



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

Formulating longitudinal regression models in R

Dethlefsen, Claus; Lundbye-Christensen, Søren

Publication date:
2006

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Dethlefsen, C., & Lundbye-Christensen, S. (2006). Formulating longitudinal regression models in R. Poster session presented at Valencia/ISBA Eighth World Meeting on Bayesian Statistics, Benidorm, Alicante, Spain.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- ? Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- ? You may not further distribute the material or use it for any profit-making activity or commercial gain
- ? You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Formulating Longitudinal Regression Models in R

Søren Lundbye-Christensen*, Claus Dethlefsen**

*Department of Mathematical Sciences, Aalborg University, Denmark

**Center for Cardiovascular Research, Aalborg Hospital, Aarhus University Hospital, Denmark



Introduction

Aim: To create a free software package for easy specification of generalized dynamic linear models

Alternatives: Ox + SsfPack, StructTS. But they are not easy to use, or do not support complex models

Idea: Use the glm-call in R as a template and use iterated extended Kalman smoothing

The package sspir is available from CRAN cran.r-project.org

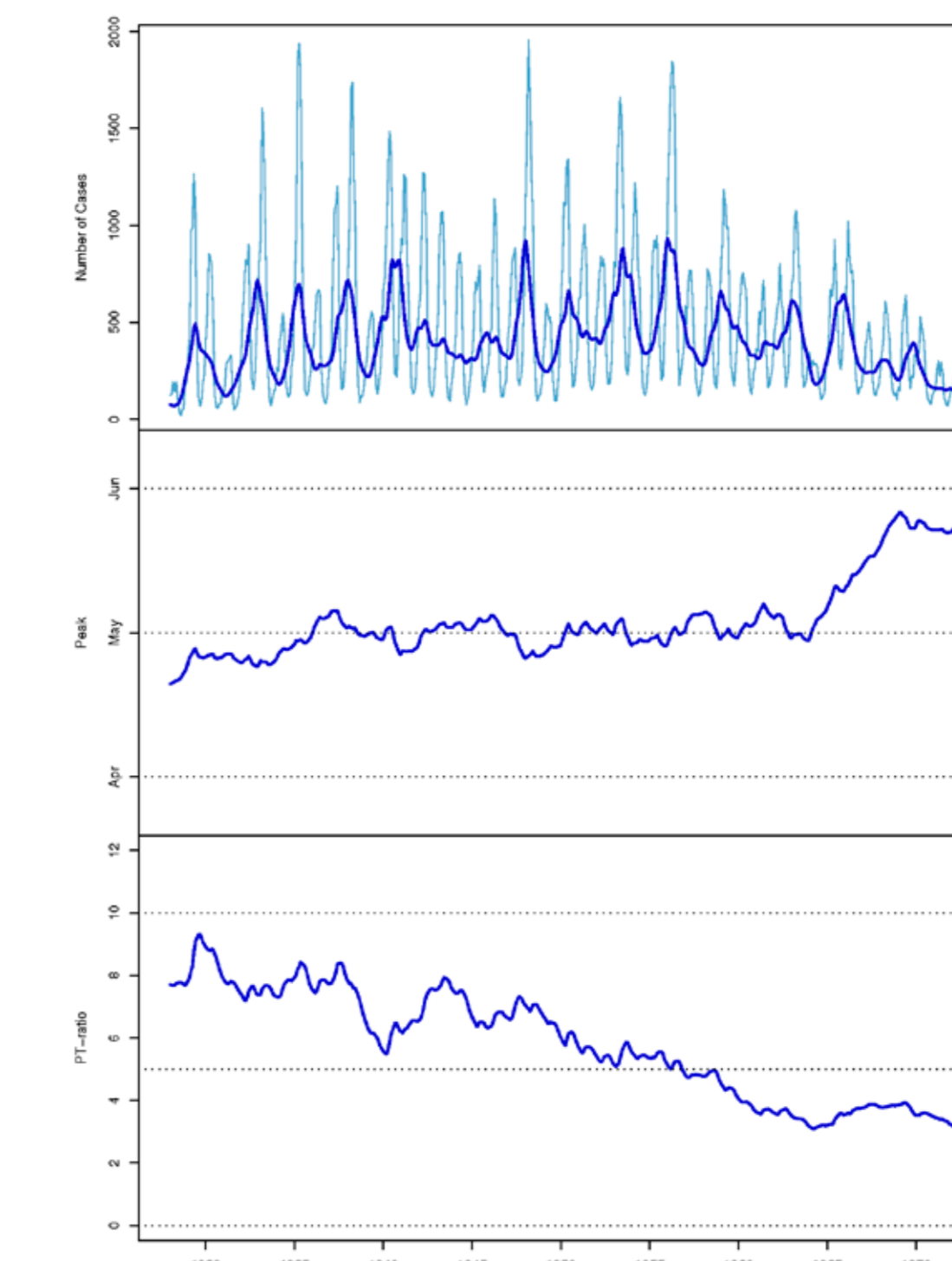
Model

$$y_t \sim \text{Exp.fam.}(\lambda_t = F_t^\top \theta_t)$$

$$\theta_t \sim N(G_t \theta_{t-1}, W_t)$$

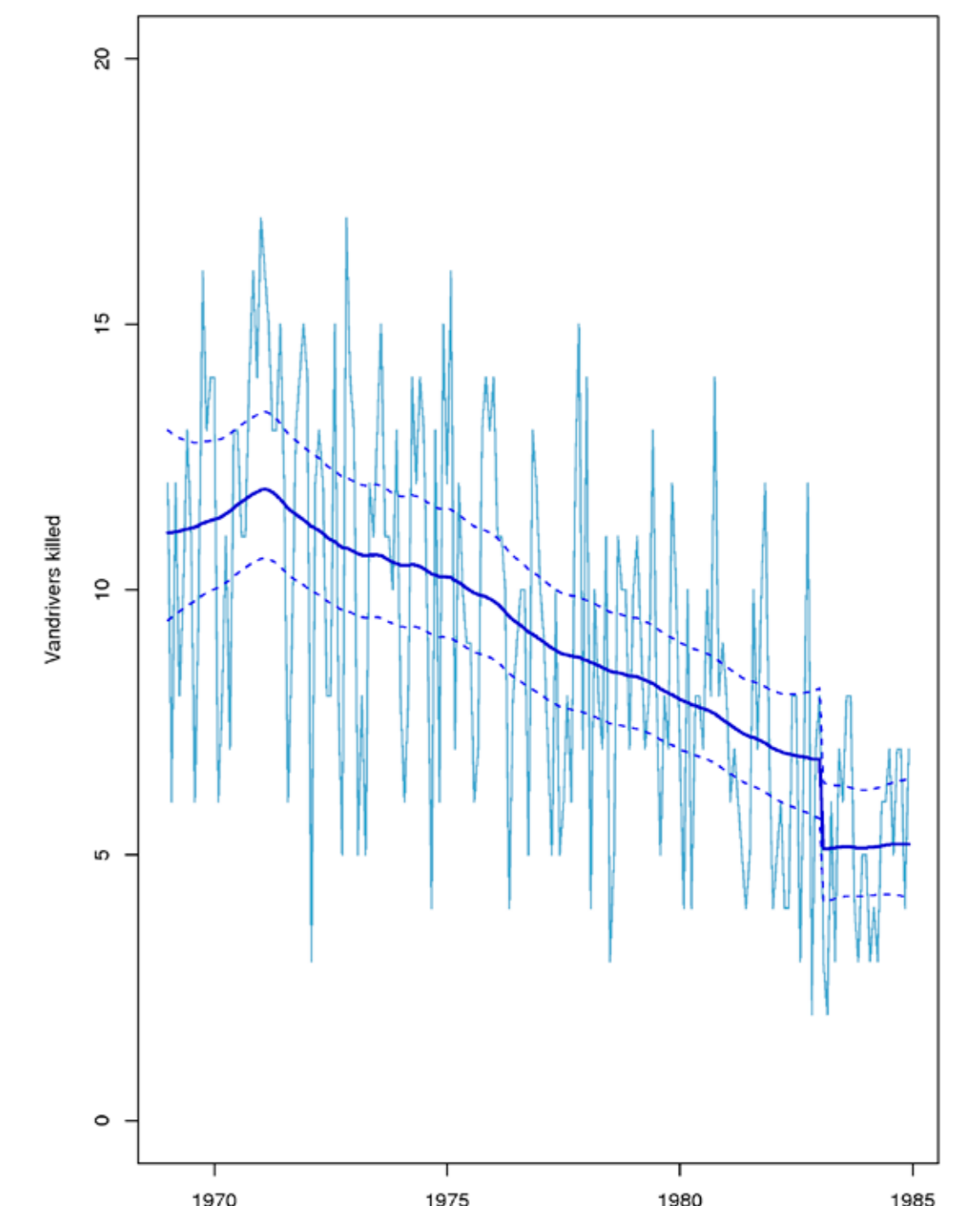
Examples

Variation in the incidence of mumps, NYC, 1927-1972



```
ssm ( mumps ~ -1 + tvar(polytime(index,1)) +
      tvar(polytrig(index,12,1)),
      family=poisson(link=log),
      phi=phi.start,
      C0 = diag(4)
    )
```

Monthly numbers of light goods van drivers killed in road accidents from 1969 to 1984 in UK



```
ssm( y ~ tvar(1) + seatbelt +
      sumseason(time,12),
      family=poisson(link="log"),
      data=vandrivs,
      phi = c(1,0.0004945505),
      C0=diag(13)*1000
    )
```

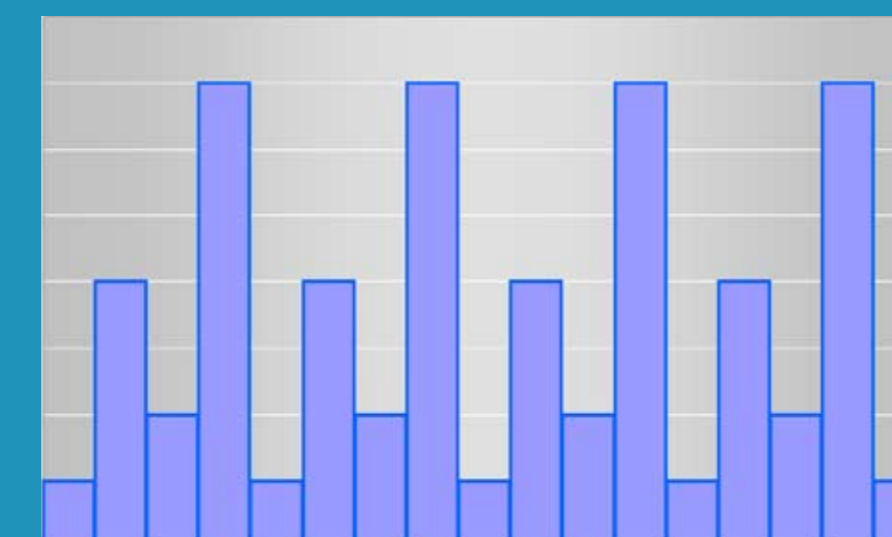
STATIC

Time trend
polytime(time,1)

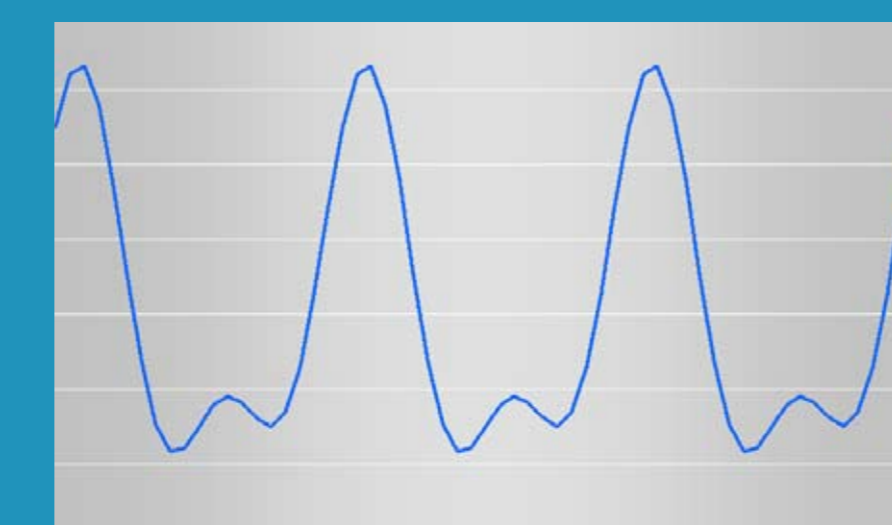
Covariates/interventions



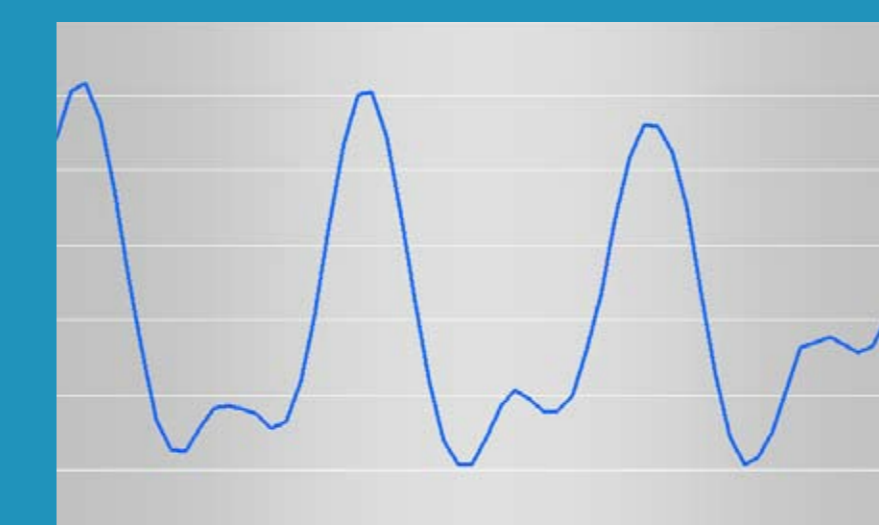
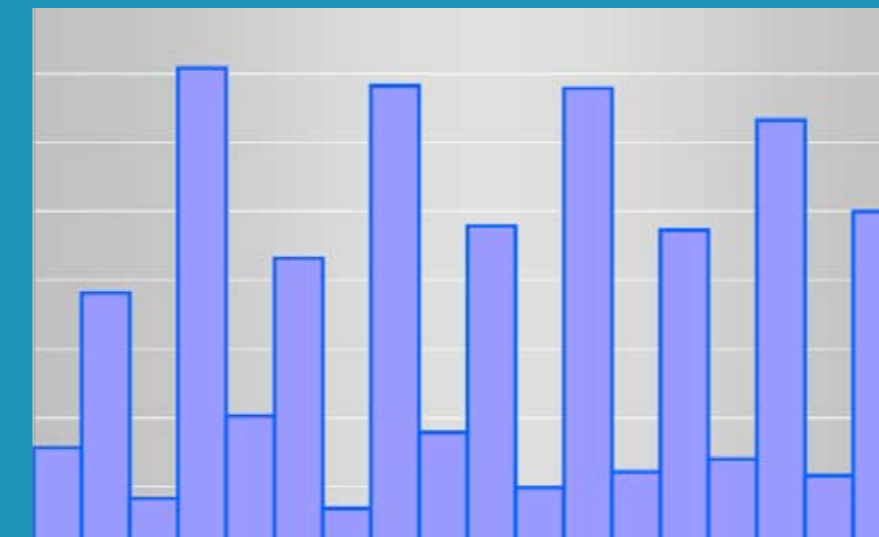
Seasonal variation
Unstructured
sumseason (time,4)



Harmonic
polytrig(time,2)



DYNAMIC



References

Dethlefsen, Lundbye-Christensen (2006). Formulating state space models in R with focus on longitudinal regression models. JStatSoft.

Durbin, Koopman (2000). Time series analysis by state space methods. OUP.

Hipel, McLeod (1994). Time series modeling of water resources and environmental systems.

Ripley (2002). Time series in R 1.5.0. R News.

West, Harrison, Migon (1985). Dynamic generalized linear models and Bayesian forecasting. JASA.

See also the poster

"Assessing seasonality in count data"

Discussion

Estimation of hyper-parameters is not implemented

The approach is very general, but slow

Random walk evolution of dynamic parameters vs. AR(1)

Multi process state space models can be built easily