Data: Respiratory deaths of children under five. Study conducted in São Paulo, Brazil, 1994-1997.

In a period of four years, the daily respiratory deaths of children under five in the city of São Paulo were observed. As explanatory variables, the humidity, the temperature, and a variety of pollutant concentrations were daily recorded. In detail, the data set contains the following variables:

TEMPO	Enumeration of days	
DIASEM	Week day	1 : Monday 2 : Tuesday 3 : Wednesday 4 : Thursday 5 : Friday 6 : Saturday 7 : Sunday
SEGUNDA	Indicator for Monday	1 : Monday 0 : not Monday
TERCA, QUARTA, QUINTA, SEXTA, SABADO	Analogous indicators for	Tuesday, Wednesday, Thursday, Friday, Saturday
OTHRES5	Number of other death causes than respiratory.	
TMIN	Minimum temperature in C.	
TMIN.2	Two-day-lag of TMIN.	
UMID	Relative humidity in %.	
PMME	Concentration of particular matter $\leq 10\mu m (PM_{10})$ in $\mu g/m^3$.	
PMME.2	Two-day-lag of PMME.	
SO2ME	Concentration of SO_2 in $\mu g/m^3$.	
SO2ME.2	Two-day-lag of SO2ME.	
COME	Concentration of CO in ppm.	
COME.2	Two-day-lag of COME.	
O3ME	Concentration of O_3 in $\mu g/m^3$.	
O3ME.2	Two-day-lag of O3ME.	
RES5	Number of respiratory deaths.	

Sample Size: 1461

Model: Generalized Additive Model

Dependent Variable: RES5 (Type: Count Data.)

Predictor:

- "Core-Model", as e.g. in Singer et al. (2002):

$$\begin{aligned}\eta &= \beta_0 + f_1(\text{TEMPO}) + f_2(\text{TMIN.2}) + f_3(\text{UMID}) + \\ &+ \beta_1 \cdot \text{SEGUNDA} + \ldots + \beta_6 \cdot \text{SABADO} + \beta_7 \cdot \text{OTHRES5}\end{aligned}$$

- To evaluate the influence of the pollutants, they are subsequently added to the core-model, usually with a two-day lag (Singer et al., 2002).

$$\tilde{\eta} = \eta + f_4(\text{SO2ME.2}).$$

Example:

Assume for illustration that the simple model

$$ln[E(\text{RES5})] = f(\text{TEMPO})$$

is fitted, where the function f is obtained with splines. One obtains a fitted curve for $RES5$ over time as in Figure 1 (The monotone curve is robustified against outliers in the predictor space).

Link: $g(\mu) = \ln(\mu)$, i.e. $h(\eta) = \exp(\eta)$.

Literature working with this data:

Conceição, G.M.S., Miraglia, S.G.E.K., Kishi, H.S., Saldiva, P.H.N. and Singer, J.M. (2001). Air pollution and children mortality: a time series study in São Paulo, Brazil. *Environ Health Perspect*, **109**, 347-350, *http://ehpnet1.niehs.nih.gov/members/ 2001/suppl-3/347-350conceicao/conceicao-full.html*

Singer, J.M., de André, C.D.S., Lima, L.P., and Conceição, G.M.S. (2002). Association between atmospheric pollution and mortality in São Paulo, Brazil: regression models and analysis strategy. In Y. Dodge (Ed.), *Statistical Data Analysis based on the L1 norm an related methods*, pp 439–450. Birkhäuser, Berlin.

Einbeck, J., de André, C.D.S., and Singer, J.M. (2004). Local smoothing with robustness against outlying predictors. *Environmetrics* **15, 541–554**.

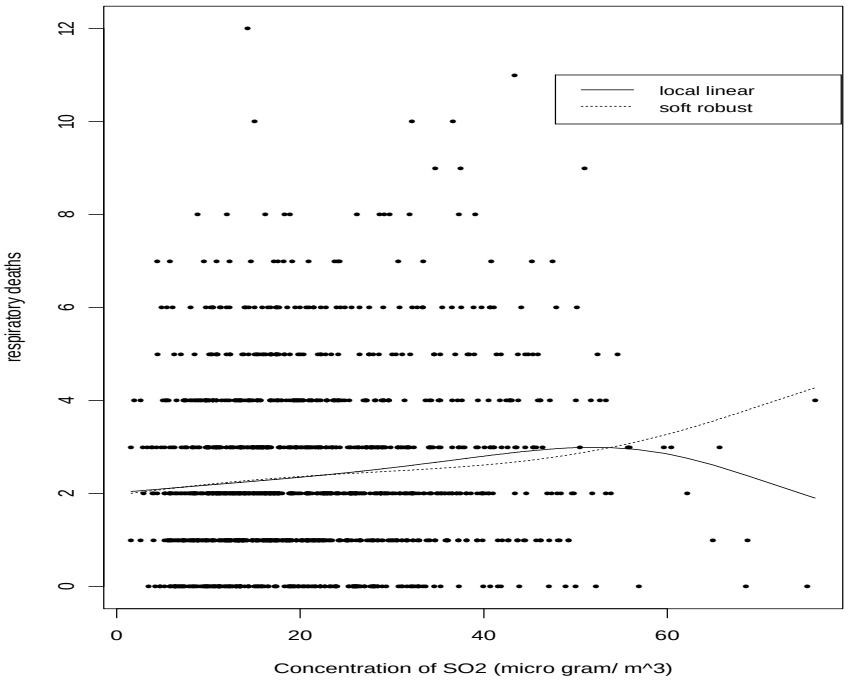

Figure 1: Local linear and soft robustified smoother through respiratory data.