A space-time conditional intensity model for invasive meningococcal disease occurrence

Sebastian Meyer[1,2,*], Johannes Elias[3], and Michael Höhle[4,2,]**

[1]Department of Psychiatry and Psychotherapy, Ludwig-Maximilians-Universität, München, Germany
[2]Department of Statistics, Ludwig-Maximilians-Universität, München, Germany
[3]German Reference Centre for Meningococci, University of Würzburg, Würzburg, Germany
[4]Department for Infectious Disease Epidemiology, Robert Koch Institute, Berlin, Germany
*email: Sebastian.Meyer@med.uni-muenchen.de
**email: HoehleM@rki.de

SUMMARY: A novel point process model continuous in space-time is proposed for quantifying the transmission dynamics of the two most common meningococcal antigenic sequence types observed in Germany 2002–2008. Modelling is based on the conditional intensity function (CIF) which is described by a superposition of additive and multiplicative components. As an epidemiological interesting finding, spread behaviour was shown to depend on type in addition to age: basic reproduction numbers were 0.25 (95% CI 0.19-0.34) and 0.11 (95% CI 0.07-0.17) for types B:P1.7-2,4:F1-5 and C:P1.5,2:F3-3, respectively. Altogether, the proposed methodology represents a comprehensive and universal regression framework for the modelling, simulation and inference of self-exciting spatio-temporal point processes based on the CIF. Usability of the modelling in biometric practice is promoted by an implementation in the R package `surveillance`.

KEY WORDS: Conditional intensity function; Infectious disease surveillance data; Spatio-temporal point process; Stochastic epidemic modelling.

1. Introduction

Infectious diseases – such as influenza, gastroenteritis, and the "swine flu" among humans, or foot and mouth disease, the "bird flu", and classical swine fever among animals – are a matter of tremendous public concern especially gaining attention in case of outbreaks. The present work concentrates on stochastic modelling and associated inference for spatio-temporal epidemic point referenced data motivated by the analysis of routinely collected invasive meningococcal disease (IMD) data. IMD is a life-threatening human bacterial disease mostly manifesting as meningitis or sepsis. Its pathogenic agent, *Neisseria meningitidis* (aka *meningococcus*), can be transmitted by large droplet secretions from the respiratory tract of colonized or infected humans. The only reservoir of meningococci is the human (mostly nasopharyngeal) mucosa (Rosenstein et al., 2001). Data on cases of IMD related to the two most common meningococcal finetypes B:P1.7-2,4:F1-5 and C:P1.5,2:F3-3 in Germany 2002–2008 are obtained from the German Reference Centre for Meningococci (Nationales Referenzzentrum für Meningokokken, NRZM). Here, a 'finetype' represents a unique combination of serogroup, sequence type of variable region 1 and 2 of the outer membrane protein PorA, and sequence type of the variable region of the outer membrane protein FetA. One specific question of interest for the researchers at the NRZM is whether the two finetypes (in what follows abbreviated B and C) exhibit different spatio-temporal behaviour.

The postal code of the patient's home address was the spatial resolution available for our analysis. Despite being spatially discrete we consider centroids of postal code areas as quasi-continuous in space when looking at entire Germany. As usual with infectious diseases, the actual time point of infection is unknown for the IMD cases. Therefore, we define the beginning of illness and infectivity as the date of specimen sampling.
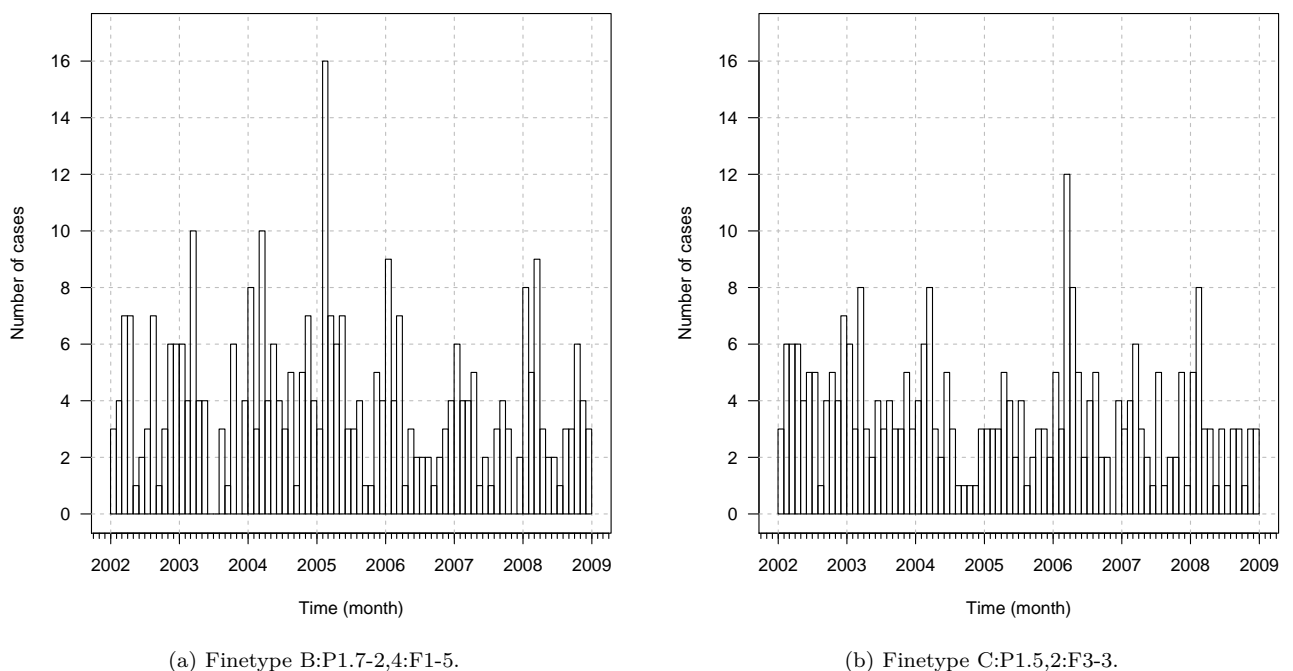
(a) Finetype B:P1.7-2,4:F1-5.

(b) Finetype C:P1.5,2:F3-3.

Figure 1. Monthly numbers of IMD cases for both finetypes separately.

All in all, $n = 636$ infections with finetypes B (336) and C (300) have been registered. Figure 1 shows the monthly numbers of IMD cases for each finetype. Cases of IMD predominantly occur during winter and early spring, which can be seen from more or less pronounced peaks in the figure. Specifically, a connection between outbreaks of meningococci and influenza is hypothesized. For example, Jensen et al. (2004) found an association between the influenza detection rate and the number of IMD cases during the same week in temporal analysis of data from Northern Jutland County in Denmark, during 1980–1999.

Figure 2 presents the spatial distributions of the two finetypes based on the postcodes of the patients' residences. Over the 7-year period some cases shared the same postal code, therefore, the area of each point in the figure is drawn proportional to the number of cases at its location. For the serogroup B finetype in (a) the highest point multiplicity is 16, whereas for the serogroup C finetype in (b) this number is 4. In connection with the temporal occurrence of the events shown in Figure 1, the spatial distribution suggests that IMD is an endemic disease, i.e. cases can occur at any time and at any location. The maps also show the population densities of the districts, which can be assumed to be roughly proportional to the population at risk of infection. Spatial heterogeneity of the observed point patterns thus partially arises from spatial variation in the population density. Not surprisingly, the intensity of points in metropolitan areas like Berlin, Munich or the Ruhr is higher. Animated graphics of the space-time locations of infections give more insight into the epidemic character of the finetypes, and can be found as Web Animation 1. Here, it appears as if finetype B exhibits a more stationary pattern than finetype C – in the sense that infections cluster more in space and time. It is supposed, yet not proven, that this phenomenon is due to differences in the mucosal immune reaction elicited; specifically, finetype B might be more successful than C in evading mucosal clearance.

Quantifying the dynamics of IMD would be an important step in the finetype characterisation of IMD. We want to perform such an investigation in a spatio-temporal manner and therefore
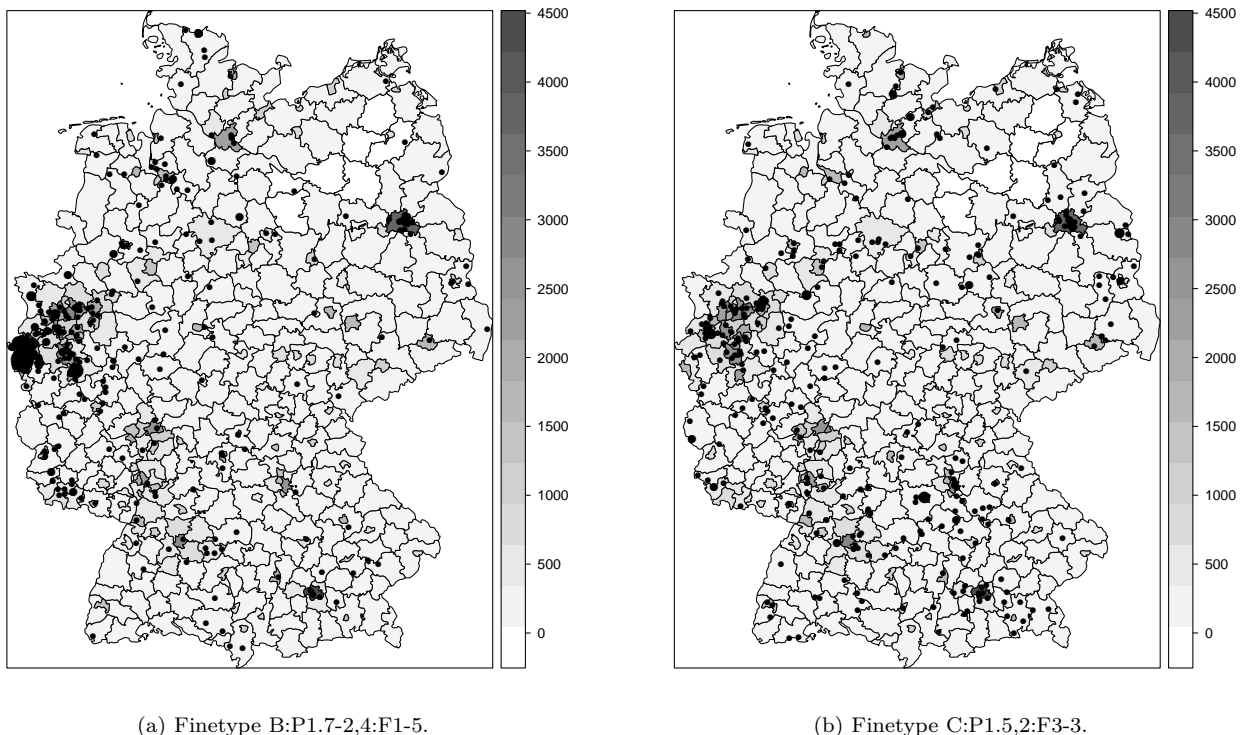
(a) Finetype B:P1.7-2,4:F1-5.

(b) Finetype C:P1.5,2:F3-3.

Figure 2. Spatial point patterns of the cases of meningococci by finetype during the years 2002–2008. The area of each dot is proportional to the number of cases at its location. Also shown are the population densities (inhabitants per km^2) of Germany's districts (source: Federal Statistical Office (DESTATIS) (2009)).

use spatio-temporal point processes as modelling framework. Specifically, we want to establish a regression framework allowing us to quantify the transmission dynamics of IMD and its dependency on covariates. Point process modelling has in the context of epidemics been used in a discrete spatial setting in, e.g., Neal and Roberts (2004), Diggle (2006), Scheel et al. (2007) and Jewell et al. (2009). Spatio-temporal epidemic modelling in an explicit continuous spatial setting, however, is rare with Diggle et al. (2005) being one of the few examples of covariate adjusted modelling. One explanation is the balancing between optimal spatial resolution of the data and confidentiality of cases.

Recently, there have been suggestions for splitting the dynamics of infectious diseases into endemic and epidemic components; see Held et al. (2005) for a discrete spatial – discrete time perspective and Höhle (2009) for a discrete spatial – continuous time perspective. For the continuous spatial – continuous time setting, similar modelling approaches have been seen in the analysis of earthquake data, see e.g. Ogata (1998, 1999). Other areas of application are the modelling of forest fires (Peng et al., 2005), residential burglaries (Mohler et al., 2010), and the analysis of bird nesting patterns (Diggle et al., 2009). Altogether, our proposed modelling provides a unifying regression framework – beyond epidemics – for the modelling, inference and simulation of spatio-temporal point processes.

This article is organized as follows: Section 2 presents the spatio-temporal two-component epidemic model based on the CIF, whereas Sections 3 and 4 discuss inference and simulation for the proposed model. Section 5 analyses the IMD data, and a discussion in Section 6 finalizes the article.

3

2. Spatio-Temporal Two-Component CIF Model

In the following text, we propose a novel additive-multiplicative model for the conditional intensity function of an infectious disease process continuous in space-time with events occuring in a prespecified observation period $[0, T]$, $T > 0$, and observation region $W \subset \mathbb{R}^2$. The CIF $\lambda^*(t, \boldsymbol{s})$ represents the instantaneous rate or hazard for events at time t and location $\boldsymbol{s}$ given all the observations up to time t (the asterisk notation shall represent the conditioning on the random past history of the process).

The basic framework of the proposed model is to superimpose endemic and epidemic components in order to model the IMD surveillance data – an idea similar to the two-component spatial SIR model (Höhle, 2009):

$$\lambda^*(t, \boldsymbol{s}) = h(t, \boldsymbol{s}) + e^*(t, \boldsymbol{s}) \qquad (t > 0, \boldsymbol{s} \in W) \, .$$

The epidemic component $e^*(t, \boldsymbol{s})$ represents the spread of the disease by person-to-person contact. The endemic component $h(t, \boldsymbol{s})$ models otherwise imported cases and is – contrary to the epidemic component – independent of the internal history of the process.

2.1 *Specification of the Endemic Component $h(t, s)$*

The endemic component is of the multiplicative form $h(t, \boldsymbol{s}) = \rho(t, \boldsymbol{s}) \exp(\boldsymbol{\beta}' \boldsymbol{z}(t, \boldsymbol{s}))$, where $\rho(t, \boldsymbol{s})$ is a known spatio-temporal intensity offset, e.g. the population density at time t in the district containing the location $\boldsymbol{s}$, such that the endemic rate of infection is proportional to the population density. Furthermore, $\boldsymbol{z}(t, \boldsymbol{s})$ is a linear predictor of endemic covariates, e.g., this could be a temporal trend or exogenous covariates resulting from another jointly evolving point process. For example, in the IMD application, an endemic covariate is the number of influenza cases on a week × district grid (possibly time-lagged). Altogether, the endemic component is modelled as a piecewise constant function on some spatio-temporal grid resulting from a decomposition of the time period $(0, T]$ and the observation region W. The consecutive time intervals of this decomposition (e.g. weeks) are denoted by $C_1, \ldots, C_D \subset (0, T]$, and the spatial tiles (e.g. districts) are denoted by $A_1, \ldots, A_M \subset W$. Let the functions $\tau(t)$ and $\xi(\boldsymbol{s})$ return the indices of the temporal and spatial grid units containing time point t and coordinate $\boldsymbol{s}$, respectively. Then, the endemic component can be written as

$$h(t, \boldsymbol{s}) = \rho_{\tau(t), \xi(\boldsymbol{s})} \exp\left(\boldsymbol{\beta}' \boldsymbol{z}_{\tau(t), \xi(\boldsymbol{s})}\right) , \tag{1}$$

where $\rho_{\tau(t), \xi(\boldsymbol{s})}$ is the known interval- and tile-specific offset and $\{\boldsymbol{z}_{\tau, \xi} : \tau \in \{1, \ldots, D\}, \xi \in \{1, \ldots, M\}\}$ is a collection of covariates on the spatio-temporal grid $\{C_1, \ldots, C_D\} \times \{A_1, \ldots, A_M\}$.

2.2 *Specification of the Epidemic Component $e^*(t, s)$*

The self-exciting component of the model essentially provides a description of the infection pressure at a space-time location $(t, \boldsymbol{s})$ caused by each infectious individual. This infectivity of an infectious individual j, denoted by $e_j(t, \boldsymbol{s})$, corresponds to the inhomogeneous rate of a Poisson process, the realisations of which are the space-time locations of infected individuals. This so called triggering function is factorized into separate effects of marks, elapsed time, and relative location:

$$e_j(t, \boldsymbol{s}) = e^{\eta_j} \, g(t - t_j) \, f(\boldsymbol{s} - \boldsymbol{s}_j) \, , \qquad (t > t_j) \tag{2}$$

where $(t_j, \boldsymbol{s}_j)$ is the infection time and location of individual j, $\eta_j = \gamma_0 + \boldsymbol{\gamma}' \boldsymbol{m}_j$ is a linear predictor based on the vector of unpredictable marks $\boldsymbol{m}_j$ attached to the infected individual, and g and f are positive temporal and spatial interaction functions, respectively. The effects $\boldsymbol{\gamma}$ of marks reflect that different individuals might cause more or less secondary cases, depending on individual characteristics.

4

The interaction functions describe the decay of infectivity with an increasing spatial or temporal distance from the infection source. In infectious disease applications, f is often taken to be a radially symmetric kernel corresponding to an isotropic spread of the disease, such that $f(\boldsymbol{s} - \boldsymbol{s}_j) \equiv f(\|\boldsymbol{s} - \boldsymbol{s}_j\|)$. A typical example is to let f be the kernel of a bivariate normal density with zero mean and diagonal covariance matrix. The temporal interaction function could be chosen as $g(t) = e^{-\alpha t}$, $t > 0$, $\alpha > 0$, representing an exponential temporal decay of infectivity (Hawkes, 1971).

The resulting epidemic component $e^*(t, \boldsymbol{s})$ is the sum of the contributions (2) of all infectious individuals at time t and location $\boldsymbol{s}$. Formally,

$$
\begin{aligned}
e^*(t, \boldsymbol{s}) &= \int_{(0,t) \times W \times \mathcal{M}} \mathbb{1}_{(0,\varepsilon]}(t - \tilde{t})\, \mathbb{1}_{[0,\delta]}(\|\boldsymbol{s} - \tilde{\boldsymbol{s}}\|)\, e^{\eta_j}\, g(t - \tilde{t})\, f(\boldsymbol{s} - \tilde{\boldsymbol{s}})\, N(\mathrm{d}\tilde{t} \times \mathrm{d}\tilde{\boldsymbol{s}} \times \mathrm{d}\tilde{\boldsymbol{m}})\,, \\
&= \sum_{j \in I^*(t,\boldsymbol{s})} e^{\eta_j}\, g(t - t_j)\, f(\boldsymbol{s} - \boldsymbol{s}_j)\,,
\end{aligned}
\tag{3}
$$

where $\mathcal{M}$ is the mark space, N is the time-space-mark point process counting the infections and $I^*(t, \boldsymbol{s}) := \left\{ j \in \{1, \dots, N_g(t-)\} : \mathbb{1}_{(0,\varepsilon]}(t - t_j) = 1 \ \wedge \ \mathbb{1}_{[0,\delta]}(\|\boldsymbol{s} - \boldsymbol{s}_j\|) = 1 \right\}$ is the history-dependent set of infectives at time t and location $\boldsymbol{s}$, where $N_g(t-) = N((0,t) \times W \times \mathcal{M})$. In the above, the hyperparameters $\varepsilon, \delta > 0$ are introduced as known *maximum* temporal and spatial interaction ranges. A past event only influences the process at time t and location $\boldsymbol{s}$, if both indicator functions are true, i.e. if it occurred at most ε time units ago at a location within distance δ.

2.3 *Characteristics of the Model*

Altogether, the proposed CIF model for a self-exciting spatio-temporal point process with components (1) and (3) is

$$
\lambda^*(t, \boldsymbol{s}) = \rho_{\tau(t),\xi(\boldsymbol{s})} \exp\left(\boldsymbol{\beta}' \boldsymbol{z}_{\tau(t),\xi(\boldsymbol{s})} \right) + \sum_{j \in I^*(t,\boldsymbol{s})} e^{\eta_j}\, g(t - t_j)\, f(\boldsymbol{s} - \boldsymbol{s}_j)\,,
$$

which we shall call `twinstim` to indicate a *two*-component *s*patio-*t*emporal (conditional) *i*ntensity *m*odel. For the proposed model an interesting quantity is the individual-specific mean number μ_j of infections caused by individual j inside its spatio-temporal range of interaction:

$$
\begin{aligned}
\mu_j &= \int_0^\infty \int_{\mathbb{R}^2} e_j(t, \boldsymbol{s})\, \mathbb{1}_{(0,\varepsilon]}(t - t_j)\, \mathbb{1}_{[0,\delta]}(\|\boldsymbol{s} - \boldsymbol{s}_j\|)\, \mathrm{d}t\, \mathrm{d}\boldsymbol{s} \\
&= e^{\eta_j} \cdot \int_0^\varepsilon g(t)\, \mathrm{d}t \cdot \int_{b(\boldsymbol{0},\delta)} f(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s}\,.
\end{aligned}
\tag{4}
$$

Here, $b(\boldsymbol{0}, \delta)$ denotes the disc centred at (0,0)' with radius δ. The integration domain $\mathbb{R}_+ \times \mathbb{R}^2$ above stems from the theoretical point of view that the point process occurs in unlimited time and space. In practice this is not observable, but individuals near the border would be attributed a truncated value of μ_j if integrating over W – or, similarly, $[0, T]$ – only. Such edge effects are overcome by (4), which also simplifies interpretation by providing a quantity similar to the basic reproduction number R_0 known from classical epidemic modelling. Specifically, the number μ_j offers an intuitive way of interpreting the parameters $\boldsymbol{\gamma}$ in the linear predictor η_j, because they can be handled as usual in Poisson regression models: a unit positive change in a specific continuous mark m_{jl} multiplies the mean number of infections by the corresponding parameter e^{γ_l}.

2.4 *Extension: Type-Specific* `twinstim`

Although the model of the previous subsection allows for a finetype-specific infectivity through the vector of unpredictable marks $\boldsymbol{m}_j$, it is not applicable for a joint modelling of both finetypes. This

is because finetypes do not change during transmission. Therefore, the point process model will be extended to a marked version suitable for the specific application of IMD and point patterns with different event types in general.

Denote by $\mathcal{K} = \{1, \ldots, K\} \subset \mathbb{N}$ the set of possible event types. Define an indicator matrix $\boldsymbol{Q} = (q_{k,l})_{k,l \in \mathcal{K}}$, where $q_{k,l} \in \{0; 1\}$, which determines the possible ways of transmission. If $q_{k,l}$ equals 1, an infective type k event can cause an event of type l. For instance, the IMD data would require $\boldsymbol{Q} = \boldsymbol{I}_2$, because the transmission is finetype-specific. A marked spatio-temporal point process on $(0, T] \times W \times \mathcal{K}$ is then defined by the following model for the CIF:

$$
\begin{aligned}
\lambda^*(t, \boldsymbol{s}, \kappa) &= h(t, \boldsymbol{s}, \kappa) + e^*(t, \boldsymbol{s}, \kappa) && (5) \\
h(t, \boldsymbol{s}, \kappa) &= \rho_{\tau(t), \xi(\boldsymbol{s})} \exp\left(\beta_0(\kappa) + \boldsymbol{\beta}' \boldsymbol{z}_{\tau(t), \xi(\boldsymbol{s})}\right) \\
e^*(t, \boldsymbol{s}, \kappa) &= \sum_{j \in I^*(t, \boldsymbol{s}, \kappa)} e_j(t, \boldsymbol{s}) \\
e_j(t, \boldsymbol{s}) &= \exp(\eta_j) \cdot g(t - t_j | \kappa_j) \cdot f(\boldsymbol{s} - \boldsymbol{s}_j | \kappa_j) \\
I^*(t, \boldsymbol{s}, \kappa) &= \left\{ j \in \{1, \ldots, N_g(t-)\} : \mathbb{1}_{(0, \varepsilon]}(t - t_j) = 1 \;\wedge\; \mathbb{1}_{[0, \delta]}(\|\boldsymbol{s} - \boldsymbol{s}_j\|) = 1 \;\wedge\; q_{\kappa_j, \kappa} = 1 \right\}.
\end{aligned}
$$

Here, the transmission indicators from the matrix $\boldsymbol{Q}$ have been integrated into $I^*(t, \boldsymbol{s}, \kappa)$. Note that the event type κ_j is now part of the vector $\boldsymbol{m}_j$, which enables type-specific epidemic intercepts as well as type interactions with individual covariates in the linear predictor η_j. The new endemic intercept $\beta_0(\kappa)$ either represents a type-specific endemic intercept, i.e. $\beta_0(\kappa) = \sum_{k=1}^{K} \beta_{0,k} \mathbb{1}_{\{k=\kappa\}}(\kappa) = \beta_{0,\kappa}$, or contains only a single global intercept $\beta_0(\kappa) = \beta_0$, corresponding to the hypothesis $\beta_0 = \beta_{0,1} = \cdots = \beta_{0,K}$. For the remainder of the endemic predictor, the model assumes independence of κ, which means that the effect of endemic covariates is homogeneous over the event types. However, the history-dependent set $I^*(t, \boldsymbol{s}, \kappa)$ of infective individuals now accounts for the transmission regime $\boldsymbol{Q}$ between the event types, and the interaction functions are allowed to depend on the type of the infective event as well.

3. Statistical Inference

This section deals with likelihood inference for the parameters of the CIF in (5) based on the observed marked spatio-temporal point pattern $\boldsymbol{x} = \{(t_i, \boldsymbol{s}_i, \boldsymbol{m}_i) : i = 1, \ldots, n\}$, where the event type κ_i is part of the vector of marks $\boldsymbol{m}_i$, and n is the number of events, i.e. a realisation of $N_g(T)$. The parameter vector in question is $\boldsymbol{\theta} = (\boldsymbol{\beta}_0', \boldsymbol{\beta}', \boldsymbol{\gamma}', \boldsymbol{\sigma}', \boldsymbol{\alpha}')'$, where $\boldsymbol{\sigma}$ and $\boldsymbol{\alpha}$ are the parameter vectors of the spatial and temporal interaction functions $f_{\boldsymbol{\sigma}}$ and $g_{\boldsymbol{\alpha}}$, respectively.

In our framework, no attempt is made to model unpredictable marks like gender and age but they are taken as given predictor variables in models of the CIF. In this case, the log-likelihood of the underlying point process N on $[0; T] \times W \times \mathcal{M}$ may be conveniently written as (Daley and Vere-Jones, 2003)

$$
\sum_{i=1}^{n} \log \lambda_{\boldsymbol{\theta}}^*(t_i, \boldsymbol{s}_i, \kappa_i) - \int_0^T \int_W \sum_{\kappa \in \mathcal{K}} \lambda_{\boldsymbol{\theta}}^*(t, \boldsymbol{s}, \kappa) \, dt \, d\boldsymbol{s} \, .
$$

The components of the above sum can be directly calculated for a specific value of the parameter vector $\boldsymbol{\theta}$ after having determined the set $I^*(t_i, \boldsymbol{s}_i, \kappa_i)$ of potential sources of infection for the ith event. Furthermore, the integrations of the endemic and epidemic components of the CIF can be performed separately due to their additive superposition. Recalling that the endemic component is a piecewise constant function on the spatio-temporal grid $\{C_1, \ldots, C_D\} \times \{A_1, \ldots, A_M\}$, its integral

6

is in fact a sum over this grid of smallest observed units in space-time:

$$\int_0^T \int_W \sum_{\kappa \in \mathcal{K}} h_{\boldsymbol{\theta}}(t, \boldsymbol{s}, \kappa) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{s} = \left(\sum_{\kappa \in \mathcal{K}} \exp\left(\beta_0(\kappa)\right) \right) \cdot \sum_{\tau=1}^{D} \sum_{\xi=1}^{M} |C_\tau| |A_\xi| \rho_{\tau,\xi} \exp\left(\boldsymbol{\beta}' \boldsymbol{z}_{\tau,\xi}\right) . \tag{6}$$

The integrated epidemic component can be simplified by moving the indicators of the function $I^*(t, \boldsymbol{s}, \kappa)$ back into the sum:

$$\int_0^T \int_W \sum_{\kappa \in \mathcal{K}} e_{\boldsymbol{\theta}}^*(t, \boldsymbol{s}, \kappa) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{s}$$

$$= \int_0^T \int_W \sum_{\kappa \in \mathcal{K}} \sum_{j=1}^{n} \mathbb{1}_{(0,\varepsilon]}(t - t_j) \, \mathbb{1}_{[0,\delta]}(\|\boldsymbol{s} - \boldsymbol{s}_j\|) \, q_{\kappa_j, \kappa} \, e^{\eta_j} \, g_{\boldsymbol{\alpha}}(t - t_j | \kappa_j) \, f_{\boldsymbol{\sigma}}(\boldsymbol{s} - \boldsymbol{s}_j | \kappa_j) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{s}$$

$$= \sum_{j=1}^{n} q_{\kappa_j, \cdot} \, e^{\eta_j} \left(\int_0^{\min\{T - t_j; \varepsilon\}} g_{\boldsymbol{\alpha}}(t | \kappa_j) \, \mathrm{d}t \right) \left(\int_{R_j} f_{\boldsymbol{\sigma}}(\boldsymbol{s} | \kappa_j) \, \mathrm{d}\boldsymbol{s} \right) . \tag{7}$$

Here, $q_{\kappa_j, \cdot} := \sum_{\kappa \in \mathcal{K}} q_{\kappa_j, \kappa}$ is the number of different event types that can be triggered by an event of type κ_j, and $R_j := \left\{ W \cap b(\boldsymbol{s}_j; \delta) \right\} - \boldsymbol{s}_j$ is the spatial interaction region of the jth event centred at its location.

The evaluation of the two-dimensional integral over the domains R_j is the most sophisticated task of the log-likelihood evaluation. Meyer (2009) compared accuracy and speed of different cubature rules for performing the numerical integration. Here, the two-dimensional midpoint rule (see e.g. Stroud, 1971) proved to be best suited for the task. In contrast, the evaluation of the definite integral over the temporal interaction function is analytically accessible for typical choices of $g_{\boldsymbol{\alpha}}$.

Altogether, an analytical maximisation of the above log-likelihood is not feasible, and a numerical optimisation routine such as BFGS (see e.g. Nocedal and Wright, 1999, Section 8.1) is required. Here, it is advantageous to know the score function $s(\boldsymbol{\theta})$, which is derived in Web Appendix A. Uncertainty of the parameter estimates is deduced from the expected Fisher information $\mathcal{I}(\boldsymbol{\theta})$ as estimated by the "optional variation process" adapted to the marked spatio-temporal setting – see Web Appendix B for details. Significance of specific model parameters can be investigated by Wald or likelihood ratio tests and model selection is performed based on Akaike's information criterion (AIC).

4. Simulation Algorithm

In general, the usability of a model class is greatly improved by the ability to simulate from a specific model. For instance, it enables model checking and parametric bootstrap. For evolutionary point processes specified by their CIF, *Ogata's modified thinning algorithm* (Daley and Vere-Jones, 2003, Algorithm 7.5.V.) provides a convenient and exact way to simulate realisations of the process. The algorithm requires piecewise upper bounds for the intensity $\lambda_g^*(t)$ of the ground process $N_g(t) := N((0, t] \times W \times \mathcal{K})$. This intensity is determined as

$$\lambda_g^*(t) = \int_W \sum_{\kappa \in \mathcal{K}} \lambda^*(t, \boldsymbol{s}, \kappa) \, \mathrm{d}\boldsymbol{s} = \left(\sum_{\kappa \in \mathcal{K}} e^{\beta_0(\kappa)} \right) \left(\sum_{\xi=1}^{M} |A_\xi| \, \rho_{\tau(t),\xi} \, e^{\boldsymbol{\beta}' \boldsymbol{z}_{\tau(t),\xi}} \right)$$

$$+ \sum_{j=1}^{N_g(t-)} \left(\sum_{\kappa \in \mathcal{K}} q_{\kappa_j, \kappa} \right) e^{\eta_j} \, \mathbb{1}_{(0,\varepsilon]}(t - t_j) \, g(t - t_j | \kappa_j) \int_{R_j} f(\boldsymbol{s} | \kappa_j) \, \mathrm{d}\boldsymbol{s} .$$

This function is bounded above by the CIF $\overline{\lambda}_g^*(t)$, which is defined by replacing $g(t|\kappa)$ by the *constant* temporal interaction function $\overline{g}(t|\kappa) = \max_{u>0} g(u|\kappa)$. This CIF is piecewise constant in time as it only jumps at time points where any of the endemic covariates in $\boldsymbol{z}_{\tau(t),\xi}$ in any tile ξ changes its value, or when the set of currently infectious individuals changes, i.e. whenever a new event occurs or a previous event stops triggering.

Given a parameter vector $\boldsymbol{\theta}$, the ranges of interaction ε and δ, as well as a sampling scheme for the marks $\boldsymbol{m}_j$, the time point of the next infection starting from the current time $t = t_0$ can be generated as follows: Draw an exponentially distributed random variate Δ with rate $\overline{\lambda}_g^*(t_0)$. The simulated value of Δ is a proposal for the waiting time to the next event, i.e. the next time point of infection might be $\tilde{t} = t_0 + \Delta$. However, this proposal is not valid if the rate $\overline{\lambda}_g^*(t)$ had changed between t_0 and $\tilde{t}$. In this case, time is set to the first changepoint after t_0 and a new Δ is simulated. Eventually, a proposed time point $\tilde{t}$ is valid. It is then accepted with probability $\lambda_g^*(\tilde{t})/\overline{\lambda}_g^*(\tilde{t})$. If it is rejected, time is set to $t = \tilde{t}$ and a new waiting time Δ is simulated as above. If it is accepted, location $\tilde{\boldsymbol{s}}$ and type $\tilde{\kappa}$ of the event have to be simulated. At first, the source of infection is sampled with probabilities proportional to the respective components of $\lambda_g^*(\tilde{t})$:

$$\mathbb{P}(\text{endemic source}) \cdot \lambda_g^*(\tilde{t}) = \left(\sum_{\kappa \in \mathcal{K}} e^{\beta_0(\kappa)}\right) \left(\sum_{\xi=1}^{M} |A_\xi| \, \rho_{\tau(\tilde{t}),\xi} \, e^{\boldsymbol{\beta}' \boldsymbol{z}_{\tau(\tilde{t}),\xi}}\right) \tag{8}$$

$$\mathbb{P}(\text{source} = \text{event } j) \cdot \lambda_g^*(\tilde{t}) = \left(\sum_{\kappa \in \mathcal{K}} q_{\kappa_j,\kappa}\right) e^{\eta_j} \mathbb{1}_{(0,\varepsilon]}(\tilde{t} - t_j) \, g(\tilde{t} - t_j|\kappa_j) \int_{R_j} f(\boldsymbol{s}|\kappa_j) \, \mathrm{d}\boldsymbol{s} \, ,$$

for $j \in \{1, \ldots, N_g(\tilde{t}-)\}$. On the one hand, if the new event has an endemic source, then $\mathbb{P}(\tilde{\kappa} = k) \propto \exp(\beta_0(k))$, $k \in \mathcal{K}$, and $\mathbb{P}(\tilde{\boldsymbol{s}} \in A_\xi) \propto |A_\xi| \, \rho_{\tau(\tilde{t}),\xi} \, e^{\boldsymbol{\beta}' \boldsymbol{z}_{\tau(\tilde{t}),\xi}}$, $\xi = 1, \ldots, M$. In the sampled tile $A_{\tilde{\xi}}$, the location $\tilde{\boldsymbol{s}}$ is uniformly distributed. On the other hand, if the new event was triggered by the previous event j, then $\tilde{\kappa} \sim U(\{k : q_{\kappa_j,k} = 1\})$, and $\tilde{\boldsymbol{s}} = \boldsymbol{s}_j + \boldsymbol{v}$, where $\boldsymbol{v}$ is drawn from the density $f(\boldsymbol{s}|\kappa_j)/\int_{R_j} f(\boldsymbol{s}|\kappa_j) \, \mathrm{d}s$ on R_j, e.g. using rejection sampling.

A scheme of the described algorithm can be found as Web Appendix C.

5. Application to the IMD Data

Although visual comparisons between the finetypes and heuristic comparisons of the estimates of separate finetype-specific models are possible, this does not allow to assess potential differences statistically. We thus conduct a joint analysis of the two finetypes by the marked `twinstim` of Section 2.4. We perform model selection for the joint point pattern of 630 cases of IMD with complete age and gender information by using AIC to compare all models with the CIF composed by subsets of the following terms:

- Endemic component: common or finetype-specific intercept, linear time trend, time-of-year effects (one or two harmonics), and linear effect of weekly number of influenza cases registered in the district of a point (no time lag, lags 0 and 1, lags 0–2, or lags 0–3) taken from the SurvStat database (Robert Koch-Institut, 2009).
- Epidemic component: gender, age (categorized as 0-2, 3-18 and $\geqslant 19$ years), finetype, and age-finetype interaction.

As an offset in the endemic component, we use the district-specific population density $\rho_{\xi(\boldsymbol{s})}$ (inhabitants per km^2). A fixed hyperparameter of $\varepsilon = 30$ days is assumed – this maximal temporal interaction range is consistent with the range used in, e.g., Zangwill et al. (1997). Because the

Table 1

Parameter estimates for the endemic (top) and epidemic (bottom) component of the model with the lowest AIC (AIC=18968). The p-values correspond to Wald tests.

	Estimate	Std. Error	z value	$\mathbb{P}(|Z| > |z|)$
β_0	-20.3652	0.0872	-233.53	$< 2 \cdot 10^{-16}$
β_{trend}	-0.0493	0.0223	-2.21	0.027
β_{sin}	0.2618	0.0649	4.03	$5.5 \cdot 10^{-05}$
β_{cos}	0.2668	0.0644	4.14	$3.4 \cdot 10^{-05}$
γ_0	-12.5746	0.3128	-40.21	$< 2 \cdot 10^{-16}$
$\gamma_{3\text{-}18}$	0.6463	0.3195	2.02	0.04310
$\gamma_{\geqslant 19}$	-0.1868	0.4321	-0.43	0.66558
γ_{C}	-0.8496	0.2574	-3.30	0.00097
$\log \sigma$	2.8287	0.0819		

number of supposedly direct transmissions in the IMD dataset is humble, we will furthermore assume a constant temporal interaction function g (i.e. constant spread within the ε days) in order to not overparametrize the epidemic component. The spatial hyperparameter is fixed at $\delta = 200$ km – this parameter needs only to be large enough not to influence the estimation of the actual spatial interaction function f.

To restrict the model search, and hence computing time, we first performed the search for all 600 models ($2 \cdot 2 \cdot 3 \cdot 5$ configurations of the endemic component and $2 \cdot 5$ configurations of the epidemic component) with constant spatial interaction function f. Hereafter, the top 10 models of this search were investigated further with two Gaussian spatial interaction functions: one with joint variance parameter and one with finetype-specific variance parameter.

The CIF of the resulting AIC-best model obtained by this search was $\lambda_{\boldsymbol{\theta}}^*(t, \boldsymbol{s}, \kappa) =$

$$\rho_{\xi(\boldsymbol{s})} \cdot \exp\left(\beta_0 + \beta_{\text{trend}} \frac{\lfloor t \rfloor}{365} + \beta_{\text{sin}} \sin\left(\lfloor t \rfloor \frac{2\pi}{365}\right) + \beta_{\text{cos}} \cos\left(\lfloor t \rfloor \frac{2\pi}{365}\right)\right)$$

$$+ \sum_{j \in I^*(t, \boldsymbol{s}, \kappa)} q_{\kappa_j, \kappa} \, \exp\left(\gamma_0 + \gamma_{3\text{-}18} \mathbb{1}_{[3,18]}(\text{age}_j) + \gamma_{\geqslant 19} \mathbb{1}_{[19,\infty)}(\text{age}_j) + \gamma_{\text{C}} \mathbb{1}_{\{C\}}(\kappa_j)\right) \, f_\sigma(\boldsymbol{s} - \boldsymbol{s}_j).$$

Here, $(t, \boldsymbol{s}, \kappa)$ denotes days since 31 December 2001, coordinate in ETRS89 (kilometre scale) and finetype. With $\lfloor t \rfloor$ we denote monday of week $\tau(t)$, i.e. the lower bound of time intervals $C_1, \ldots, C_D$. In the linear predictor of the epidemic component, age group 0-2 and type B serve as reference categories. The corresponding parameter estimates of the best model, now fitted to the 635 cases with available age, are found in Table 1.

Thus, there appears to be no noteworthy difference in the endemic behaviour of the two types: a linear downward time trend superimposed with one harmonic best describes the endemic behaviour of the point pattern (see Figure 3(a)). An additional effect of past numbers of influenza cases does not improve the model. In contrast, there is an effect of past IMD cases, i.e. the process is indeed self-exciting. Comparing the endemic-only model with the model enriched by an epidemic intercept only, greatly improves the fit (ΔAIC=202.84). In the epidemic component, there is a detectable dependence on marks with type C being less aggressive than type B. Figure 3(b) shows the resulting finetype-specific spatial interaction functions which for type C is $e^{\hat{\gamma}_C} \cdot 100\% = 43\%$ of type B. Finally, there is a significant age difference in the infectivity of cases: the highest potential is found in the 3-18 year old, which could be interpreted as the kindergarten and school-aged children having a higher contact behaviour than e.g. adults.

Based on the selected model, basic reproduction numbers of $\hat{\mu}_{\text{B}} = 0.25$ (95% CI 0.19-0.34) vs. $\hat{\mu}_{\text{C}} = 0.11$ (95% CI 0.07-0.17) are obtained by calculating the type-specific expectation of (4) over the empirical distribution function of the additional covariates in the epidemic predictor (here: age
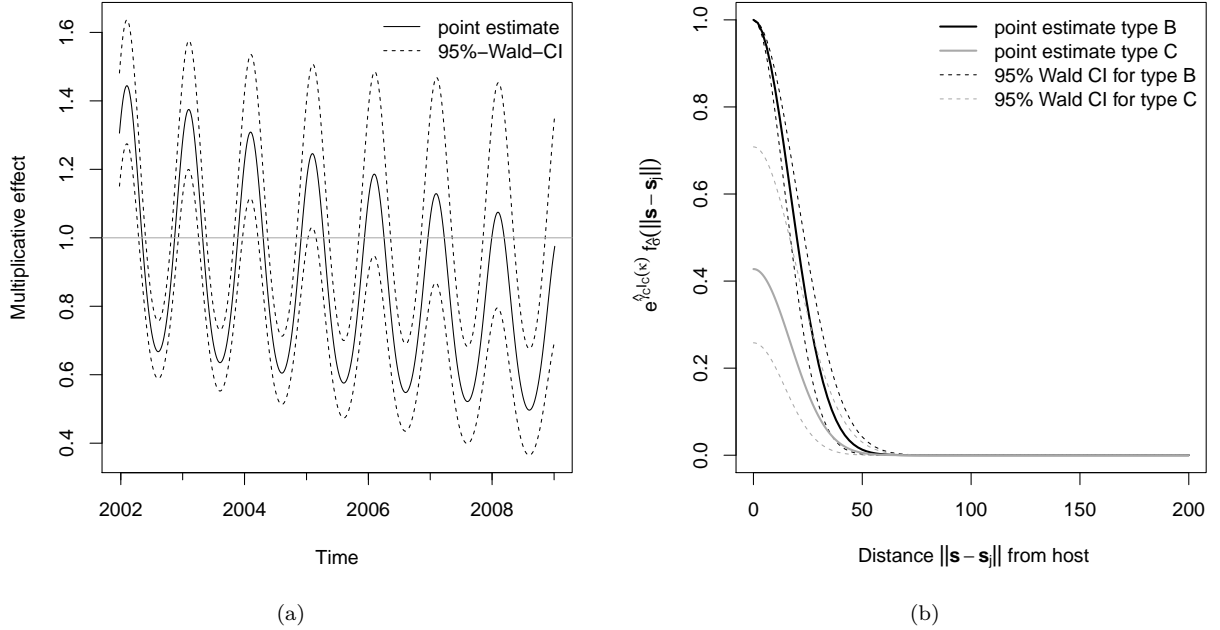
Figure 3. (a) Trend and seasonal component of the fitted model; one observes the typical IMD peak in late February and minimum in August. Furthermore, (b) shows the spatial interaction function multiplied by the type modifier illustrating the higher epidemic potential of type B.

group). The confidence intervals are given as the 0.025 and 0.975 quantiles of samples obtained by re-computing $\hat{\mu}_B$ and $\hat{\mu}_C$ for 999 additional coefficient vectors drawn from the asymptotic multivariate normal distribution of the parameter estimates in Table 1. The confidence intervals thus indicate a higher epidemic potential of the serogroup B finetype. Note that these numbers are lower than what one would expect from the literature, e.g. Trotter et al. (2005) report an R_0 estimate of 1.36 for *serogroup* C. Two explanations account for this discrepancy: firstly, our estimation is based on transmission between cases with invasive disease and not between asymptomatic carriers, who are not represented in disease surveillance data. Secondly, use of an endemic component means that our R_0 estimates are destined to be lower, because sporadic cases do not contribute to the number of secondary cases. Still, our estimates provide realistic lower bounds for carriage reproduction numbers.

To inspect the goodness-of-fit of the selected spatio-temporal point process model, we follow the suggestion by Ogata (1988) (see also Rathbun, 1996) by computing $Y_i = \hat{\Lambda}_g^*(t_i) - \hat{\Lambda}_g^*(t_{i-1})$, $i = 2, \ldots, n$, where $\hat{\Lambda}_g^*(t)$ is the fitted cumulative intensity function of the ground process. If the estimated CIF describes the true CIF well, then $U_i = 1 - \exp(-Y_i) \overset{\text{iid}}{\sim} U(0,1)$. Figure 4(a) contains a plot of the cumulative density function (CDF) of the observed U_i and for comparison the CDF of the $U(0,1)$-distribution together with error bounds computed by inverting the one sample Kolmogorov-Smirnov test. The fit appears good, but noticable deviations for $u_i < 0.15$ can be observed, which we suspect to occur due to the tie-breaking strategy of subtracting $\epsilon = 0.01$ days from ties. As observations are on a per-day basis and thus are interval censored we re-estimated the model for a data set where ties were broken by subtracting a $U(0,1)$-distributed random number from each observation time. Figure 4(b) shows the improved fit of this analysis – the relative changes in the parameter estimates are minor.
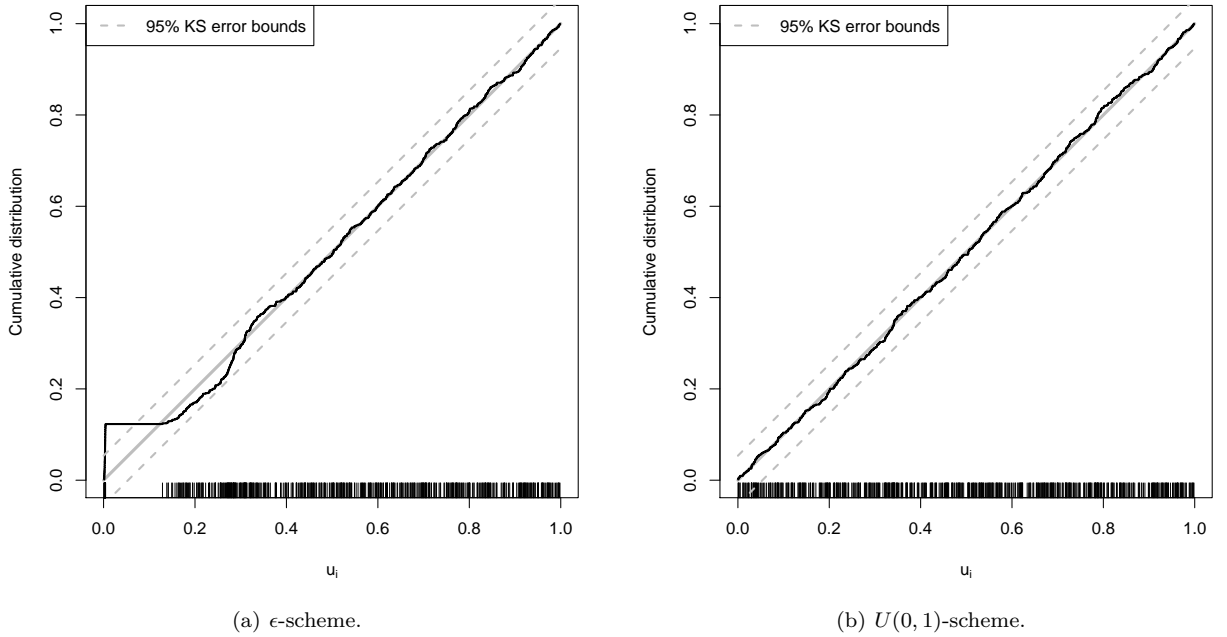
10

(a) ϵ-scheme.

(b) $U(0,1)$-scheme.

Figure 4. CDF of the observed U_i together with 95% Kolmogorov-Smirnov error bounds for data with tie breaking according to the (a) ϵ scheme and (b) $U(0,1)$ scheme.

Another way of assessing the goodness-of-fit is by simulation from the fitted CIF. Figure 5 shows the observed 7-year incidences (per 100,000 inhabitants) of the 413 districts for both finetypes together. In order to identify extreme observations that are not explained by the selected model, we simulated 100 realisations of the process and determined the 2.5% and 97.5% quantiles of the district-specific 7-year incidences. In the figure, districts with observed incidences outside the simulated 95%-range are marked by triangles. Many of the 17 districts with an excess are found around the city Aachen at the border to the Netherlands. The deviation from the model could thus be explained by edge effects hiding potential transmissions across the border.

Altogether, we are led to the conclusion that the proposed model provides a useful description of the spread of IMD. It allows a quantification that the serogroup B finetype has a higher epidemic potential than the serogroup C finetype and shows age difference in spread behaviour. A sensitivity analysis confirmed robustness of these results for increasing values of δ. Order and significance of the finetype difference in the epidemic component remained stable for ε in the range of 1-5 weeks to 1-4 months. Age group results were slightly more varying: the 3-18 year olds remain having the highest epidemic potential, but from $\varepsilon > 35$ days on, the oldest age group comes in second. The sensitivity analysis also showed, that there is too little information to estimate ε from the IMD data – we are thus forced to fix the hyperparameter at a biological plausible value.

6. Discussion

We presented a comprehensive framework for modelling, inference and simulation for infectious disease occurrence data. In the case of IMD, the infected individual is effectively removed from the transmission network once the disease becomes manifest. Secondary cases are thought to acquire the infective strain either from the case during incubation or from asymptotic carriers close to the case. Although marks attached to the case can naturally not account for the latter mode of
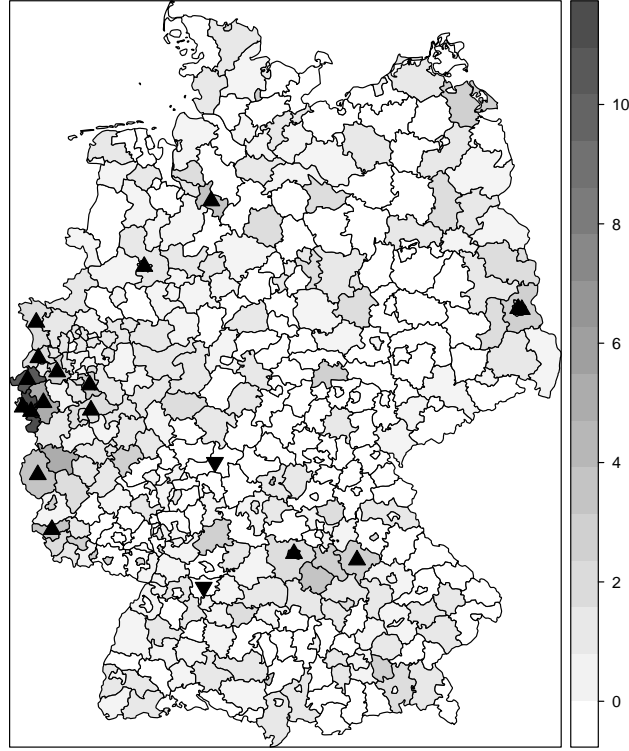
Figure 5. Observed incidence (per 100,000 inhabitants) during 2002-2008 for both finetypes together. Triangles pointing up (down) indicate districts with a higher (lower) incidence than explained by 100 simulations from the model.

transmission, they represent a valid proxy for the transmission network of the case when analysing surveillance data, which typically lack information regarding carriage.

Despite use of disease surveillance data, we were able to quantify differences in IMD transmission dynamics based on age and finetype. That the modelling requires an epidemic component is of epidemiologic interest in its own, as this shows that IMD incidence goes beyond sporadic occurrences. To our knowledge, our analysis is the first report of finetype-specific differences in spread tendencies. Contrary to previous analyses we were not able to find a significant connection between IMD and concurrent number of influenza cases. The spatial spread appeared to happen at a rather small scale – a scale which the usual district resolution data collected as part of the German Infection Protection Act does not allow to analyse. Thus our work is also a contribution to the controversy between patient privacy and the need for high-resolution data to gain new epidemiological insights. One important question in this debate is how good a proxy the patient's residence is for his general whereabouts.

Even though our CIF modelling is similar in form to the proposal in Höhle (2009), the *continuous space* of the IMD application makes epidemic modelling conceptually different. The classical SIR model framework does not apply in this situation, because events do not originate from a predefined population and individuals can not be partitioned into model compartments anymore. Thus, including population density becomes important and one needs to distinguish between covariate information of events and covariates of the surrounding environment within which the process occurs. Furthermore, likelihood inference is complicated by requiring an additional integration over space

12

for complex polygons. Finally, the now proposed space-time interaction functions are completely general in form and thus provide an advantage over the previous linear basis decomposition and resulting parameter constraints.

An issue currently not dealt with in our estimation are edge effects, i.e. data are only available for Germany, but infections occur outside the observation window. For example, Elias et al. (2010) investigate the contribution of cross-border spread to increased incidence of IMD in the German region of Aachen neighbouring the Netherlands. A cross-border effect is indeed detected by our simulation in Figure 5 where the Aachen region has higher observed incidences than can be explained by our model. Hence, the actual disease clusters are wider than observed in Germany, which potentially causes underestimation of the epidemic weight. Edge correction for inference in spatio-temporal point processes is, however, still an open methodological issue.

An additional strength of the proposed modelling is that it offers a parametric framework for conducting prospective change-point analysis in spatio-temporal point processes typical in disease surveillance: Within the framework of stochastic process control one could e.g. use likelihood ratio detectors to monitor the time point where inclusion of an epidemic component is necessary to describe the observed data. This would correspond in idea to the time series setting investigated in Höhle and Paul (2008) or the homogeneous spatio-temporal Poisson process setting of Assunçáo and Correa (2009).

The presented methods for inference and simulation of `twinstim` models are available as part of the `R` package `surveillance` (Höhle et al., 2011; Höhle, 2007) available from the Comprehensive `R` Archive Network.

7. Supplementary Materials

The Web Animation referenced in Section 1 and the Web Appendices referenced in Sections 3 and 4 are available under the Paper Information link at the *Biometrics* website `http://www.biometrics.tibs.org/`.

References

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

Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes*, volume I: Elementary Theory and Methods of *Probability and its Applications*. Springer-Verlag, New York, 2nd edition.

Diggle, P., Rowlingson, B., and li Su, T. (2005). Point process methodology for on-line spatio-temporal disease surveillance. *Environmetrics* **16,** 423–434.

Diggle, P. J. (2006). Spatio-temporal point processes, partial likelihood, foot and mouth disease. *Statistical Methods in Medical Research* **15,** 325–336.

Diggle, P. J., Kaimi, I., and Abellana, R. (2009). Partial-likelihood analysis of spatio-temporal point-process data. *Biometrics* **66,** 347–354.

Elias, J., Schouls, L. M., van de Pol, I., Keijzers, W. C., Martin, D. R., Glennie, A., Oster, P., Frosch, M., Vogel, U., and van der Ende, A. (2010). Vaccine preventability of meningococcal clone, Greater Aachen Region, Germany. *Emerging Infectious Diseases* **16,** 465–472.

Federal Statistical Office (DESTATIS) (2009). Gemeindeverzeichnis GV 2000. Districts as of 31/12/2008. Data as of 31/12/2007.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58,** 83–90.

Held, L., Höhle, M., and 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,** 571–582.

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

Höhle, M., Meyer, S., and Paul, M. (2011). `surveillance: Temporal and spatio-temporal modeling and monitoring of epidemic phenomena`. R package version 1.3-1.

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

Jensen, E. S., Lundbye-Christensen, S., Samuelsson, S., Sørensen, H. T., and Schønheyder, H. C. (2004). A 20-year ecological study of the temporal association between influenza and meningococcal. *European Journal of Epidemiology* **19,** 181–187.

Jewell, C. P., Kypraios, T., Neal, P., and Roberts, G. O. (2009). Bayesian analysis for emerging infectious diseases. *Bayesian Analysis* **4,** 465–496.

Meyer, S. (2009). Spatio-temporal infectious disease epidemiology based on point processes. Master's thesis, Department of Statistics, Ludwig-Maximilians-Universität, München. Available as `http://epub.ub.uni-muenchen.de/11703/`.

Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P., and Tita, G. E. (2010). Self-exciting point process modeling of crime. Technical report, Department of Mathematics and Computer Science, Santa Clara University.

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

Nocedal, J. and Wright, S. J. (1999). *Numerical Optimization.* Springer.

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

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics* **50,** 379–402.

Ogata, Y. (1999). Seismicity analysis through point-process modeling: A review. *Pure and Applied Geophysics* **155,** 471–507.

Peng, R. D., Schoenberg, F. P., and Woods, J. A. (2005). A space-time conditional intensity model for evaluating a wildfire hazard index. *Journal of the American Statistical Association* **100,** 26–35.

Rathbun, S. L. (1996). Asymptotic properties of the maximum likelihood estimator for spatio-temporal point processes. *Journal of Statistical Planning and Inference* **51,** 55–74.

Robert Koch-Institut (2009). SurvStat@RKI. `http://www3.rki.de/SurvStat`.

Rosenstein, N. E., Perkins, B. A., Stephens, D. S., Popovic, T., and Hughes, J. M. (2001). Meningococcal Disease. *The New England Journal of Medicine* **344,** 1378–1388.

Scheel, I., Aldrin, M., Frigessi, A., and Jansen, P. A. (2007). A stochastic model for infectious salmon anemia (ISA) in Atlantic salmon farming. *Journal of the Royal Society, Interface* **4,** 699–706.

Stroud, A. H. (1971). *Approximate Calculation of Multiple Integrals.* Prentice Hall, Englewood Cliffs.

Trotter, C., Gay, G. J., and Edmunds, W. J. (2005). Dynamic models of meningococcal carriage, disease, and the impact of serogroup c conjugate vaccination. *American Journal of Epidemiology* **162,** 89–100.

Zangwill, K. M., Schuchat, A., Riedo, F. X., Pinner, R. W., Koo, D. T., Reeves, M. W., and Wenger, J. D. (1997). School-based clusters of meningococcal disease in the United States. *JAMA* **277,** 389–395.