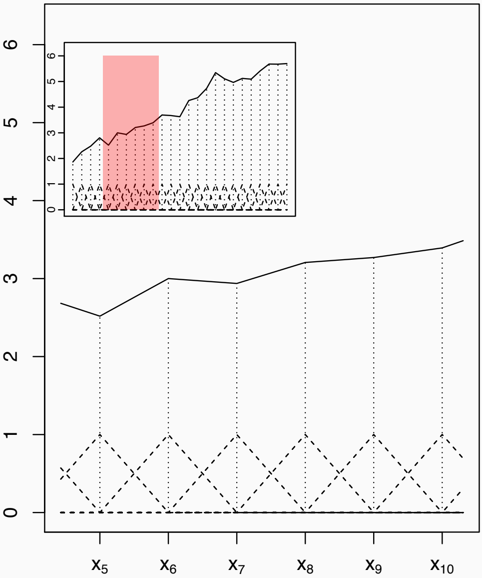
- join together to make up the overall curve, $s(x)$
- go in the design matrix, $\mathbf{X}$

penalties

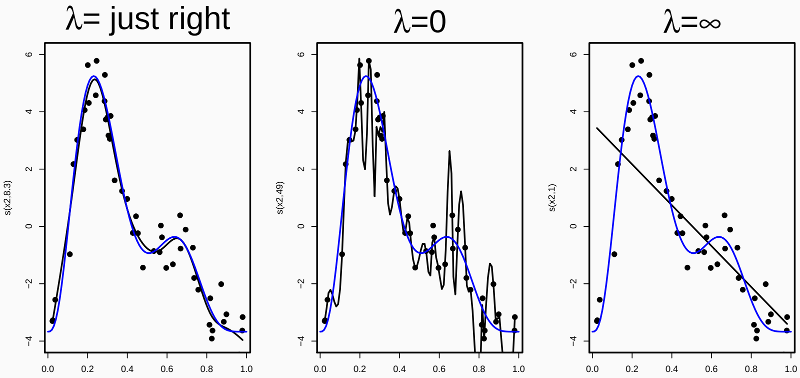
- controls how the line behaves
- penalty matrix $\mathbf{S}$ (structure)
- smoothing parameter λ (how wiggly)

"basis-penalty smoothers"

Basis functions



Penalty and smoothing parameter



SPDE approach

Lindgren et al, 2011

Journal of the
Royal Statistical Society

SERIES B
Statistical
Methodology

J. R. Statist. Soc. B (2011)
73, *Part* 4, *pp.* 423–498

An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach

Finn Lindgren and Håvard Rue

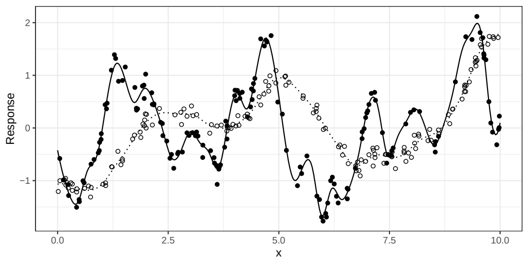
Norwegian University of Science and Technology, Trondheim, Norway

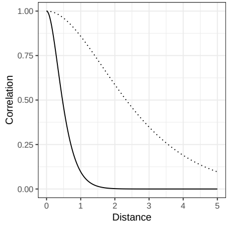
and Johan Lindström

Lund University, Sweden

Two Selling points of SPDE approach

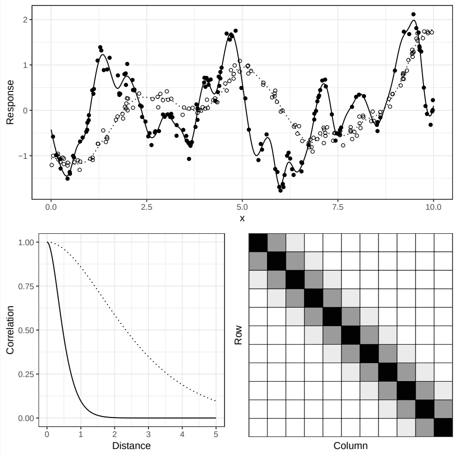
1. A certain kind of SPDE produces curves that have Matérn correlation functions, a popular choice of correlation function.





Two Selling points of SPDE approach

2. The precision matrix Q can be computed very efficiently, a bottleneck in many other correlation approaches.



SPDE method comes in 2 bits...

finite elements

- join together to make up the overall curve
- go in the predictor matrix, $\mathbf{A}$

precision

- controls how the curve behaves
- precision matrix $\mathbf{Q}$ (structure)
- correlation parameter κ (how wiggly)
- variance parameter τ

Wait a minute...

SPDE method comes in 2 bits...

~~finite elements~~

basis functions

- join together to make up the overall curve
- go in the predictor design matrix, $\mathbf{A} = \mathbf{X}$

~~precision~~

penalty

- controls how the curve behaves
- precision matrix $\mathbf{Q} = \mathbf{S}$ (structure)
- correlation parameter κ (how wiggly)
- variance parameter τ
- $\kappa, \tau \rightsquigarrow \lambda$

SPDE is a basis-penalty smoother

The SPDE has a corresponding smoothness penalty.

Pick same basis, then

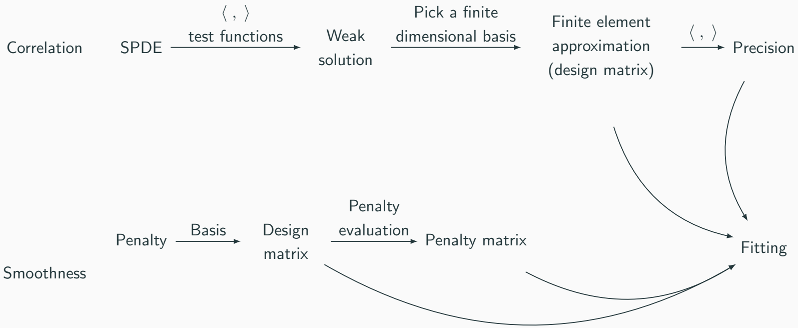
$$\mathbf{A} = \mathbf{X}$$

Use SPDE and corresponding penalty, then

$$\mathbf{Q} = \mathbf{S}$$

See our paper for more mathematical details.

The SPDE method is a basis-penalty smoother



Practical implications

What about R-INLA and mgcv?

- Both are ways to fit this same model.
- Both `R-INLA` and `mgcv` do approximations:
  - `R-INLA` does a Laplace approximation to get around MCMC.
  - `mgcv` does REML using Laplace approximation too.
- `mgcv` does empirical Bayes, so uses point estimates for hyperparameters (smoothing parameters, variance components).
- (`R-INLA` can do this too!)

What can we do with this?

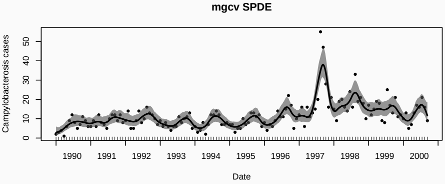
- Writing SPDE models as basis-penalty smoothers means we can use them elsewhere
- Anywhere structured random effects can be used!
- `mgcv`, `TMB`, Stan, etc
- We can combine them with other approaches!

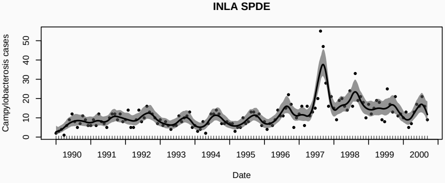
mgcv code

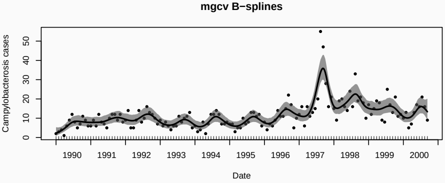
Setup using R-INLA, fit with mgcv

```r
## Simplified code-snippets
## Full code in supplementary materials

# get finite element matrices
inlamats <- inla.mesh.fem(mesh)
c1 <- as.matrix(inlamats$c1)
g1 <- as.matrix(inlamats$g1)
g2 <- as.matrix(inlamats$g2)

# construct the design matrix
X <- as.matrix(inla.spde.make.A(mesh, observed_locs))

# build the penalty / precision matrix
S <- kappa^4 * c1 + 2*kappa^2 * g1 + g2

# fit
b <- gam(y ~ X - 1,
         paraPen = list(X = list(S)),
         method = "REML")
```

1D comparison







Spatial smoothing comparison

Take-home

A modelling framework emerges

Use basis-penalty or basis-precision as a common currency.

1. **Choose a covariance model**: explicitly (as in kriging, etc), via smoothness penalty (e.g., splines), or with an SPDE;
2. **Approximate the precision matrix** Q: reduce dimension or induce sparsity in Q;
3. **Draw approximate inference using a software implementation**: e.g., with `mgcv`, MCMC (e.g., Stan; `JAGS`), `R-INLA`, `lme4` or `TMB`.

More info?

DOI 10.1007/s13253-019-00377-z

Understanding the Stochastic Partial Differential Equation Approach to Smoothing

David L. Miller, Richard Glennie, and Andrew E. Seaton

Correlation and smoothness are terms used to describe a wide variety of random quantities. In time, space, and many other domains, they both imply the same idea: quantities that occur closer together are more similar than those further apart. Two popular statistical models that represent this idea are basis-penalty smoothers (Wood in Texts in statistical science, CRC Press, Boca Raton, 2017) and stochastic partial differential equations (SPDEs) (Lindgren et al. in J R Stat Soc Series B (Stat Methodol) 73(4):423–498, 2011). In this paper, we discuss how the SPDE can be interpreted as a smoothing penalty and can be fitted using the `R` package `mgcv`, allowing practitioners with existing knowledge of smoothing penalties to better understand the implementation and theory behind the SPDE approach.

Supplementary materials accompanying this paper appear online.

Key Words: Smoothing; Stochastic partial differential equations; Generalized additive model; Spatial modelling; Basis-penalty smoothing.

Slides: `https://github.com/dill/SPDE-talk-IBS`