GAM for modelling excess mortality of infectious diseases

Mikael Franko, SLU

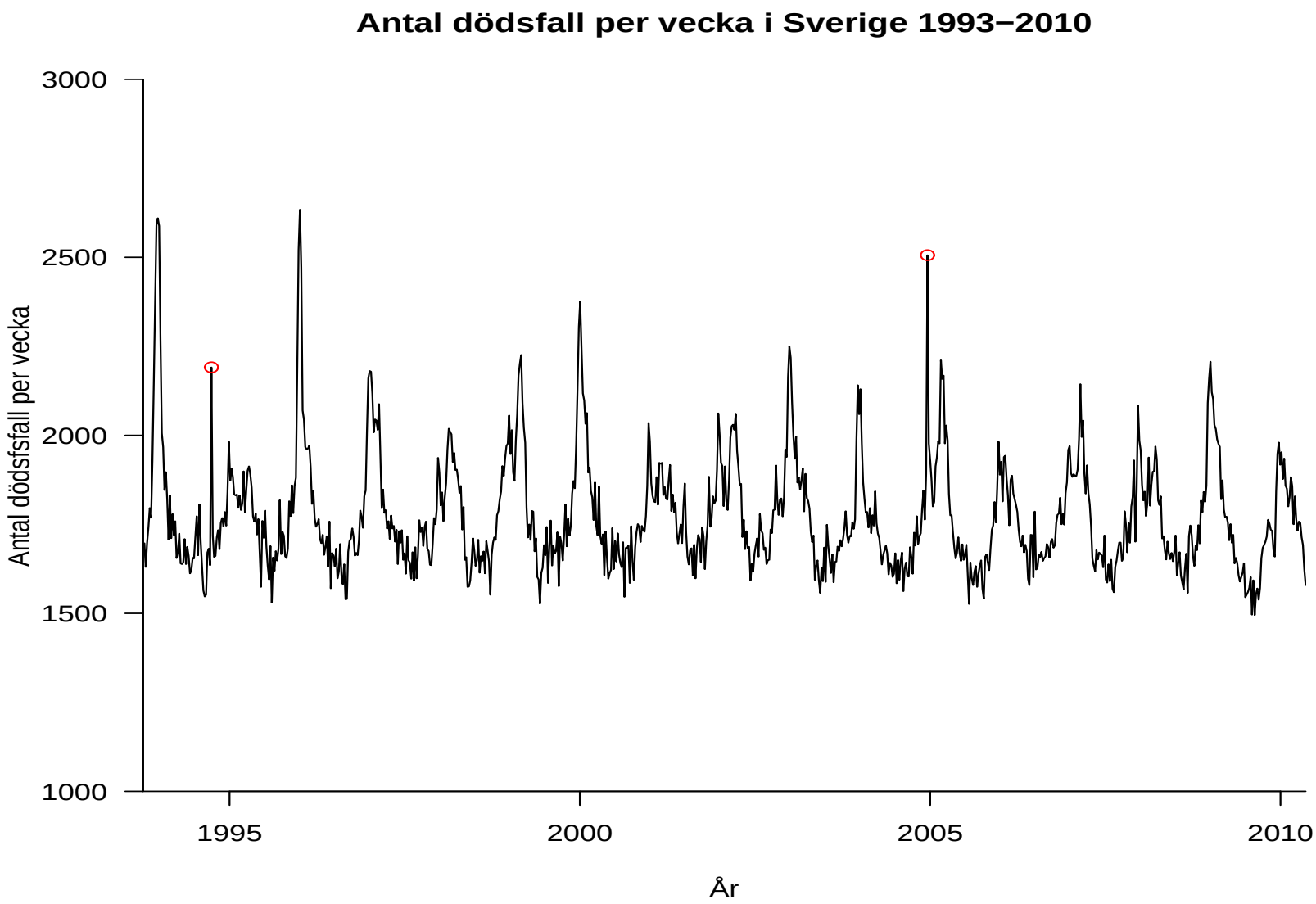
Katarina Brus Sjölander, Kjell-Olof Hedlund, Sandra Rubinovà and Annika Linde, Swedish Institute for Infectious Disease Control

23 August 2016

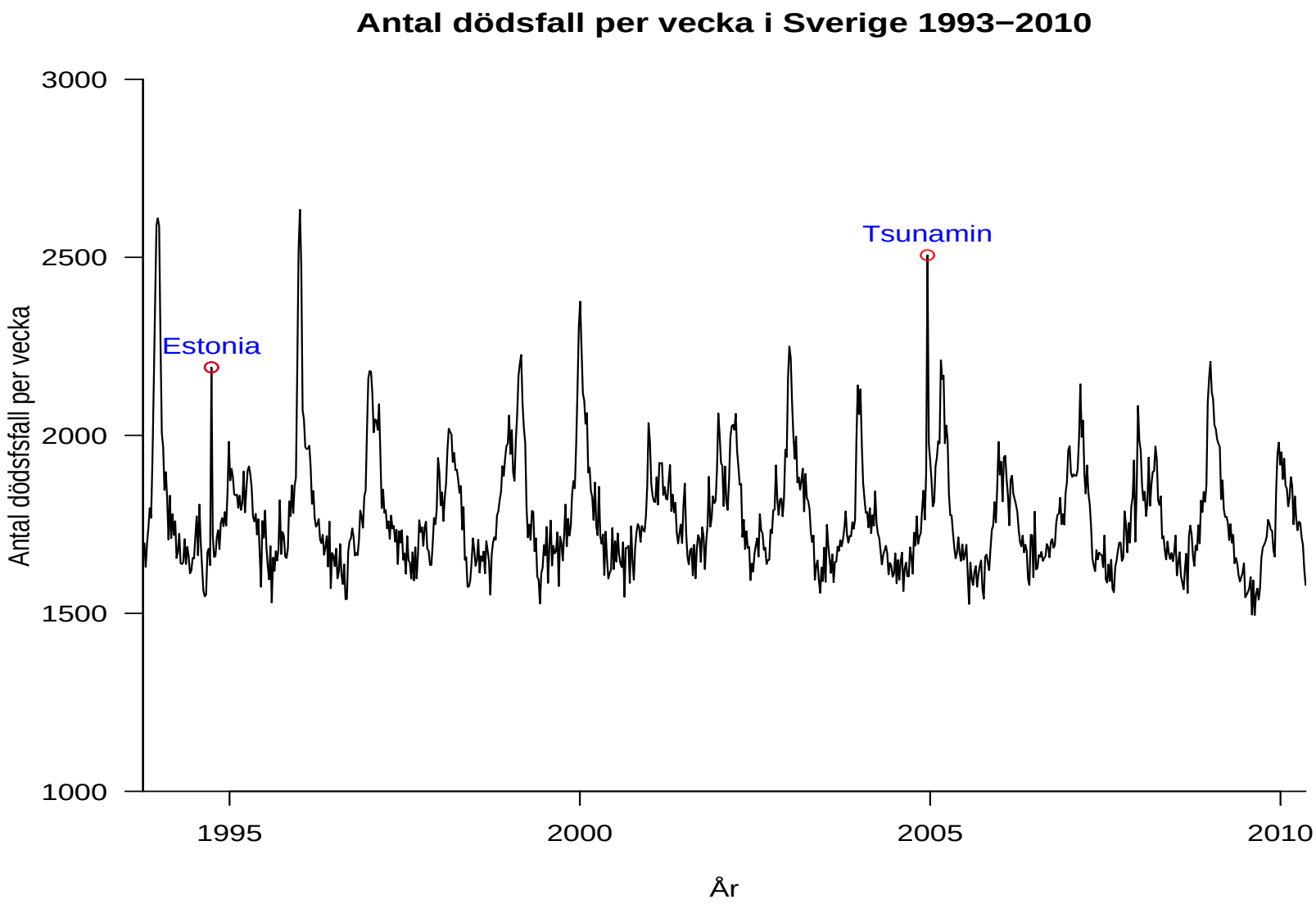
Background

- In the mid 19th century, it was observed that mortality increased in a population during severe influenza seasons.
- Even if influenza is rarely a direct cause of death it can still cause serious complications like pneumonia and meningitis, which can be lethal especially among the elderly.
- The concept *excess mortality* was introduced in the 1960s as the difference between observed mortality and estimated mortality in the absence of influenza (baseline mortality).
- Baseline mortality is often estimated using regression and time series models.

Background

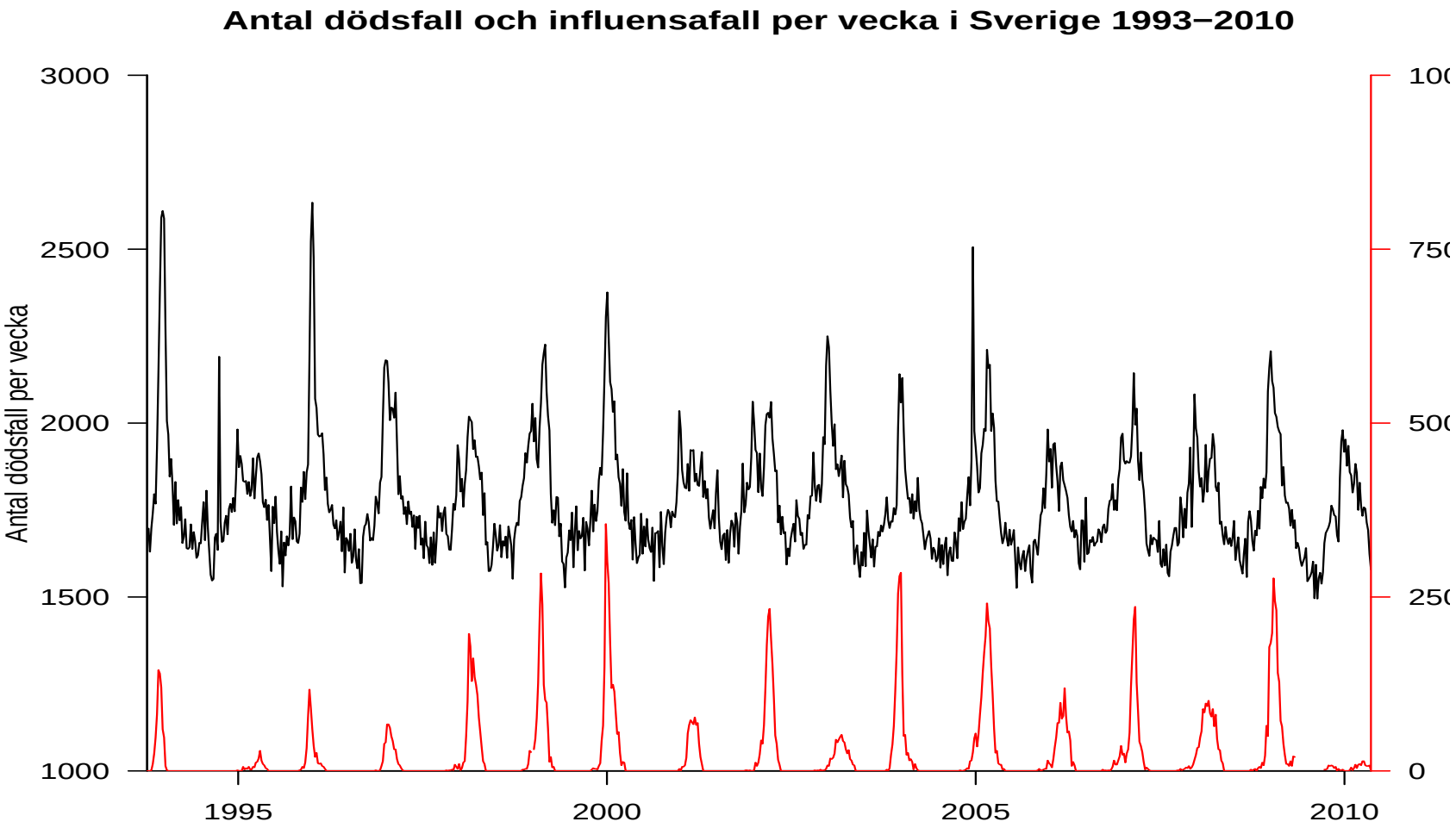

Antal dödsfall per vecka i Sverige 1993–2010
Antal dödsfsfall per vecka
3000
2500
2000
1500
1000
1995
2000
2005
2010
År

Background


Antal dödsfall per vecka i Sverige 1993–2010
Antal dödsfall per vecka
3000
2500
2000
1500
1000
Estonia
Tsunamin
1995
2000
2005
2010
År

Background

All laboratory verified cases of influenza has been reported to the Swedish Institute for Infectious Disease Control every week from week number 40 to week number 20 since 1993.



Serfling model

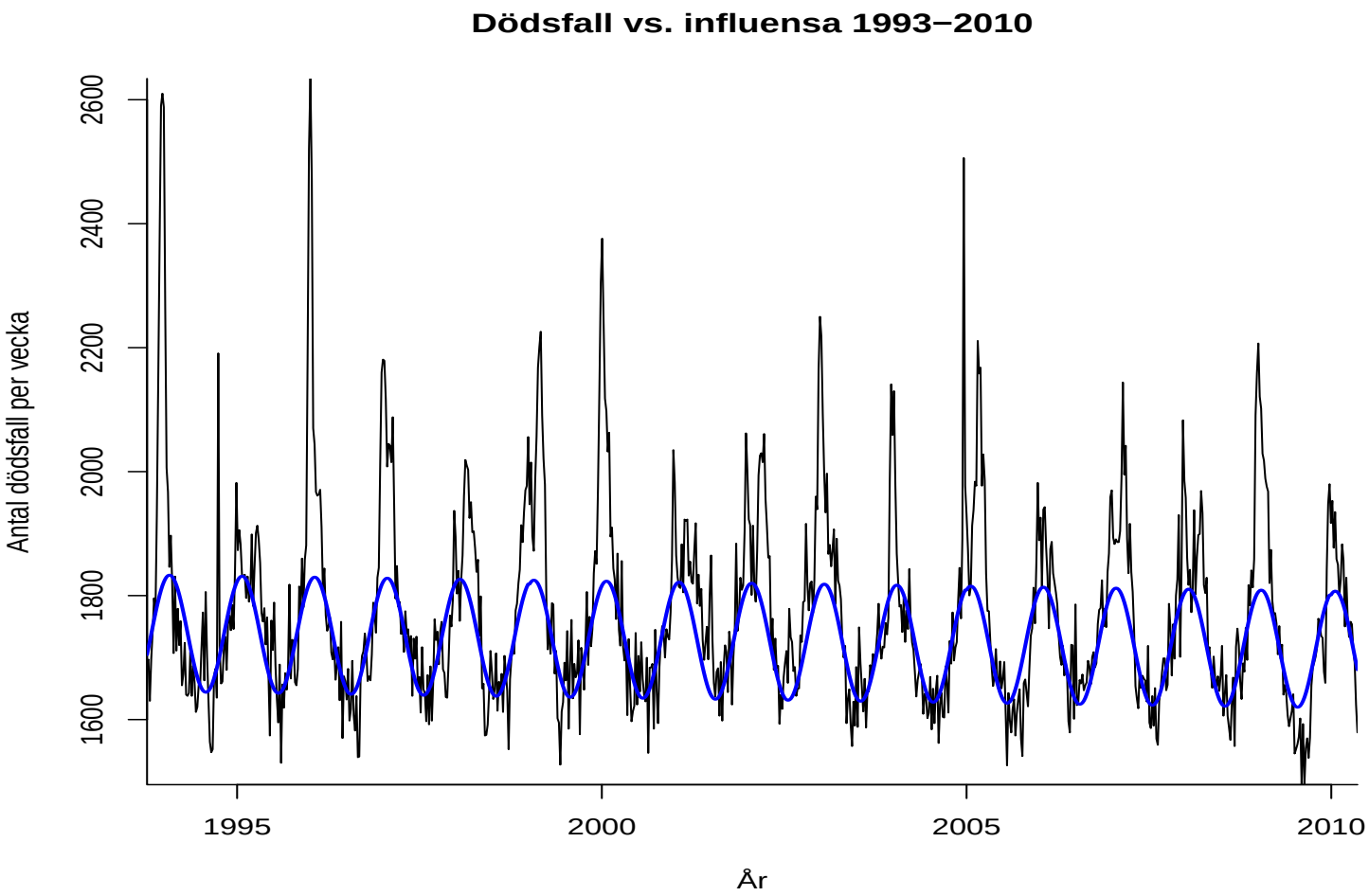
Serfling (1963) assumed that baseline mortality can be described as

$$y_t = \beta_0 + \beta_1 t + \beta_2 \sin\left(2\pi \frac{w_t}{52}\right) + \beta_3 \cos\left(2\pi \frac{w_t}{52}\right) + \varepsilon_t$$

where y_t is the number of deaths in week t with number w_t and ε_t are normally distributed autocorrelated error terms.

The parameters β_0, β_1, β_2 and β_3 are estimated using influenza free weeks.

Serfling model



Excess mortality: 3300 deaths per season

Serfling model

Pros

- Simple!
- Detailed infection data are not needed.

Cons

- Harmonic functions does not capture real mortality very well.
- All causes of excess mortality are indistinguishable.

Regression I

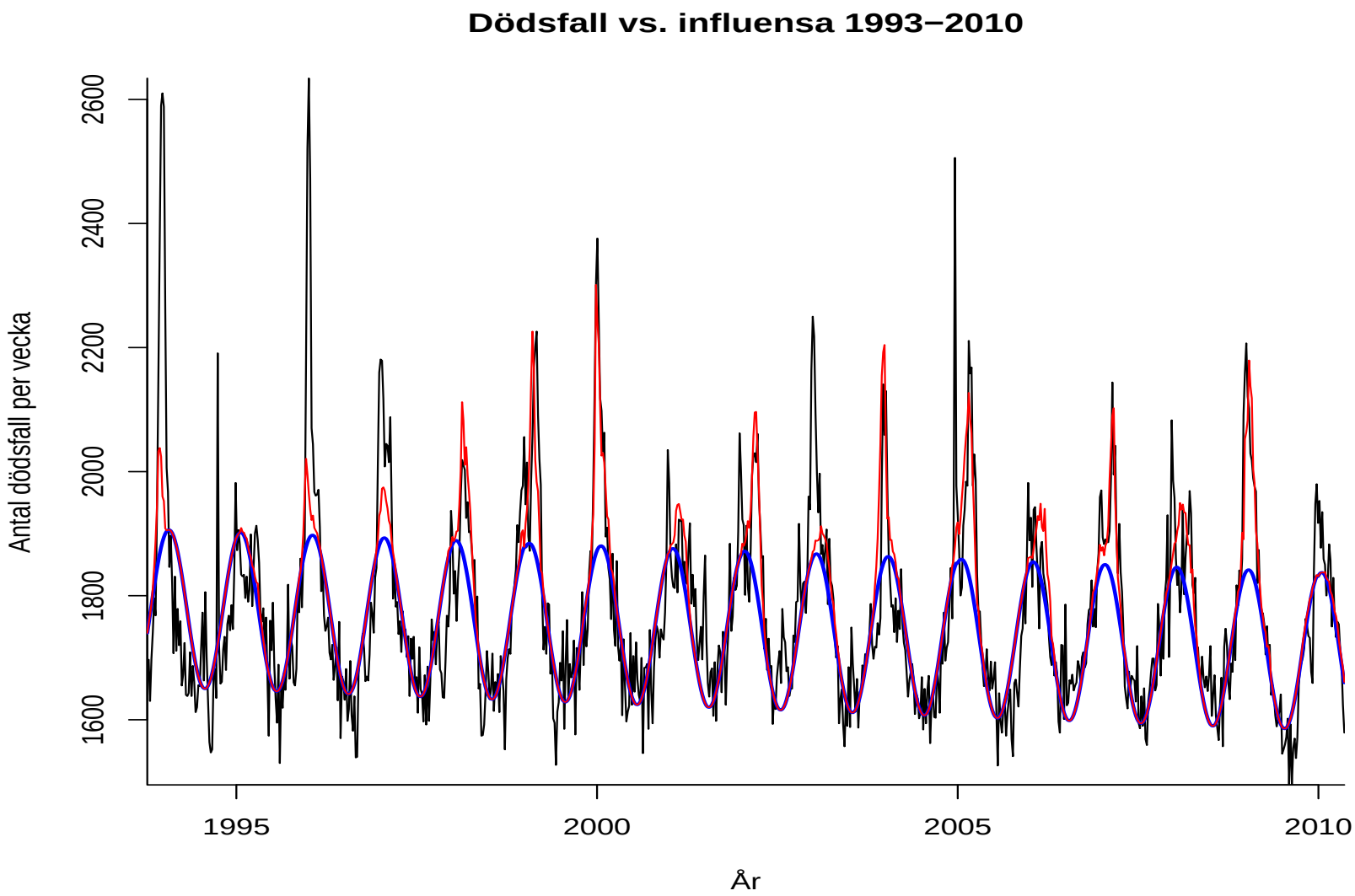
The Serfling model can be extended to incorporate influenza data.

$$y_t = \beta_0 + \beta_1 t + \beta_2 \sin\left(2\pi \frac{w_t}{52}\right) + \beta_3 \cos\left(2\pi \frac{w_t}{52}\right) + \beta_4 x_t + \varepsilon_t$$

where x_t is the number of reported influenza cases in week t.

Parameters are estimated using the whole time series.

Regression I



Excess mortality: 2650 deaths per season.

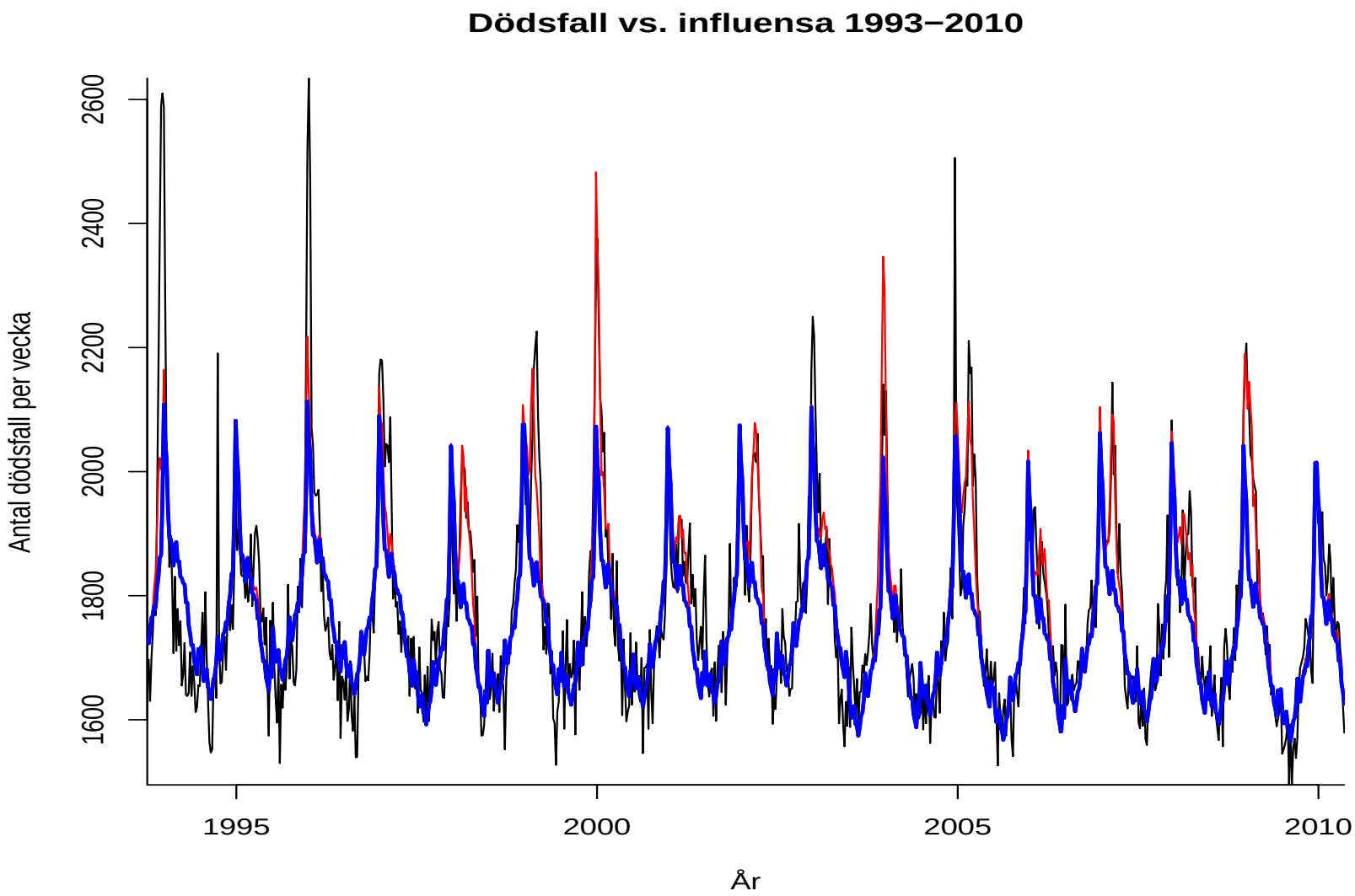
Regression II

Categorical variables for trend and seasonal variation.

$$y_t = \beta_{0s_t} + \beta_{1w_t} + \beta_2 x_t + \varepsilon_t$$

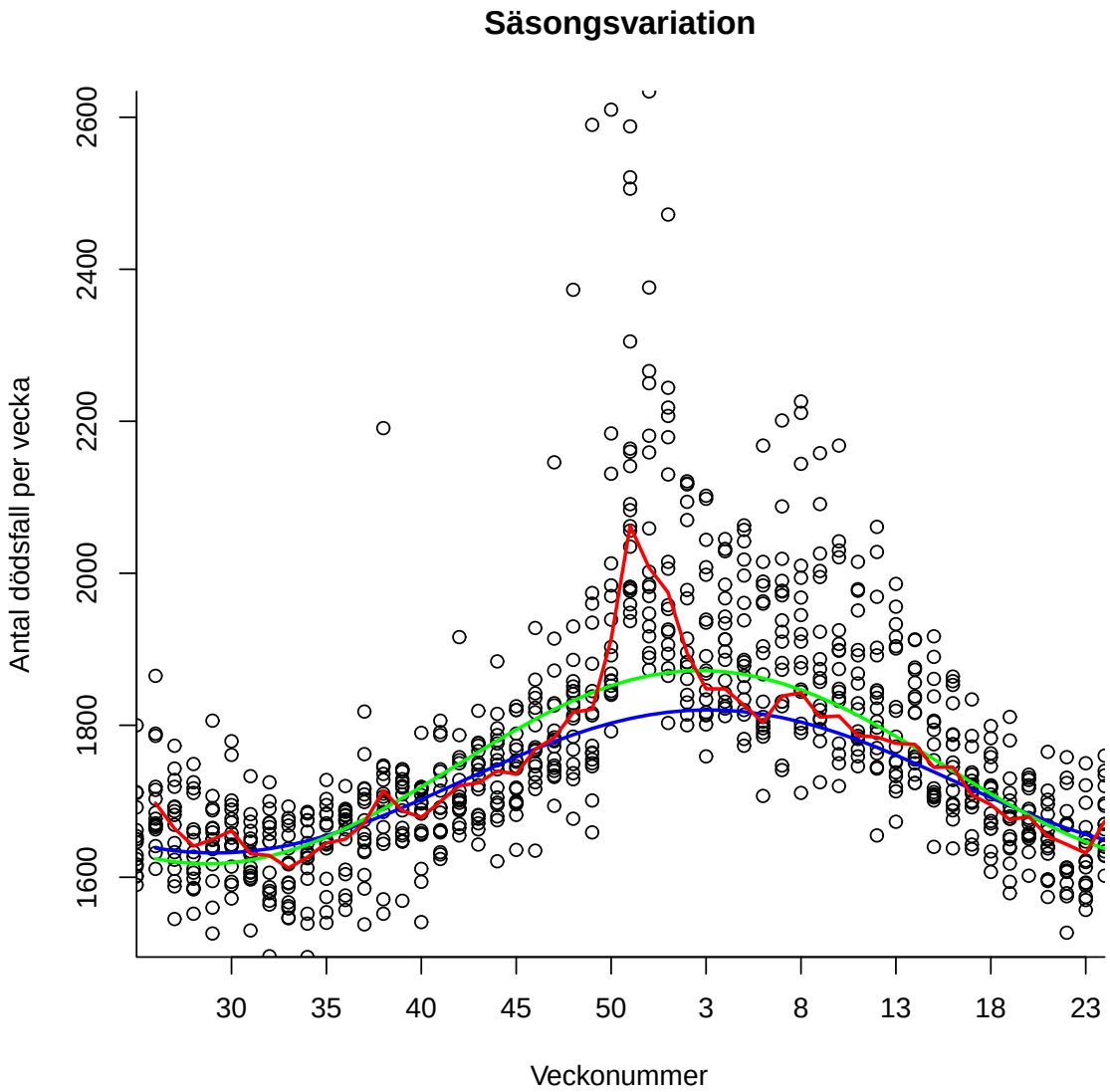
where s_t is season for week t.

Regression II



Excess mortality: 2400 deaths per season

Seasonal variation

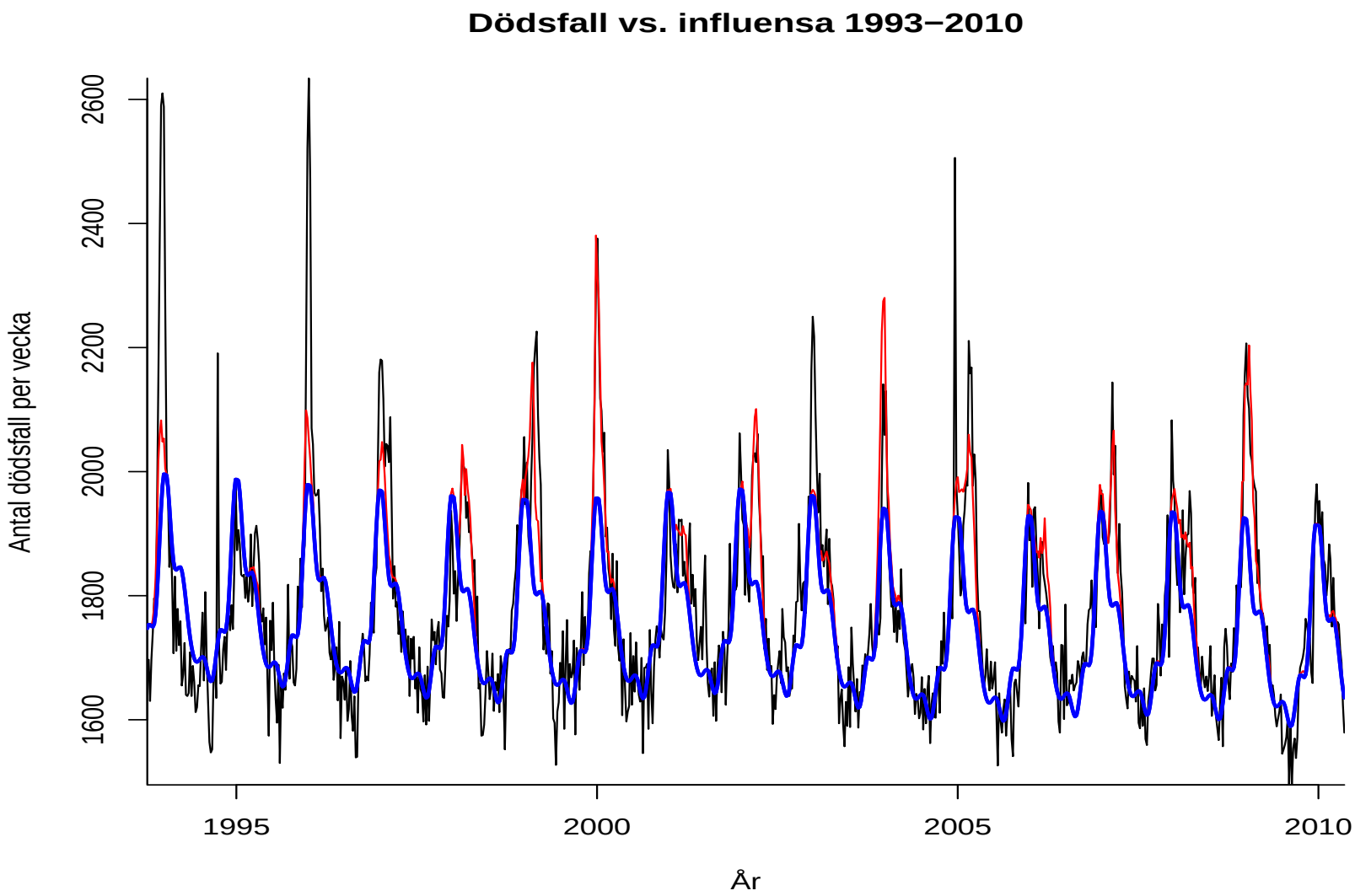


GAM model

$$y_t = \beta_0 + s_1(t) + s_2(w_t) + \beta_1 x_t + \varepsilon_t$$

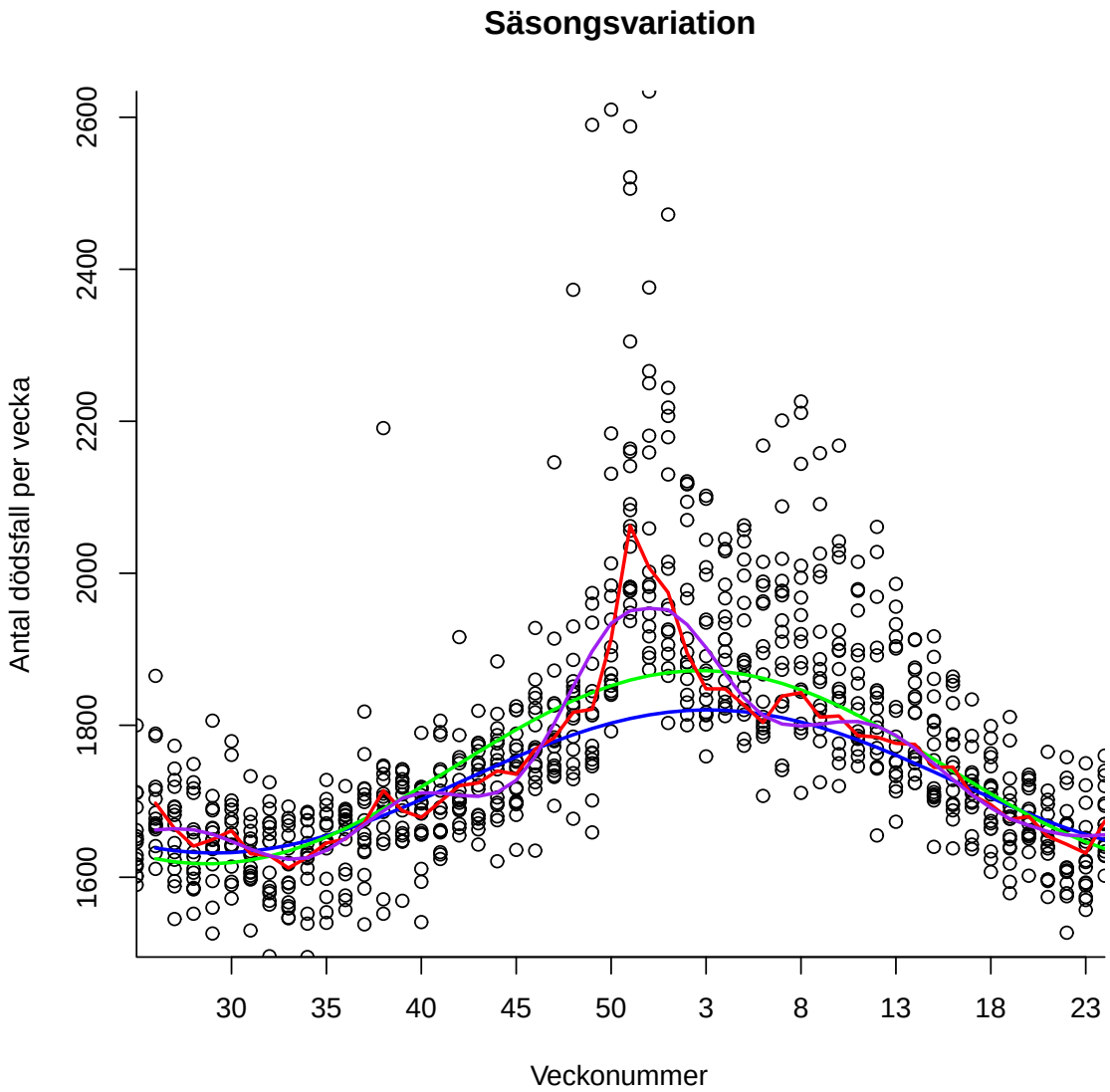
where s_1 and s_2 are arbitrary smooth spline functions.

GAM model



Excess mortality: 2500 deaths per season

Seasonal variation



Excess mortality with respect to influenza, RS-virus and norovirus

- RS-virus (Respiratory Syncytial Virus) is an infection of the respiratory tract that mostly affects infants and young children. It can cause bronchiolitis and pneumonia in infants younger than one year.
- Norovirus causes gastroenteritis with severe vomiting (vinterkräksjukan) that can lead to dehydration and weakened immune system, especially among the elderly.

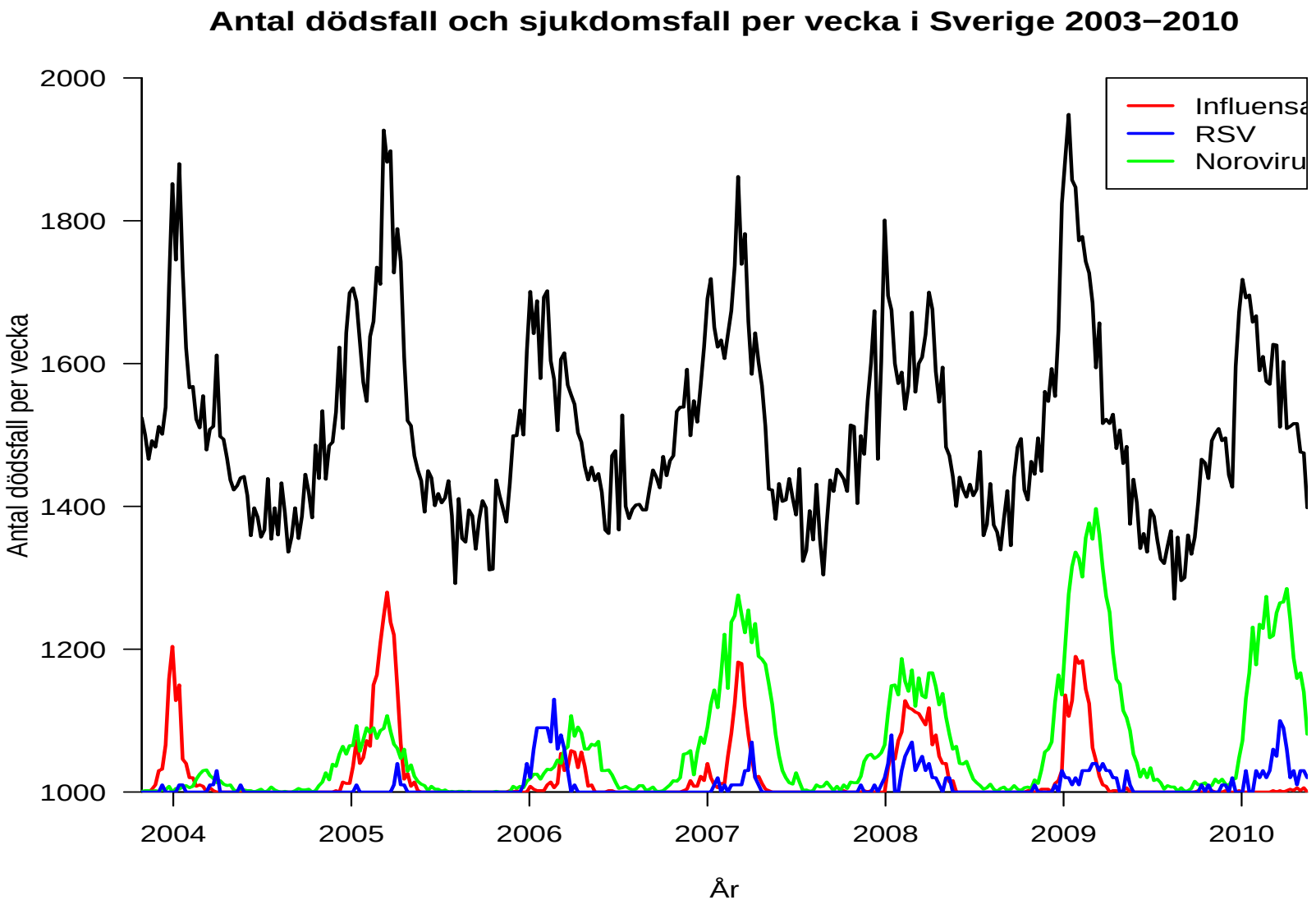
Excess mortality with respect to influenza, RS-virus and norovirus

The following data were used in this study:

1. The number of deaths per week
2. The number of laboratory verified cases of influenza per week
3. The number of laboratory verified cases of RS-virus per week
4. The number of laboratory verified cases of norovirus per week

only for those older than 65 years during the period from week 43 in 2003 to week 19 in 2010.
Population weighted mean temperature in Sweden was also included.

Excess mortality with respect to influenza, RS-virus and norovirus



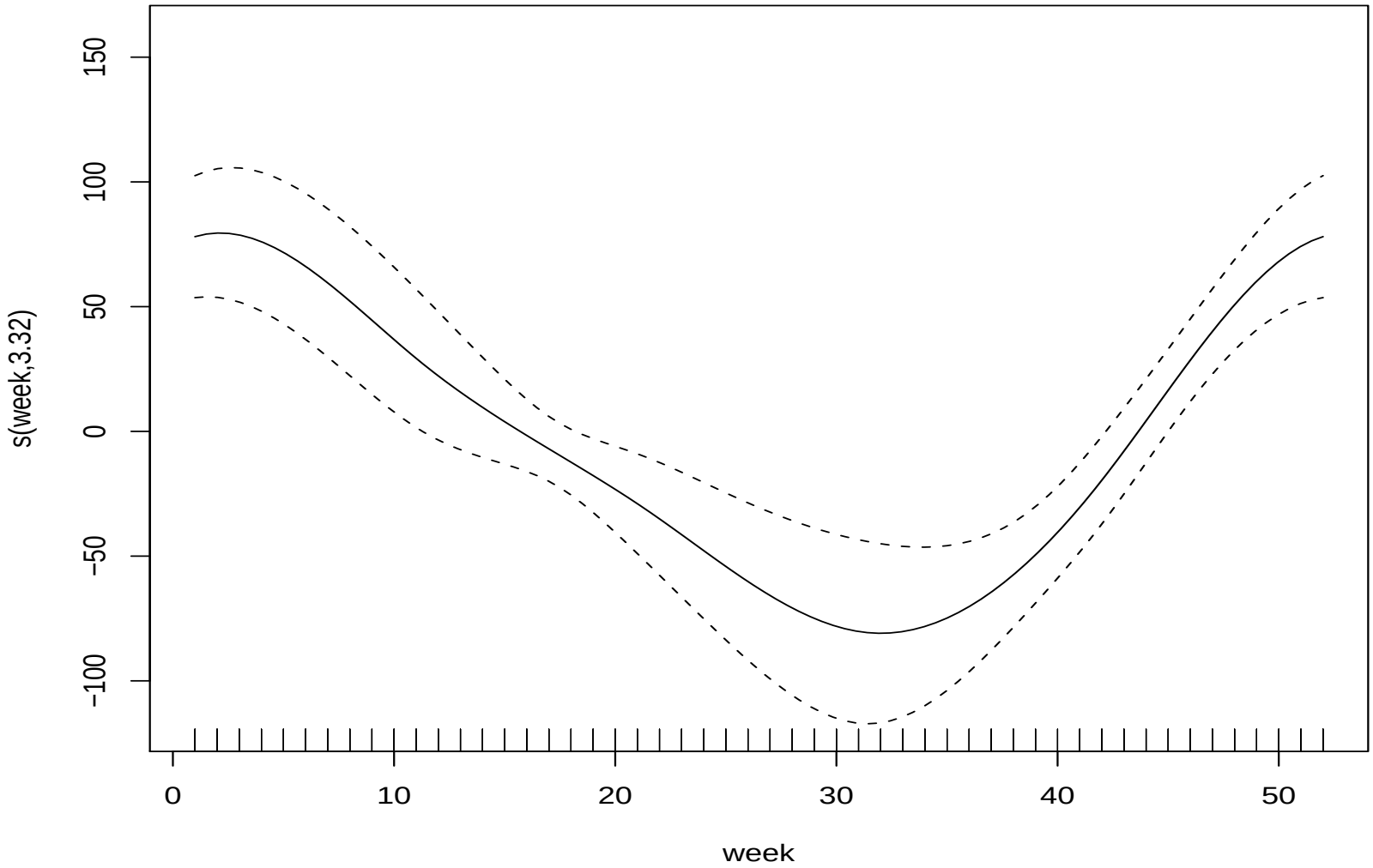
GAM model

$$
\begin{aligned}
y_t &\sim \text{Po}(\mu_t + \varepsilon_t) \\
\mu_t &= \beta_0 + s_1(t) + s_2(w_t) + \beta_{1s_t} x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + s_3(z_t)
\end{aligned}
$$

where y_t is the number of deaths, x_{1t}, x_{2t} and x_{3t} are the number of cases of influenza, RS-virus and norovirus, respectively, z_t is mean temperature, w_t is week number, s_t is season and s_1, s_2 and s_3 are arbitrary smooth spline functions. Autocorrelated errors were also included, partly to account for overdispersion.

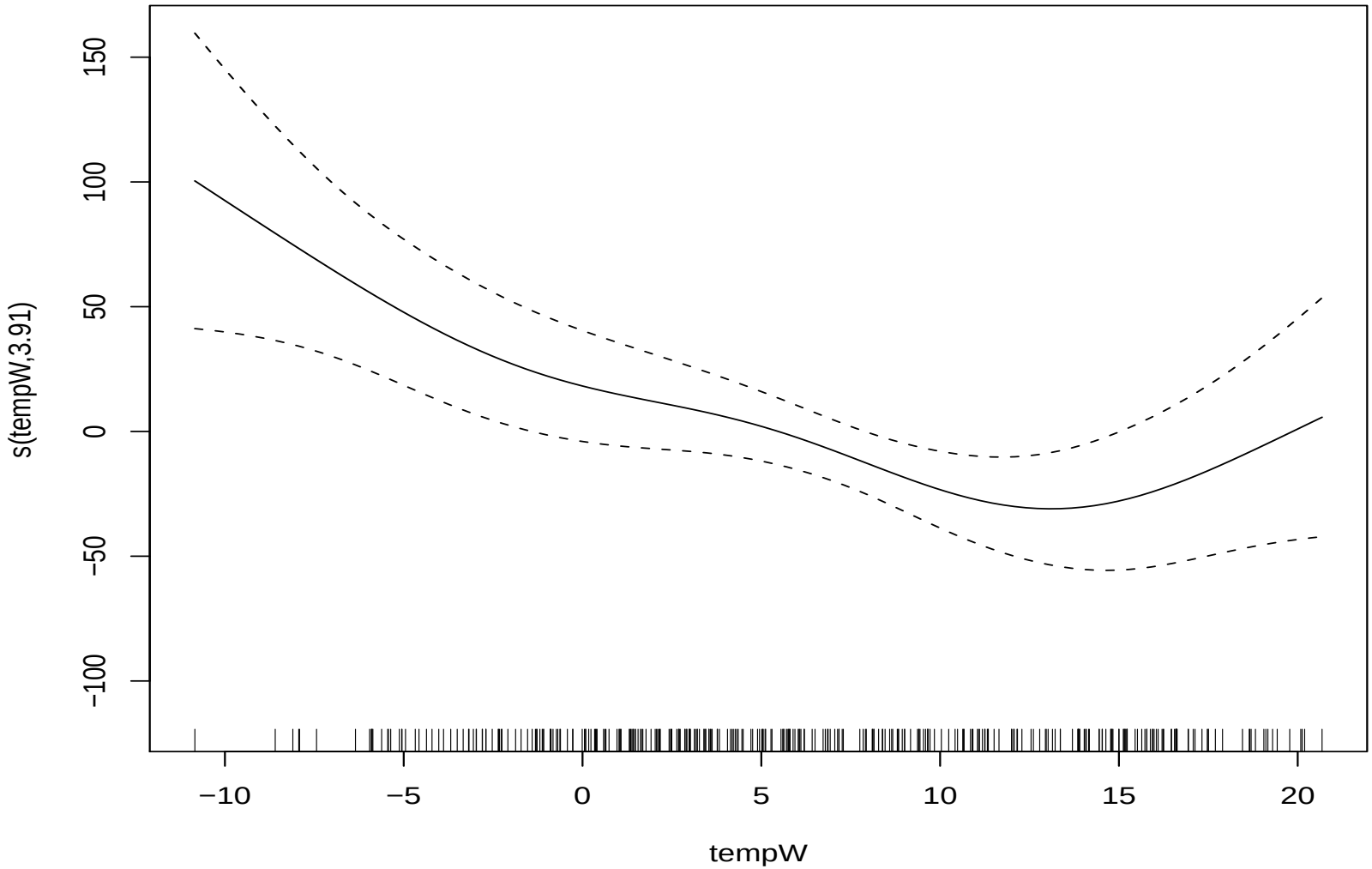
GAM model

Spline function with respect to week

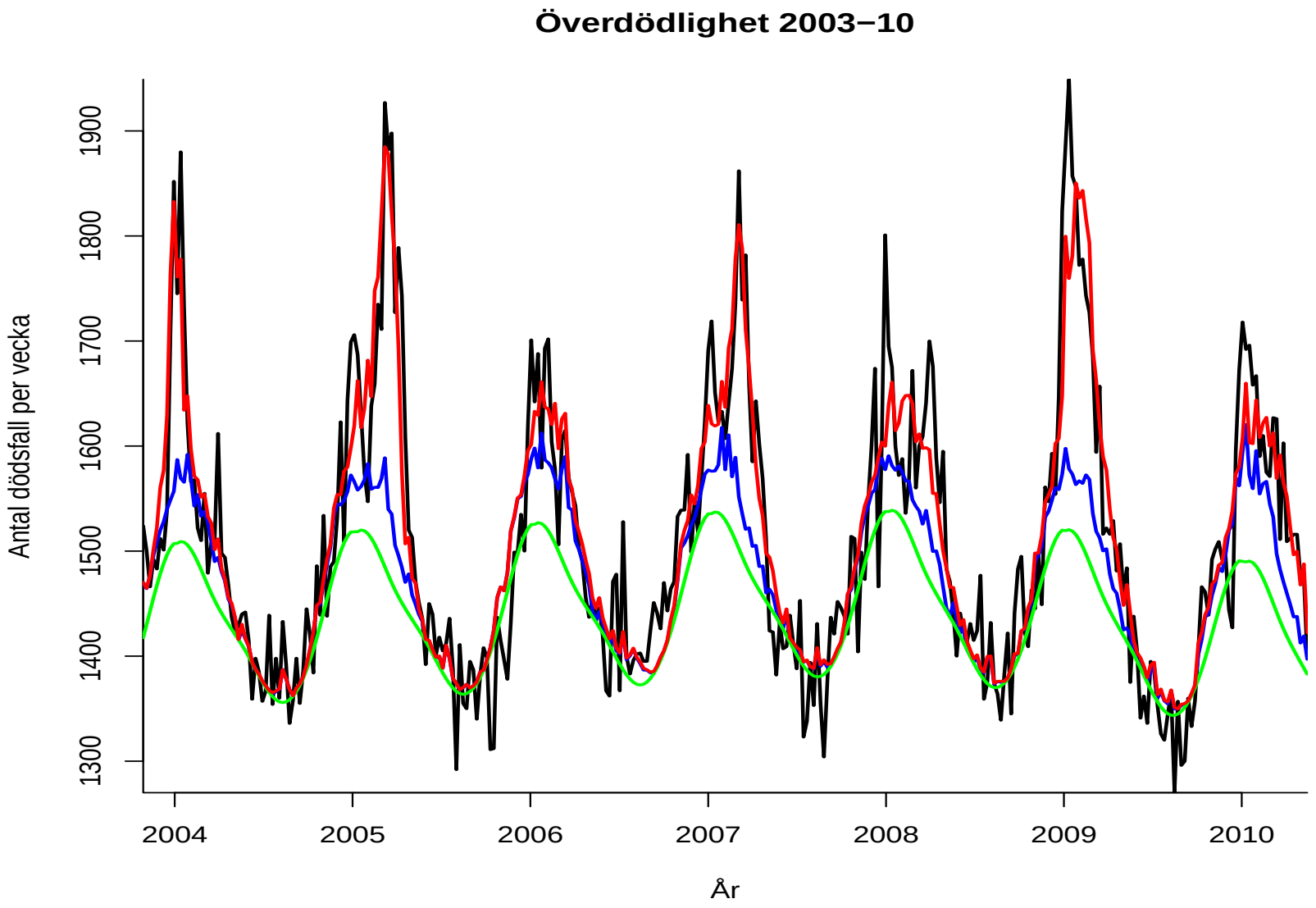


GAM model

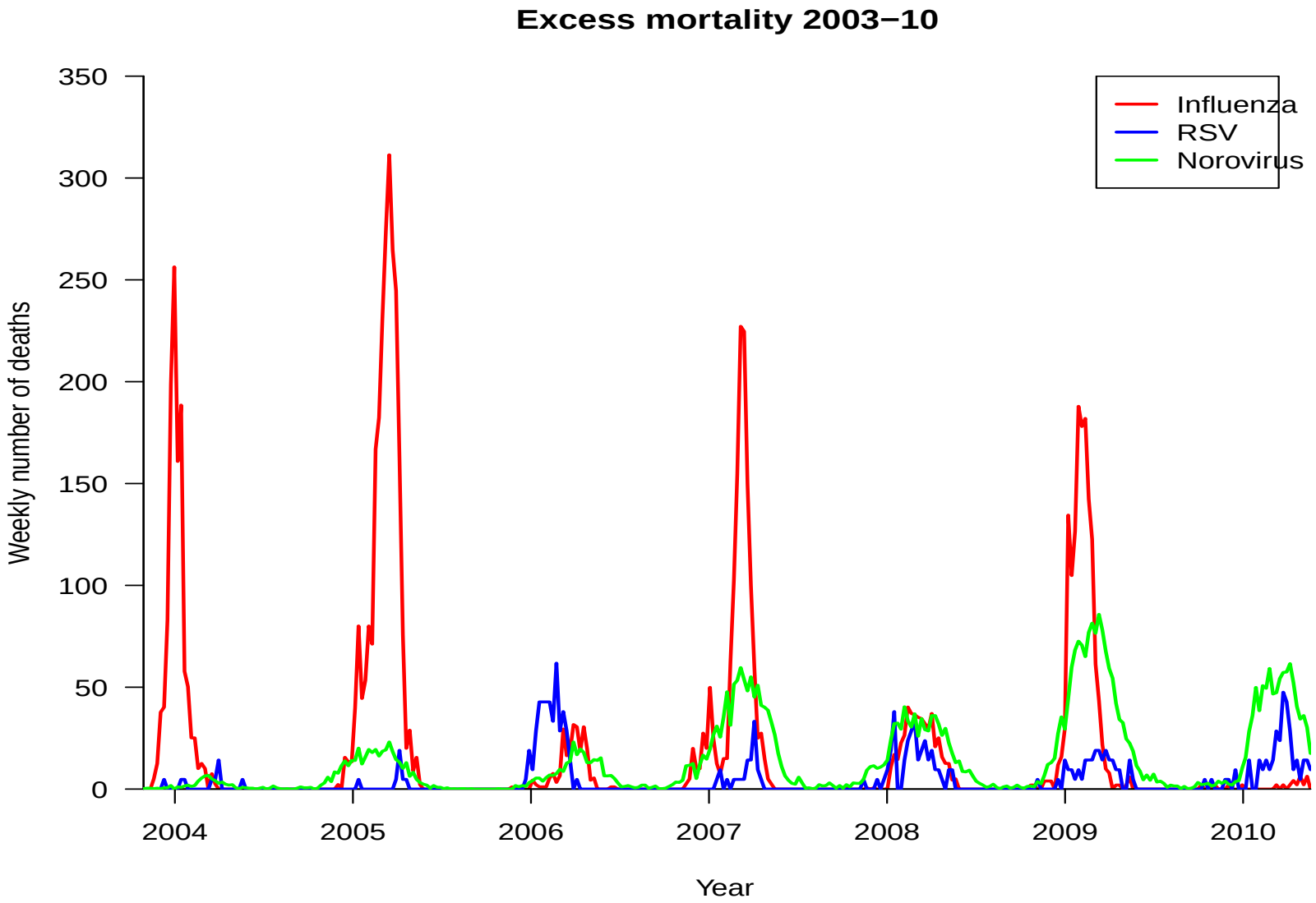
Spline function with respect to temperature



GAM model


Överdödlighet 2003−10
Antal dödsfall per vecka
1300
1400
1500
1600
1700
1800
1900
2004
2005
2006
2007
2008
2009
2010
År

GAM model


Excess mortality 2003–10
Influenza
RSV
Norovirus
Weekly number of deaths
0
50
100
150
200
250
300
350
2004
2005
2006
2007
2008
2009
2010
Year

Estimated excess mortality

Season	Influenza		RS-virus		Norovirus	
2003/04	1180	(720,1640)	40	(0,80)	60	(10,110)
2004/05	2400	(1730,3080)	40	(0,70)	380	(70,700)
2005/06	240	(-290,770)	440	(30,860)	280	(50,500)
2006/07	1380	(800,1950)	110	(10,220)	980	(180,1770)
2007/08	470	(-320,1260)	310	(20,610)	750	(140,1370)
2008/09	1410	(740,2080)	240	(10,470)	1350	(250,2440)
2009/10	40	(0,70)	340	(20,660)	930	(170,1690)
Average	1020	(800,1240)	220	(10,430)	680	(120,1230)

Conclusions

- RS-virus and norovirus contribute to excess mortality among the elderly, but to a lesser degree than influenza.
- Variable severity is necessary to capture impact of influenza, possibly of RSV and norovirus too.
- Longer time series are needed to get better estimates.