gamlss

for statistical modelling

Generalized additive models for location scale and shape

Bob Rigby

Mikis Stasinopoulos

LONDON metropolitan university

Summary of GAMLSS talk

1- Introduction to GAMLSS modelling

2- The gamlss package in R

3- Fitting discrete distributions

4- Fitting continuous distributions

6- Smoothing models

8- Centile estimation

9-End

Introduction to GAMLSS modelling

1.1 Introduction to statistical modelling

Statistical modelling:

Model fitting,

Hypothesis testing,

Model selection,

Model diagnostic checking, plots and statistics,

Prediction.

Objective:

If hypothesis testing, use formal statistical tests.

If prediction, use model selection criteria, e.g. AIC.

1.2 History

Regression type of models :

(explanatory variables) X →y (response variables)

Regression analysis

$$y = Xb + e$$

Generalised Linear models: Nelder and Weddeburn (1972)

$g(\mu) = Xb$ where $y \sim$ Exponential family(μ, σ)

Continuous: Normal, Gamma, Inverse Gaussian

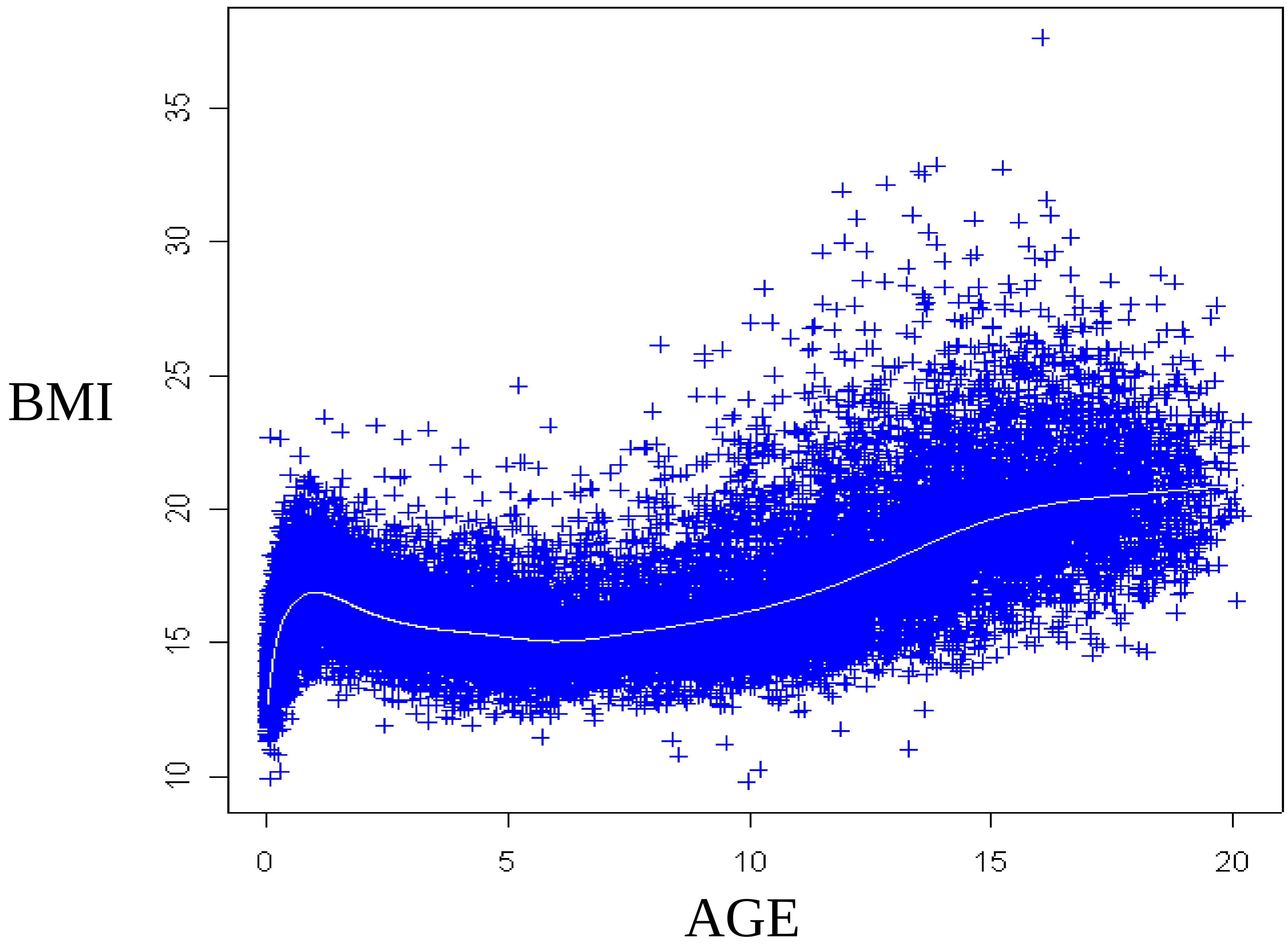
discrete : Binomial, Poisson

Generalised Additive Models Hastie and Tibshirani (1990)

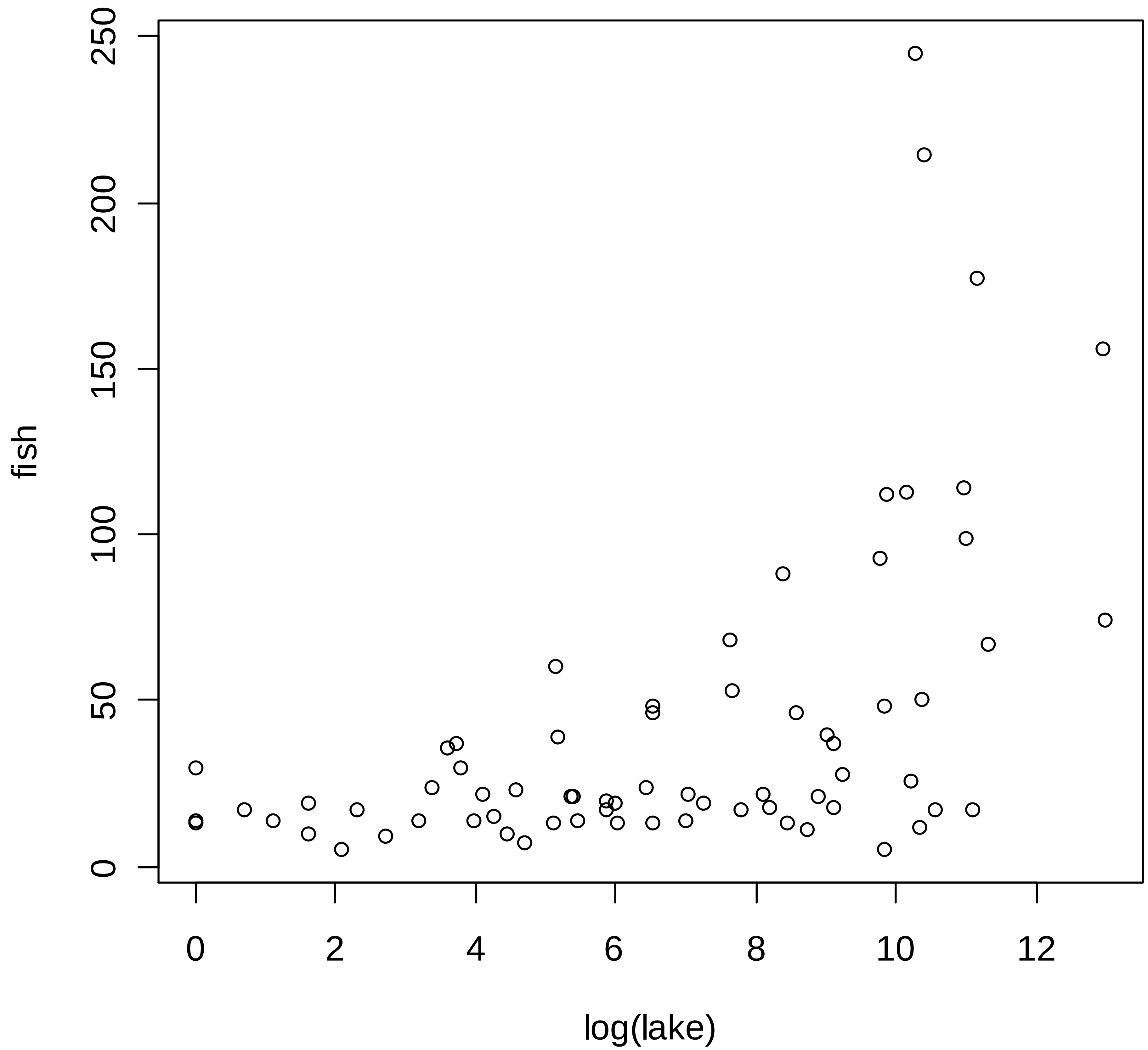
$$g(\mu) = Xb + \sum h(x)$$

1.3 Motivating Examples:

BMI against AGE for Dutch girls



Example 2: The fish species data



gamlss
for statistical modelling

1.4 GAMLSS model

$Y \sim D(\mu, \sigma, \nu, \tau)$ where D is any distribution and

Known link

Predictor

Linear terms

$$g_1(\boldsymbol{\mu}) = \boldsymbol{\eta}_1 = \mathbf{X}_1\boldsymbol{\beta}_1 + \sum_{j=1}^{J_1} \mathbf{Z}_{j1}\boldsymbol{\gamma}_{j1}$$

$$g_2(\boldsymbol{\sigma}) = \boldsymbol{\eta}_2 = \mathbf{X}_2\boldsymbol{\beta}_2 + \sum_{j=1}^{J_2} \mathbf{Z}_{j2}\boldsymbol{\gamma}_{j2}$$

$$g_3(\boldsymbol{\nu}) = \boldsymbol{\eta}_3 = \mathbf{X}_3\boldsymbol{\beta}_3 + \sum_{j=1}^{J_3} \mathbf{Z}_{j3}\boldsymbol{\gamma}_{j3}$$

$$g_4(\boldsymbol{\tau}) = \boldsymbol{\eta}_4 = \mathbf{X}_4\boldsymbol{\beta}_4 + \sum_{j=1}^{J_4} \mathbf{Z}_{j4}\boldsymbol{\gamma}_{j4}.$$

Additive terms

Random effects

Here $\gamma_{jk} \sim N_{q_{jk}}\left(0, G_{jk}^{-1}\right)$ and $G_{jk} = G_{jk}(\lambda)$

Semi-parametric GAMLSS model

$Y \sim D(\mu, \sigma, \nu, \tau)$ where D is any distribution and

$$g_1(\mu) = \mathbf{X}_1\beta_1 + \sum_{j=1}^{J_1} h_{j1}(\mathbf{x}_{j1})$$

$$g_2(\sigma) = \mathbf{X}_2\beta_2 + \sum_{j=1}^{J_2} h_{j2}(\mathbf{x}_{j2})$$

$$g_3(\nu) = \mathbf{X}_3\beta_3 + \sum_{j=1}^{J_3} h_{j3}(\mathbf{x}_{j3})$$

$$g_4(\tau) = \mathbf{X}_4\beta_4 + \sum_{j=1}^{J_4} h_{j4}(\mathbf{x}_{j4}).$$

Nonlinear semi-parametric GAMLSS model

$Y \sim D(\mu, \sigma, \nu, \tau)$ where D is any distribution and

$$
\begin{aligned}
g_1(\mu) &= h_1(\mathbf{X}_1, \beta_1) + \sum_{j=1}^{J_1} h_{j1}(\mathbf{x}_{j1}) \\
g_2(\sigma) &= h_2(\mathbf{X}_2, \beta_2) + \sum_{j=1}^{J_2} h_{j2}(\mathbf{x}_{j2}) \\
g_3(\nu) &= h_3(\mathbf{X}_3, \beta_3) + \sum_{j=1}^{J_3} h_{j3}(\mathbf{x}_{j3}) \\
g_4(\tau) &= h_4(\mathbf{X}_4, \beta_4) + \sum_{j=1}^{J_4} h_{j4}(\mathbf{x}_{j4}).
\end{aligned}
$$

Nonlinear parametric GAMLSS model

$Y \sim D(\mu, \sigma, \nu, \tau)$ where D is any distribution and

$$
\begin{aligned}
g_1(\mu) &= h_1(\mathbf{X}_1, \beta_1) \\
g_2(\sigma) &= h_2(\mathbf{X}_2, \beta_2) \\
g_3(\nu) &= h_3(\mathbf{X}_3, \beta_3) \\
g_4(\tau) &= h_4(\mathbf{X}_4, \beta_4)
\end{aligned}
$$

Parametric GAMLSS model

$Y \sim D(\mu, \sigma, \nu, \tau)$ where D is any distribution and

$$g_1(\mu) = \mathbf{X}_1 \boldsymbol{\beta}_1$$

$$g_2(\sigma) = \mathbf{X}_2 \boldsymbol{\beta}_2$$

$$g_3(\nu) = \mathbf{X}_3 \boldsymbol{\beta}_3$$

$$g_4(\tau) = \mathbf{X}_4 \boldsymbol{\beta}_4$$

1.5 Population distributions for Y

1.4.1 General comments

1) A wide range of discrete and continuous distributions implemented, including highly skew and kurtotic distributions

2) Easy implementation of new distributions

3) Different parameterisations of a distribution can be implemented

4) Truncated distributions and censored data easily implemented

Table 1.2: Implemented GAMLSS distributions (with default link functions)

Distributions	R Name	μ	σ	ν	τ
Beta	BE()	logit	logit	-	-
Beta Inflated (at 0 and 1)	BEINF()	logit	logit	log	log
Beta Binomial	BB()	logit	log	-	-
Binomial	BI()	logit	-	-	-
Box-Cox Cole and Green	BCCG()	identity	log	identity	-
Box-Cox Power Exponential	BCPE()	identity	log	identity	log
Box-Cox-t	BCT()	identity	log	identity	log
Delaporte	DEL()	log	log	logit	-
Gamma	GA()	log	log	-	-
Gumbel	GU()	identity	log	-	-
Inverse Gaussian	IG()	log	log	-	-
Johnson's SU	JSU()	identity	log	identity	log
Johnson's original SU	JSUo()	identity	log	identity	log
Logistic	LO()	identity	log	-	-
Log Normal	LNO()	log	log	fixed	-
Multinomial	MULTIN()	log	log	log	log
Negative Binomial type I	NBI()	log	log	-	-
Negative Binomial type II	NBII()	log	log	-	-
NET	NET()	identity	log	fixed	fixed
Normal	NO()	identity	log	-	-
Poisson	PO()	log	-	-	-
Poisson inverse Gaussian	PIG()	log	log	-	-
Power Exponential	PE()	identity	log	log	-
Reverse Gumbel	RG()	identity	log	-	-
Sichel	SICHEL()	log	log	identity	-
Skew Exponential Power	SEP()	identity	log	identity	log
t Family	TF()	identity	log	log	-
Weibull	WEI()	log	log	-	-
Weibull (PH)	WEI2()	log	log	-	-
Zero inflated poisson	ZIP	log	logit	-	-

1.5 Additive terms

Each parameter of the distribution μ, σ, ν, τ is modelled using terms in explanatory variables x

Parametric additive terms

- Linear and interaction terms for variables and factors.
- Polynomials, inverse polynomials,
  piecewise polynomials (with fixed knots),
  fractional polynomials (Royston and Altman, 1994)
- Non-linear parametric terms

Smoothing and random effects additive terms

- Additive smoothing terms
  - ✓ loess (Cleveland *et al*., 1993)
  - ✓ cubic splines (Green and Silverman, 1994)
  - ✓ P-splines (Eilers and Marx, 1996)
  - ✓ varying coefficient models (Hastie and Tibshirani, 1993)
- Random effects (overdispersion, simple random effects, random coefficients)
- Parameter driven Time Series (random walks)

Additive terms in GAMLSS

Each parameter of the distribution μ, σ, ν, τ is modelled using (additive) terms in explanatory variables x

Model	Formula	gamlss code
simple linear	$\beta_0 + \beta_1 x$	x
polynomial	$\beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$	poly(x,3)
fractional polynomial	$\beta_0 + \beta_1 x^{p_1} + \beta_2 x^{p_2}$	fp(x,2)
	for $p \in (-2, -1, 0, 0.5, 1, 2, 3)$	
e.g.	$\beta_0 + \beta_1 x^{0.5} + \beta_2 x^{-2}$	bfp(0.5,-2)
power polynomial	$\beta_0 + \beta_1 x^{p_1} + \beta_2 x^{p_2}$	pp(x,2)
loess smoother		lo(x,3)
p-spline smoother		ps(x,3)
cubic spline smoother		cs(x,3)

2.0 The R packages

gamlss package

gamlss.nl package

gamlss.tr package

gamlss.dist package

2.1 The gamlss package

- The gamlss() function creates a gamlss object
- Methods for a gamlss object

  AIC(), addterm(), coef(), deviance(), fitted(), formula(), lot(), print(), predict(), residuals(), update(),
- Others functions

  centiles(), fitted.plot(), GAIC(), gamlss.scope(), par.plot(), lperd(), pdf.plot(), prof.plot(), prof.term(), Q.stats(), refit(), rqres.plot(), stepGAIC(), term.plot()

2.3 The gamlss family of distributions

Many distributions available

Each distribution has:

dNO : the pdf function

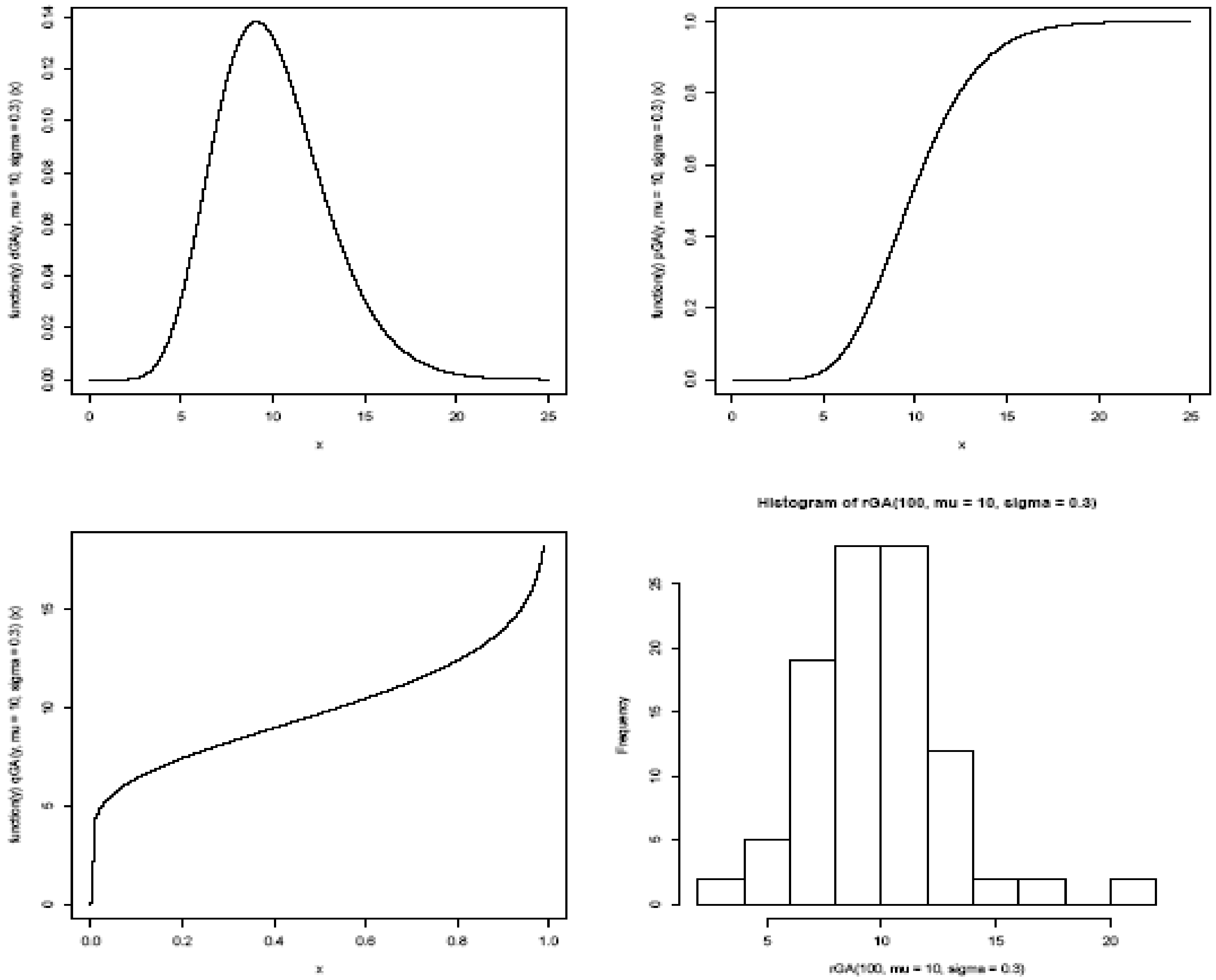
pNO : the cdf function

qNO : the inverse cdf function

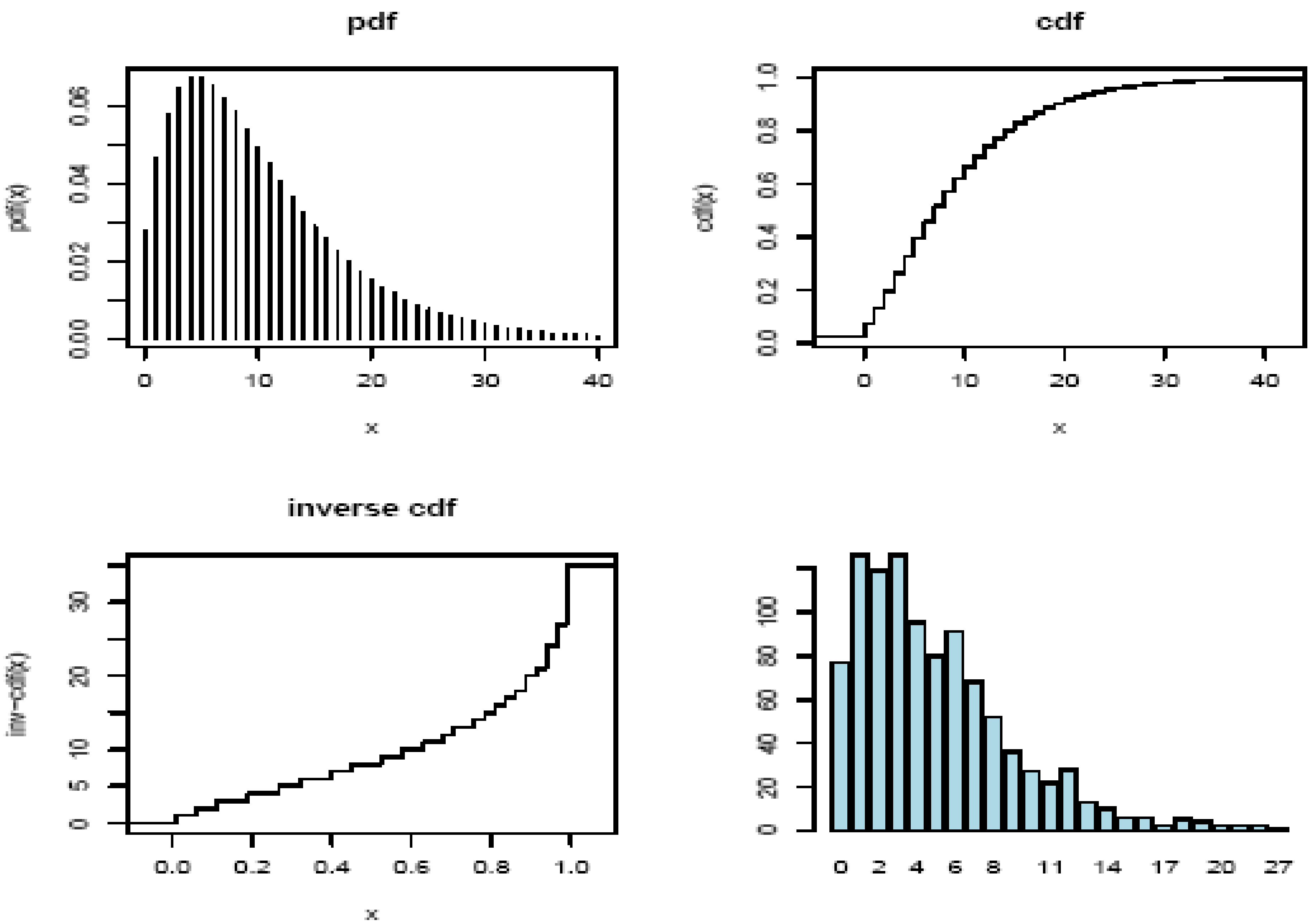
rNO : random generating function

NO : the fitting function

Plotting a continuous distribution



Plotting a discrete distribution



Demonstration

```
library(gamlss)
data(abdom)
plot(y~x, col="blue", xlab="age", ylab="circumference", data=abdom)
# fitting polynomials
abd0<-gamlss(y~poly(x,3), family=NO, data=abdom)
abd00<-gamlss(y~x+I(x^2)+I(x^3), family=NO, data=abdom)
abd00
summary(abd00)
# fitting cubic splines
abd1<-gamlss(y~cs(x,df=3), family=NO, data=abdom)
abd1
summary(abd1)
#  fitted sigma
fitted(abd1,"sigma")[1]
predict(abd1,what="sigma", type="response")[1] [1]
```

3.0 Fitting discrete distributions

One parameter distributions

BI	Binomial
PO	Poisson

Two parameter distributions

BB	Beta-Binomial
NBI	Negative Binomial type I
NBII	Negative Binomial type II
PIG	Poisson-Inverse Gaussian
ZIP	Zero inflated Poisson

Three parameter distributions

SICHEL	Sichel
DEL	Delaporte

Example 1: Computer failure count data

```
> failure <- c(4, 0, 0, 0, 3, 2, 0, 0, 6, 7, 6, 2, 1, 11, 6, 1,
+ 2, 1, 1, 2, 0, 2, 2, 1, 0, 12, 8, 4, 5, 0, 5, 4, 1, 0, 8,
+ 2, 5, 2, 1, 12, 8, 9, 10, 17, 2, 3, 4, 8, 1, 2, 5, 1, 2,
+ 2, 3, 1, 2, 0, 2, 1, 6, 3, 3, 6, 11, 10, 4, 3, 0, 2, 4, 2,
+ 1, 5, 3, 3, 2, 5, 3, 4, 1, 3, 6, 4, 4, 5, 2, 10, 4, 1, 5,
+ 6, 9, 7, 3, 1, 3, 0, 2, 2, 1, 4, 2, 13, 0, 2, 1, 1, 0, 3,
+ 16, 22, 5, 1, 2, 4, 7, 8, 6, 11, 3, 0, 4, 7, 8, 4, 4, 5)

> mPO <- histDist(failure, "PO", main = "PO")
GAMLSS-RS iteration 1: Global Deviance = 769.9487
GAMLSS-RS iteration 2: Global Deviance = 769.9487
```

Fitting discrete distributions to computer failure data

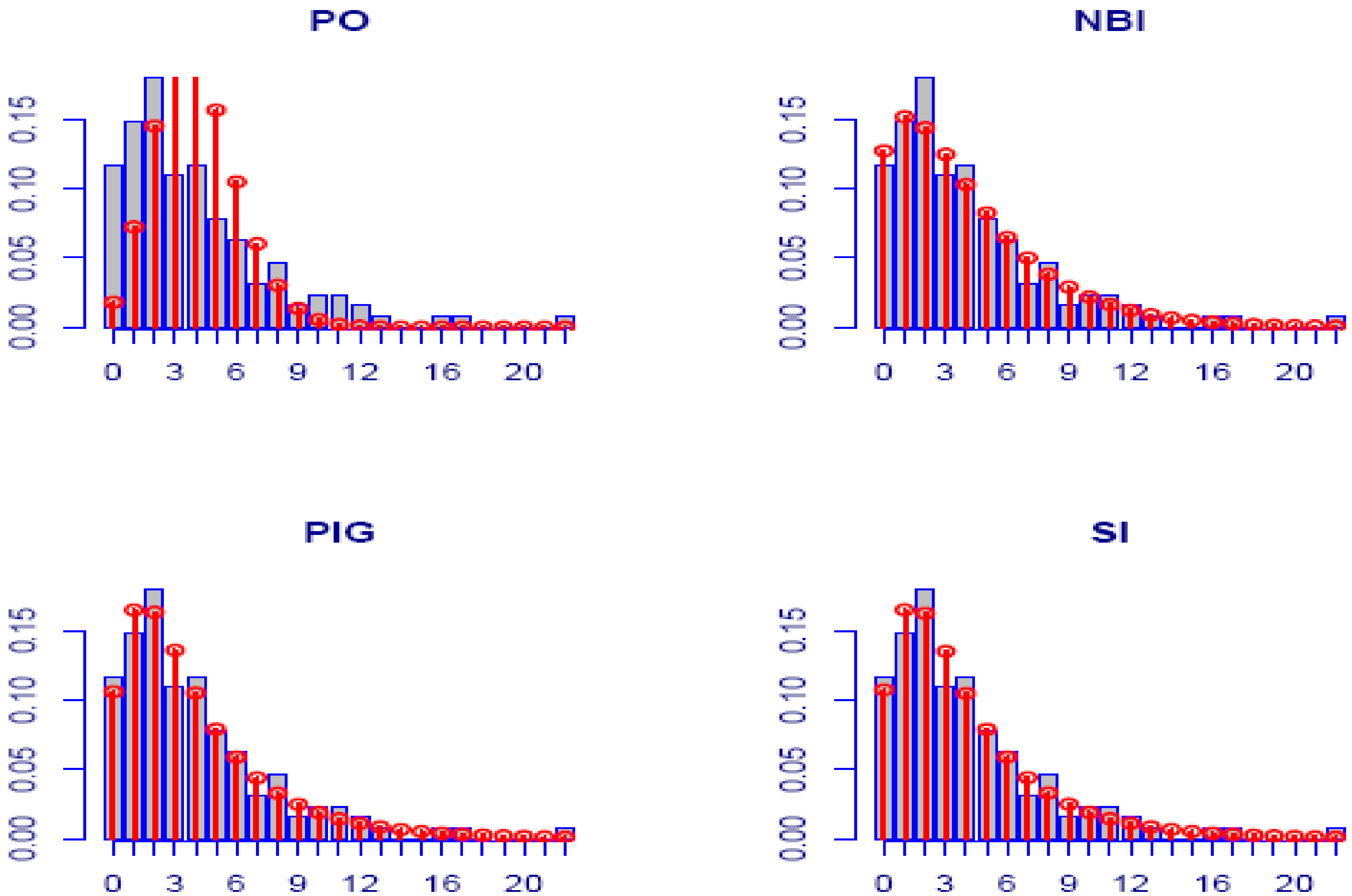


Table of Akaike information criterion (AIC)

>AIC(mPO, mNBI, mPIG, mSI)

	df	AIC
mPIG	2	636.4160
mNBI	2	636.8405
mSI	3	638.3314
mPO	1	771.9487

From the GAIC table above we conclude that the PIG model is the appropriate model here, (although NBI model is close).

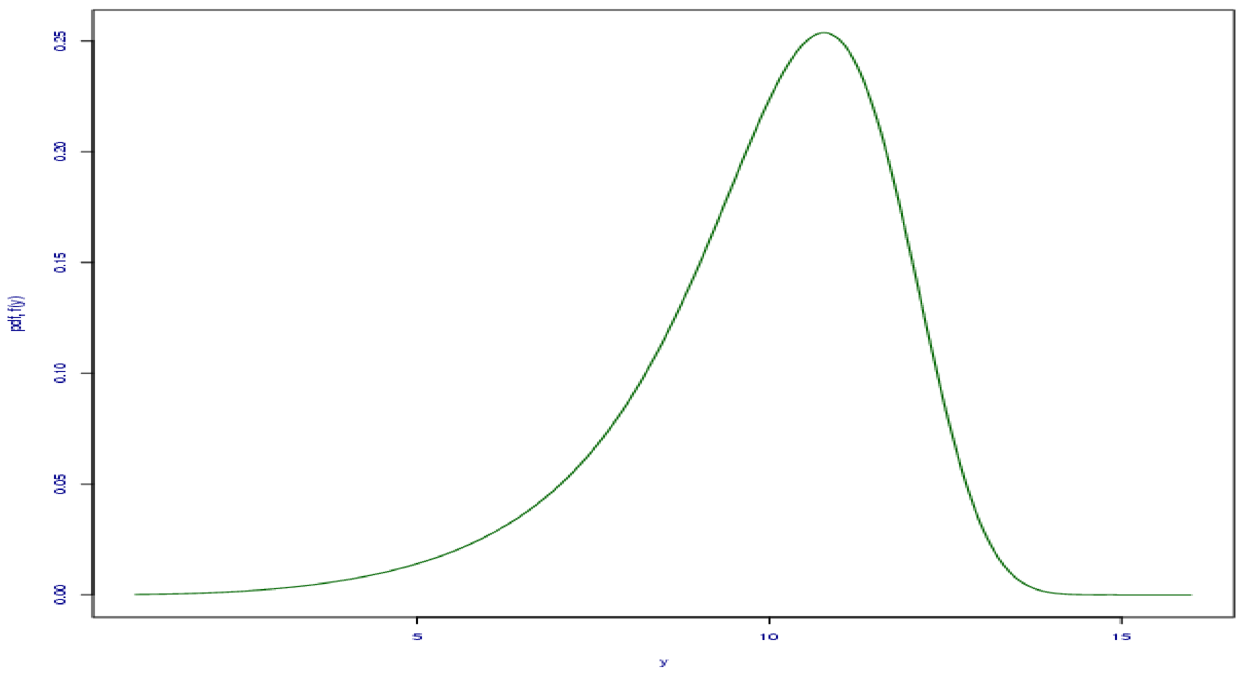
4.0 Fitting continuous distributions

Four parameters

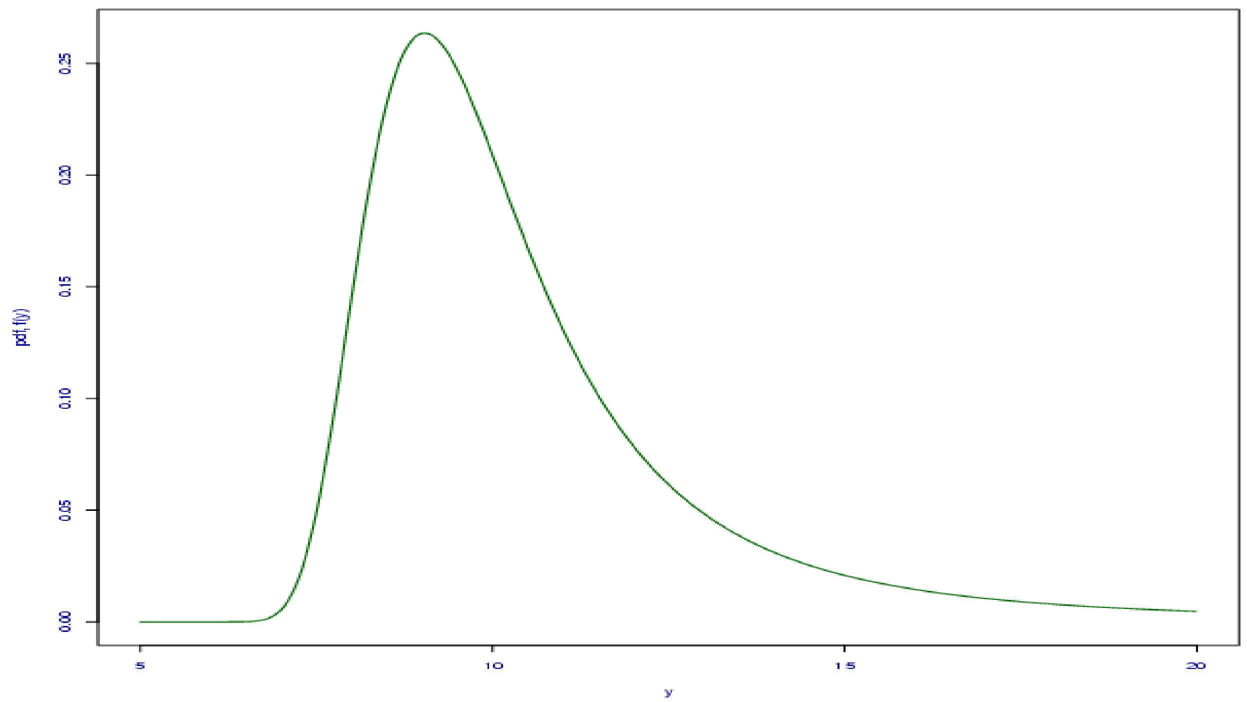
μ	location
σ	scale
ν	skewness
τ	kurtosis

Skewness and kurtosis

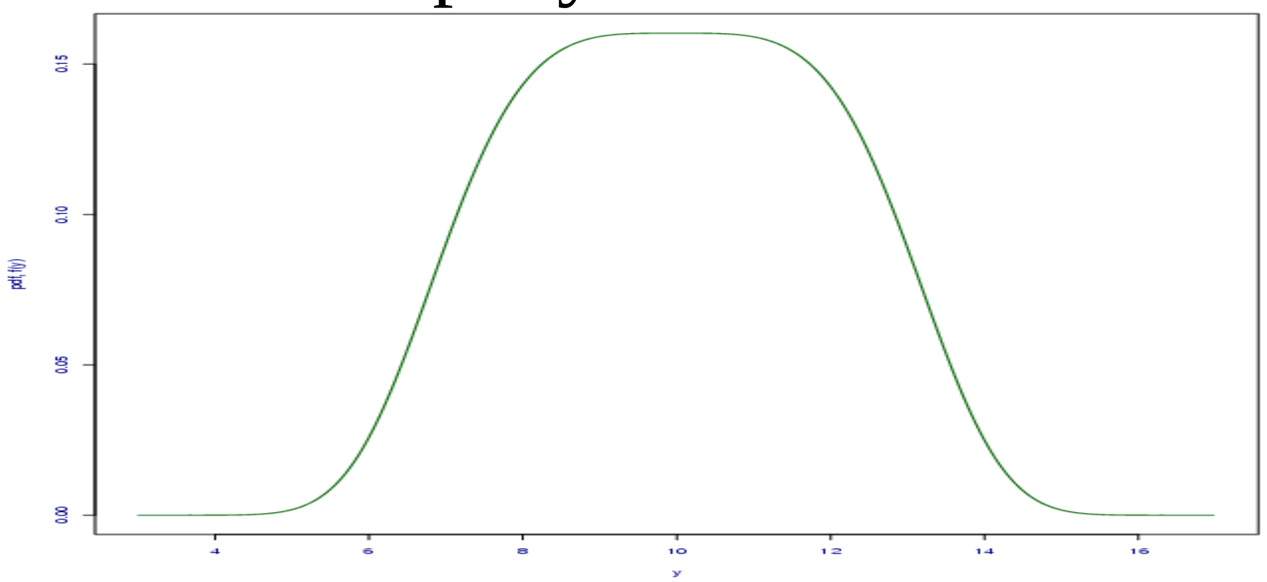
negative skewness



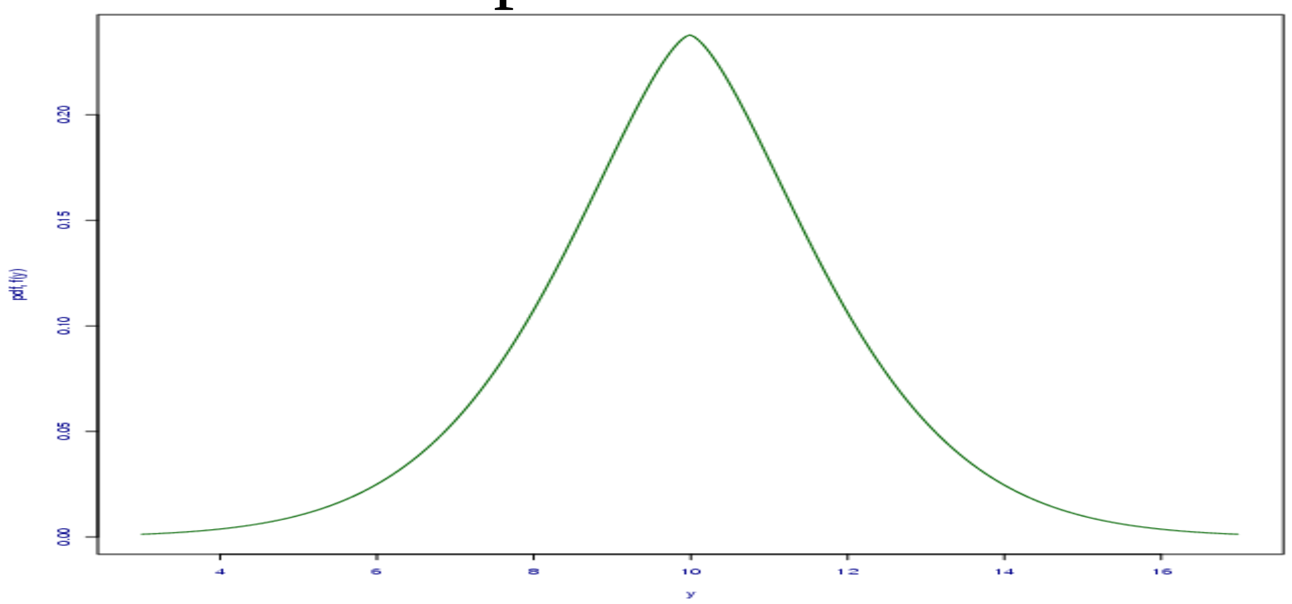
positive skewness



platykurtosis



leptokurtosis



Continuous distributions for *Y*

Two parameter distributions

BE	Beta
GA	Gamma
GU	Gumbel
LO	Logistic
LNO	Log Normal
NO	Normal
IG	Inverse Gaussian
RG	Reverse Gumbel
WEI	Weibull (also WEI2, WEI3)

Continuous distributions for *Y*

Three parameter distributions

BCCG	Box-Cox Normal
PE	Power Exponential
TF	*t* family

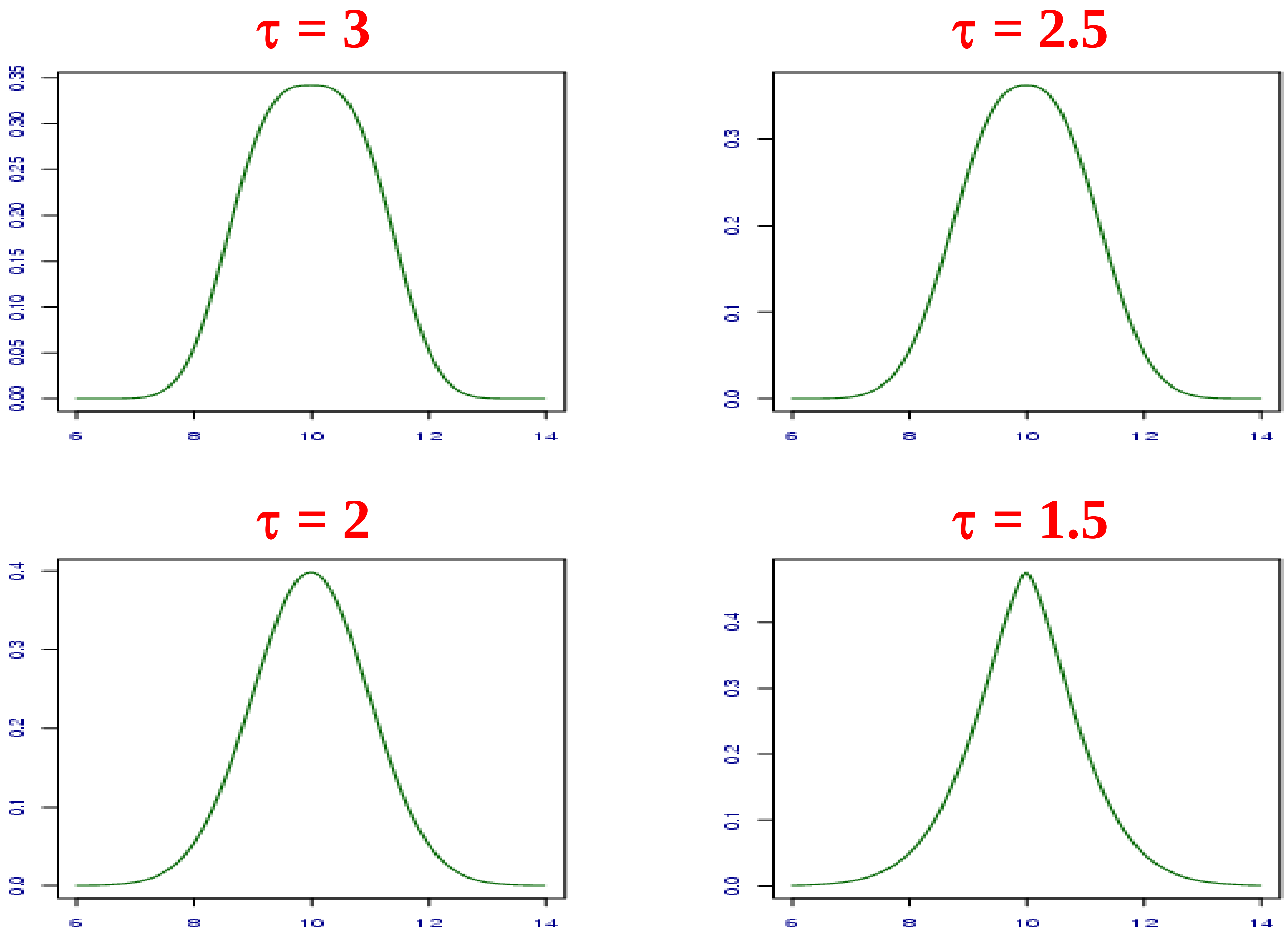
Continuous distributions for *Y*

Four parameter distributions

BCT	Box-Cox *t*
BCPE	Box-Cox Power Exponential
JSU	Johnson Su
SHASH	Sinh Arc Shinh
SEP	Skew Exponential Power
ST3	Skew *t*
BEINF	Beta inflated

BCPE distribution for $\nu = 1$, symmetric case

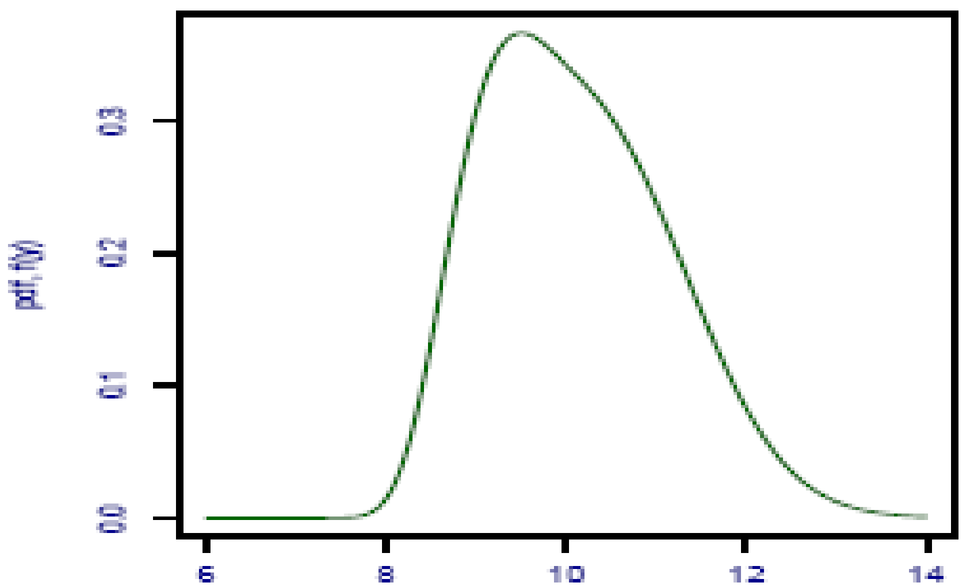
[$\mu = 10$, $\sigma = 0.1$, $\nu = 1$, $\tau = 3$, 2.5, 2, 1.5]



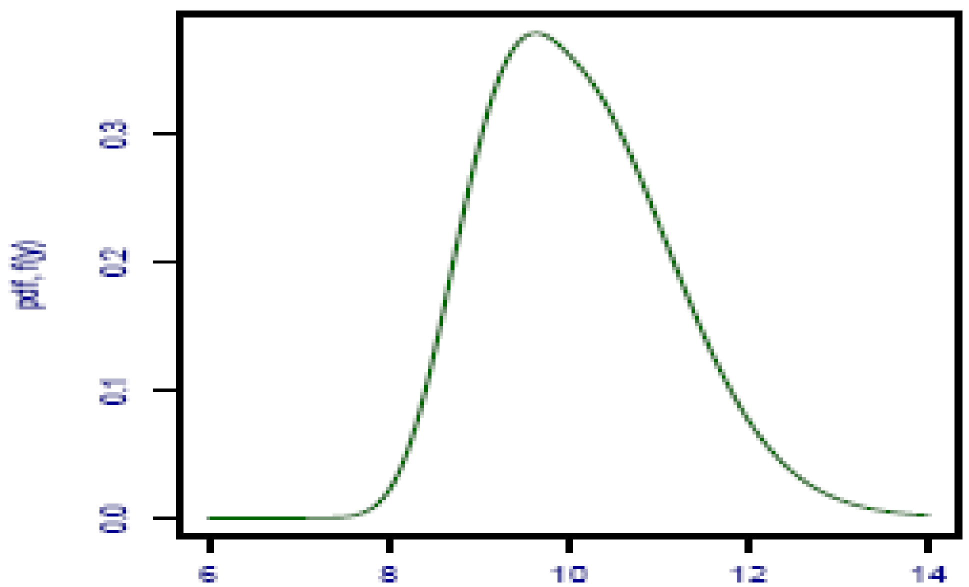
BCPE distribution for $\nu < 1$, positively skew case

$[\mu = 10\ ,\ \sigma = 0.1,\ \nu = -1,\ \tau = 3,\ \ 2.5,\ \ 2,\ \ 1.5]$

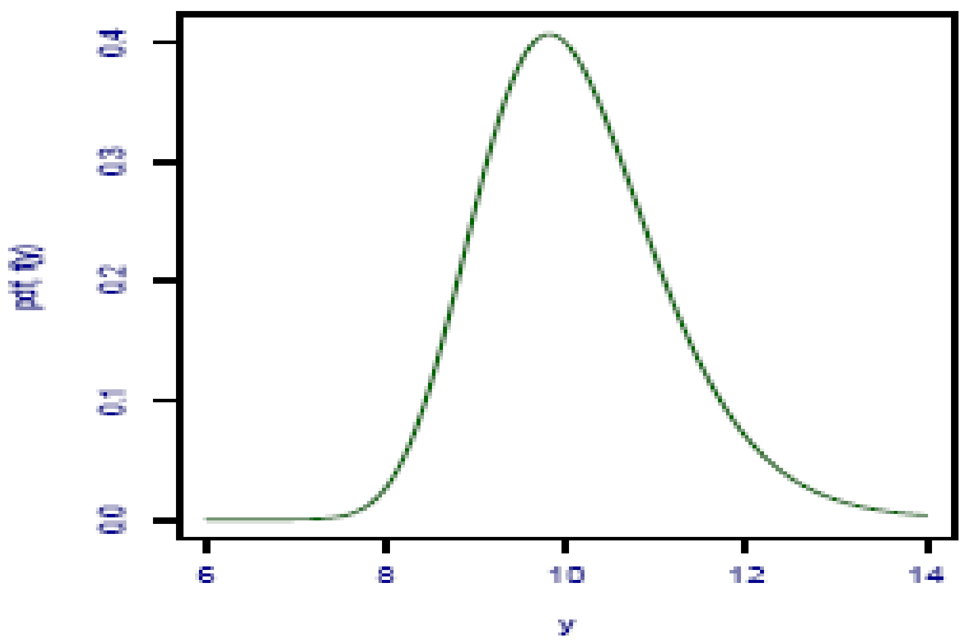
$\tau = 3$



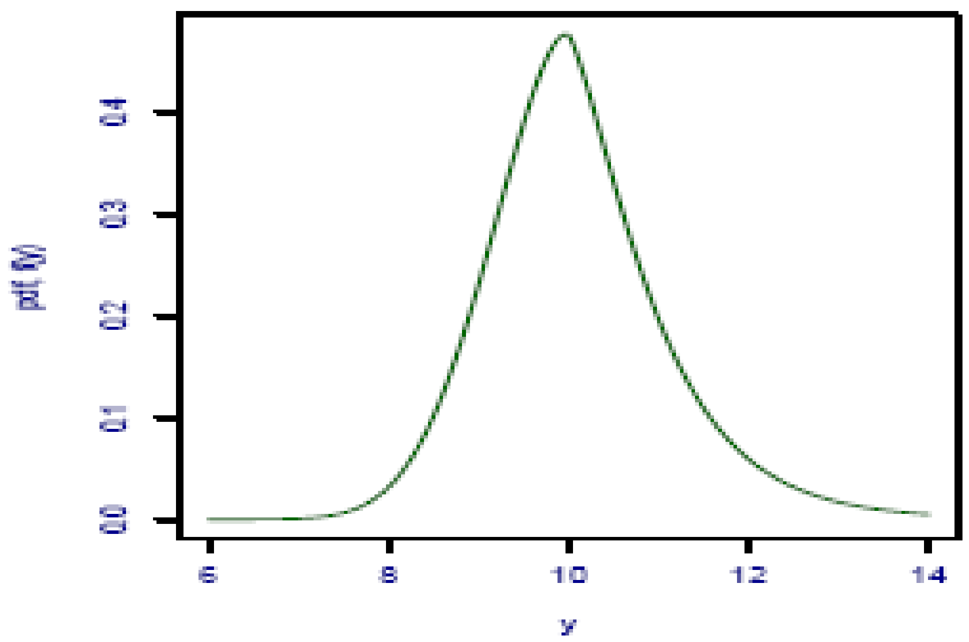
$\tau = 2.5$



$\tau = 2$



$\tau = 1.5$



The Parzen data

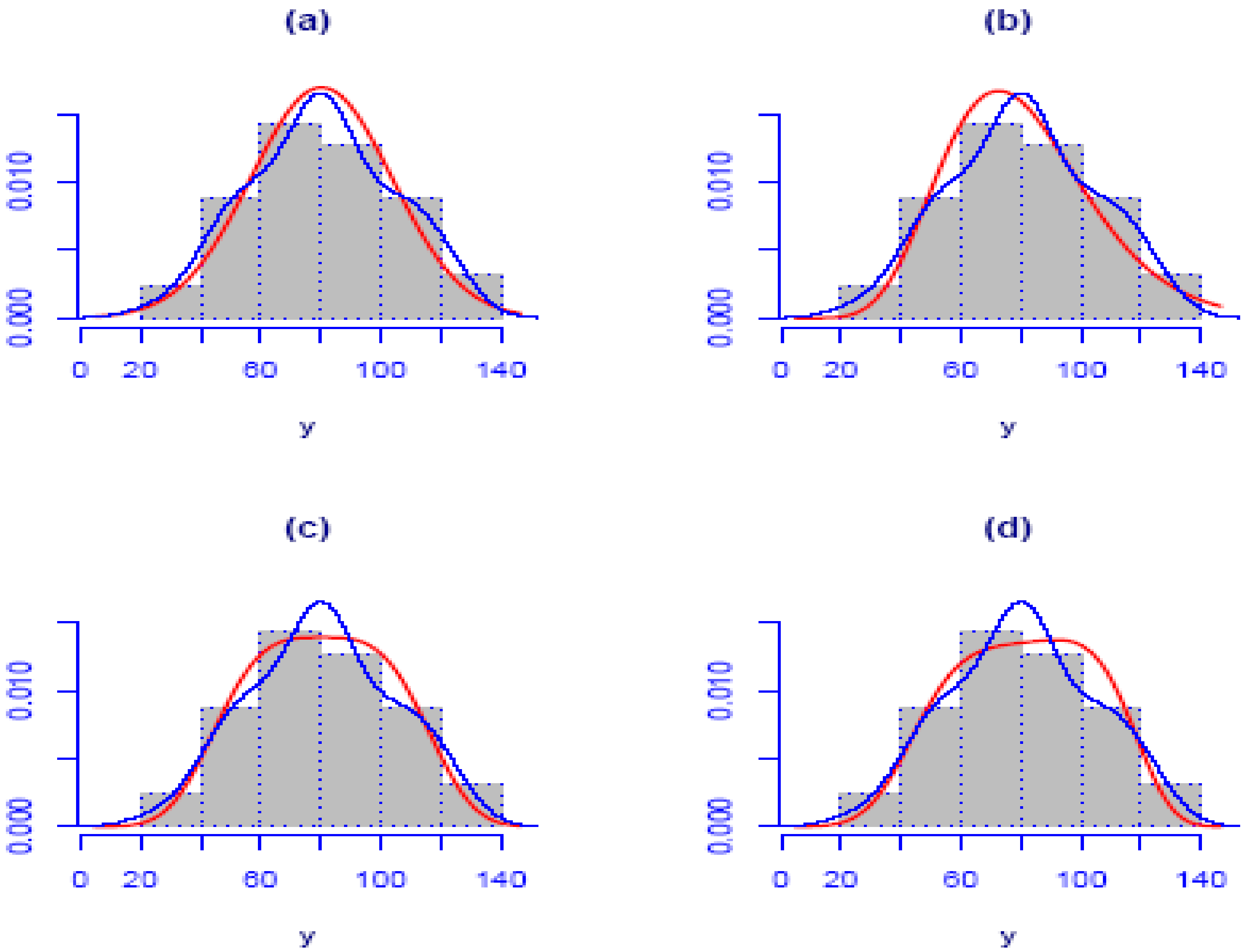
snowfall data, Parzen (1979): the annual snowfall in Buffalo, NY (inches) for the 63 years, from 1910 to 1972 inclusive.

```
> snowfall <- c(126.4, 82.4, 78.1, 51.1, 90.9, 76.2, 104.5, 87.4,
+     110.5, 25, 69.3, 53.5, 39.8, 63.6, 46.7, 72.9, 79.7, 83.6,
+     80.7, 60.3, 79, 74.4, 49.6, 54.7, 71.8, 49.1, 103.9, 51.6,
+     82.4, 83.6, 77.8, 79.3, 89.6, 85.5, 58, 120.7, 110.5, 65.4,
+     39.9, 40.1, 88.7, 71.4, 83, 55.9, 89.9, 84.8, 105.2, 113.7,
+     124.7, 114.5, 115.6, 102.4, 101.4, 89.8, 71.5, 70.9, 98.3,
+     55.5, 66.1, 78.4, 120.5, 97, 110)
```

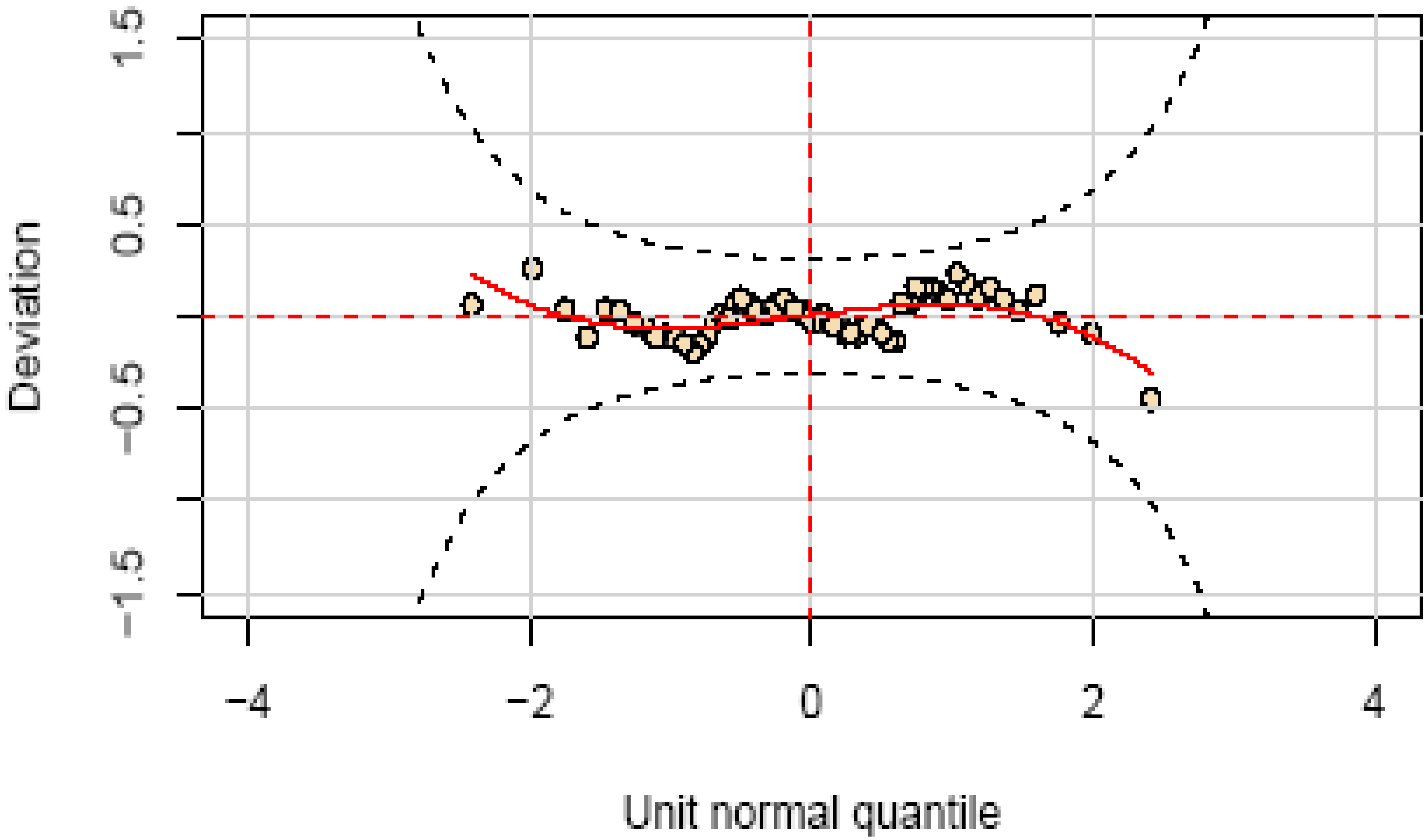
Parzen data: analysis

```
> op <- par(mfrow = c(2, 2))
> mNO <- histDist(snowfall, "NO", density = TRUE, main = "(a)",
+       ymax = 0.017)
> mGA <- histDist(snowfall, "GA", density = TRUE, main = "(b)",
+       ymax = 0.017)
> mPE <- histDist(snowfall, "PE", density = TRUE, main = "(c)",
+       ymax = 0.017)
> mBCPE <- histDist(snowfall, "BCPE", density = TRUE, main = "(d)",
+       ymax = 0.017)
> par(op)

> AIC(mNO, mGA, mPE, mBCPE)

      df      AIC
mNO    2 580.7331
mPE    3 581.3780
mBCPE  4 583.2114
mGA    2 583.8153
```

Parzen data: plots



Parzen data: worm plot

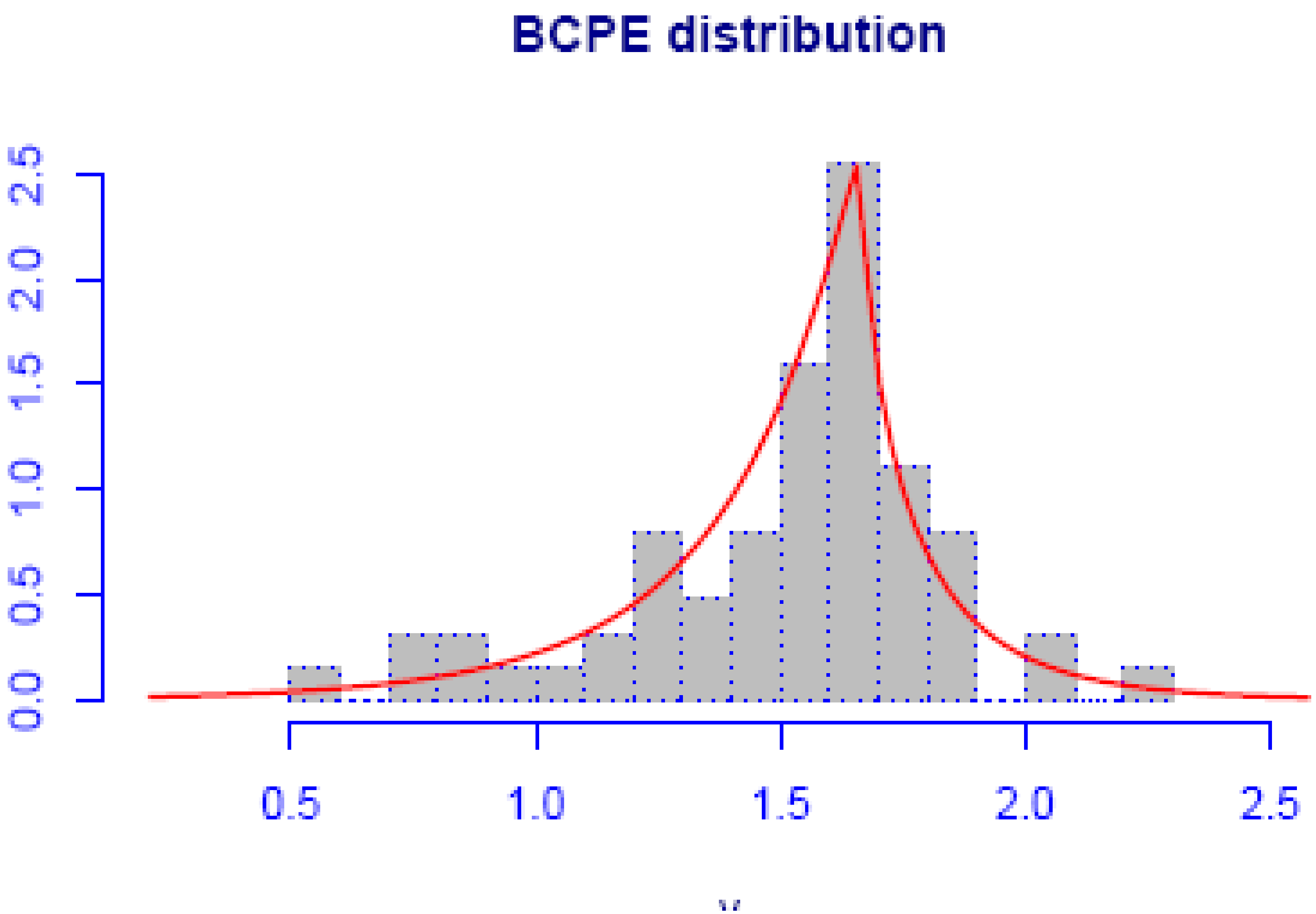


strength of glass fibres data

Strength of glass fibres, measured at the National Physical Laboratory, England, see Smith and Naylor (1987),

```
> strength <- c(0.55, 0.74, 0.77, 0.81, 0.84, 0.93, 1.04, 1.11,
+     1.13, 1.24, 1.25, 1.27, 1.28, 1.29, 1.3, 1.36, 1.39, 1.42,
+     1.48, 1.48, 1.49, 1.49, 1.5, 1.5, 1.51, 1.52, 1.53, 1.54,
+     1.55, 1.55, 1.58, 1.59, 1.6, 1.61, 1.61, 1.61, 1.61, 1.62,
+     1.62, 1.63, 1.64, 1.66, 1.66, 1.66, 1.67, 1.68, 1.68, 1.69,
+     1.7, 1.7, 1.73, 1.76, 1.76, 1.77, 1.78, 1.81, 1.82, 1.84,
+     1.84, 1.89, 2, 2.01, 2.24)
```

glass data: plot

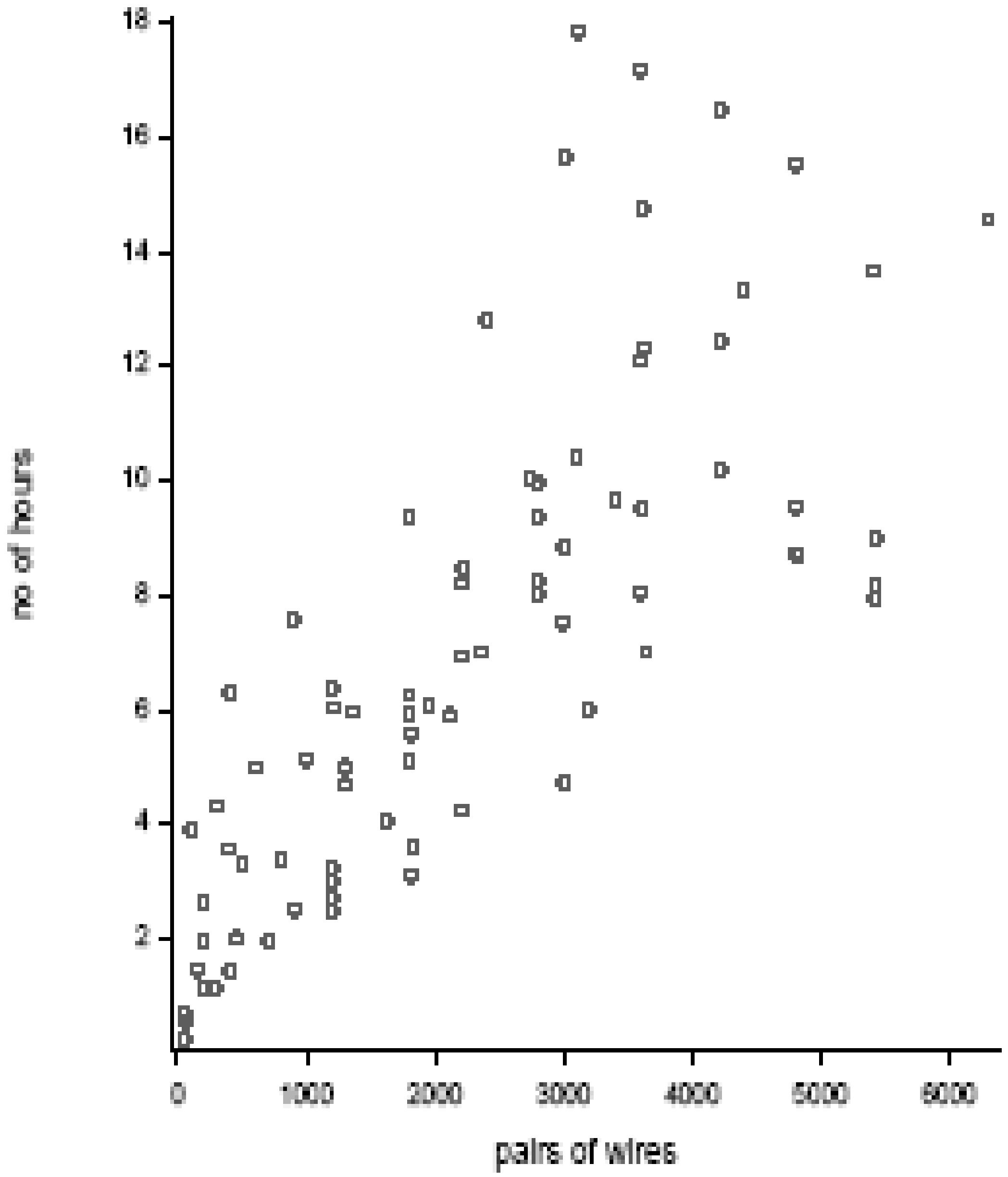


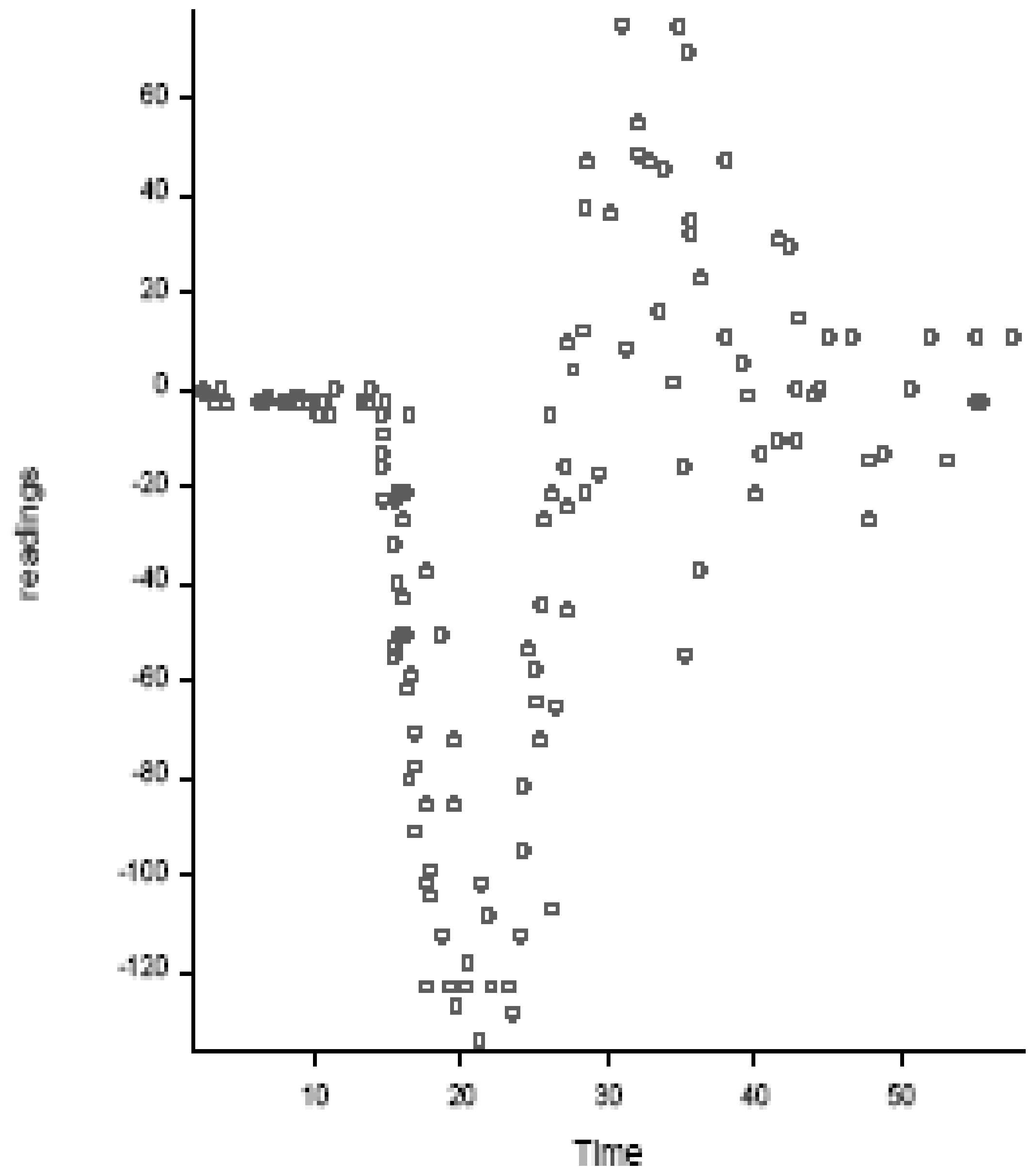
6.0 Smoothing models

Introduction

Smoothing in GAMLSS

Scatter plot smoothers





Running polynomial smoothers

Unweighted
- running mean
- running lines
- running quadratics
- running cubic

Weighted
- Kernel smoothers
- Loess

Running mean smoother

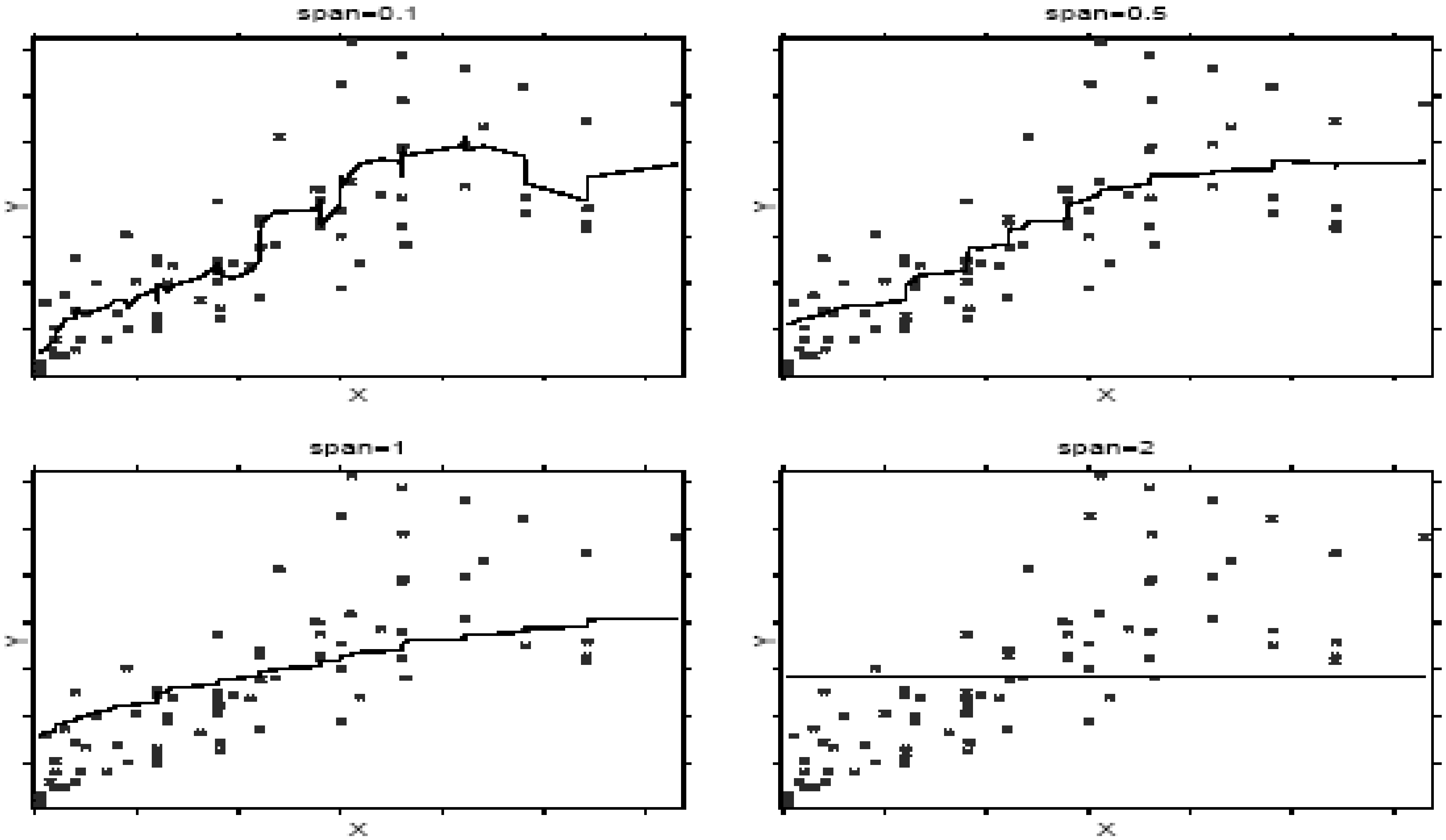


Figure 2.2: Cable data: Running mean smoothers with span=0.1,0.5,1, and 2

Running line smoother

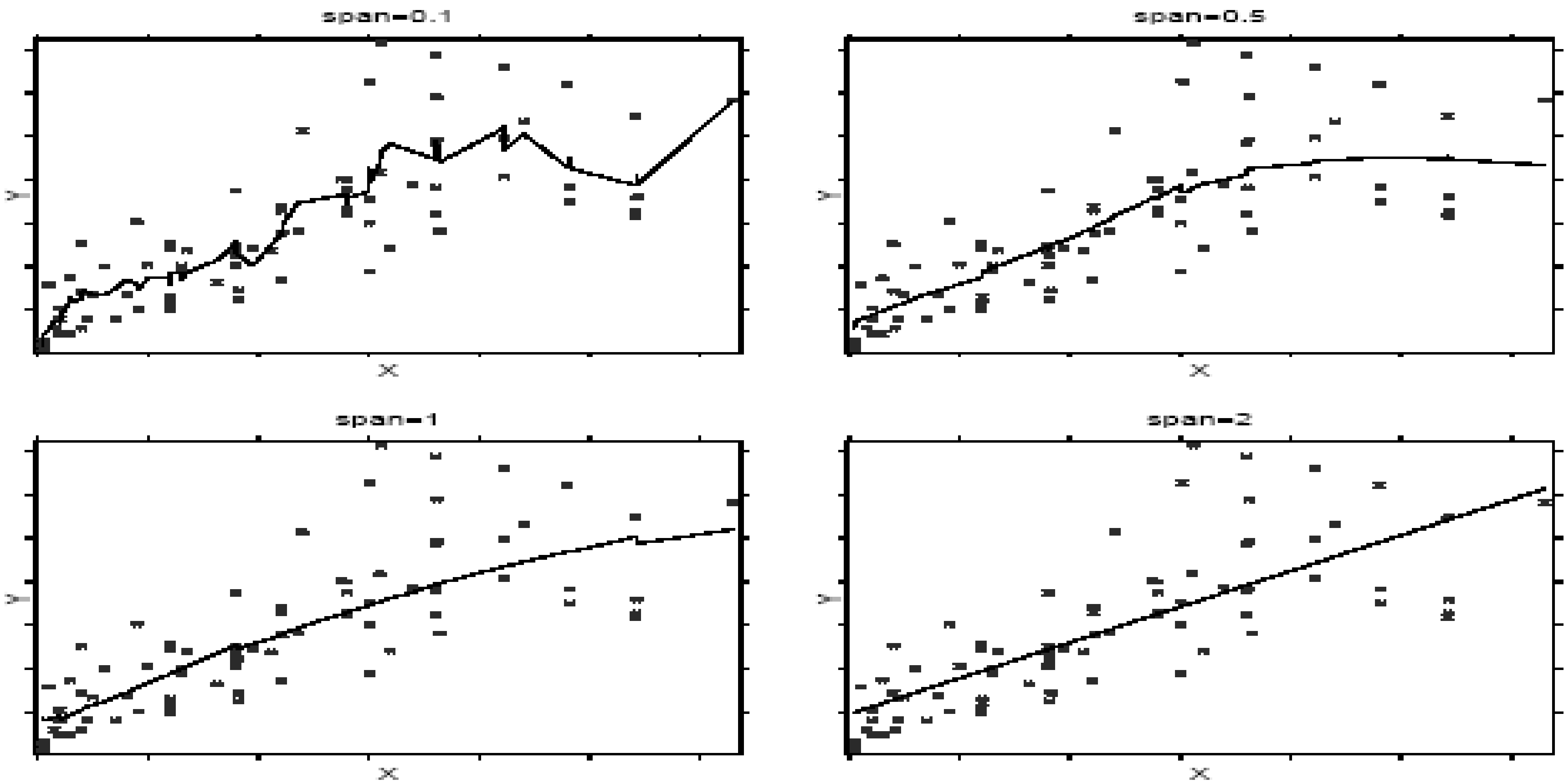


Figure 2.3: Cable data: Running line smoothers with span=0.1,0.5,1, and 2

gamlss

for statistical modelling

Comparison of different running smoothers

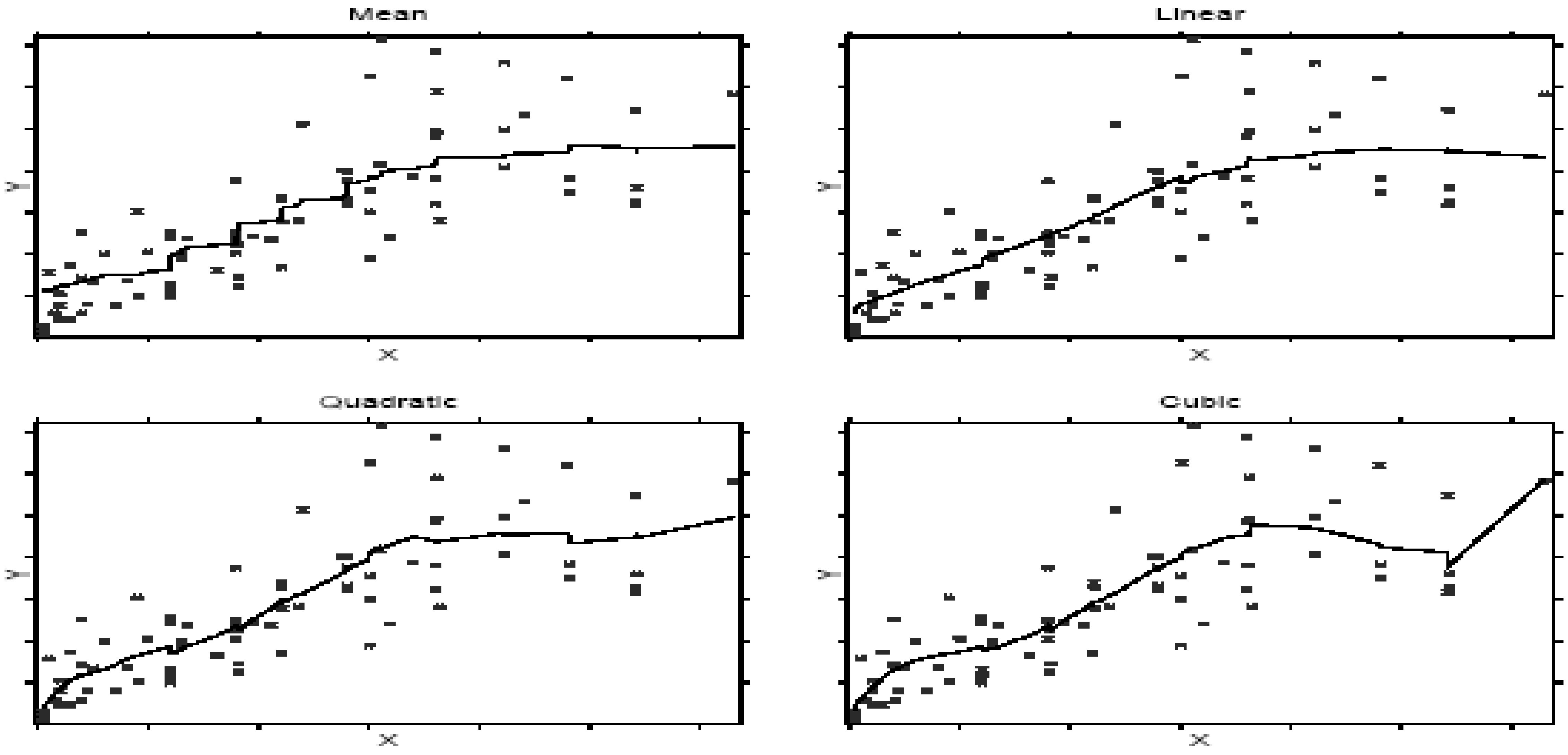


Figure 2.4: Cable data: Running mean, linear, quadratic and cubic smoothers with span=0.5

Kernel smoothers: weighed running means

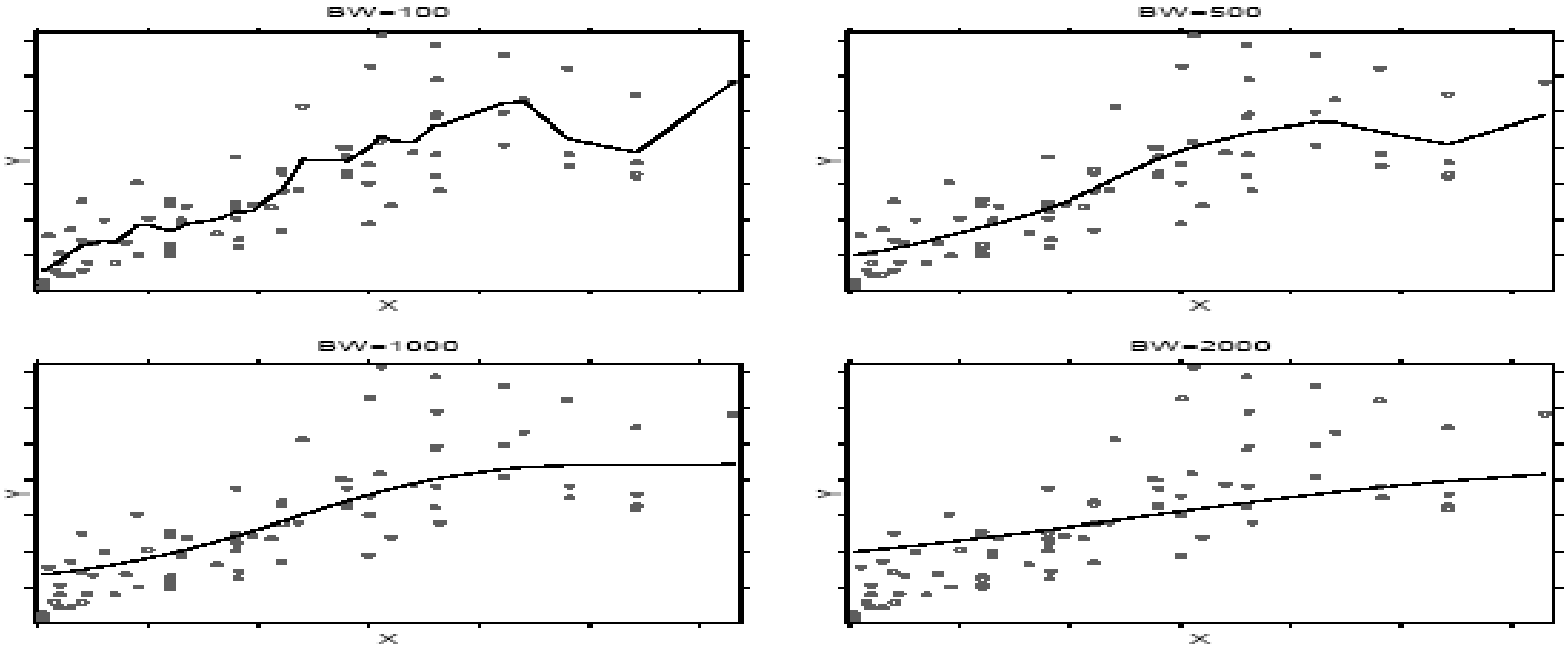


Figure 2.10: Cable data: Gaussian kernel curves with different bandwidths (BW)

Locally Weighted Running-line Smoothers (loess)

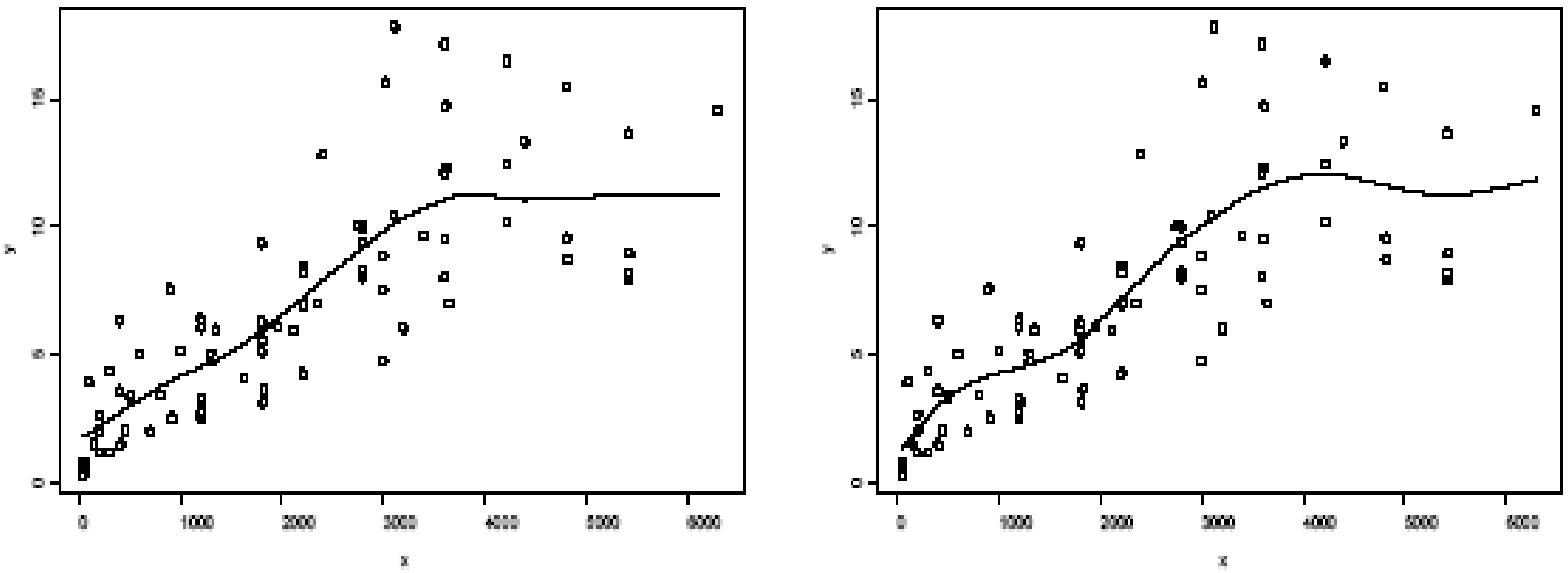


Figure 2.11: a) loess curves using the cable data: a loess with Gaussian kernel, linear fit and span=0.5. , (b) loess curve using Gaussian kernel, quadratic fit and span=0.5.

cubic splines and penalised splines

$$\eta(x) = \beta_{00} + \beta_{01}x + \beta_{02}x^2 + \beta_{03}x^3 + \sum_{k=1}^{K} \beta_k (x - \gamma_k)^3 H(x > \gamma_k)$$

The Rent data

R: rent response variable, the monthly net rent in DM, i.e. the monthly rent minus calculated or estimated utility cost.

F: floor space in square meters

A: year of construction

SP: a two level factor indicating whether the location is above average (550 observations) or not (1419 observations)

SM: a two level factor indicating whether the location is below average (172 obs.) or not (1797 obs.)

B: a two level factor indicating whether there is a bathroom (1925 obs.) or not (44 obs.)

H: a two level factor indicating whether there is central heating (1580 obs.) or not (389 obs.)

L: a two level factor indicating whether the kitchen equipment is above average (161 obs.) or not (1808 obs.)

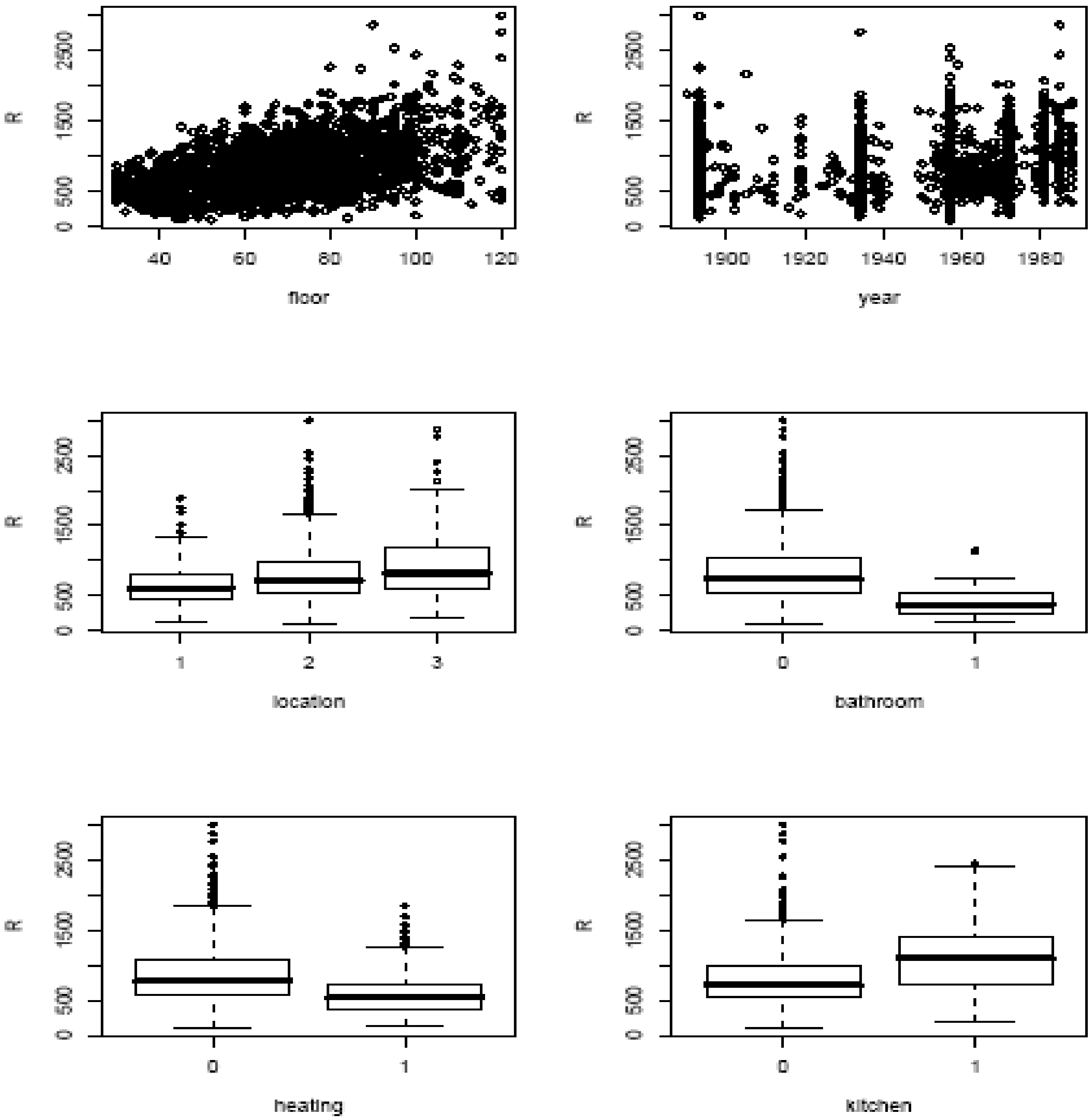


Figure 6.2: Plots of net rent (R) against each of the explanatory variables.

Cubic splines, the cs() function

```
> library(gamlss)
> data(rent)
> r1 <- gamlss(R ~ cs(Fl) + cs(A), data = rent, family = GA)
GAMLSS-RS iteration 1: Global Deviance = 27920.7
GAMLSS-RS iteration 2: Global Deviance = 27920.74
GAMLSS-RS iteration 3: Global Deviance = 27920.74
> op <- par(mfrow = c(2, 1))
> term.plot(r1, se = TRUE, partial = TRUE)
```

Cubic splines (additive plots): rent data

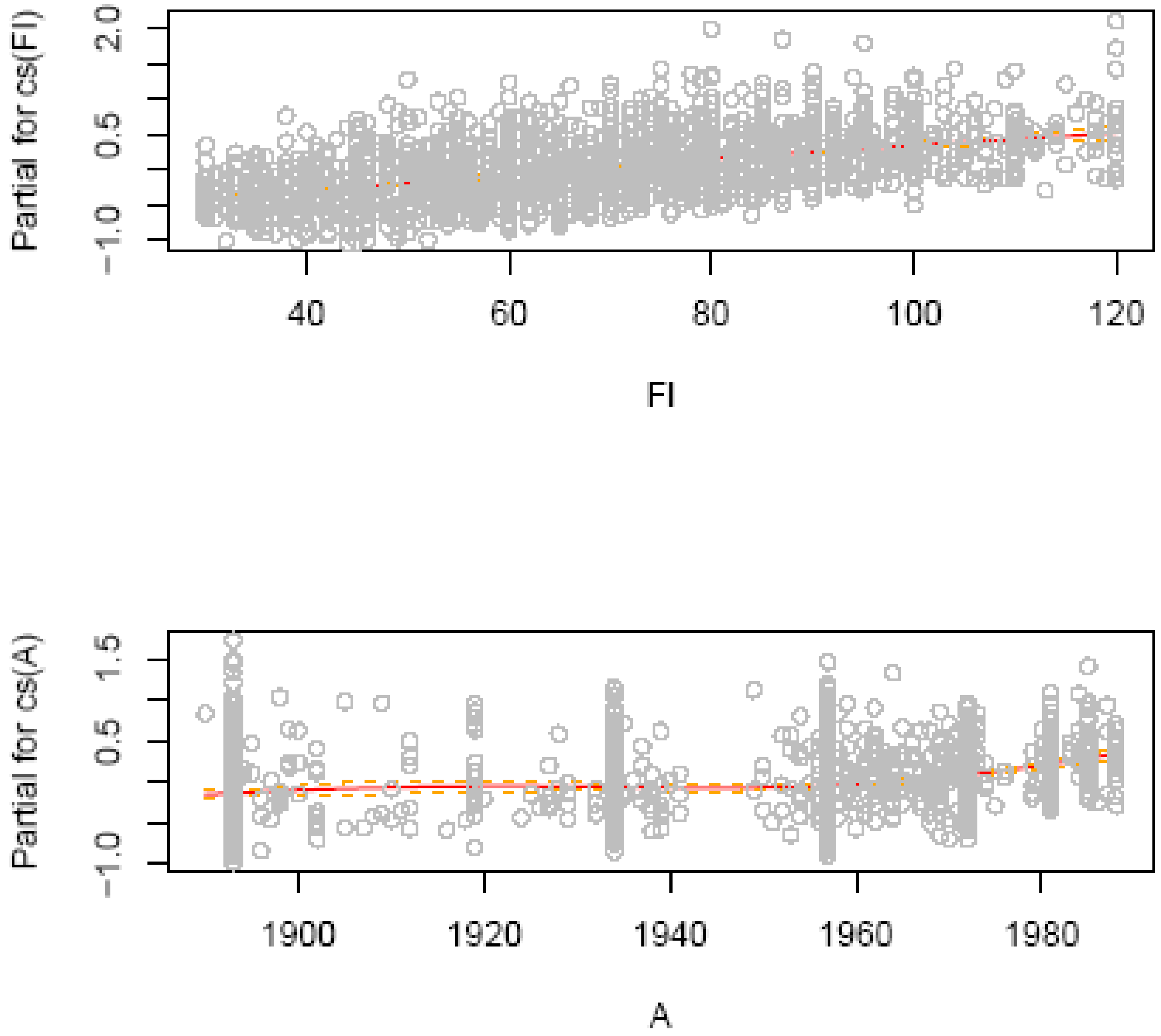


Figure 6.3: Rent data: Additive plots for the cubic splines

Cubic splines (contour plot): rent data

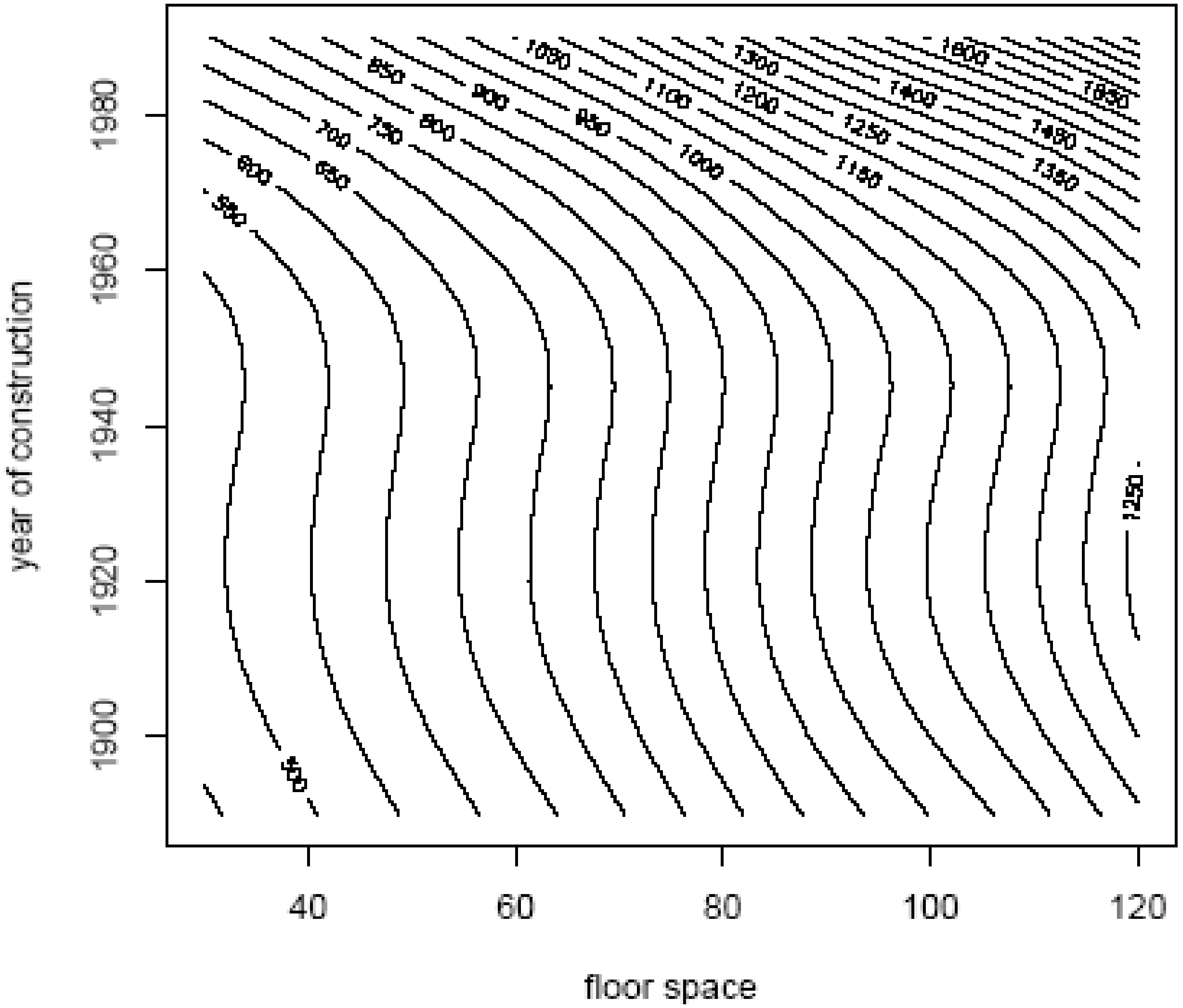


Figure 6.4: Rent data: contour plot for the fitted additive cubic splines model

Cubic splines (3d plot): rent data

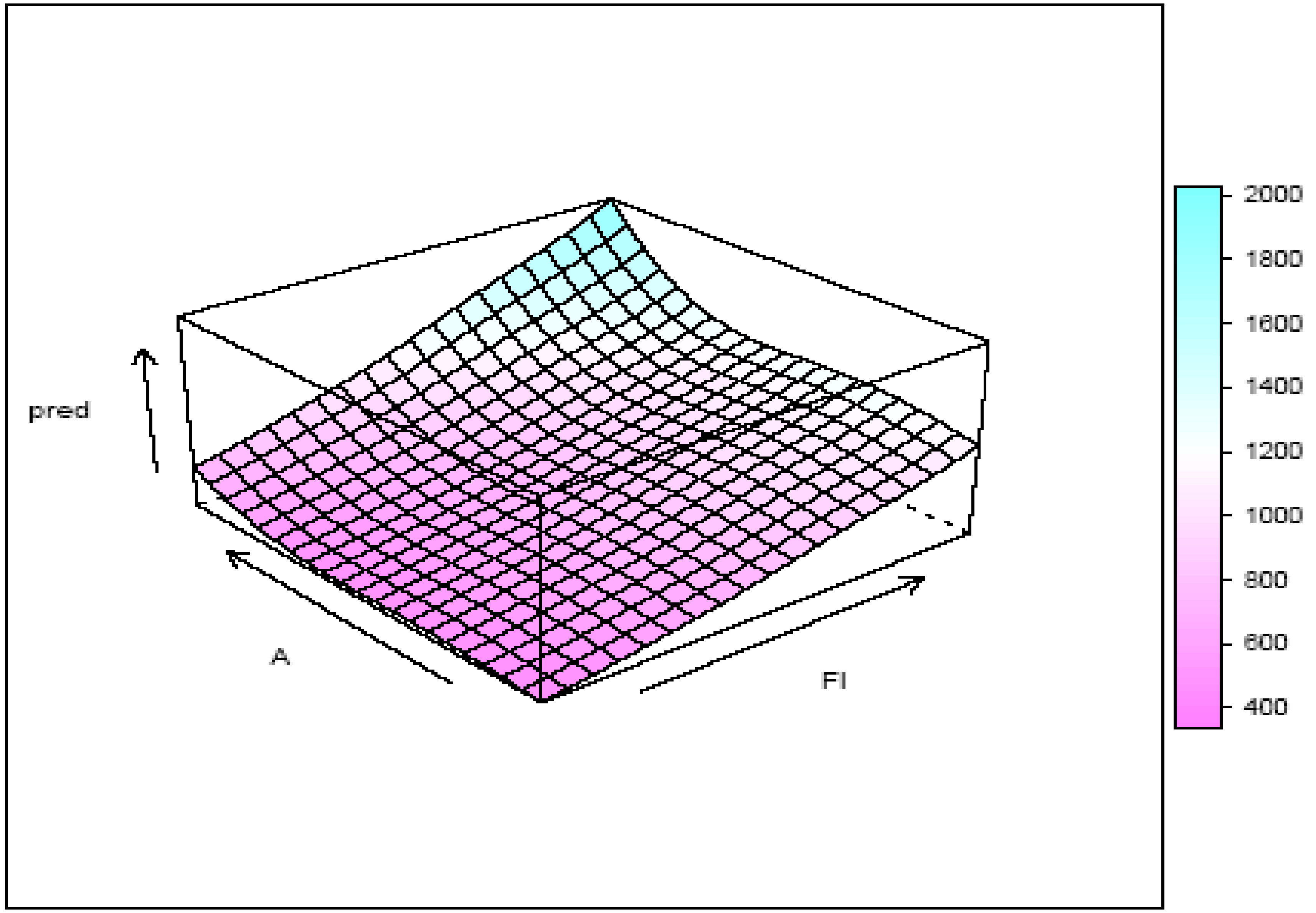


Figure 6.5: Rent data: surface plot for the fitted additive cubic splines model

loess: the rent data, plots

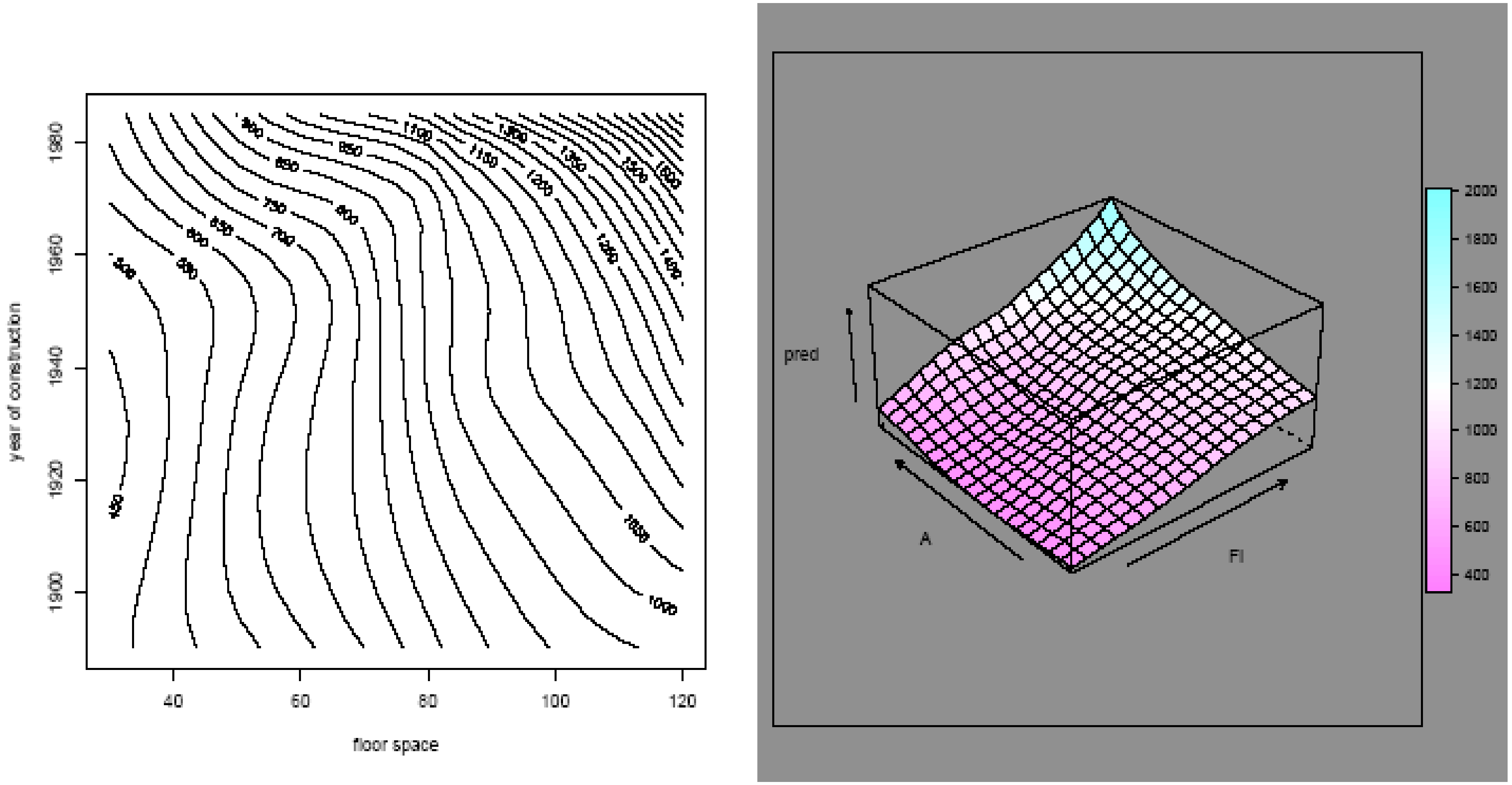


Figure 6.10: Rent data: contour and surface plot for the fitted loess surface model

7 Centile estimation

9.1 Example

9.2 Modelling the distribution

9.3 Modelling the parameters

9.4 Model selection

9.5 Modelling head circumference against age

9.6 Model diagnostics

9.7 Conclusions

7.1 Introduction to centile estimation

7.1.1 Data example

Variables

Head circumference (HEAD)

against age,

for 7040 males under 22 years,

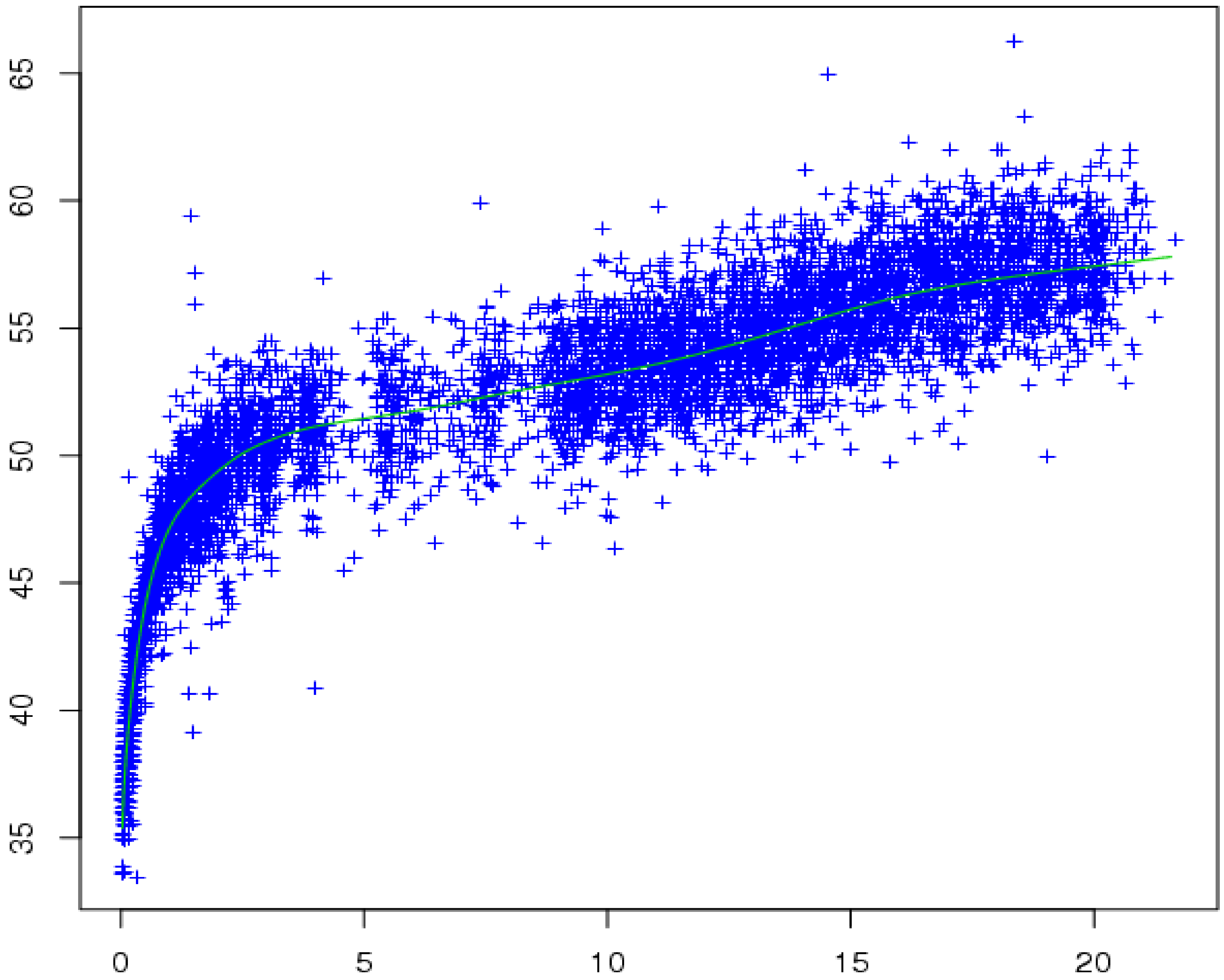
Study

cross sectional study,

'The Fourth Dutch Growth Study',

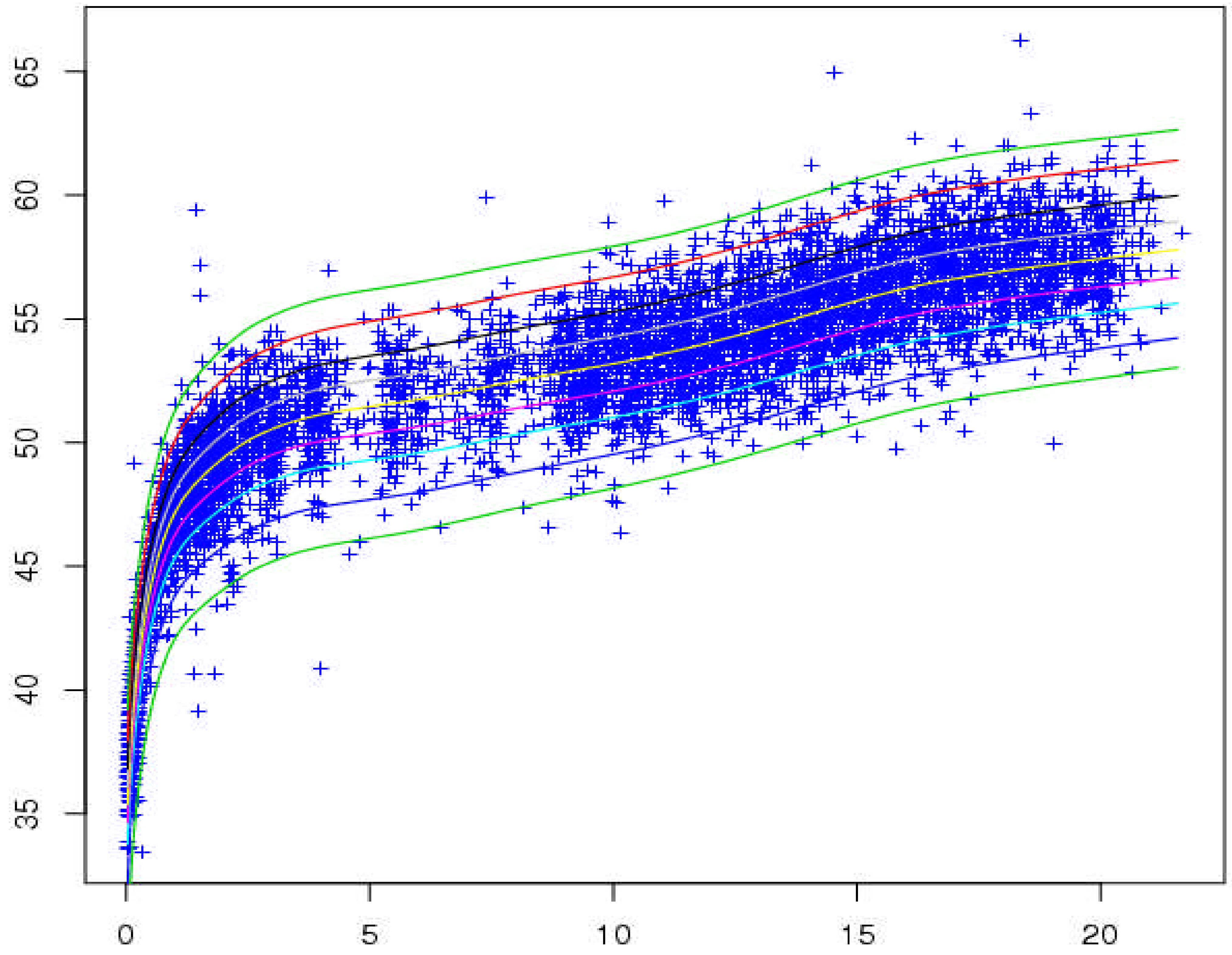
Fredriks *et al.* (2000)

HEAD against AGE for Dutch males



7.1.2 Objective

To obtain centile curves of HEAD against AGE



9.3 Model for smooth centile curves

$Y \sim D(\mu, \sigma, \nu, \tau)$ where D is any distribution, and

where $Y = HEAD$ and $x = AGE^{\xi}$ and

$$\mu = cs(x, df_{\mu})$$

$$log(\sigma) = cs(x, df_{\sigma})$$

$$\nu = cs(x, df_{\nu})$$

$$log(\tau) = cs(x, df_{\tau})$$

9.4 Automatic model selection of degrees of freedom and ξ

We need to select the five values $df_\mu, df_\sigma, df_\nu, df_\tau, \xi$

An automatic procedure is used to select the values which minimize the generalized Akaike information criterion GAIC

GAIC(#) = Deviance + #.df

where # is a penalty for each degree of freedom used in the model and df is the total degrees of freedom used in the model, Akaike (1983)

Choosing the penalty

Akaike information criterion AIC uses #=2

Schwartz Bayesian Criterion SBC uses #=log(n)=8.9

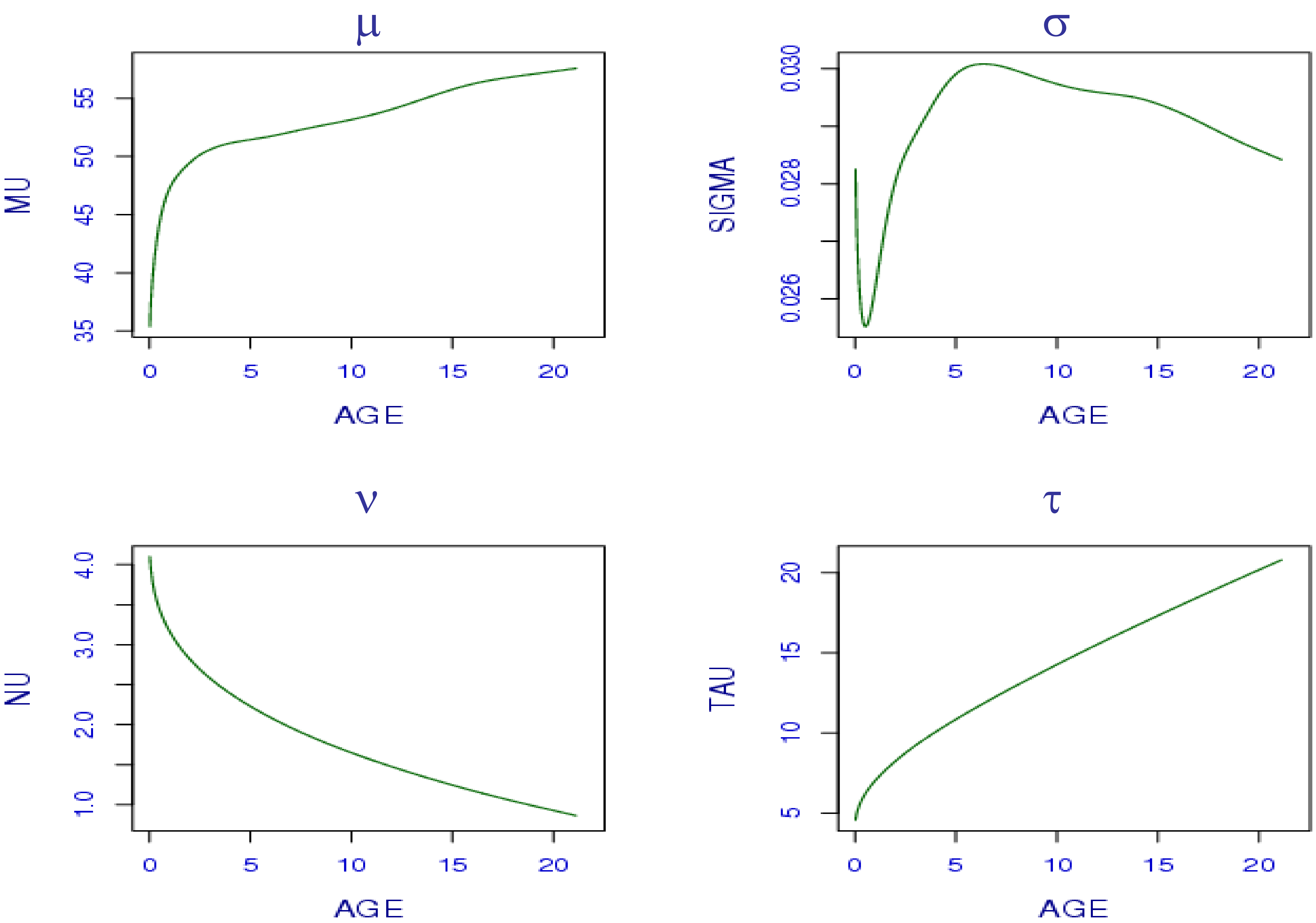
Alternative values of # can be used, e.g. 3.

Higher penalty # $\Rightarrow$ smoother but biased centiles.

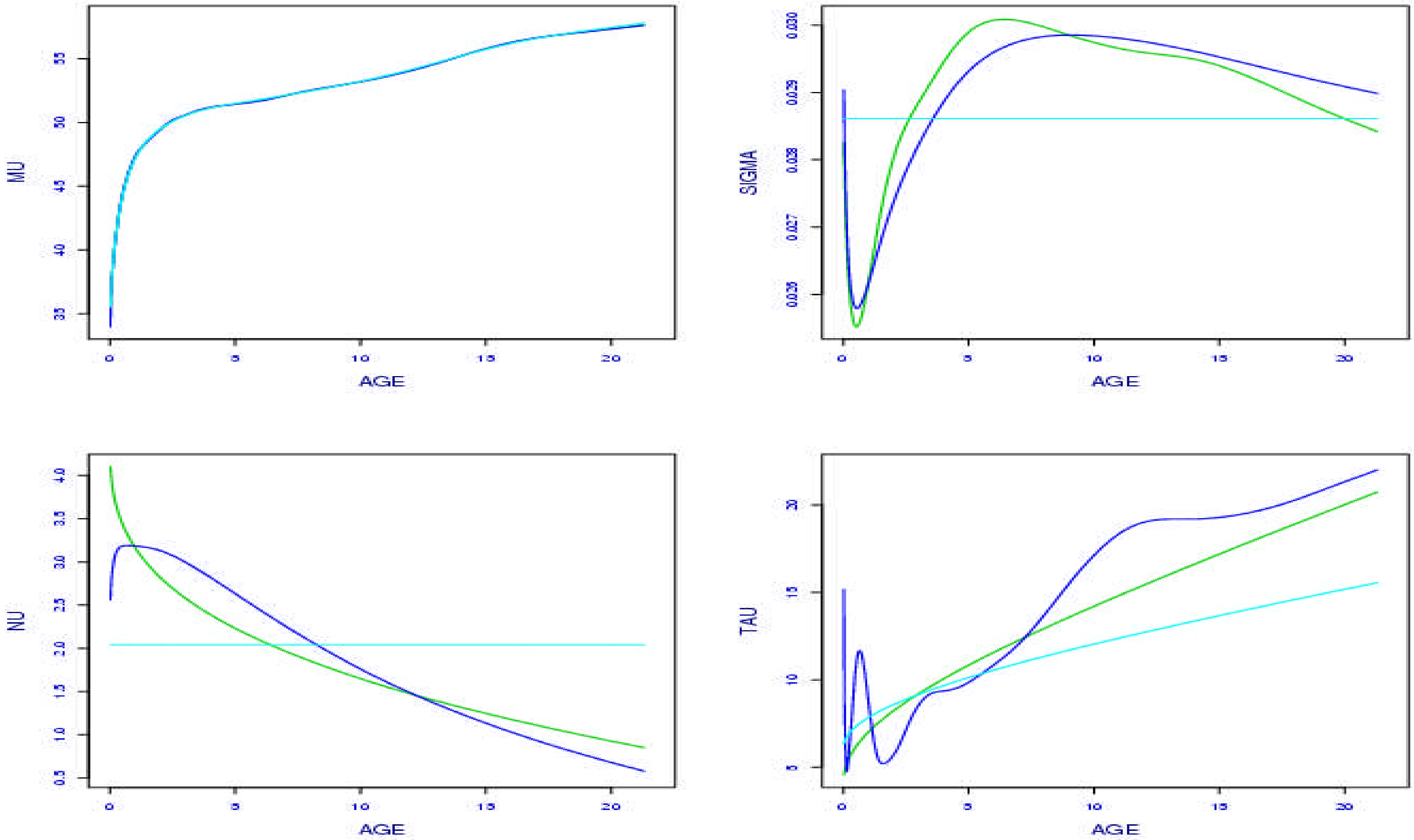
Lower penalty # $\Rightarrow$ rougher but less biased centiles.

7.5.2 Fitted parameters μ, σ, υ, τ against AGE

(for BCT model chosen with penalty # = 3)



Comparison of fitted models for # = 2, 3, 8.9

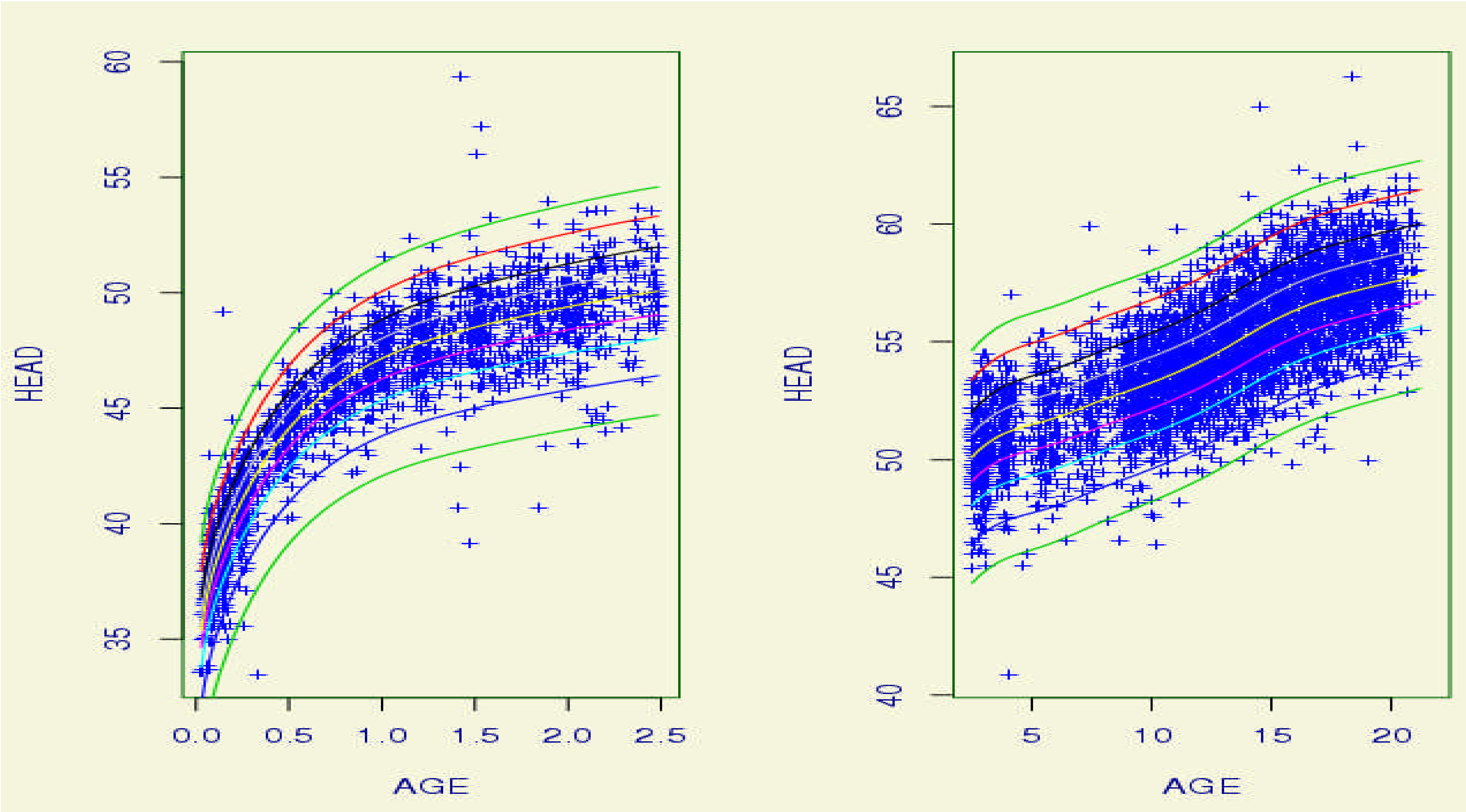


7.5.3 Choosing the distribution

distribution	GAIC(3)-26814.7	dfμ	dfσ	dfν	dfτ	ξ
NO	221.6	16.4	30	-	-	0.001
BCCG (LMS)	172.9	16.7	20	14.7	-	0.01
BCPE (LMSP)	81.7	12.2	7.9	2	2	0.34
SEP	71.7	11.7	3.7	2	2	0.40
TF	4.8	13.1	2.9	-	3.1	0.27
JSU	3.4	11.7	3.4	2	2	0.46
BCT (LMST)	0	12.3	5.7	2	2	0.33

Centiles for BCT model chosen with # = 3

(0.4, 2, 10, 25, 50, 75, 90, 98, 99.6) %



References

Akaike, H. (1983) Information measures and model selection. *Bulletin of the International Statistical Institute*, **50**, 277-290.

Box, G. E. P. and Tiao, G. C. (1962) A further look at robustness via Bayes's theorem. *Biometrika*, **49**, 419-432.

Cole, T. J. (1988) Fitting smoothed centile curves to reference data. *Journal of the Royal Statistical Society, Series A*, **151**, 385-418.

Cole, T. J. and Green, P. J. (1992) Smoothing reference centile curves: the LMS method and penalized likelihood. *Statistics in Medicine*, **11**, 1305-1319.

Fredriks, A. M. *et al.* (2000) Body index measurements in 1996-7 compared with 1980. *Archives of Childhood*, **82** , 107-112.

Hansen, B. (1994) Autoregressive conditional density estimation. *International Economic Review*, **35**, 705-730.

Heagerty, P. J. and Pepe, M. S. (1999) Semiparametric estimation of regression quantiles with application to standardizing weight for height and age in US children. *Applied Statistics*, **48**, 533-551.

Healey, M. J. R., Rasbash, J. and Yang, M. (1988) Distribution free estimation of age-related centiles. *Annals of Human Biology*, **15**, 17-22.

Johnson, N. L. (1949) Systems of frequency curves generated by methods of translation. *Biometrika*, **36**, 149-176.

Koenker, R. W. and Bassett, G. W. (1978) Regression quantiles. *Econometrica*, **46**, 33-50.

Lange, K. L., Little, R. J. A. and Taylor, J. M. G. (1989) Robust statistical modelling using the *t* distribution. *Journal of the American Statistical Association*, **84**, 881-896.

Rigby, R. A. and Stasinopoulos, D. M. (2004a) Smooth centile curves for skew and kurtotic data modelled using the Box-Cox Power Exponential distribution. (to appear in *Statistics and Medicine*).

References

gamlss

for statistical modelling

gamlss

for statistical modelling