

# Endemic-epidemic modelling with surveillance

Sebastian Meyer  
Institute of Medical Informatics, Biometry, and Epidemiology  
Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany  
23 March 2022



```
library("surveillance")
```

Some background:

- Michael Höhle had worked on statistical methods for **outbreak detection** and published a first version of surveillance on CRAN in *November 2005*
- Back then, we had R 2.2.0 and only about 600 packages on CRAN.
- Since 2011, surveillance additionally provides three classes of **regression models for infectious disease surveillance data**



---

# *Journal of Statistical Software*

April 2017, Volume 77, Issue 11.

doi: 10.18637/jss.v077.i11

---

## Spatio-Temporal Analysis of Epidemic Phenomena Using the R Package *surveillance*

Sebastian Meyer  
Friedrich-Alexander-Universität  
Erlangen-Nürnberg

Leonhard Held  
University of Zurich

Michael Höhle  
Stockholm University

---

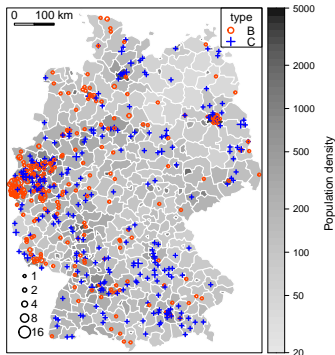
### Abstract

The availability of geocoded health data and the inherent temporal structure of communicable diseases have led to an increased interest in statistical models and software for spatio-temporal data with epidemic features. The open source R package **surveillance** can handle various levels of aggregation at which infective events have been recorded:

# Two types of supported surveillance data

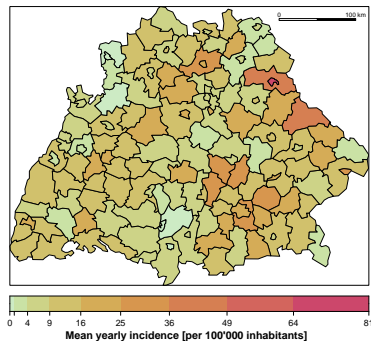
## Individual-level data

```
coordinates time sex agegrp ...
1 (4112, 3203) 1 male [3,19) ...
2 (4123, 3077) 1 male [3,19) ...
3 (4412, 2916) 6 female [19,Inf) ...
[....]
```

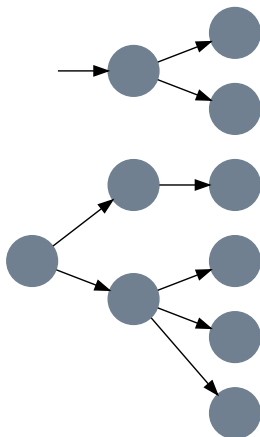


## Count time series

```
09162 09184 09179 09174 ...
2005-W01 7 0 0 1 ...
2005-W02 2 1 0 0 ...
2005-W03 1 4 0 0 ...
[....]
```



## Basic modelling concept



- Decomposed disease risk:
  - Endemic:** background risk driven by seasonality, demographics, ...
  - $\oplus$
  - Epidemic:** force of infection by previously infected individuals
- Force of infection may depend on social mixing and spatial/temporal distance to infective

# Endemic-epidemic modelling frameworks in surveillance

	twinstim	hhh4	twinSIR
<b>Data</b>	individual events in continuous space-time	event counts aggregated in space and time	individual SI[R][S] history of a fixed population
<b>Example</b>	case reports of IMD, Germany, 2002–2008	week×district counts of measles, Weser-Ems region, 2001–2002	measles outbreak among households, Hagelloch, 1861
<b>Model</b>	spatio-temporal PP	multivariate NegBin time series	multivariate temporal PP
<b>Rate</b> = <b>Endemic</b> + <b>Epidemic</b>	$\lambda(s, t)$  $V_{[s][t]} \rho_{[s][t]}$  $\sum_{j: t_j < t} \eta_j f(\ s - s_j\ ) g(t - t_j)$	$\mu_{it}$  $V_{it} e_{it}$  $\lambda_{it} Y_{i,t-1} + \phi_{it} \sum_{j \neq i} w_{ji} Y_{j,t-1}$	$\lambda_i(t) \mid \{i \in S(t)\}$  $v_i(t) \lambda_0(t)$  $\sum_{j \in I(t)} f(\ s_i - s_j\ )$

visualization

model assessment

likelihood inference

simulation

## Starting point for hhh4()

We have:

- Public health surveillance counts  $Y_{grt}$  indexed by [group,] [region,] time
- Maybe a social contact matrix  $C = (c_{g'g})$
- Maybe additional covariates: vaccination coverage, climate, socio-demographics, ...

We want: a joint model of all three data dimensions

$g$ : social mixing patterns between age groups

$r$ : spatial dynamics through human travel

$t$ : temporal dependencies inherent to communicable diseases

## An age-stratified, spatio-temporal, endemic-epidemic model

$$Y_{grt} | Y_{\cdot, \cdot, t-1} \sim \text{NegBin}(\mu_{grt}, \psi_{gr})$$

$$\mu_{grt} = v_{grt} + \phi_{grt} \sum_{g', r'} c_{g'g} w_{r'r} Y_{g', r', t-1}$$

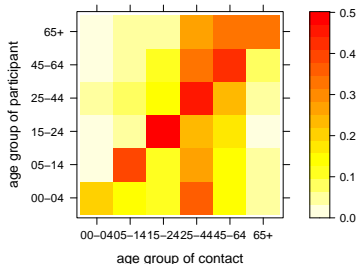


# An age-stratified, spatio-temporal, endemic-epidemic model

$$Y_{grt} | Y_{\cdot, \cdot, t-1} \sim \text{NegBin}(\mu_{grt}, \psi_{gr})$$

$$\mu_{grt} = \nu_{grt} + \phi_{grt} \sum_{g', r'} c_{g'g} w_{r'r} Y_{g', r', t-1}$$

Contact matrix ( $c_{g'g}$ ) for  $g' \rightarrow g$ ,  
for example from POLYMOD

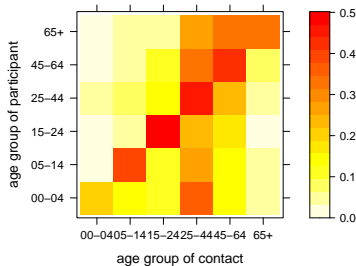


# An age-stratified, spatio-temporal, endemic-epidemic model

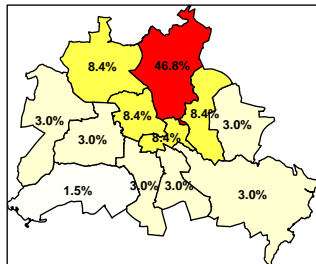
$$Y_{grt} | Y_{\cdot, \cdot, t-1} \sim \text{NegBin}(\mu_{grt}, \psi_{gr})$$

$$\mu_{grt} = \nu_{grt} + \phi_{grt} \sum_{g', r'} c_{g'g} w_{r'r} Y_{g', r', t-1}$$

Contact matrix ( $c_{g'g}$ ) for  $g' \rightarrow g$ ,  
for example from POLYMOD



Spatial weights for  $r' \rightarrow r$ , e.g.,  
power-law decay  $w_{r'r} = (o_{r'r} + 1)^{-p}$

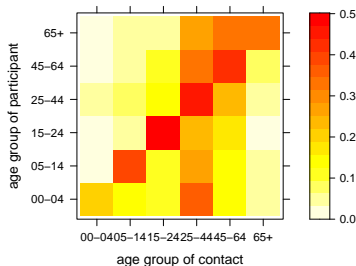


# An age-stratified, spatio-temporal, endemic-epidemic model

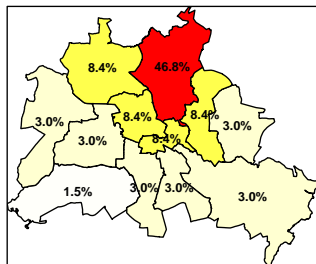
$$Y_{grt} | Y_{\cdot, \cdot, t-1} \sim \text{NegBin}(\mu_{grt}, \psi_{gr})$$

$$\mu_{grt} = v_{grt} + \phi_{grt} \sum_{g', r'} c_{g'g} w_{r'r} Y_{g', r', t-1}$$

Contact matrix ( $c_{g'g}$ ) for  $g' \rightarrow g$ ,  
for example from POLYMOD



Spatial weights for  $r' \rightarrow r$ , e.g.,  
power-law decay  $w_{r'r} = (o_{r'r} + 1)^{-\rho}$

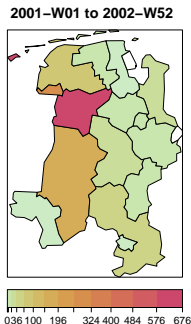


Log-linear  
predictors  
 $v_{grt}$  and  $\phi_{grt}$

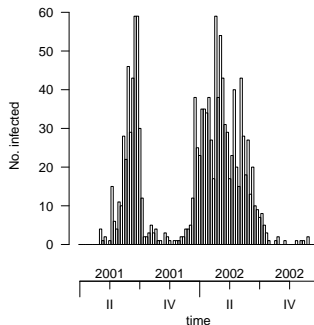
- Population
- Seasonality
- Group-specific susceptibility
- (Covariates)

## A simple example: data("measlesWeserEms")

```
plot(measlesWeserEms,  
     type = observed ~ unit)
```



```
plot(measlesWeserEms,  
     type = observed ~ time)
```



Following slides: univariate model for the (right-hand) overall time series

```
measles1 <- aggregate(measlesWeserEms, by = "unit")
```

## Univariate hhh4() model

Number of cases at time  $t$

$$Y_t | Y_{t-1} \sim \text{NegBin}(\mu_t, \psi)$$

Endemic-Epidemic decomposition of disease risk:

$$\mu_t = v_t + \lambda Y_{t-1}$$

Seasonality via sine-cosine terms in log-linear predictor  $v_t$ :

$$\log(v_t) = \alpha + \gamma \sin(2\pi t/52) + \delta \cos(2\pi t/52)$$

```
(endemic <- addSeason2formula(~1))  
~1 + sin(2 * pi * t/52) + cos(2 * pi * t/52)
```

## Univariate hhh4() model

```
measlesModel <- list(end = list(f = endemic), # "nu"
                    ar = list(f = ~1),      # "lambda"
                    family = "NegBin1")
measlesFit <- hhh4(measles1, control = measlesModel)
summary(measlesFit, amplitudeShift = TRUE, idx2Exp = TRUE)
```

Call:

```
hhh4(stsObj = measles1, control = measlesModel)
```

Coefficients:

	Estimate	Std. Error
exp(ar.1)	0.90977	0.09211
exp(end.1)	0.90641	0.20787
end.A(2 * pi * t/52)	1.43597	0.37315
end.s(2 * pi * t/52)	-0.77073	0.19739
overdisp	0.34459	0.07951

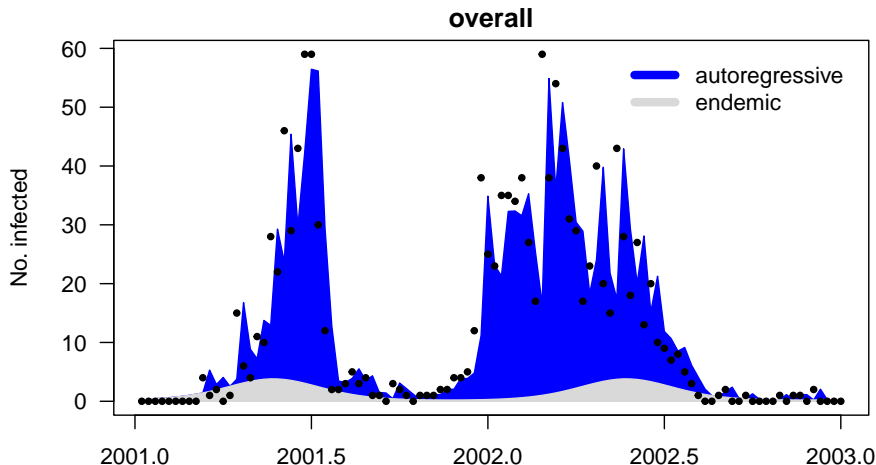
Log-likelihood: -271.32

AIC: 552.64

BIC: 565.82

## Fitted mean components

```
plot(measlesFit, type = "fitted")
```



## Methods for "hhh4" objects

Display	Extract	Modify	Other
print	nobs	update	predict
summary	coef		simulate
plot	fixef		pit
	ranef		scores
	vcov		calibrationTest
	confint		all.equal
	coeflist		oneStepAhead
	logLik		
	residuals		
	terms		
	formula		
	getNEweights		