suRv's up: Statistical software for the surveillance of infectious diseases

Michael Höhle[1]

[1]Department for Infectious Disease Epidemiology
Robert Koch Institute, Berlin, Germany

InFER2011
The University of Warwick, UK
30 Mar 2011

ROBERT KOCH INSTITUT

Outline

Sections 3 and 4 contain large programming contributions from Sebastian Meyer, München, Germany

Outline

Infectious disease modelling and monitoring in practice (1)

- Early detection of a rapid increase in the incidence of a notifiable infectious disease depends on a good surveillance system and skilled epidemiologists to analyse the data.
- Statistical software helps the good epidemiologist become even better:
  - Models and their associated inference can be used to gain insights into disease dynamics.
  - Large data streams can be monitored simultaneously to quickly detect incidence changes.
  - Reports of these investigations can be generated automatically.
- There appears to be a lack of general purpose software for routine infectious disease modelling and monitoring.

Infectious disease modelling and monitoring in practice (2)

Good software alone does not do the job. It also takes:

- An understanding of the statistical aspects of the models and their selection
- A fair compromise between resources available and cost for acquisition, installation and training.
- Didactic documentation of the software containing illustrations of its use on relevant datasets
- A convincing answer to the question: *what's the added value?*

Aim of this talk

Introduce an open source R package for the visualization, modelling and monitoring of routinely collected public health surveillance data

The R package surveillance (1)

- Prospective monitoring for univariate count data time series:
  - `farrington` – Farrington et al. (1996)
  - `cusum` – Rossi et al. (1999) and extensions
  - `rogerson` – Rogerson and Yamada (2004)
  - `bayes` – H. (2007)[1]
  - `glrnb` – H. and Paul (2008)
  - `outbreakP` – Frisén et al. (2009)
- Prospective changepoint detection for categorical time series:
  - `pairedbinCUSUM` – surgical performance (Steiner et al., 2000)
  - `categoricalCUSUM` – binomial-, beta-binomial-, multinomial logit- and Bradley-Terry modelling (H., 2010)

[1] For a Bayesian GAM extension: Talk A1 by J. Manitz (Wed 11:15-11:45).

The R package surveillance (2)

- Retrospective count data time series models:
  - `hhh` – Held et al. (2005); Paul et al. (2008)
  - `hhh4` – Paul and Held (2011)[2][3]
  - `twins` – Held et al. (2006)
- Spatio-Temporal point process modelling and monitoring:
  - `twinSIR` – H. (2010)
  - `twinstim` – Meyer et al. (2010)
  - `stcd` – continuous space - continuous time cluster detection (Assunção and Correa, 2009)

[2] Plenary talk VI "Spatio-temporal modelling of infectious disease surveillance data", by L. Held, Thursday 16:15-17.15.

[3] See poster 'Spatio-temporal modelling of infectious disease surveillance data using the R-package `surveillance`' by M. Paul.

The R package surveillance (3)

- Motivation: Provide data structure and implementation framework for methodological developments in surveillance
- Spin-off: Tool for epidemiologists and others working in applied infectious disease modelling and monitoring
- Availability: CRAN, current development version from

  `http://surveillance.r-forge.r-project.org/`

  under the GNU General Public License (GPL) v. 2.0.
- To install the development version 1.3.0 under R version 2.12:

```
> install.packages("surveillance", repos = "http://r-forge.r-project.org")
> library("surveillance")
```

Use of surveillance in practice

A number of public health institutions and projects use the package, especially for outbreak detection:

- Computer Assisted Search For Epidemics (CASE) project by the Swedish Institute for Infectious Disease Control (SMI) – Cakici et al. (2010)
- Project on understanding Disease Risks from Livestock Movement in the Greater Mekong Subregion (Anonymous, 2011)
- Governmental Institute of Public Health, Lower Saxony, Germany (Hulth et al., 2010)
- Finish National Institute for Health and Welfare

This talk will...

- Focus on statistical modelling, because modelling provides insight into the spatio-temporal disease dynamics and thus is a precondition for monitoring.
- Differentiate methods based on the spatial and temporal resolution of the available data $\rightarrow$ trend towards more accurate temporal and geocoded information
- Give a software oriented view with less focus on mathematical detail.

Outline

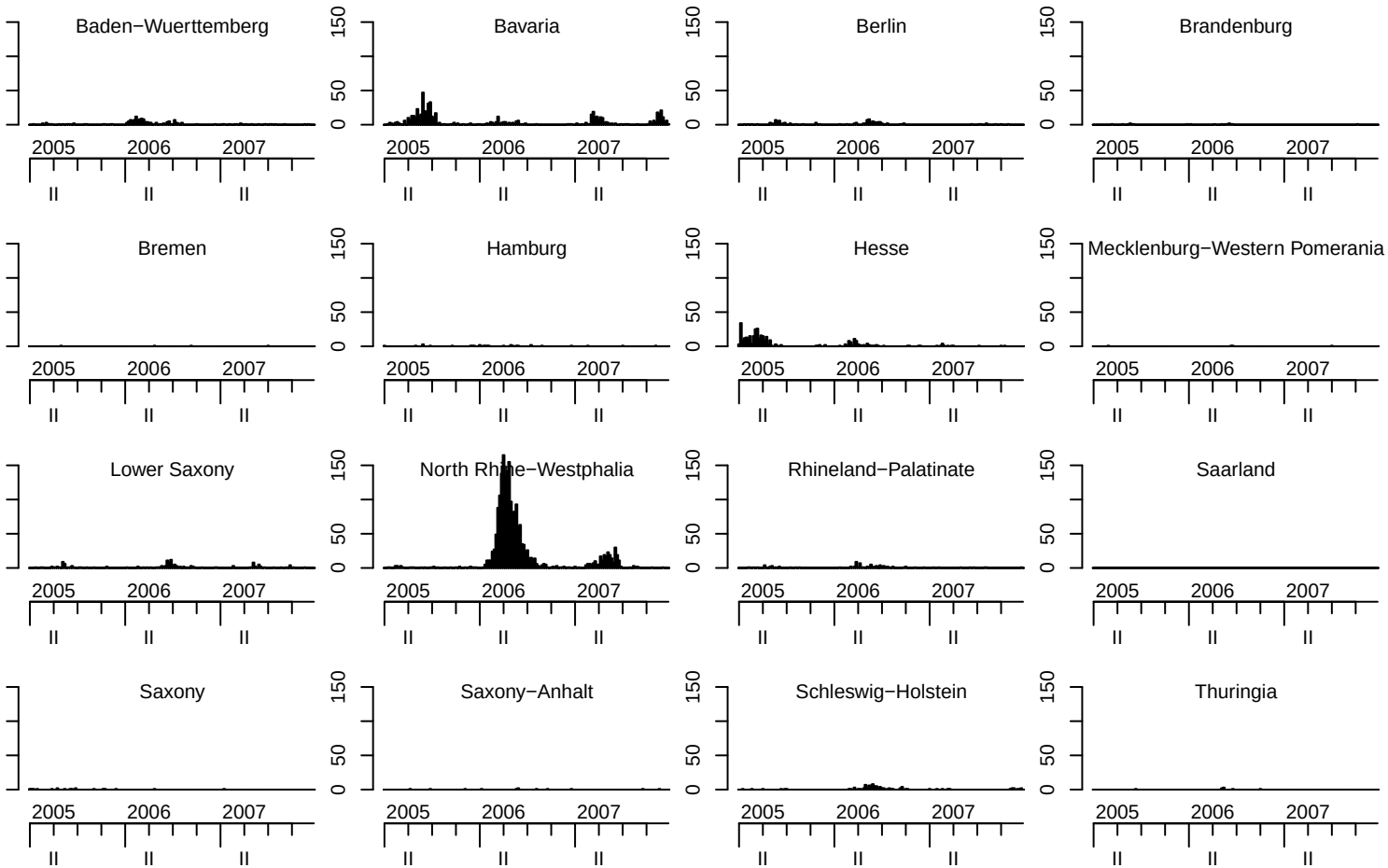
Surveillance data as multivariate time series of counts (1)

- Data from surveillance systems is, after suitable preprocessing, available as multivariate time series of counts $\{y_{it}; i = 1, \ldots, m, t = 1, \ldots, n\}$.
- The `surveillance` class for such data is the `sts` class.

```
> data("measlesDE")
> measlesDE

-- An object of class sts --
 freq:                 52
 start:                 2005 1
 dim(observed):         156 16

 Head of observed:
      Baden-Wuerttemberg Bavaria Berlin Brandenburg Bremen Hamburg Hesse
 [1,]                  0       0      0           0      0       1     3
      Mecklenburg-Western Pomerania Lower Saxony North Rhine-Westphalia
 [1,]                            0            0                      0
      Rhineland-Palatinate Saarland Saxony Saxony-Anhalt Schleswig-Holstein
 [1,]                    0        0      1             0                   0
      Thuringia
 [1,]         0

 ...
```

Surveillance data as multivariate time series of counts (2)

```
> plot(measlesDE, type = observed ~ time | unit)
```



Outbreak detection using Farrington's algorithm (1)

- The Farrington et al. (1996) algorithm is an early statistical algorithm to quickly detect emerging outbreaks in univariate count data time series based on an overdispersed Poisson GLM with intercept and trend:

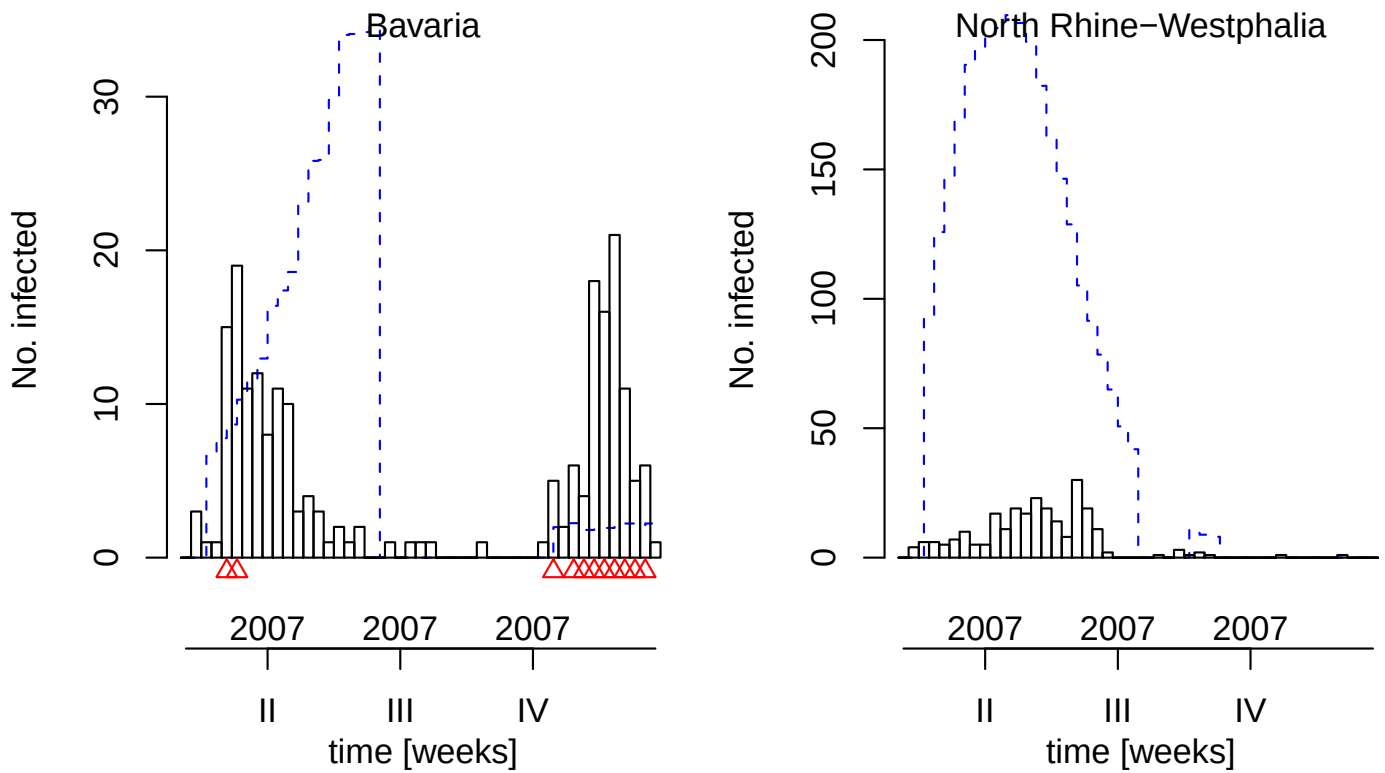
$$E(y_t) = \mu_t = \exp(\beta_0 + \beta_1 \cdot t),$$
$$Var(y_t) = \phi \cdot \mu_t.$$

- For each time point, the GLM is fitted and a $(1 - \alpha) \cdot 100\%$ prediction interval is computed.
- If the current observation is larger than the upper limit of the prediction interval, then an alarm is generated.

Outbreak detection using Farrington's algorithm (2)

The implementation in `surveillance`:

```
> twostates <- measlesDE[, c("Bavaria", "North Rhine-Westphalia")]
> s <- farrington(twostates, control = list(range = 110:nrow(twostates),
+     b = 2, w = 5))
```



Outline

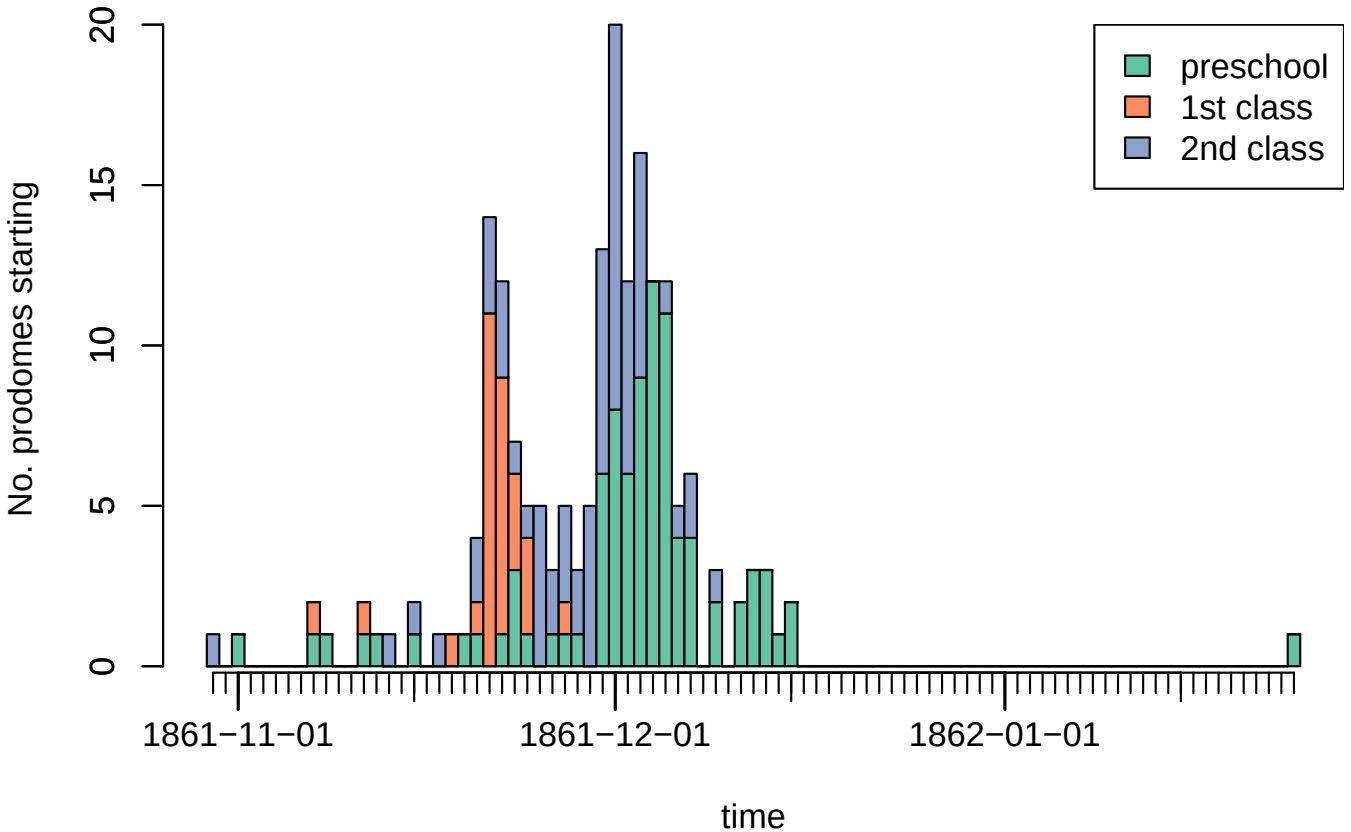
Hagelloch measles epidemic

- Measles outbreak among 188 children (age$\leq$14) during the winter 1861 in the German city of Hagelloch (near Tübingen).
- Detailed data including school class, household and household location of cases is available (Pfeilsticker, 1863; Oesterle, 1992).
- Data is prototypical for single outbreak data, but routine monitoring aspects are present.

```
> data("hagelloch")
> summary(hagelloch)

 AN SIR EPIDEMIC
   Time range: 0 -- 91.8102561747641
   Number of individuals: 188
   1 initially infected individuals:
     "184"
   0 never infected individuals
   Size of the epidemic: 187
   ...
```

Hagelloch measles epidemic – epidemic curve

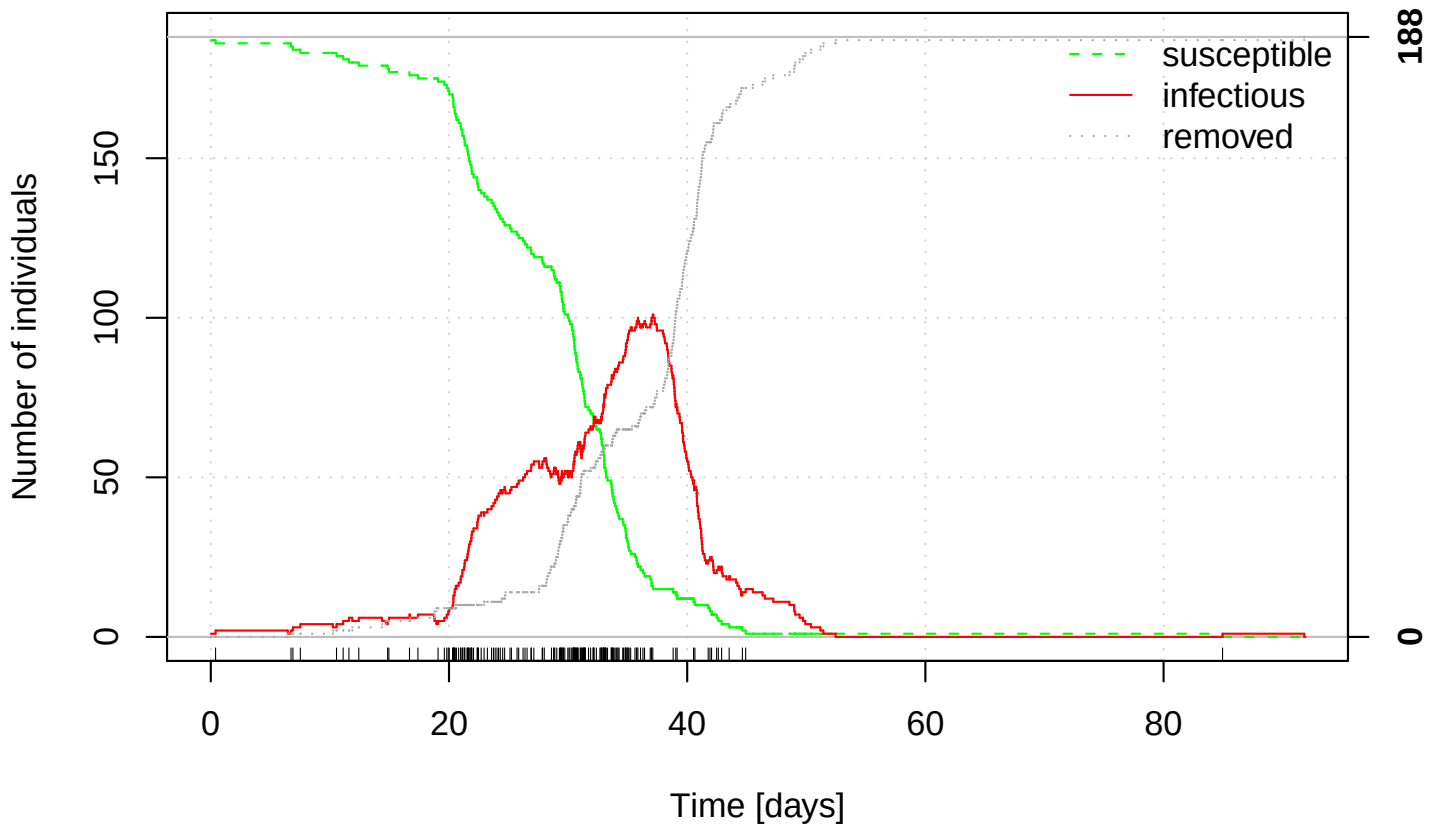
```
> library("epitools")
> hagelloch.df <- as.data.frame(hagelloch[hagelloch$BLOCK ==
+     1, ])
> with(hagelloch.df, epicurve.dates(PRO, strata = CL))
```



Hagelloch measles epidemic – SIR states

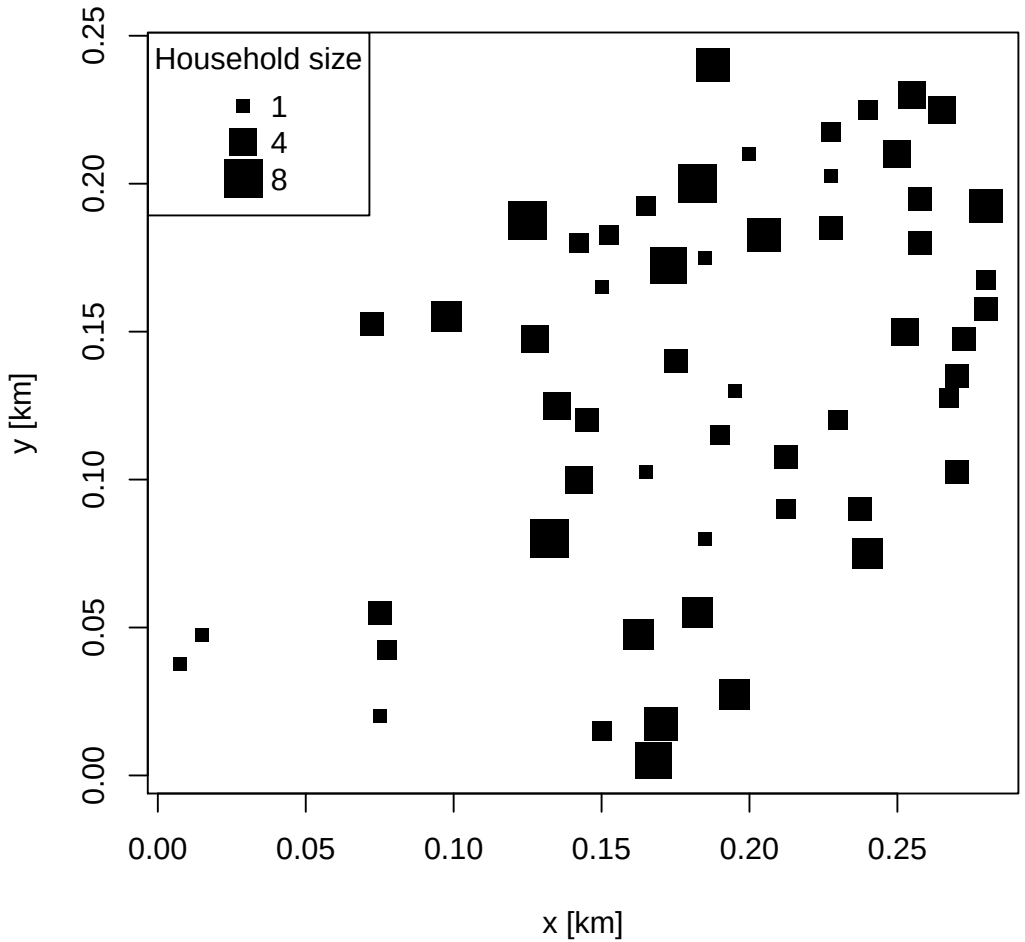
Visualization of epidemic as transition between states of a Susceptible-Infectious-Recovered (SIR) model:

```
> plot(hagelloch)
```



Hagelloch measles epidemic – household locations

```
> coordinates(hagelloch.df) <- c("x.loc", "y.loc")
> plot(hagelloch.df, cex = sqrt(multiplicity.sp(hagelloch.df)))
```



The twinSIR Model (1)

- Continuous time stochastic process with conditional intensity functions parametrized by covariates
- Set of possible event locations is known beforehand
- Risk of infection consists of two components
  - endemic component: Time to infection from external sources modelled by a Cox model
  - epidemic component: Similar to heterogeneous SIR model with, e.g., distance weighting of infectives
- The two component modelling is inspired by a model for univariate disease count data time series (Held et al., 2006).

The twinSIR Model (2)

Conditional intensity function $\lambda_i(t|\mathcal{H}_t)$ for a state change from susceptible to infectious of individual i:

$$\lambda_i(t|\mathcal{H}_t) = Y_i(t) \cdot [h_i(t) + e_i(t|\mathcal{H}_t)],$$

$$h_i(t) = \exp\left(h_0(t) + \mathbf{z}_i(t)'\boldsymbol{\beta}\right),$$

$$e_i(t|\mathcal{H}_t) = \sum_{j \in I(t)} f_{ij} = \sum_{m=1}^{q} \alpha_m \, x_{im}(t) = \mathbf{x}_i(t)'\boldsymbol{\alpha}, \quad \boldsymbol{\alpha} \geq \mathbf{0}$$

Notation: $\mathcal{H}_t$ is the history of the process, $Y_i(t)$ the at risk indicator, $h_0(t)$ the baseline hazard, $\mathbf{z}_i(t)$ a time depending endemic covariate vector, $I(t)$ the set of infectious at t and $f_{ij} \geq 0$ is a function of the covariates of i and j, e.g., the Euclidean distance between i and j described by a spline-basis.

The twinSIR Model (3)

- H. (2009) describes penalized likelihood inference and model selection using one-sided AIC when $h_0(t)$ is given as a penalized spline.
- The function `twinSIR` provides a model formula oriented implementation for model fitting and residual analysis.
- We will analyse the Hagelloch data using class room and household indicators as in Neal and Roberts (2004), but with an additional endemic component handling several outbreaks:

$$\lambda_i(t|\mathcal{H}_t) = Y_i(t) \cdot [\exp(\beta_0) + \alpha_H x_{i,H}(t) + \alpha_{c1} x_{i,c1}(t) + \alpha_{c2} x_{i,c2}(t) + \alpha_{local} x_{i,local}(t)],$$

where $x_{i,\cdot}$ is an indicator of how many infectious are, e.g., in the same sample household as i.

Analysing the Hagelloch data (2)

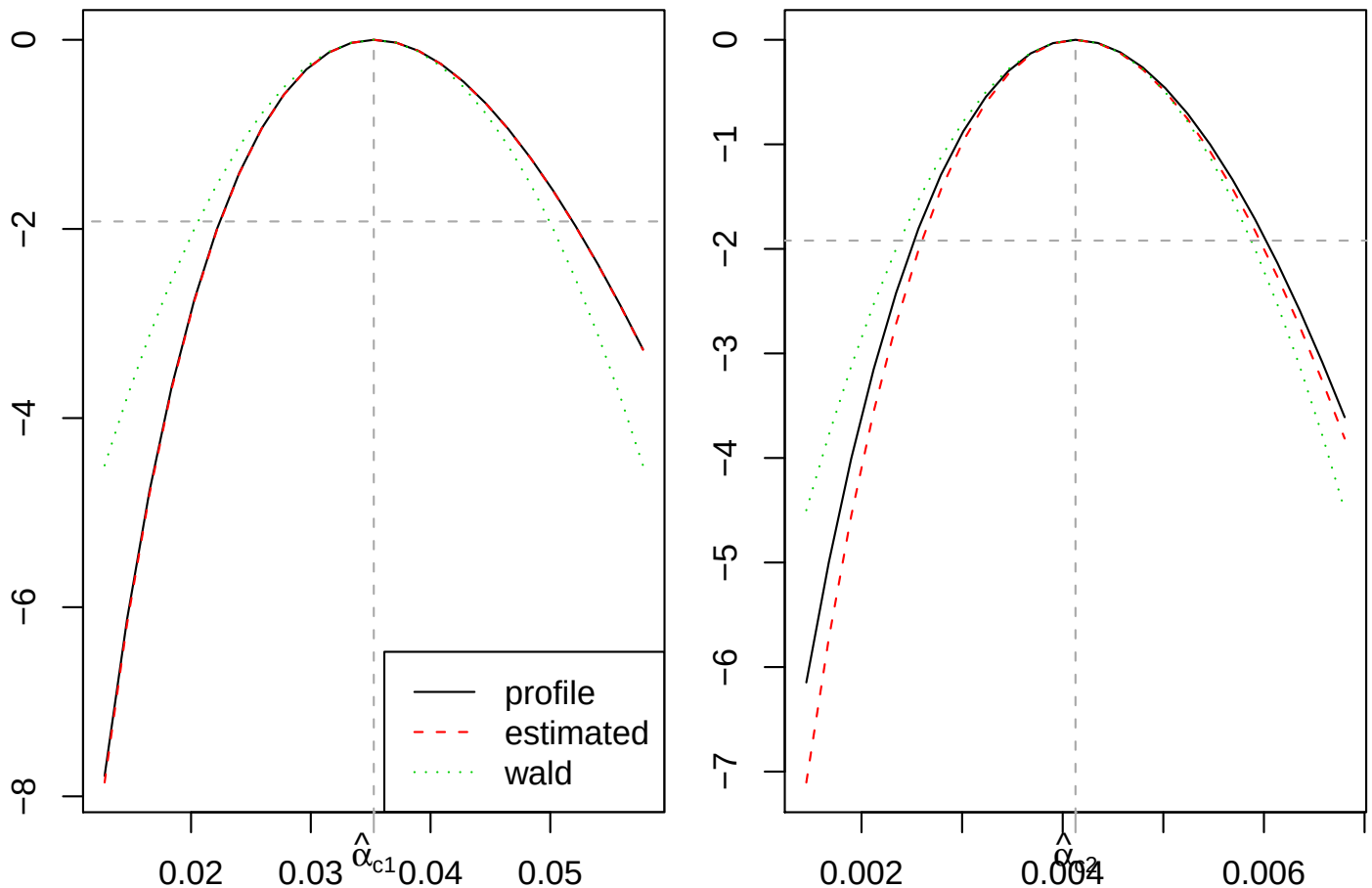
Result with surveillance:

```
> m <- twinSIR(~c1 + c2 + household + local, data = hagelloch)

> summary(m)

Call:
 twinSIR(formula = ~c1 + c2 + household + local, data = hagelloch)

 Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
 c1                 0.0352609  0.0074936   4.705 2.53e-06 ***
 c2                 0.0041271  0.0008918   4.628 3.70e-06 ***
 household          0.0300355  0.0062813   4.782 1.74e-06 ***
 local              0.0036818  0.0007191   5.120 3.06e-07 ***
 cox(logbaseline) -6.2155462  0.4345730 -14.303  < 2e-16 ***
 ---
 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

 Total number of infections:  187

 One-sided AIC: 1286.2        (simulated penalty weights)
 Log-likelihood: -640.1
 Number of log-likelihood evaluations: 68
```

Analysing the Hagelloch data (3)

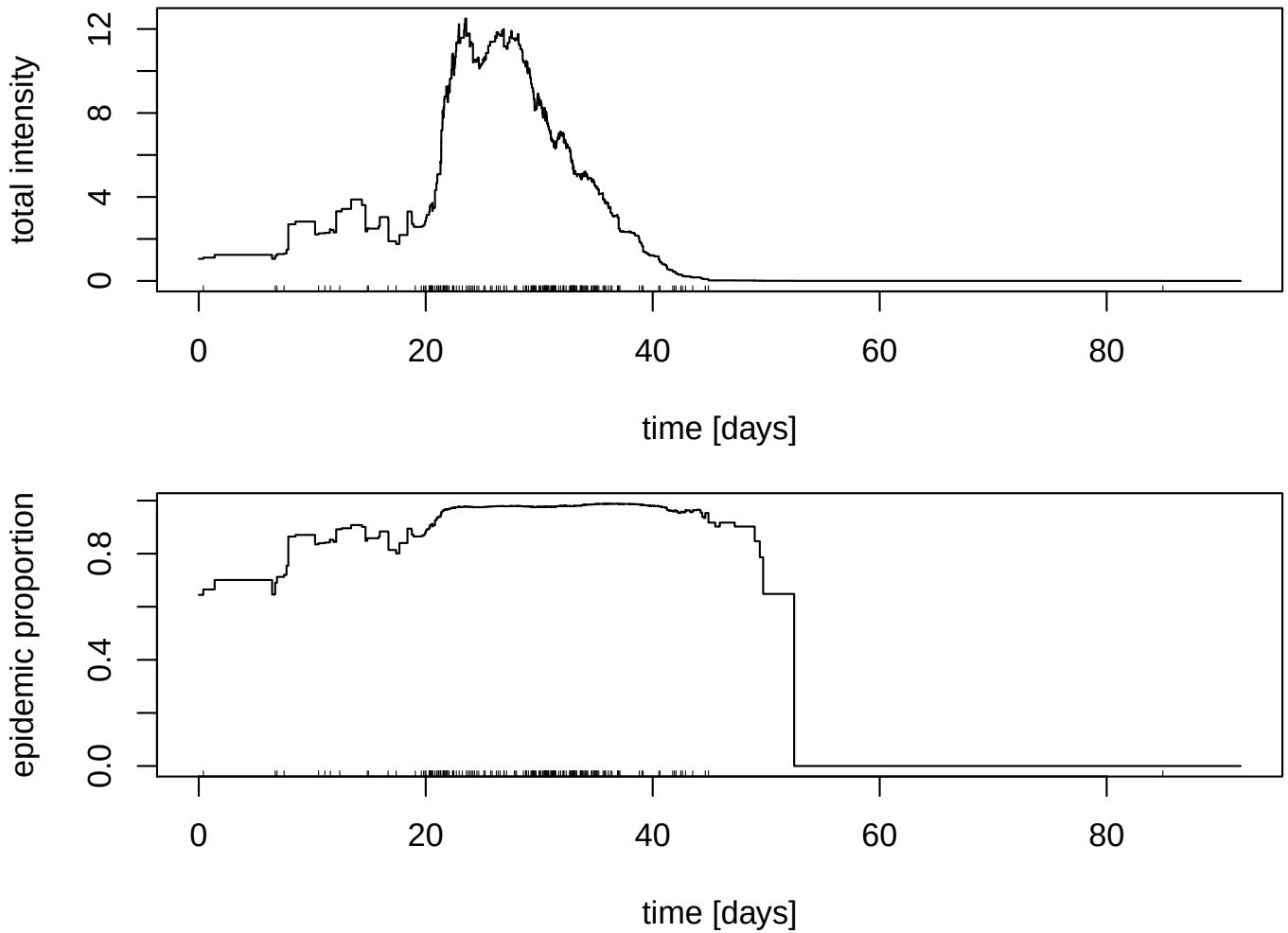
Behaviour of the loglikelihood near the MLE:

```
> prof <- profile(m,list(c(match("c1",names(coef(m))),NA,NA,25),
+                        c(match("c2",names(coef(m))),NA,NA,25)))
```



Analysing the Hagelloch data (3)

```
> plot(m, what = "total intensity")
> plot(m, what = "epidemic proportion")
```



Outline

Example: Invasive meningococcal disease (IMD)

- IMD is a life-threatening infectious disease triggered by the bacterium *Neisseria meningitidis* (aka *meningococcus*)
- Two most common finetypes in Germany in 2002–2008: 336 cases of *B:P1.7-2,4:F1-5*, 300 cases of *C:P1.5,2:F3-3*
- Case variables: date, coordinates of residence postcode, age, gender
- Scientific question: Do the finetypes spread differently?
- Statistical task: Quantify the transmission dynamics using two-component modelling of the conditional intensity function

Additive-multiplicative continuous space-time intensity model

$$\lambda^*(t, \mathbf{s}) \;=\; h(t, \mathbf{s}) \;+\; e^*(t, \mathbf{s})$$

Additive-multiplicative continuous space-time intensity model

$$\lambda^*(t, \mathbf{s}) \;=\; h(t, \mathbf{s}) \;+\; e^*(t, \mathbf{s})$$

Multiplicative endemic component

$$h(t, \mathbf{s}) = \exp\left(o_{\xi(\mathbf{s})} + \boldsymbol{\beta}' \mathbf{z}_{\tau(t),\xi(\mathbf{s})}\right)$$

- Piecewise constant function on a spatio-temporal grid $\{C_1, \ldots, C_D\} \times \{A_1, \ldots, A_M\}$ with time interval index $\tau(t)$, region index $\xi(\mathbf{s})$
- Region-specific offset $o_{\xi(\mathbf{s})}$, e.g. the log-population density
- Endemic linear predictor $\boldsymbol{\beta}' \mathbf{z}_{\tau(t),\xi(\mathbf{s})}$ includes discretised time trend and exogenous effects, e.g. influenza cases

Additive-multiplicative continuous space-time intensity model

$$\lambda^*(t,\mathbf{s}) \;=\; h(t,\mathbf{s}) \;+\; e^*(t,\mathbf{s})$$

Additive epidemic (self-exciting) component

$$e^*(t,\mathbf{s}) = \sum_{j \in I^*(t,\mathbf{s};\varepsilon,\delta)} e^{\eta_j}\, g_{\boldsymbol{\alpha}}(t-t_j)\, f_{\boldsymbol{\sigma}}(\mathbf{s}-\mathbf{s}_j)$$

- Individual infectivity weighting through linear predictor $\eta_j = \boldsymbol{\gamma}'\mathbf{m}_j$ based on the vector of unpredictable marks
- Positive parametric interaction functions, e.g. $f_\sigma(\mathbf{s}) = \exp\left(-\frac{\|\mathbf{s}\|^2}{2\sigma^2}\right)$ and $g_\alpha(t) = e^{-\alpha t}$
- Set of active infectives depends on fixed maximum temporal and spatial interaction ranges ε and δ

Marked extension with event type

- Motivation: joint modelling of both finetypes of IMD
- Additional dimension $\mathcal{K} = \{1, \ldots, K\}$ for event type $\kappa \in \mathcal{K}$

Marked CIF

$$\begin{aligned}\lambda^*(t, \mathbf{s}, \kappa) &= \exp\left(\beta_{0,\kappa} + o_{\xi(\mathbf{s})} + \boldsymbol{\beta}'\mathbf{z}_{\tau(t),\xi(\mathbf{s})}\right) \\ &\quad + \sum_{j \in I^*(t,\mathbf{s};\varepsilon,\delta)} q_{\kappa_j,\kappa}\, e^{\eta_j}\, g_{\boldsymbol{\alpha}}(t - t_j|\kappa_j)\, f_{\boldsymbol{\sigma}}(\mathbf{s} - \mathbf{s}_j|\kappa_j)\end{aligned}$$

- Type-specific endemic intercept
- Type-specific transmission, $q_{k,l} \in \{0, 1\}$, $k, l \in \mathcal{K}$
- Type-specific effect modification $\eta_j = \boldsymbol{\gamma}'\mathbf{m}_j$, κ_j is part of $\mathbf{m}_j$
- Type-specific interaction functions, e.g. variances σ^2_κ

Basic reproduction number

- An important quantity in epidemic modelling is the mean number of offspring each case generates.
- Since offspring are generated in time according to an inhomogeneous Poisson process we define

Basic reproduction number

$$\mu_i = e^{\eta_i} \cdot \left[\int_0^{\varepsilon} g_{\boldsymbol{\alpha}}(t)\,\mathrm{d}t\right] \cdot \left[\int_{b(\mathbf{0},\delta)} f_{\boldsymbol{\sigma}}(\mathbf{s})\,\mathrm{d}\mathbf{s}\right], \quad i = 1, \ldots, N.$$

- Type specific reproduction numbers are obtained by averaging the μ_i's for each type.

The epidataCS S3 class

`events` `SpatialPointsDataFrame` object containing information about the space-time location of the individual cases and their covariates.

`W` `SpatialPolygons` object describing the observational region.

`stgrid` Spatial partitioning, A, of W and a temporal partitioning, C, of $[0, T]$ together forming a space-time partition with $|A| \times |C|$ cells containing the endemic covariates as a `data.frame`.

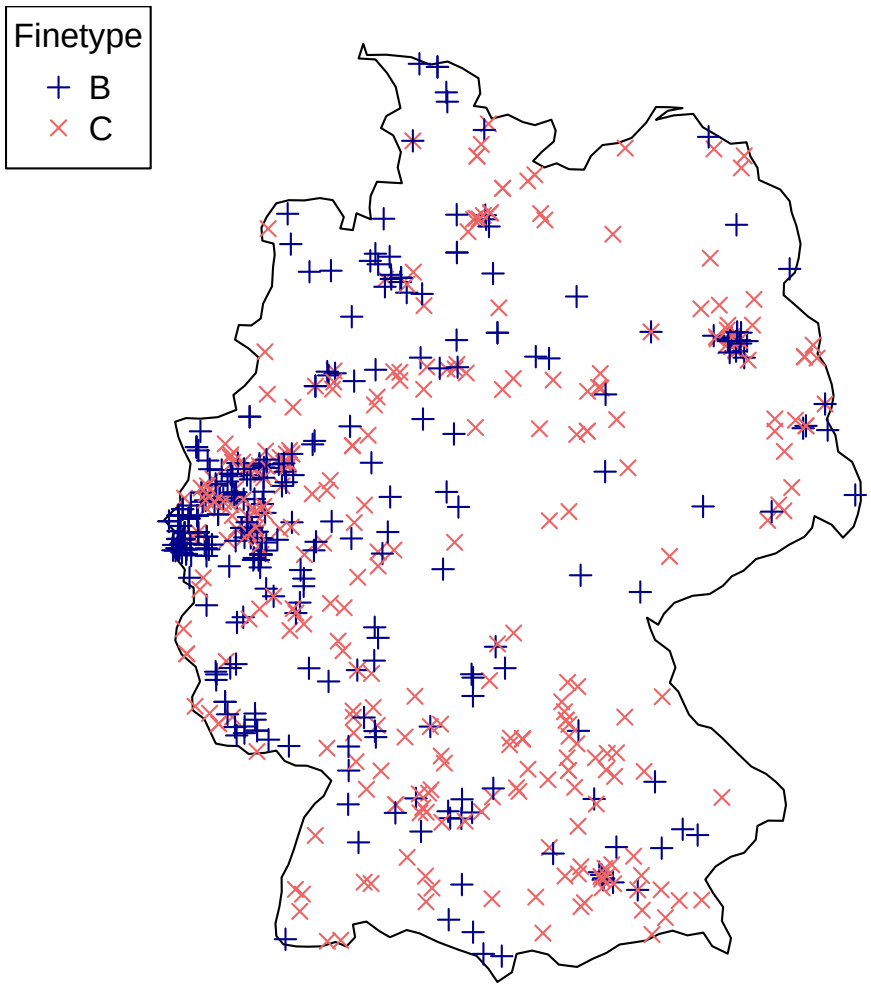
`qmatrix` $\mathcal{K} \times \mathcal{K}$ indicator matrix defining possible interaction between different types.

Visualization of IMD data (1)

```
> data("imdepi")
> class(imdepi)

[1] "epidataCS" "list"

> imdepi

 History of an epidemic
 Observation period: 0 -- 2562
 Observation window (bounding box): [4034, 4670] x [2687, 3543]
 Spatio-temporal grid (not shown): 366 time blocks, 413 tiles
 Types of events: 'B' 'C'
 Overall number of events: 636

           coordinates ID  time  tile type eps.t eps.s    sex   agegrp BLOCK start popdensity
 103 (4112.19, 3202.79)  1  0.21 05554    B    30   200   male   [3,19)     1     0         261
 402 (4122.51, 3076.97)  2  0.71 05382    C    30   200   male   [3,19)     1     0         519
 ...
```

Visualization of IMD data (1)

```
> with(imdepi, { plot(W) ; plot(events,add=TRUE)})
```



Visualization of IMD data (3)

Spatio-temporal visualization of disease occurrence using the `animation` package Xie (2010). Produces animated GIF files or Flash animations:

```
> animate(imdepi)
```

Analysis of IMD data (1)

```
> m <- twinstim(endemic  = ~1 + offset(log(popdensity)) + I(start/365) +
+                 sin(start * 2 * pi/365) + cos(start * 2 * pi/365),
+                 epidemic = ~1 + agegrp + type,
+                 siaf = siaf_log1, data = imdepi, subset = allEpiCovNonNA,
+                 optim.args = optim.args, finetune=FALSE,
+                 nCub = 36,
+                 typeSpecificEndemicIntercept = FALSE)

> toLatex(summary(m))
```

	Estimate	Std. Error	z value	$\mathbb{P}(\lvert Z\rvert > \lvert z\rvert)$
h.(Intercept)	−20.3815	0.0874	−233.25	$< 2 \cdot 10^{-16}$
h.I(start/365)	−0.0442	0.0223	−1.98	0.047
h.sin(start*2*pi/365)	0.2643	0.0648	4.08	$4.5 \cdot 10^{-05}$
h.cos(start*2*pi/365)	0.2627	0.0642	4.09	$4.3 \cdot 10^{-05}$
e.(Intercept)	−12.5277	0.3236	−38.71	$< 2 \cdot 10^{-16}$
e.agegrp[3,19)	0.7054	0.3266	2.16	0.03078
e.agegrp[19,Inf)	−0.2076	0.4577	−0.45	0.65013
e.typeC	−0.8900	0.2658	−3.35	0.00081
e.siaf.1	2.7708	0.0817	33.93	$< 2 \cdot 10^{-16}$
AIC:	18837			
Log-likelihood:	−9409			

Analysis of IMD data (2)

R_0 calculations:

```
> R0events <- R0(m, newevents = marks(imdepi))
> (R0.hat <- tapply(R0events, marks(imdepi)$type, mean,
+     na.rm = TRUE))
```

```
     B      C
0.24150 0.10171
```

To obtain CIs, simulate coefficients from asymptotic normal:

```
> boot <- replicate(999, {
+     m$coefficients <- mvrnorm(1, mu = coef(m), Sigma = vcov(m))
+     R0events <- R0(m, newevents = marks(imdepi))
+     tapply(R0events, marks(imdepi)$type, mean, na.rm = TRUE)
+ })
```

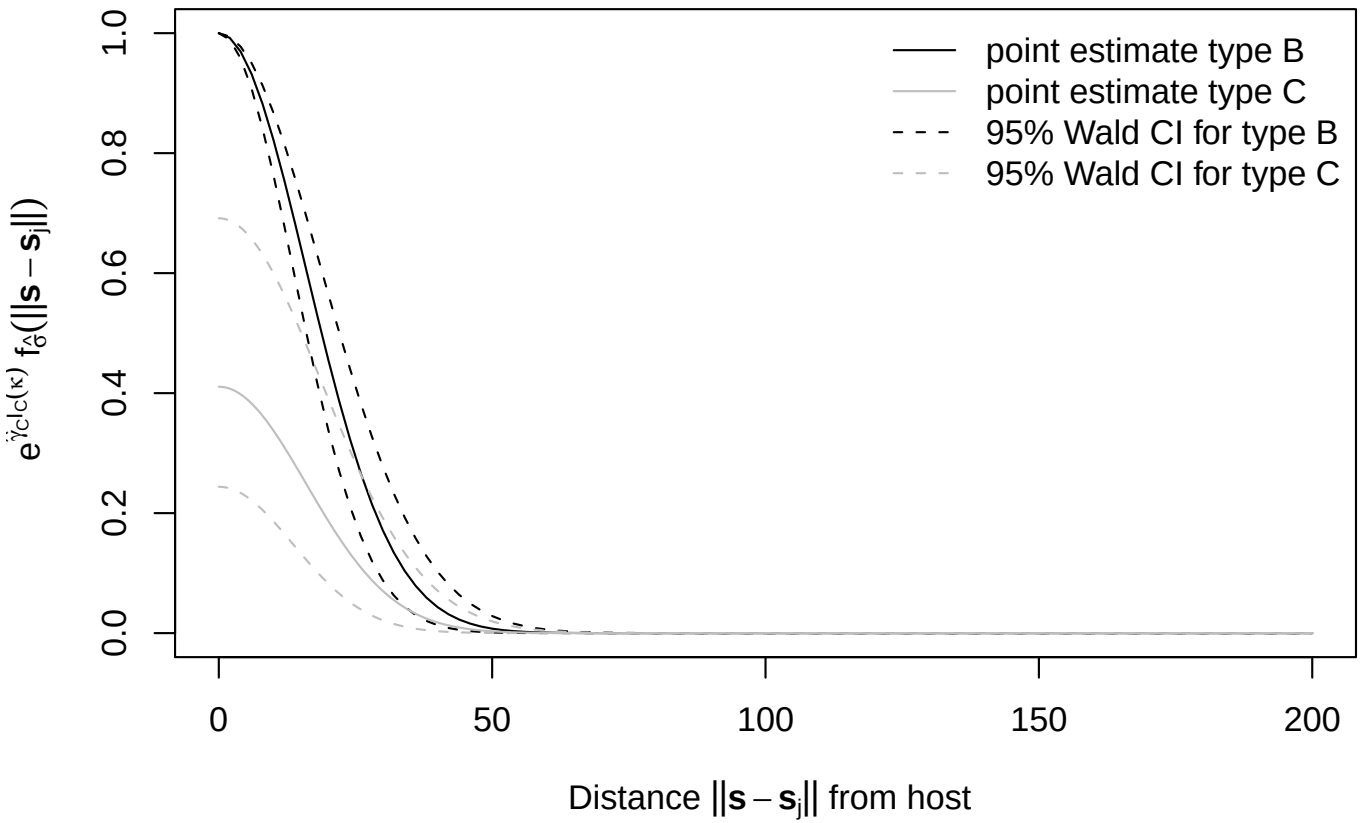
```
> t(apply(cbind(R0.hat, boot), 1, quantile, p = c(0.025,
+     0.975)))
```

```
     2.5%   97.5%
B 0.185715 0.32314
C 0.065356 0.16601
```

Analysis of IMD data (3)

Visualization of the fitted spatial interaction function:

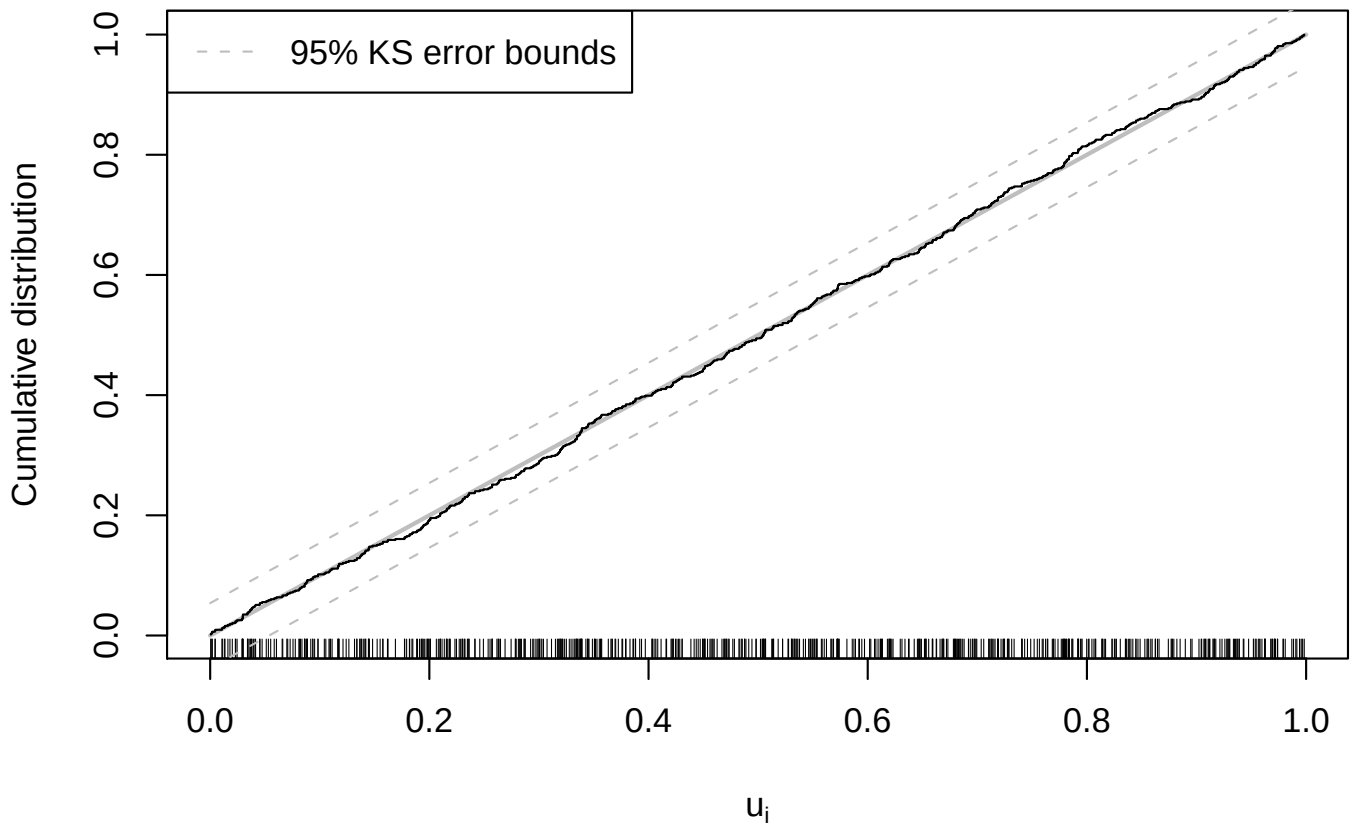
```
> plotiaf(m, iaf = "siaf", cols = "black")
```



Analysis of IMD data (5)

Residual plot as in Ogata (1988) by considering Cox-Snell-like residuals based on the fitted cumulative intensity function:

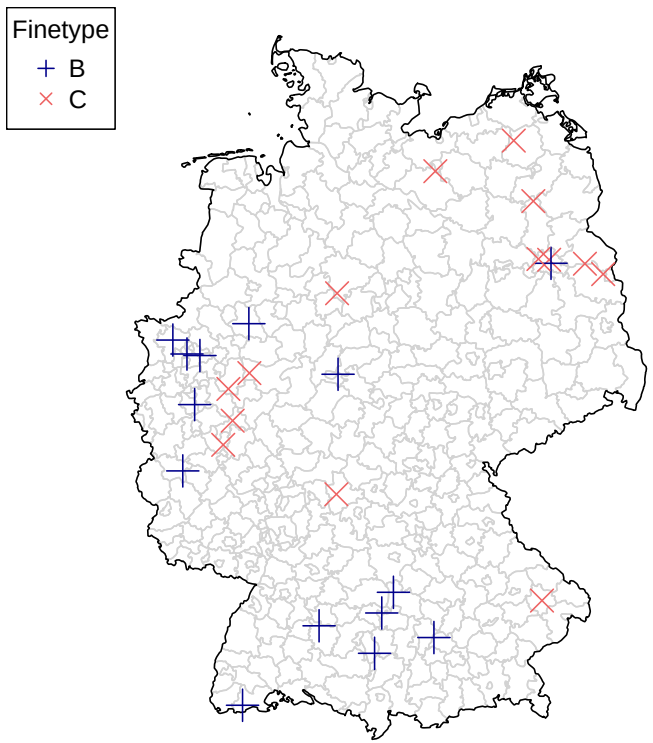
```
> residuals(m, plot = TRUE)
```



Analysis of IMD data (4)

Simulation from a fitted twinstim model using an adaptation of *Ogata's modified thinning* algorithm for marked point processes:

```
> stpp <- simulate(m, nsim = 1,
+     data = imdepi, tiles = kreise,
+     t0 = 200, T = 300,
+     W = W, nCub = 24)

> summary(stpp)$typeTable

 B  C
14 14

> with(stpp, {
+     plot(kreise)
+     plot(W, add = TRUE)
+     plot(events, add = TRUE)
+ })
```



Outline

Discussion (1)

- Talk presented steps towards R functionality for general modelling and monitoring of routine collected infectious disease data.
- `twinSIR` and `twinstim` provide data structures and inference framework for small- to medium-sized georeferenced event data (beyond epidemics!)
- Combining database, R, Sweave/odfWeave and LaTeX/OpenOffice allows for easy generation of reports
- R similarly can analyse outbreak investigation studies using logistic, conditional logistic and exact logistic regression.

Acknowledgements

Persons:

- Michaela Paul, Andrea Riebler and Leonhard Held, Institute of Social and Preventive Medicine, University of Zurich, Switzerland
- Sebastian Meyer and Valentin Wimmer, Ludwig-Maximilians-Universität München, Germany, and Mathias Hofmann, Technical University of Munich, Germany

Financial Support:

- German Science Foundation (DFG, 2003-2006)
- Munich Center of Health Sciences (2007-2010)

Literature I

Anonymous, 2011. GMS trade information. Website, `http://trade.animalhealthresearch.asia`.

Assunção, R., Correa, T., 2009. Surveillance to detect emerging space-time clusters. Computational Statistics & Data Analysis 53 (8), 2817–2830.

Cakici, B., Hebing, K., Grünewald, M., Saretok, P., Hulth, A., 2010. Case: a framework for computer supported outbreak detection. BMC Medical Informatics and Decision Making 10 (14).

Farrington, C., Andrews, N., Beale, A., Catchpole, M., 1996. A statistical algorithm for the early detection of outbreaks of infectious disease. Journal of the Royal Statistical Society, Series A 159, 547–563.

Frisén, M., Andersson, E., Schiöler, L., 2009. Robust outbreak surveillance of epidemics in sweden. Statistics in Medicine 28, 476–493.

Held, L., Hofmann, M., Höhle, M., Schmid, V., 2006. A two component model for counts of infectious diseases. Biostatistics 7, 422–437.

Held, L., Höhle, M., Hofmann, M., 2005. A statistical framework for the analysis of multivariate infectious disease surveillance data. Statistical Modelling 5, 187–199.

Höhle, M., 2007. surveillance: An R package for the monitoring of infectious diseases. Computational Statistics 22 (4), 571–582.

Literature II

Höhle, M., 2009. Additive-multiplicative regression models for spatio-temporal epidemics. Biometrical Journal 51 (6), 961–978.

Höhle, M., 2010. Changepoint detection in categorical time series. In: Kneib, T., Tutz, G. (Eds.), Statistical Modelling and Regression Structures – Festschrift in Honour of Ludwig Fahrmeir. Springer, pp. 377–397.

Höhle, M., Paul, M., 2008. Count data regression charts for the monitoring of surveillance time series. Computational Statistics & Data Analysis 52 (9), 4357–4368.

Hulth, A., Andrews, N., Ethelberg, S., Dreesman, J., Faensen, D., van Pelt, W., Schnitzler, J., 2010. Practical usage of computer-supported outbreak detection in five european countries. Eurosurveillance 15 (36).

Meyer, S., Elias, J., Höhle, M., 2010. A space-time conditional intensity model for infectious disease occurence. Tech. Rep. 95, Department of Statistics, Ludwig-Maximilians-Universität München, available as http://epub.ub.uni-muenchen.de/11898/.

Neal, P., Roberts, G. O., 2004. Statistical inference and model selection for the 1861 Hagelloch measles epidemic. Biostatistics 5 (2), 249–261.

Oesterle, H., 1992. Statistische reanalyse einer masernepidemie 1861 in hagelloch. Ph.D. thesis, Eberhard-Karls Universität, Tübingen.

Literature III

Ogata, Y., 1988. Statistical models for earthquake occurences and residual analysis for point processes. Journal of the American Statistical Association 83 (401), 9–27.

Paul, M., Held, L., 2011. Predictive assessment of a non-linear random effects model for multivariate time series of infectious disease counts. Statistics in MedicineEpub ahead of print DOI: 10.1002/sim.4177.

Paul, M., Held, L., Toschke, A. M., 2008. Multivariate modelling of infectious disease surveillance data. Statistics in Medicine 27, 6250–6267.

Pfeilsticker, A., 1863. Beiträge zur pathologie der masern mit besonderer berücksichtgung der statistischen verhältnisse. Ph.D. thesis, Eberhard-Karls Universität, Tübingen, available as `http://www.archive.org/details/beitrgezurpatho00pfeigoog`.

Rogerson, P., Yamada, I., 2004. Approaches to syndromic surveillance when data consist of small regional counts. Morbidity and Mortality Weekly Report 53, 79–85.

Rossi, G., Lampugnani, L., Marchi, M., 1999. An approximate CUSUM procedure for surveillance of health events. Statistics in Medicine 18, 2111–2122.

Steiner, S. H., Cook, R. J., Farewell, V. T., Treasure, T., 2000. Monitoring surgical performance using risk-adjusted cumulative sum charts. Biostatistics 1 (4), 441–452.

Xie, Y., 2010. animation: Demonstrate Animations in Statistics. R package version 1.1-3.
URL `http://CRAN.R-project.org/package=animation`