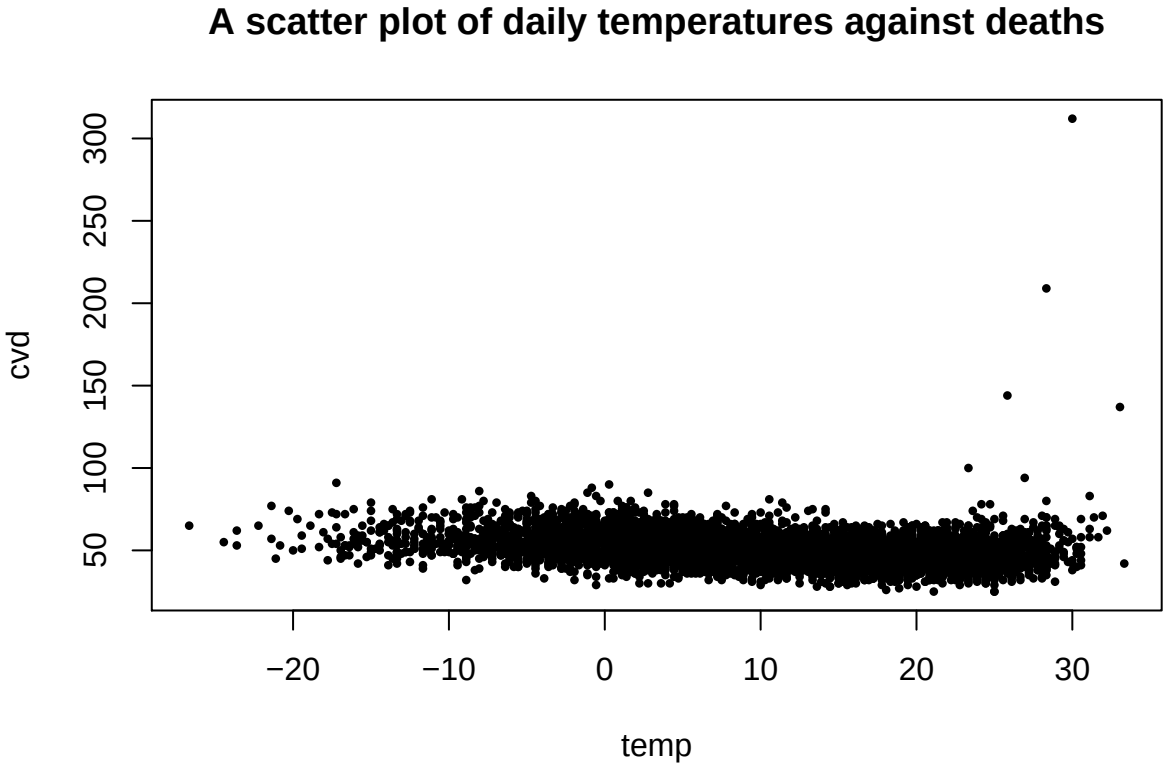
Reproducible report example

Ivan C. Hanigan

Some exploratory analysis

In this section we do some exploratory analysis of the NMMAPS data for deaths in Chicago 1987-2000. The code, messages and intermediary results are hidden in the resulting report document.

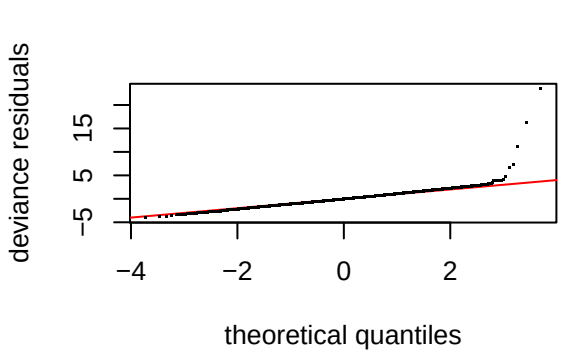
We made a simple scatter plot shown below

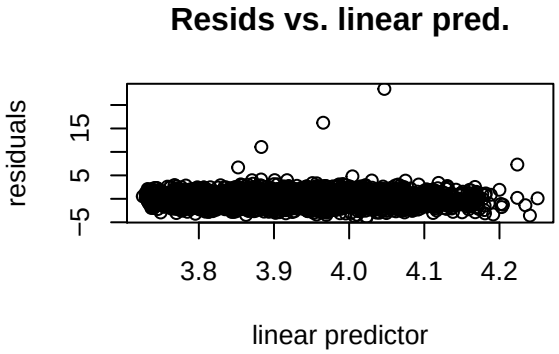


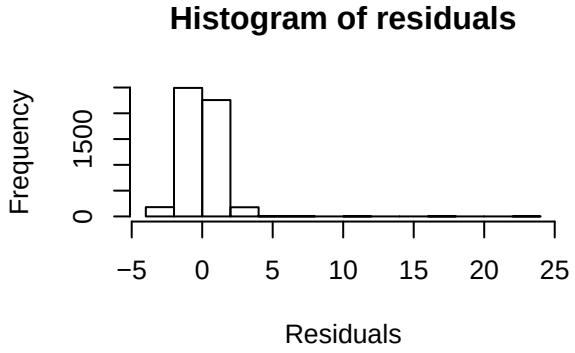
We ran some exploratory models. A Poisson GAM with smooth functions on temperature and time was compared to a linear fit on temperature.

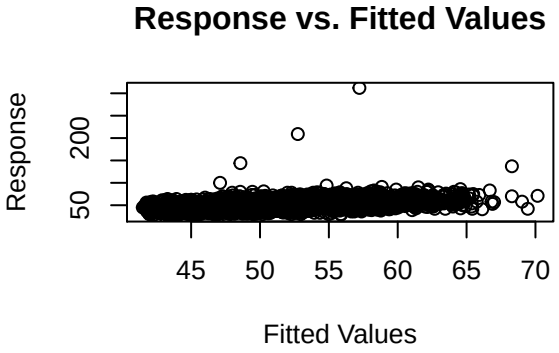
```
##
## Family: poisson
## Link function: log
##
## Formula:
## cvd ~ s(temp) + s(time)
##
## Parametric coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) 3.925191   0.001969    1993   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df Chi.sq p-value
```

```
## s(temp) 8.474  8.901   1289  <2e-16 ***
## s(time) 8.719  8.977   1098  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.228   Deviance explained = 25.5%
## UBRE = 0.43229  Scale est. = 1         n = 5114
```

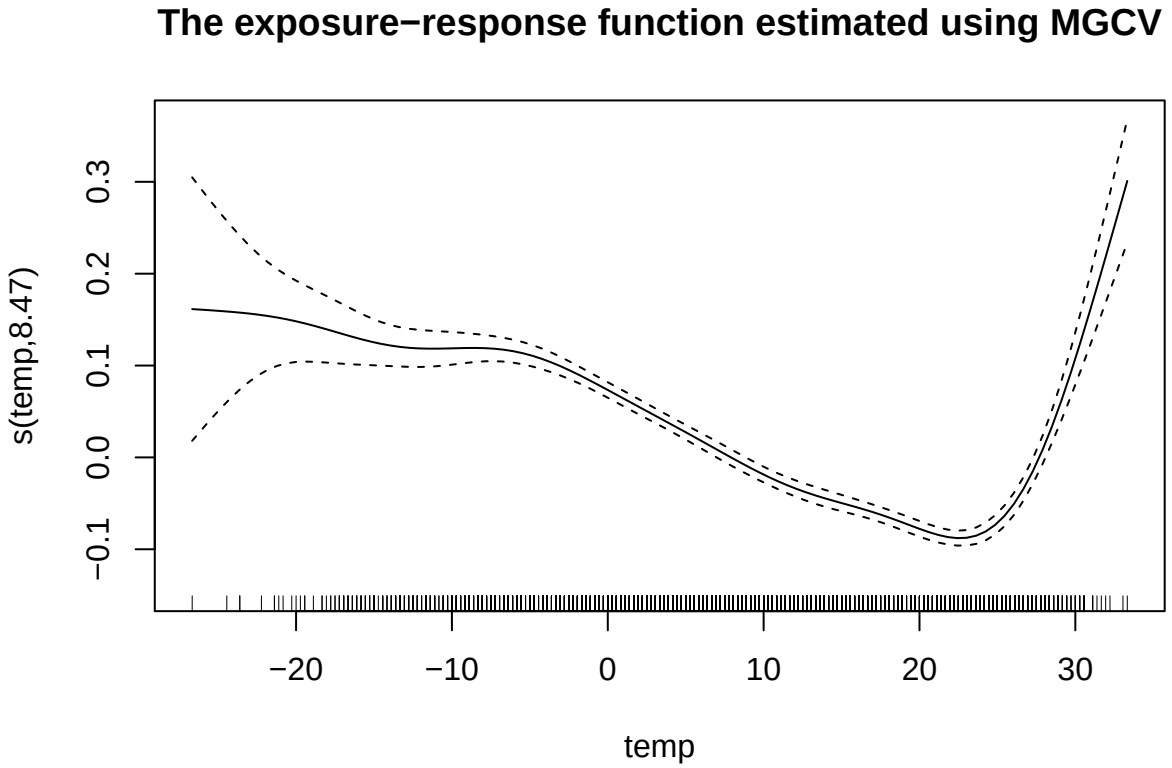









```
##
## Method: UBRE   Optimizer: outer newton
## full convergence after 6 iterations.
## Gradient range [4.417538e-09,2.033826e-07]
## (score 0.4322868 & scale 1).
## Hessian positive definite, eigenvalue range [9.239887e-05,0.0001626577].
## Model rank =  19 / 19
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##            k'   edf k-index p-value
## s(temp) 9.000 8.474   0.994    0.38
## s(time) 9.000 8.719   0.795    0.00
```



The result can be automatically inserted to the text. This model has a delta AIC of -289.8 (smoothed minus linear term).