RESEARCH ARTICLE

# Modeling latent spatio-temporal disease incidence using penalized composite link models

**Dae-Jin Lee**[1]*, **María Durbán**[2], **Diego Ayma**[3], **Jan Van de Kassteele**[4]

**1** BCAM - Basque Center for Applied Mathematics, Bilbao, Bizkaia, Spain, **2** Department of Statistics, Universidad Carlos III de Madrid, Leganés, Madrid, Spain, **3** Facultad de Ciencias, Universidad Católica Norte, Antofagasta, Chile, **4** RIVM - National Institute for Public Health and the Environment, Bilthoven, Utrecht, The Netherlands

* dlee@bcamath.org

## Abstract

Epidemiological data are frequently recorded at coarse spatio-temporal resolutions to protect confidential information or to summarize it in a compact manner. However, the detailed patterns followed by the source data, which may be of interest to researchers and public health officials, are overlooked. We propose to use the penalized composite link model (Eilers PCH (2007)), combined with spatio-temporal P-splines methodology (Lee D.-J., Durban M (2011)) to estimate the underlying trend within data that have been aggregated not only in space, but also in time. Model estimation is carried out within a generalized linear mixed model framework, and sophisticated algorithms are used to speed up computations that otherwise would be unfeasible. The model is then used to analyze data obtained during the largest outbreak of Q-fever in the Netherlands.

## Introduction

In recent decades, the development of spatio-temporal statistical methods in Epidemiology, the cornerstone of Public Health, has been remarkable. Such development has been mainly driven by advances in geographic information systems (GISs), access to reliable health data registers, and the availability of powerful software capabilities to process and analyze large amounts of data. The methodological contributions to the analysis of spatio-temporal health data come from several interdisciplinary researchers, whose backgrounds are mostly related to Statistics, Geography, Environmental Sciences, and Epidemiology.

Within the diversity of of epidemiological research, disease mapping has attracted much interest in Public Health as it helps to visualize disease incidence of mortality risk patterns in a specific area. To this end, appropriate statistical methods have been applied to health data, which are usually recorded per spatial units, to provide smoothed disease incidences per unit. Smoothing is performed to obtain more stable and less noisy estimates of the incidence rates associated with each unit [1], which helps to determine meaningful spatial patterns. By adding the temporal dimension to this context, it is possible to examine the evolution of disease

incidence in each unit, during a certain period of time (generally divided in years), but it implies a challenge for smoothing data, in terms of computational time and storage, even more if disaggregation is also needed in time. Several techniques have been proposed for the spatio-temporal smoothing of health data; most of which are developed under an empirical Bayes approach, where B-splines are used [2, 3] or a hierarchical Bayesian framework where conditional autoregressive (CAR) structures are included [4–6]. Regarding this last approach, methods that use integrated nested Laplace approximations (INLA, [7]) have recently been proposed (see [8–10]; among others).

All the works mentioned above provide smoothed estimates that change over time within a unit and across units within time interval, although each estimated value is assigned to the whole unit. Furthermore, most of them can be extended to include covariates or explanatory variables, which must have the same spatio-temporal resolution as health data. Therefore, they restrict the incorporation of population information at fine-scale and other relevant risk factors recorded at a finer resolution (see the *modifiable areal unit problem*; [11, Ch. 29]). The challenge here is not only that the map is dynamic, but the fact that, in order to obtain a better insight of the evolution of the incidence, detailed maps are needed at a finer spatial and temporal scale than the one provided (for example, grid instead of areas and weeks instead of months) while maintaining the coherence with the aggregated observed counts. To overcome the limitations of previous works, we propose the use of the *spatio-temporal penalized* composite link model(ST-PCLM) to estimate the latent distribution of spatio-temporally grouped count data. Our proposal is based on the penalized composite link model of [12], combined with a spatio-temporal P-spline methodology [13], to obtain smoothed estimates at a finer resolution from data aggregated over space and/or time. Moreover, the approach allows to include population information at fine-scale and specific random effects or further correlation structure, if necessary. This could be in the form of unstructured variation (including random effects per spatial unit) or structured variation by assuming a known variance-covariance matrix for the spatial random effects, for example using a CAR model as in [14].

As in any latent observation problem, there are infinitely possible solutions that fit the observed data. In that sense, the model is not identifiable, and the use of some prior knowledge or constraints (based on experience or common sense) is required for the problem to be well posed. Because we want to estimate a distribution, it is reasonable to expect a relatively smooth result, so we assume here that the underlying spatio-temporal process behind aggregated data is smooth and stationary (although this last hypothesis could also be relaxed by using multidimensional adaptive smoothing as in [15]). The flexibility of the model is given by the use of B-spline bases and a discrete penalty on the regression coefficients, following the P-spline methodology [16]. Smoothness is controlled by three smoothing parameters (two for the spatial dimension and one for the temporal dimension) that have to be estimated along with the overall trend at the fine scale. To perform the spatio-temporal disaggregation, the ST-PCLM uses a composition matrix that links both coarse and fine resolutions, and which is expressed as a Kronecker product of two marginal composition matrices (one acting at a spatial level and another at temporal level).

We can find several techniques that allow the spatial disaggregation of health data. For example, the use of an interpolation process from empirical Bayes estimates [17], (generalized) Poisson kriging methods [18], log-Gaussian Cox processes [19], and spatial composite link mixed models [20] have been suggested to produce a continuous smooth surface (across the study area) from regional health data. However, as far as we know, there is no appropriate model to address the problem of disaggregation of health data in both space and time (although there are some works about spatio-temporal disaggregation in a Gaussian context; see, for example, [21–25]). A common among social scientist is geographical microsimulation

(see for example [26]). Geographical microsimulation models simulate populations in given geographical areas, with population characteristics as close as possible to their real counterparts. But a drawback may be that a representative sample of individuals at the fine-scale data required by those methods is not always available. The methodology presented here allows to create detailed dynamic maps for disease incidence data. Publicly available maps in aggregated form over space and time are the only information required, as well as geographical locations and time points where a finer resolution prediction is required.

## Materials and methods

### Q fever data

Q fever is a widespread zoonotic disease caused by the bacterium *Coxiella burnetii*. C. burnetii transmission to humans is mainly associated with ruminants such as cattle, sheep, and goats. During parturition or abortion of infected animals, high numbers of *C. burnetii* are shed within the amniotic fluids and the placenta. These organisms end up in the environment, where they may survive for long periods due to their resistance to heat, dryness, and many common disinfectants. Humans are often highly susceptible to the disease, and very few organisms may be necessary to cause infection. More information about this infectious disease is provided in [27].

The Southern Netherlands faced large outbreaks of human Q fever from 2007 to 2010 [28]. In this country, the local municipal health services (MHSs) are responsible for registering all confirmed diagnoses of acute Q fever. The information collected is then entered into the electronic national infectious diseases surveillance database. Due to confidentiality, these data are not publicly available and, in some instances, maybe provided in an aggregated form. In this case, the data were made available monthly at the municipality level.

Fig 1 shows the temporal distribution of Q fever cases (in months) from January 2007 to July 2010. A total of 3806 acute Q fever cases were registered in this period: 192 in 2007, 980 in 2008, 2309 in 2009, and 325 in 2010. The epidemic peaks of each year were observed every spring, specifically during May. This coincides with the birth period of small ruminants (sheep and goats), a fact that was pointed out in several studies about those exceptionally large Q fever outbreaks in the Netherlands (see, for example, [28, 29]). Since the largest outbreak was observed during 2009, we have studied the distribution of Q fever incidence in that year.

Fig 2a shows postal code areas affected by human Q fever (red points) in 2009. However, the number of cases was only publicly available at the municipality level. Since dairy goat and sheep farms are not evenly distributed across the Netherlands, Q-fever infection did not occur in all areas of the country. In this study we have focused on a $60 \times 60$km area in the south of the Netherlands (see black square in Fig 2a) with a total of 72 municipalities. The total number of Q fever cases reported in these municipalities was 1798. Taking into account the number of inhabitants in each municipality, we can calculate Q fever incidence (per 100000 inhabitants). Fig 2b shows the spatial distribution of the resulting Q fever incidence (aggregated over months in 2009), with higher incidence values observed around the municipalities of Landerd (1439.676), Lith (562.546), and Heusden (295.006). Both Figs 1 and 2 show a smooth distribution of counts in space and time which makes our initial assumption of a smooth latent distribution plausible.

### The spatio-temporal penalized composite link model

The Spatio-Temporal Penalized Composite Link Model (ST-PCLM) is the name of the model we propose for the spatio-temporal disaggregation of epidemiological data. ST-PCLM combines the penalized composite link model (PCLM) of [12] and spatio-temporal smoothing
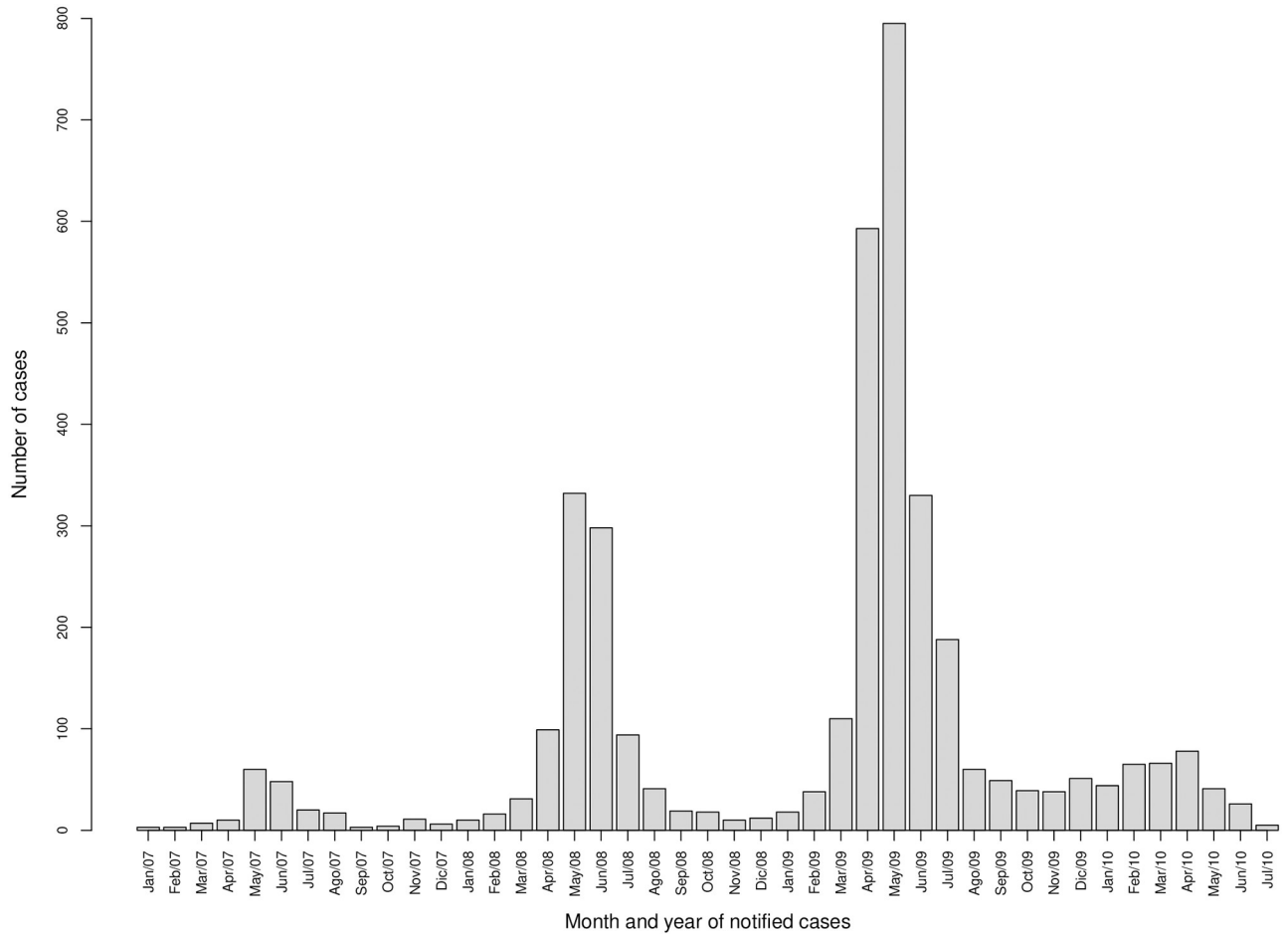
**Fig 1. Human Q fever cases in the Netherlands grouped per months, from January 2007 to July 2010.**

with P-splines [13]. The PCLM generalizes the model proposed by [30] and allows the estimation of a smooth trend, from grouped data, at a finer resolution. The idea behind it is as follows:

$$\boxed{\dfrac{\text{Distribution}}{\boldsymbol{\gamma}}} \rightarrow \boxed{\dfrac{\text{Composition}}{\boldsymbol{\mu} = \boldsymbol{C}\boldsymbol{\gamma}}} \rightarrow \boxed{\dfrac{\text{Sampling}}{\mathbf{y} \frown Poiss(\boldsymbol{\mu})}}$$

We observe data $\mathbf{y}$ from a Poisson distribution with mean $\boldsymbol{\mu}$, then we assume that $\boldsymbol{\mu}$ is a composition of the ungrouped distribution $\boldsymbol{\gamma}$. The approach assumes that the underlying (ungrouped) distribution is smooth, but otherwise let the data determine their shape.

As PCLM applies to the unidimentional scenario, we develop it in the spatio-temporal context. Then, we have included the ideas presented in [13] to reformulate the new spatio-temporal PCLM under a generalized linear mixed model (GLMM) framework, i.e., ST-PCLM. Hereafter, we will show the development of the ST-PCLM and a computationally efficient way to estimate the ST-PCLM parameters.

**The model.** Let $y_{it}$, $i = 1, \ldots, n$, $t = 1, \ldots, T_a$, denote count data recorded over $n$ non-overlapping spatial units $v_i$, which constitute the area of interest, at $T_a$ time periods. Suppose that

## a) Q fever outbreaks in 2009

## b) Q fever incidence in study area



**Fig 2. Map of human Q fever cases in the Netherlands, 2009.** Left: red points indicate the residential addresses of human cases (2309 in total). Right: study area in the south of the Netherlands showing (raw) incidence (per 100000 inhabitants) of Q fever by municipality in 2009. Source: geo-referenced data are obtained from https://www.cbs.nl/en-gb/onze-diensten/open-data/statline-as-open-data/cartography. a) Q fever outbreaks in 2009. b) Q fever incidence in study area.
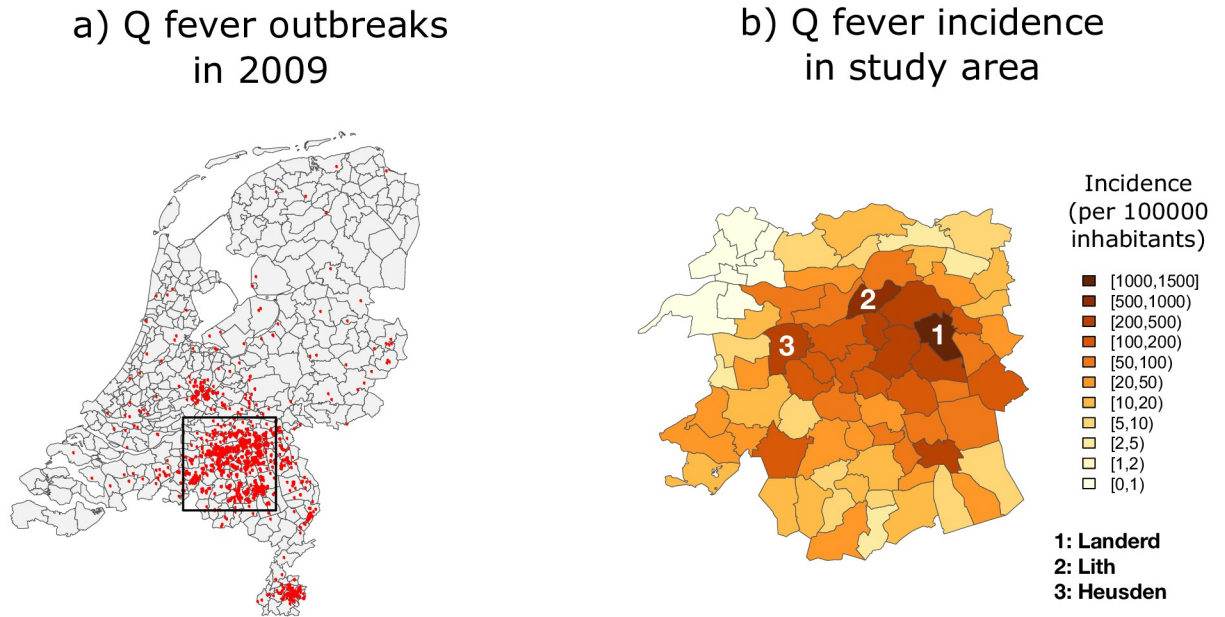
https://doi.org/10.1371/journal.pone.0263711.g002

we want to estimate the latent distribution of the vector of counts at a spatio-temporal support that is a refinement of the original one. The fine support is determined by three coordinates: $x_1 = (x_{11}, \ldots, x_{1m})'$ and $x_2 = (x_{21}, \ldots, x_{2m})'$, with $m > n$, which represent the longitude and latitude coordinates of the spatial refinement, respectively; and $x_3 = (x_{31}, \ldots, x_{3T_f})'$, with $T_f > T_a$, which represents the coordinates of the temporal refinement. We assume that the spatial refinement remains fixed at each instant of time in $x_3$, but the method can easily be extended to relax this assumption. Assuming that $y$ is Poisson distributed with mean vector $\mu$, the ST-PCLM is given by:

$$\mu = \mathbf{C}_{st}\gamma_f = \mathbf{C}_{st}\exp\{f_{st}(x_1, x_2, x_3)\}, \tag{1}$$

where $\gamma_f$ denotes the fine-scale latent mean, $\mathbf{C}_{st}$ is the composition matrix that describes how these latent means are combined to yield $\mu$, and $f_{st}(x_1, x_2, x_3)$ represents the fine-scale spatio-temporal trend. Several approaches can be used to model that function, e.g.: Gaussian random fields, spatio-temporal kriging or penalized splines, among others. We propose the last option since smoothness might be a reasonable property to ask for: if we are estimating an unknown distribution, it seems natural to assume smoothness since we do not know much about the underlying process (the use of the other techniques mentioned is out of the scope of this paper and will be subject of further work). We assume that the non-separable function $f_{st}(x_1, x_2, x_3)$ may be represented by the product of simpler functions $f_{x_1}(x_1), f_{x_2}(x_2)$, and $f_{x_3}(x_3)$. Therefore, the basis representation of the function would be given by:

$$f_{st}(x_1, x_2, x_3) = \mathbf{B}_{st}\theta = \{\mathbf{B}_3 \otimes (\mathbf{B}_2 \square \mathbf{B}_1)\}\theta, \tag{2}$$

where $\mathbf{B}_1 = \mathbf{B}(x_1)$, $\mathbf{B}_2 = \mathbf{B}(x_2)$, and $\mathbf{B}_3 = \mathbf{B}(x_3)$ are univariate B-spline bases [16] of dimensions $m \times c_1$, $m \times c_2$ and $T_f \times c_3$ respectively. The matrix operators $\otimes$ and $\square$ on the right-hand side of

Eq (2) represent Kronecker and Box products (row tensor products), respectively [31] (if the spatial refinement changes over time, Box product would be used instead of Kronecker product). To achieve smoothness, we use an anisotropic penalty (based on second order differences constructed separately for each independent variable) over the vector of regression coefficients $\boldsymbol{\theta}$. The form of this matrix is as follows:

$$\mathbf{P}_{st} = \lambda_1 \mathbf{I}_{c_3} \otimes \mathbf{I}_{c_2} \otimes \mathbf{P}_1 + \lambda_2 \mathbf{I}_{c_3} \otimes \mathbf{P}_2 \otimes \mathbf{I}_{c_1} + \lambda_3 \mathbf{P}_3 \otimes \mathbf{I}_{c_2} \otimes \mathbf{I}_{c_1}, \tag{3}$$

where $\mathbf{I}_{c_d}$ denotes an identity matrix of dimension $c_d \times c_d$, $\lambda_d$ is the smoothing parameter that controls the amount of smoothing along the coordinate $x_d$, and $\mathbf{P}_d = \mathbf{D}_d' \mathbf{D}_d$ is the marginal penalty matrix based on $\mathbf{D}_d$, which computes $q_d$-th differences, i.e., $\Delta^{q_d} \boldsymbol{\theta} = \mathbf{D}_d \boldsymbol{\theta}$, for $d = 1, 2, 3$. As we mentioned earlier, the matrix in Eq (3) has an anisotropic property in the sense that allows a different amount of smoothing for each coordinate.

Since we are assuming that the spatial refinement is fixed over the temporal refinement, the composition matrix $\mathbf{C}_{st}$ in model (1) is obtained as:

$$\mathbf{C}_{st} = \mathbf{C}_t \otimes \mathbf{C}_s, \tag{4}$$

where $\mathbf{C}_s$ and $\mathbf{C}_t$ are the spatial and temporal composition matrices of dimensions $n \times m$ and $T_a \times T_f$, respectively. The structure of these matrices depends on the type and level of aggregation. For example, if we want to estimate the latent distribution at a fine spatial grid (over a study area), the entries of the spatial composition matrix will be:

$$[\mathbf{C}_s]_{ij} = \begin{cases} 1 & \text{if } (x_{1j}, x_{2j}) \text{ belongs to unit } v_i \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $(x_{1j}, x_{2j})$ are the cell centroid coordinates of the fine grid, for $i = 1, \ldots, n$ and $j = 1, \ldots, m$. Another option to construct $\mathbf{C}_s$ is to consider its entries as the area proportions that each grid cell shares with a specific unit. The temporal composition matrix $\mathbf{C}_t$ is used to disaggregate coarse time intervals into detailed time periods (for example, from years or trimesters to months, weeks, or days). In our application, we show the structure that $\mathbf{C}_t$ will have, specifically for our purposes. Note that if $\mathbf{C}_s = \mathbf{I}_n$, $\mathbf{C}_t = \mathbf{I}_{T_a}$, and the unit centroids are used as spatial coordinates, the presented methodology is reduced to the Poisson version of the proposal given by [13] for smoothing of spatio-temporal count data.

When spatial data are recorded over a coarse rectangular grid, an appropriate definition for $\mathbf{B}_{st}$ in Eq (2) is $\mathbf{B}_{st} = \mathbf{B}_3 \otimes (\mathbf{B}_2 \otimes \mathbf{B}_1)$, where the spatial refinement correspond to the cell centroid coordinates of a fine grid. The penalty matrix $\mathbf{P}_{st}$ in Eq (3) is still valid in this context, because its definition is independent of data structure. Moreover, the spatial composition matrix would be given by $\mathbf{C}_s = \mathbf{C}_2 \otimes \mathbf{C}_1$, where each $\mathbf{C}_d$ is constructed according to the disaggregation of the coarse grid cells into small ones, for $d = 1, 2$. Although the Kronecker structure in (4) is a computational advantage, our model can cope with more complex situations where the spatial disaggregation changes from one time point to another (for example, census tracks changes along the years) by simply changing the structure of $\mathbf{C}_{st}$.

**Mixed model representation.** Now, we show how to incorporate fine-scale population information into the model (1). To do so, we must reformulate the model as a Generalized Linear Mixed Model (GLMM) by following the proposal of [13] as it is briefly described below.

In the P-splines literature, [13] show a nice representation of the spatio-temporal trend as a mixed model. Following their approach, the term $f_{st}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$ in model (2) can be rewritten as $f_{st}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = \mathbf{B}_{st}\boldsymbol{\theta} = \boldsymbol{x}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\alpha}$. Thus, we have:

$$\boldsymbol{\mu} = \mathbf{C}_{st}\boldsymbol{\gamma}_f = \mathbf{C}_{st}\{\exp(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\alpha} + \log(\boldsymbol{e}_f))\}, \text{ with } \boldsymbol{\alpha} \sim \mathcal{N}(\mathbf{0}, \mathbf{G}), \tag{6}$$

where $\mathbf{X}$ and $\mathbf{Z}$ are fixed and random effects matrices, and $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ are their associated coefficients, respectively. Random effects have covariance matrix $\mathbf{G}$ that depends on the smoothing parameters. The model includes an offset term $\log(\boldsymbol{e}_f)$ that allows the analysis of mortality or incidence rates instead of counts. The vector $\boldsymbol{e}_f$ could be, for example, fine-scale population information or expected number of deaths. If we only have the offset term at the coarse scale, i.e., $\log(\boldsymbol{e}) = (\log(e_{11}), ..., \log(e_{n1}), ..., \log(e_{1T_1}), ...., \log(e_{nT_a}))'$, a naive approach to estimate $\boldsymbol{e}_f$ assumes that the elements of $\boldsymbol{e}$ are evenly distributed throughout the fine resolution. Therefore, we can compute these naive estimates as $\widehat{\boldsymbol{e}}_{\text{naive}} = \mathbf{C}_{st}^{-}\boldsymbol{e}$, where $\mathbf{C}_{st}^{-}$ denotes the Moore-Penrose inverse of $\mathbf{C}_{st}$.

The construction of the matrices $\mathbf{X}$, $\mathbf{Z}$, and $\mathbf{G}$ in model (6) are obtained by using the singular value decomposition (SVD) of each discrete penalty matrix $\mathbf{P}_d$ in Eq (3), for $d$ = 1, 2, 3. The mixed model matrices are given by:

$$
\begin{aligned}
\mathbf{X} \quad &= \mathbf{X}_3 \otimes (\mathbf{X}_2 \square \mathbf{X}_1), \\
\mathbf{Z} \quad &= [\mathbf{Z}_3 \otimes (\mathbf{X}_2 \square \mathbf{X}_1) : \mathbf{X}_3 \otimes (\mathbf{Z}_2 \square \mathbf{X}_1) : \mathbf{X}_3 \otimes (\mathbf{X}_2 \square \mathbf{Z}_1) : \\
&\quad \mathbf{Z}_3 \otimes (\mathbf{Z}_2 \square \mathbf{X}_1) : \mathbf{Z}_3 \otimes (\mathbf{X}_2 \square \mathbf{Z}_1) : \mathbf{X}_3 \otimes (\mathbf{Z}_2 \square \mathbf{Z}_1) : \mathbf{Z}_3 \otimes (\mathbf{Z}_2 \square \mathbf{Z}_1)],
\end{aligned}
\tag{7}
$$

where $\mathbf{X}_d = \mathbf{B}_d\,\mathbf{U}_{dn}$ and $\mathbf{Z}_d = \mathbf{B}_d\,\mathbf{U}_{ds}$ are constructed from the matrices of singular vectors corresponding to null and non-null singular values of the SVD of $\mathbf{P}_d$, $\mathbf{U}_{dn}$ and $\mathbf{U}_{ds}$, respectively, for $d$ = 1, 2, 3. Denoting $\tilde{\boldsymbol{\Sigma}}_d$ as the diagonal matrix with the non-null singular values of the SVD of $\mathbf{P}_d$ in the main diagonal, the inverse of the covariance matrix $\mathbf{G}$ becomes the block-diagonal matrix:

$$
\begin{aligned}
\mathbf{G}^{-1} = \text{blockdiag} \quad &(\lambda_3 \mathbf{F}_{3u}, \lambda_2 \mathbf{F}_{2u}, \lambda_1 \mathbf{F}_{1u}, \\
&\lambda_2 \mathbf{F}_{22} + \lambda_3 \mathbf{F}_{32}, \lambda_1 \mathbf{F}_{12} + \lambda_3 \mathbf{F}_{31}, \lambda_1 \mathbf{F}_{11} + \lambda_2 \mathbf{F}_{21}, \\
&\lambda_1 \mathbf{F}_{1t} + \lambda_2 \mathbf{F}_{2t} + \lambda_3 \mathbf{F}_{3t}),
\end{aligned}
\tag{8}
$$

where:

$$
\begin{aligned}
\mathbf{F}_{1u} &= \mathbf{I}_{q_3} \otimes \mathbf{I}_{q_2} \otimes \tilde{\boldsymbol{\Sigma}}_1, & \mathbf{F}_{2u} &= \mathbf{I}_{q_3} \otimes \tilde{\boldsymbol{\Sigma}}_2 \otimes \mathbf{I}_{q_1}, & \mathbf{F}_{3u} &= \tilde{\boldsymbol{\Sigma}}_3 \otimes \mathbf{I}_{q_2} \otimes \mathbf{I}_{q_1}, \\
\mathbf{F}_{11} &= \mathbf{I}_{q_3} \otimes \mathbf{I}_{c_2-q_2} \otimes \tilde{\boldsymbol{\Sigma}}_1, & \mathbf{F}_{12} &= \mathbf{I}_{c_3-q_3} \otimes \mathbf{I}_{q_2} \otimes \tilde{\boldsymbol{\Sigma}}_1, & \mathbf{F}_{21} &= \mathbf{I}_{q_3} \otimes \tilde{\boldsymbol{\Sigma}}_2 \otimes \mathbf{I}_{c_1-q_1}, \\
\mathbf{F}_{22} &= \mathbf{I}_{c_3-q_3} \otimes \tilde{\boldsymbol{\Sigma}}_2 \otimes \mathbf{I}_{q_1}, & \mathbf{F}_{31} &= \tilde{\boldsymbol{\Sigma}}_3 \otimes \mathbf{I}_{q_2} \otimes \mathbf{I}_{c_1-q_1}, & \mathbf{F}_{32} &= \tilde{\boldsymbol{\Sigma}}_3 \otimes \mathbf{I}_{c_2-q_2} \otimes \mathbf{I}_{q_1}, \\
\mathbf{F}_{1t} &= \mathbf{I}_{c_3-q_3} \otimes \mathbf{I}_{c_2-q_2} \otimes \tilde{\boldsymbol{\Sigma}}_1, & \mathbf{F}_{2t} &= \mathbf{I}_{c_3-q_3} \otimes \tilde{\boldsymbol{\Sigma}}_2 \otimes \mathbf{I}_{c_1-q_1}, & \mathbf{F}_{3t} &= \tilde{\boldsymbol{\Sigma}}_3 \otimes \mathbf{I}_{c_2-q_2} \otimes \mathbf{I}_{c_1-q_1}.
\end{aligned}
$$

If data are spatially recorded over a rectangular coarse grid, the corresponding mixed model matrices are obtained as in Eq (7), but replacing the Box products $\square$ by Kronecker products $\otimes$. The formulation of $\mathbf{G}^{-1}$ remains the same.

**Parameter estimation.** Once the ST-PCLM defined in (6) is in the GLMM framework, it is possible to estimate its parameters. This estimation procedure was presented by [20] in a spatial disaggregation context, and it involves two interrelated stages: (a) estimation of fixed coefficients and random effects ($\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$); and (b) estimation of smoothing parameters ($\lambda_1$, $\lambda_2$, and $\lambda_3$). The penalized quasi-likelihood (PQL) methods of [32] are used for stage (a), and the restricted (or residual) maximum likelihood (REML, [32, 33]) is used for stage (b) as a numerical optimization criterion for smoothing parameter selection. Technical details are provided in [20] and, thus, we only describe here the necessary results.

For given values of $\lambda_1$, $\lambda_2$, and $\lambda_3$, the estimation of the fixed and random effects coefficients of the model (6) are obtained by maximizing the following approximate penalized log-likelihood:

$$
\boldsymbol{y}'\log(\boldsymbol{\mu}) - \mathbf{1}'\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{\alpha}'\mathbf{G}^{-1}\boldsymbol{\alpha},
\tag{9}
$$

where $\mathbf{1}$ denotes a vector of ones of length $n \cdot T_a$ and $\boldsymbol{\mu} = \mathbf{C}_{st} \, \boldsymbol{\gamma}_f$, with $\boldsymbol{\gamma}_f = \boldsymbol{e}_f{}^* \exp(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\alpha})$. Differentiation of Eq (9) with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ leads to the score equations:

$$\begin{aligned} \check{\mathbf{X}}'(\boldsymbol{y} - \boldsymbol{\mu}) &= \mathbf{0}, \\ \check{\mathbf{Z}}'(\boldsymbol{y} - \boldsymbol{\mu}) &= \mathbf{G}^{-1}\boldsymbol{\alpha}, \end{aligned} \tag{10}$$

where $\check{\mathbf{X}} = \mathbf{W}^{-1}\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{X}$ and $\check{\mathbf{Z}} = \mathbf{W}^{-1}\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{Z}$ are called 'working' mixed model matrices, since $\mathbf{W}$ = $\mathrm{diag}(\boldsymbol{\mu})$ and $\boldsymbol{\Gamma} = \mathrm{diag}(\boldsymbol{\gamma}_f)$ change during the estimation procedure. Defining the working vector as:

$$\boldsymbol{z} = \check{\mathbf{X}}\boldsymbol{\beta} + \check{\mathbf{Z}}\boldsymbol{\alpha} + \mathbf{W}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}),$$

the solution of the score equations in (10) via Fisher scoring algorithm is expressed as the iterative solution of the system:

$$\begin{bmatrix} \check{\mathbf{X}}'\mathbf{W}\check{\mathbf{X}} & \check{\mathbf{X}}'\mathbf{W}\check{\mathbf{Z}}\mathbf{G} \\ \check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{X}} & \mathbf{I} + \check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{Z}}\mathbf{G} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} \check{\mathbf{X}}'\mathbf{W}\boldsymbol{z} \\ \check{\mathbf{Z}}'\mathbf{W}\boldsymbol{z} \end{bmatrix}, \tag{11}$$

where $\boldsymbol{b} = \mathbf{G}^{-1}\boldsymbol{\alpha}$. This yields to a modified version of the standard mixed model estimators:

$$\widehat{\boldsymbol{\beta}} = (\check{\mathbf{X}}'\mathbf{V}^{-1}\check{\mathbf{X}})^{-1}\check{\mathbf{X}}'\mathbf{V}^{-1}\boldsymbol{z}, \tag{12}$$

$$\begin{aligned} \widehat{\boldsymbol{\alpha}} &= \mathbf{G}\check{\mathbf{Z}}'\mathbf{V}^{-1}(\boldsymbol{z} - \check{\mathbf{X}}\widehat{\boldsymbol{\beta}}) \\ &= \mathbf{G}\check{\mathbf{Z}}'\mathbf{N}\boldsymbol{z}, \end{aligned} \tag{13}$$

where:

$$\mathbf{V} = \mathbf{W}^{-1} + \check{\mathbf{Z}}\mathbf{G}\check{\mathbf{Z}}',$$

$$\mathbf{N} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\check{\mathbf{X}}(\check{\mathbf{X}}'\mathbf{V}^{-1}\check{\mathbf{X}})^{-1}\check{\mathbf{X}}'\mathbf{V}^{-1}. \tag{15}$$

Conditioning on the estimates (12) and (13), the smoothing parameters $\lambda_1$, $\lambda_2$, and $\lambda_3$ can be estimated by maximizing the approximate REML (see [32, Eq. 13]):

$$\begin{aligned} l^*(\mathbf{V}) &= \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}\log|\check{\mathbf{X}}'\mathbf{V}^{-1}\check{\mathbf{X}}| - \frac{1}{2}(\boldsymbol{z} - \check{\mathbf{X}}\widehat{\boldsymbol{\beta}})'\mathbf{V}^{-1}(\boldsymbol{z} - \check{\mathbf{X}}\widehat{\boldsymbol{\beta}}) \\ &= \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}\log|\check{\mathbf{X}}'\mathbf{V}^{-1}\check{\mathbf{X}}| - \frac{1}{2}\boldsymbol{z}'(\mathbf{V}^{-1} - \mathbf{V}^{-1}\check{\mathbf{X}}(\check{\mathbf{X}}'\mathbf{V}^{-1}\check{\mathbf{X}})^{-1}\check{\mathbf{X}}'\mathbf{V}^{-1})\boldsymbol{z}, \end{aligned} \tag{16}$$

where $\lambda_1$, $\lambda_2$, and $\lambda_3$ are involved in $\mathbf{V}$ through $\mathbf{G}$. Therefore, optimal estimates for the ST-PCLM parameters are obtained by iteration between Eqs (12), (13) and (16), until convergence.

Once the ST-PCLM parameter estimates at convergence are obtained, we can derive standard errors for $\widehat{\boldsymbol{\eta}} = \mathbf{X}\widehat{\boldsymbol{\beta}} + \mathbf{Z}\widehat{\boldsymbol{\alpha}}$ by using the Bayesian approximation of the variance-covariance matrix for $(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\alpha}})'$ (see [34] for further details). Thus, the approximate standard errors for $\widehat{\boldsymbol{\eta}}$ are obtained by taking the square root of $\mathbb{V}\mathrm{ar}(\widehat{\boldsymbol{\eta}})$, which is obtained as:

$$\mathbb{V}\mathrm{ar}(\widehat{\boldsymbol{\eta}}) = \mathrm{diag}([\mathbf{X}:\mathbf{Z}] \begin{bmatrix} \check{\mathbf{X}}'\mathbf{W}\check{\mathbf{X}} & \check{\mathbf{X}}'\mathbf{W}\check{\mathbf{Z}} \\ \check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{X}} & \check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{Z}} + \mathbf{G}^{-1} \end{bmatrix}^{-1} [\mathbf{X}:\mathbf{Z}]').$$

We should note that the estimation procedure described above can be computationally inefficient if we are working with large datasets and the goal is to obtain estimates at a very fine scale. Furthermore, the direct creation of the matrices $\mathbf{C}_{sf}$, $\mathbf{X}$, and $\mathbf{Z}$ in model (7) can easily

lead to storage problems. In the next Sections, we will provide solutions for the ST-PCLM estimation in terms of storage and efficiency (these solutions were not available in [20]).

**Fast algorithm for spatio-temporal penalized composite link models.** In order to efficiently compute estimates of the smoothing parameters, we have adapted the SAP (Separation of Anisotropic Penalties) algorithm of [35] to the ST-PCLM context. To avoid possible storage problems, we have used the so-called GLAM (*Generalized Linear Array Methods*, [31, 36]) in the ST-PCLM setting. These methods also provide an efficient way to compute the matrix of cross-products required for the SAP algorithm, easing the computation time of model estimation.

Under the ST-PCLM approach and conditioning on the estimates given in Eqs (12) and (13), we can estimate $\lambda_1$, $\lambda_2$, and $\lambda_3$ by numerical maximization of the approximate REML in Eq (16). The usual algorithms used to approximate the solutions, such as the one proposed by [37] (extended to the generalized case by [38]), can only deal with situations in which the variance-covariance matrix is linear on the variance parameters. In the case of spatio-temporal models with anisotropic penalties, regression parameters are affected by more than one smoothing parameter, and so standard approaches can't be used, since the corresponding variance-covariance matrix does not have the required form. [35] proposed the SAP algorithm which provides a numerical solution to REML estimates, but it is able to deal with models that have a precision matrix for the random effect vector that is linear in the inverse of the variance parameters. These precision matrices are common when penalized smooth models with anisotropic penalties are reformulated as (generalized) linear mixed models, where the smoothing parameters are seen as ratios of variance components, i.e., $\lambda_d = \frac{\phi}{\tau_d^2}$, for $d$ = 1, 2, 3. Since we are working under a Poisson framework, the dispersion parameter, $\phi$, is equal to 1. Thus, the problem is reduced to obtain estimates for the variance components $\tau_1^2$, $\tau_2^2$, and $\tau_3^2$. The reformulation of the SAP algorithm for the ST-PCLM is described below.

Following [35] we can derive closed-form expressions, from the approximate REML, for the variance components $\tau_d^2$, for $d$ = 1, 2, 3. These estimates are given by:

$$\widehat{\tau}_d^2 = \frac{\widehat{\boldsymbol{\alpha}}' \boldsymbol{\Lambda}_d \widehat{\boldsymbol{\alpha}}}{\mathrm{ed}_d}, \tag{17}$$

where:

$$\mathrm{ed}_d = \mathrm{trace}\left( \breve{\mathbf{Z}}' \mathbf{N} \breve{\mathbf{Z}} \mathbf{G} \frac{\boldsymbol{\Lambda}_d}{\tau_d^2} \mathbf{G} \right), \tag{18}$$

with $\mathbf{N}$ defined in Eq (15), and

$$\boldsymbol{\Lambda}_1 = \mathrm{blockdiag}(\mathbf{0}_{q_1 q_2 (c_3 - q_3)}, \mathbf{0}_{q_1 q_3 (c_2 - q_2)}, \mathbf{F}_{1u}, \mathbf{0}_{q_1 (c_2 - q_2)(c_3 - q_3)}, \mathbf{F}_{12}, \mathbf{F}_{11}, \mathbf{F}_{1t}),$$
$$\boldsymbol{\Lambda}_2 = \mathrm{blockdiag}(\mathbf{0}_{q_1 q_2 (c_3 - q_3)}, \mathbf{F}_{2u}, \mathbf{0}_{q_2 q_3 (c_1 - q_1)}, \mathbf{F}_{22}, \mathbf{0}_{q_2 (c_1 - q_1)(c_3 - q_3)}, \mathbf{F}_{21}, \mathbf{F}_{2t}),$$
$$\boldsymbol{\Lambda}_3 = \mathrm{blockdiag}(\mathbf{F}_{3u}, \mathbf{0}_{q_1 q_3 (c_2 - q_2)}, \mathbf{0}_{q_2 q_3 (c_1 - q_1)}, \mathbf{F}_{32}, \mathbf{F}_{31}, \mathbf{0}_{q_3 (c_1 - q_1)(c_2 - q_2)}, \mathbf{F}_{3t}).$$

The non-null submatrices of each $\boldsymbol{\Lambda}_d$ were previously defined in Section. Here the inverse of the covariance matrix $\mathbf{G}$ in model (6) is written in terms of $\tau_d^2$'s and can be decomposed as $\mathbf{G}^{-1} = \frac{1}{\tau_1^2} \boldsymbol{\Lambda}_1 + \frac{1}{\tau_2^2} \boldsymbol{\Lambda}_2 + \frac{1}{\tau_3^2} \boldsymbol{\Lambda}_3$, where the capital lambdas are defined above. The SAP algorithm for the ST-PCLM parameter estimation is given in Algorithm 1, which is an adaptation of the algorithm provided in [35, p. 945].

**Algorithm 1** SAP algorithm for the ST-PCLM parameters estimation

**Require:** Convergence tolerances $\nu_1$ and $\nu_2$ (e.g., $1 \times 10^{-6}$) and maximum number of iterations $\mathrm{maxit}_1$ and $\mathrm{maxit}_2$ (e.g., 100).

1: Set initial values for the mixed model coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$, and
   the variance components $\tau_1^2$, $\tau_2^2$, and $\tau_3^2$ (for example, $\widehat{\boldsymbol{\beta}}^{(0)} = \mathbf{0}$ with
   length $q_1\ q_2\ q_3$, $\widehat{\boldsymbol{\alpha}}^{(0)} = \mathbf{0}$ with length ($c_1\ c_2\ c_3 - q_1\ q_2\ q_3$), and
   $\widehat{\tau}_1^{2(0)} = \widehat{\tau}_2^{2(0)} = \widehat{\tau}_3^{2(0)} = 1$). Set $k = 0$.
2: **for** 1 **to** maxit$_1$ **do**
3:   Given the current estimates for the mixed model coefficients, con-
     struct the matrix of weights $\bar{\mathbf{W}}$ and the working vector $\mathbf{z}$ as follows:

$$\begin{aligned}
\boldsymbol{\Gamma} &= \text{diag}(\widehat{\gamma}_{\mathrm{f}}^{(k)}), \text{ with } \widehat{\gamma}_{\mathrm{f}}^{(k)} = \exp(\mathbf{X}\widehat{\boldsymbol{\beta}}^{(k)} + \mathbf{Z}\widehat{\boldsymbol{\alpha}}^{(k)} + \log(\boldsymbol{e}_{\mathrm{f}})); \\
\mathbf{W} &= \text{diag}(\widehat{\mu}^{(k)}), \text{ with } \widehat{\mu}^{(k)} = \mathbf{C}_{\mathrm{st}}\widehat{\gamma}_{\mathrm{f}}^{(k)}; \\
\boldsymbol{z} &= \breve{\mathbf{X}}\widehat{\boldsymbol{\beta}}^{(k)} + \breve{\mathbf{Z}}\widehat{\boldsymbol{\alpha}}^{(k)} + \mathbf{W}^{-1}(\boldsymbol{y} - \widehat{\boldsymbol{\mu}}^{(k)}) = \mathbf{W}^{-1}\mathbf{C}\boldsymbol{\Gamma}\widehat{\boldsymbol{\eta}}^{(k)} + \mathbf{W}^{-1}(\boldsymbol{y} - \widehat{\boldsymbol{\mu}}^{(k)}),
\end{aligned}$$

   with $\widehat{\boldsymbol{\eta}}^{(k)} = \mathbf{X}\widehat{\boldsymbol{\beta}}^{(k)} + \mathbf{Z}\widehat{\boldsymbol{\alpha}}^{(k)}$.
4:   **for** 1 **to** maxit$_2$ **do**
5:     Given the current estimates for the variance components, obtain
       new estimates for $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ by solving the system in Eq (11). The
       resulting estimates are denoted as $\widehat{\boldsymbol{\beta}}^{(k+1)}$ and $\widehat{\boldsymbol{\alpha}}^{(k+1)}$, respectively.
6:     Obtain new estimates for the variance components using Eq (17).
       The resulting estimates are denoted as $\widehat{\tau}_d^{2(k+1)}$, for $d = 1, 2, 3$.
7:     Compare new variance component estimates with the previous
       ones, using the following convergence criterion:

$$\frac{\sum_{d=1}^{3} |\widehat{\tau}_d^{2(k+1)} - \widehat{\tau}_d^{2(k)}|}{3} \leq \nu_1.$$

8:     If the convergence tolerance is achieved, **break**, otherwise set
       $\widehat{\tau}_d^{2(k)} = \widehat{\tau}_d^{2(k+1)}$ and repeat steps 5, 6, and 7 until convergence.
9:   **end for**
10:  Compute a new estimate for the fine-scale smooth trend vector
     using the fixed and random effects estimates obtained in the
     last iteration of step 5. The resulting vector is denoted as
     $\widehat{\boldsymbol{\eta}}^{(k+1)}$. Compare the new estimate with the previous one, using the
     following convergence criterion:

$$\frac{\|\widehat{\boldsymbol{\eta}}^{(k+1)} - \widehat{\boldsymbol{\eta}}^{(k)}\|^2}{\|\widehat{\boldsymbol{\eta}}^{(k+1)}\|^2} \leq \nu_2.$$

11:  If the convergence tolerance is achieved, **break**, otherwise set $k$
     $= k + 1$ and repeat steps 2 to 11 until convergence.
12: **end for**

Although we have suggested setting the maximum number of iterations to 100, our experi-
ence is that the number of iterations needed to achieve convergence is much smaller. In gen-
eral, we have observed that the maximum number of iterations to obtain optimal variance
components is greater than the number of iterations to obtain optimal estimates for the linear
predictor. In the analysis of Q-fever data, convergence was reached after 30 iterations.

We should note that we can efficiently compute the trace in Eq (18) by taking into account
that $\mathbf{G}\,\boldsymbol{\Lambda}_d\,\mathbf{G}$ is a diagonal matrix. Thus, we only have to compute the diagonal of $\breve{\mathbf{Z}}'\mathbf{N}\breve{\mathbf{Z}}$ to obtain

this trace. From [39, Eq. (5.3)], we have that:

$$\check{\mathbf{Z}}'\mathbf{N} = [\mathbf{0}_{(c_1c_2c_3-q_1q_2q_3)\times q_1q_2q_3} : \mathbf{I}_{(c_1c_2c_3-q_1q_2q_3)}]\begin{bmatrix} \check{\mathbf{X}}'\mathbf{W}\check{\mathbf{X}} & \check{\mathbf{X}}'\mathbf{W}\check{\mathbf{Z}}\mathbf{G} \\ \check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{X}} & \mathbf{I}+\check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{Z}}\mathbf{G} \end{bmatrix}^{-1}[\check{\mathbf{X}}:\check{\mathbf{Z}}]\mathbf{W}.$$

Therefore, the diagonal elements of matrix $\check{\mathbf{Z}}'\mathbf{N}\check{\mathbf{Z}}$ are obtained by the column-wise addition of:

$$\left([\mathbf{0}_{(c_1c_2c_3-q_1q_2q_3)\times q_1q_2q_3} : \mathbf{I}_{(c_1c_2c_3-q_1q_2q_3)}]\begin{bmatrix} \check{\mathbf{X}}'\mathbf{W}\check{\mathbf{X}} & \check{\mathbf{X}}'\mathbf{W}\check{\mathbf{Z}}\mathbf{G} \\ \check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{X}} & \mathbf{I}+\check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{Z}}\mathbf{G} \end{bmatrix}^{-1}\right)' \odot \begin{bmatrix} \check{\mathbf{X}}'\mathbf{W}\check{\mathbf{Z}} \\ \check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{Z}} \end{bmatrix}.$$

An advantage of using the adapted SAP algorithm provided above is that we can directly compute the effective dimension (ED) of model (6). This model complexity measure is given by:

$$\text{ED} = q_1q_2q_3 + \sum_{d=1}^{3}\text{ed}_d, \tag{19}$$

where each $\text{ed}_d$ is computed from Eq (18). The first term on the right-hand side of Eq (19) corresponds to the dimension of the unpenalized (or fixed) part, whereas the second corresponds to the unpenalized (or random) part of the fitted model. To visualize the later statement, note that:

$$\sum_{d=1}^{3}\text{ed}_d = \text{trace}\left(\sum_{d=1}^{3}\check{\mathbf{Z}}'\mathbf{N}\check{\mathbf{Z}}\mathbf{G}\frac{\boldsymbol{\Lambda}_d}{\tau_d^2}\mathbf{G}\right) = \text{trace}(\check{\mathbf{Z}}'\mathbf{N}\check{\mathbf{Z}}\mathbf{G}) = \text{trace}(\check{\mathbf{Z}}\mathbf{G}\check{\mathbf{Z}}'\mathbf{N}),$$

where $\check{\mathbf{Z}}\mathbf{G}\check{\mathbf{Z}}'\mathbf{N}$ is the 'hat matrix' [40] of the unpenalized part of the fitted ST-PCLM (see Eq (13)).

Note that the matrix cross-products $\check{\mathbf{X}}'\mathbf{W}\check{\mathbf{X}}, \check{\mathbf{X}}'\mathbf{W}\check{\mathbf{Z}}$, and $\check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{Z}}$ (and its transpose) are involved in the computation of the variance component estimates. They also appear in the estimation of fixed and random effects coefficients, which are obtained by solving the system in Eq (11). These and other required matrix cross-products can be efficiently computed by adapting the so-called GLAM methods to the ST-PCLM setting, which we will show in the next Section.

**GLAM methods for spatio-temporal penalized composite link models.** When we deal with estimating latent trends in multiple dimensions, we are susceptible to problems with storage and computational burden. In the case of data arranged in multidimensional grids, these problems can be overcome using the GLAM methods developed in [31, 36]. These methods are designed to avoid direct computation of matrix cross-products where Kronecker operations are involved, by using sequences of nested matrix operations. In this Section, we show the use of these methods in the ST-PCLM context.

First consider the matrix-by-vector products $\mathbf{X}\boldsymbol{\beta}$, $\mathbf{Z}\boldsymbol{\alpha}$, and $\mathbf{C}_{\text{st}}\boldsymbol{\gamma}_{\text{f}}$ that appear in line 3 of Algorithm 1. These expressions can be computed as:

$$\mathbf{X}\boldsymbol{\beta} \equiv \rho\{\mathbf{X}_3, \rho(\mathbf{R}_1, \tilde{\mathbf{Q}})\},$$

$$\mathbf{Z}\boldsymbol{\alpha} \equiv [\rho\{\mathbf{Z}_3, \rho(\mathbf{R}_1, \tilde{\mathbf{A}}_1)\} : \rho\{\mathbf{X}_3, \rho(\mathbf{R}_2, \tilde{\mathbf{A}}_2)\} : \rho\{\mathbf{X}_3, \rho(\mathbf{R}_3, \tilde{\mathbf{A}}_3)\} : \rho\{\mathbf{Z}_3, \rho(\mathbf{R}_2, \tilde{\mathbf{A}}_4)\} :$$
$$\rho\{\mathbf{Z}_3, \rho(\mathbf{R}_3, \tilde{\mathbf{A}}_5)\} : \rho\{\mathbf{X}_3, \rho(\mathbf{R}_4, \tilde{\mathbf{A}}_6)\} : \rho\{\mathbf{Z}_3, \rho(\mathbf{R}_4, \tilde{\mathbf{A}}_7)\}],$$

$$\mathbf{C}_{\text{st}}\boldsymbol{\gamma}_{\text{f}} \equiv \rho\{\mathbf{C}_{\text{t}}, \rho(\mathbf{C}_{\text{s}}, \tilde{\boldsymbol{\Gamma}})\},$$

with $\mathbf{R}_1 = \mathcal{G}(\mathbf{X}_2, \mathbf{X}_1)$, $\mathbf{R}_2 = \mathcal{G}(\mathbf{Z}_2, \mathbf{X}_1)$, $\mathbf{R}_3 = \mathcal{G}(\mathbf{X}_2, \mathbf{Z}_1)$, $\mathbf{R}_4 = \mathcal{G}(\mathbf{Z}_2, \mathbf{Z}_1)$, where $\rho$ and $\mathcal{G}$ denote the rotated $\mathcal{H}-$ transform and the row-tensor product, respectively, defined in the S1 File. The symbol $\equiv$ means that both sides have the same elements but in a different order. The matrices

$\tilde{\mathbf{Q}}$, $\tilde{\boldsymbol{\Gamma}}$, and $\tilde{\mathbf{A}}_k$, for $k = 1, \ldots, 7$, are arrangements of the vectors $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\alpha}_k$, respectively, with $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1', \ldots, \boldsymbol{\alpha}_7')'$. The dimensions of these matrices correspond to the number of columns of the first matrix where $\rho$ acts, times the number of columns of the second matrix where $\rho$ acts (i.e., $\tilde{\mathbf{Q}}$ has dimension $\mathrm{ncol}(\mathbf{R}_1) \times \mathrm{ncol}(\boldsymbol{x}_3) = q_1 q_2 \times q_3$, $\tilde{\boldsymbol{\Gamma}}$ has dimension $\mathrm{ncol}(\mathbf{C}_s) \times \mathrm{ncol}(\mathbf{C}_t) = m \times T_f$, and so on). Therefore, it holds that $\mathrm{vec}(\tilde{\mathbf{Q}}) = \boldsymbol{\beta}$, $\mathrm{vec}(\tilde{\boldsymbol{\Gamma}}) = \gamma_f$, and $\mathrm{vec}(\tilde{\mathbf{A}}_k) = \boldsymbol{\alpha}_k$, for $k = 1, \ldots, 7$.

Next consider the matrix cross-products $\check{\mathbf{X}}'\mathbf{W}\check{\mathbf{X}}, \check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{Z}}, \check{\mathbf{X}}'\mathbf{W}\check{\mathbf{Z}}, \check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{X}}$ (which is equal to $(\check{\mathbf{X}}'\mathbf{W}\check{\mathbf{Z}})'$), $\check{\mathbf{X}}'\mathbf{W}\boldsymbol{z}$, and $\check{\mathbf{Z}}'\mathbf{W}\boldsymbol{z}$ that appear in line 5 in Algorithm 1 (i.e., in the system in Eq (11)). Note first that they can be reduced as:

$$
\begin{aligned}
\check{\mathbf{X}}'\mathbf{W}\check{\mathbf{X}} &= (\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{X})'\mathbf{W}^{-1}(\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{X}), \check{\mathbf{Z}}'\mathbf{W}\check{\mathbf{Z}} = (\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{Z})'\mathbf{W}^{-1}(\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{Z}), \\
\check{\mathbf{X}}'\mathbf{W}\check{\mathbf{Z}} &= (\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{X})'\mathbf{W}^{-1}(\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{Z}), \check{\mathbf{X}}'\mathbf{W}\boldsymbol{z} = (\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{X})'\boldsymbol{z}, \\
\check{\mathbf{Z}}'\mathbf{W}\boldsymbol{z} &= (\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{Z})'\boldsymbol{z}.
\end{aligned}
$$

Therefore, we only need to compute $\mathbf{C}_{st}\,\boldsymbol{\Gamma}\,\boldsymbol{x}$ and $\mathbf{C}_{st}\,\boldsymbol{\Gamma}\,\mathbf{Z}$. These expressions are obtained as follows:

$$
\begin{aligned}
\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{X} &\equiv \rho(\mathcal{G}(\mathbf{X}_3, \mathbf{C}_t')', \rho\{\mathcal{G}(\mathbf{R}_1, \mathbf{C}_s')', \tilde{\boldsymbol{\Gamma}}\}), \\
\mathbf{C}_{st}\boldsymbol{\Gamma}\mathbf{Z} &\equiv [\rho(\mathcal{G}(\mathbf{Z}_3, \mathbf{C}_t')', \rho\{\mathcal{G}(\mathbf{R}_1, \mathbf{C}_s')', \tilde{\boldsymbol{\Gamma}}\}) : \rho(\mathcal{G}(\mathbf{X}_3, \mathbf{C}_t')', \rho\{\mathcal{G}(\mathbf{R}_2, \mathbf{C}_s')', \tilde{\boldsymbol{\Gamma}}\}) : \\
&\quad \rho(\mathcal{G}(\mathbf{X}_3, \mathbf{C}_t')', \rho\{\mathcal{G}(\mathbf{R}_3, \mathbf{C}_s')', \tilde{\boldsymbol{\Gamma}}\}) : \rho(\mathcal{G}(\mathbf{Z}_3, \mathbf{C}_t')', \rho\{\mathcal{G}(\mathbf{R}_2, \mathbf{C}_s')', \tilde{\boldsymbol{\Gamma}}\}) : \\
&\quad \rho(\mathcal{G}(\mathbf{Z}_3, \mathbf{C}_t')', \rho\{\mathcal{G}(\mathbf{R}_3, \mathbf{C}_s')', \tilde{\boldsymbol{\Gamma}}\}) : \rho(\mathcal{G}(\mathbf{X}_3, \mathbf{C}_t')', \rho\{\mathcal{G}(\mathbf{R}_4, \mathbf{C}_s')', \tilde{\boldsymbol{\Gamma}}\}) : \\
&\quad \rho(\mathcal{G}(\mathbf{Z}_3, \mathbf{C}_t')', \rho\{\mathcal{G}(\mathbf{R}_4, \mathbf{C}_s')', \tilde{\boldsymbol{\Gamma}}\})].
\end{aligned}
$$

## Results

Once the ST-PCLM has been stated, we analyzed the data related to Q fever outbreaks in the Netherlands.

### Weekly high-resolution smooth incidence maps

As mentioned above, we have focused on the distribution of Q fever incidence in the south of the Netherlands during 2009. Some of this data were analyzed in [41], the authors focused on the construction of a smooth incidence map at a fixed time, without considering the need to obtain predictions that were consistent with the aggregated data. More recently [42] focused on the spatial transmission of the disease.

Let's remember that the data are aggregated at municipality and monthly levels, and that our goal is to obtain estimates of Q fever incidence at that fine grid and at each week of 2009 (i.e., disaggregate simultaneously in space and time) to obtain better insights of the evolution of the disease. The ST-PCLM approach, allows to visualize Q fever incidence at a finer spatio-temporal resolution, and also to incorporate fine-scale population information to the estimation of the latent process. Here we use two sources of information: 1) The Q fever counts that are available in the municipalities of the study area described in Fig 2b and months of 2009 (i.e., the initial spatial and temporal scales are municipalities and months, respectively), and 2) population measured at a fine spatial grid over the study area. Population information at fine spatial grids is available from the WorldPop project. The data containing the spatial distribution of population in 2009 in the Netherlands can be downloaded from https://www.worldpop.org/geodata/summary?id=42722.
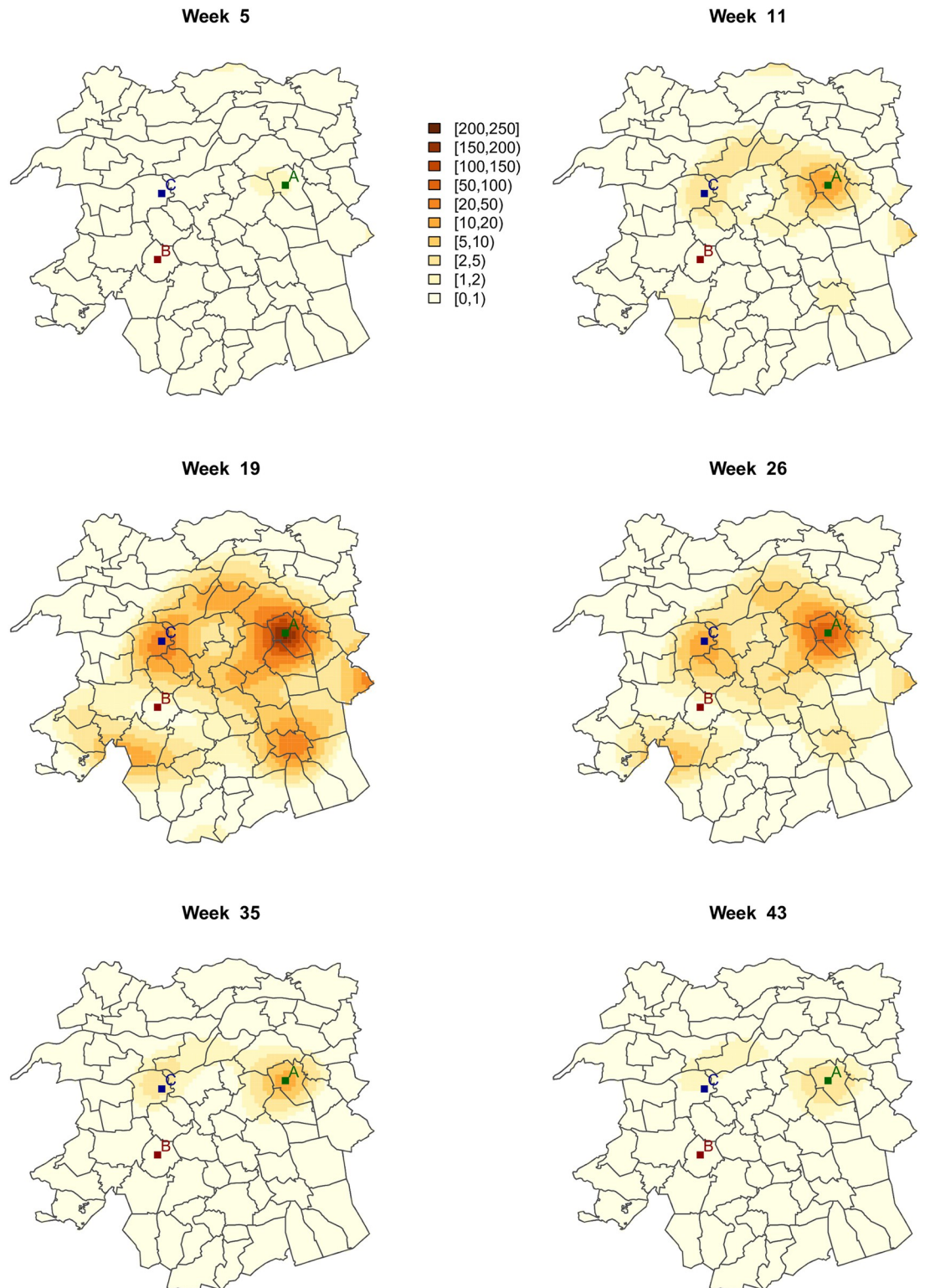
## a) Fine grid (cells of size 1000x1000 m)

## b) Population on grid



Legend:
- [7500,9000]
- [5000,7500)
- [2500,5000)
- [500,2500)
- [100,500)
- [50,100)
- [20,50)
- [10,20)
- [0,10)

**Fig 3. The map on the left shows the fine grid of 1000 × 1000 m cells in the study area shown in Fig 2b.** The map on the right shows the spatial distribution of the population on this fine grid. Source: Estimated population density per grid-cell from the WorldPop project in 2009, the Netherlands https://www.worldpop.org/geodata/summary?id=42722. a) Fine grid (cells of size 1000x1000 m). b) Population on grid.

https://doi.org/10.1371/journal.pone.0263711.g003

Fig 3a shows a fine grid composed of 4871 regular cells of size 1000 m × 1000 m. Blue dots represent the spatial coordinates of the centroids of these cells, and Fig 3b shows the spatial distribution of the population on this fine grid, which is heterogeneous across municipalities. To set up the ST-PCLM formulation, we have used: centroids of the cells described in Fig 3a as fine-scale spatial coordinates, i.e., as $x_1$ and $x_2$, $x_3 = (1, \ldots, 53)'$ (since 53 weeks were observed in 2009), second order penalties, 12 equally-spaced knots for the marginal cubic B-spline bases $\mathbf{B}_1$ and $\mathbf{B}_2$, and 8 equally-spaced knots for the marginal cubic B-spline basis $\mathbf{B}_3$. We have assumed here that the population at the fine grid is constant throughout the time period; thus, $\mathbf{e}_f$ in model (6) is considered as a vector obtained by repeating the fine-scale population fifty three times. The elements of the spatial composition matrix are obtained using Eq (5), whereas the temporal composition matrix for this case has the following form:

$$
\mathbf{C}_t = \begin{bmatrix}
1\ 1\ 1\ 1\ 1 & & & & & & & & & & & & \\
& 1\ 1\ 1\ 1 & & & & & & & & & & & \\
& & 1\ 1\ 1\ 1\ \tfrac{3}{7} & & & & & & & & & \\
& & \tfrac{4}{7}\ 1\ 1\ 1\ \tfrac{5}{7} & & & & & & & & & \\
& & & \tfrac{2}{7}\ 1\ 1\ 1\ 1\ \tfrac{1}{7} & & & & & & & & \\
& & & & \tfrac{6}{7}\ 1\ 1\ 1\ \tfrac{3}{7} & & & & & & & \\
& & & & \tfrac{4}{7}\ 1\ 1\ 1\ \tfrac{6}{7} & & & & & & & \\
& & & & & \tfrac{1}{7}\ 1\ 1\ 1\ 1\ \tfrac{2}{7} & & & & & & \\
& & & & & & \tfrac{5}{7}\ 1\ 1\ 1\ \tfrac{4}{7} & & & & & \\
& & & & & & \tfrac{3}{7}\ 1\ 1\ 1\ 1 & & & & & \\
& & & & & & & & 1\ 1\ 1\ 1\ \tfrac{2}{7} & & & \\
& & & & & & & & \tfrac{5}{7}\ 1\ 1\ 1\ 1 & & &
\end{bmatrix},
$$

where Sunday is considered the first day of the week. As opposed to the spatial composition matrix, the matrix $\mathbf{C}_t$ has some entries that are fractions. This is because some months share parts of a specific week (for example, some days of week 14 belong to March and the others to April).

Fig 4 shows the resulting ST-PCLM Q fever incidence (per 100000 inhabitants) at the target fine spatial resolution, for six selected weeks. These incidences are obtained as

**Week 5**

**Week 11**



Legend:
- ■ [200,250]
- ■ [150,200)
- ■ [100,150)
- ■ [50,100)
- ■ [20,50)
- ■ [10,20)
- ■ [5,10)
- □ [2,5)
- □ [1,2)
- □ [0,1)

**Week 19**

**Week 26**

**Week 35**

**Week 43**

**Fig 4. Smoothed Q fever incidence at a detailed spatio-temporal scale, resulting from the ST-PCLM approach, for six selected weeks.**

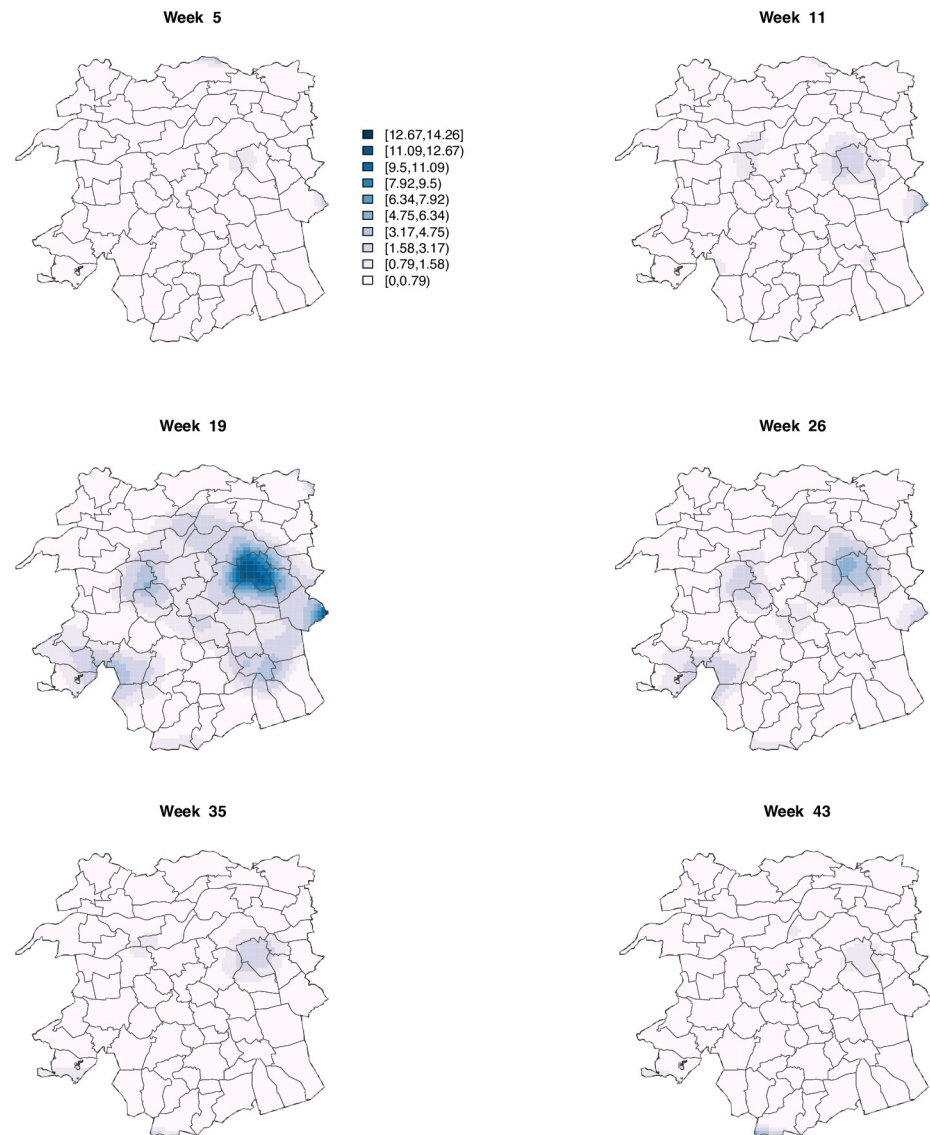**Fig 5. Approximate standard error maps associated with the smoothed Q fever incidence maps in Fig 4.**

$\widehat{\mathrm{inc}} = 100000\exp(\mathbf{X}\widehat{\boldsymbol{\beta}} + \mathbf{Z}\widehat{\boldsymbol{\alpha}})$. The evolution of the incidence varies across municipalities and weeks, where the highest incidences are observed mostly around week 19. Most of those weeks belong to April, May, and June, which have the largest number of Q fever outbreaks observed in 2009 (see Fig 1). Note also that most of the highest incidences in week 19 are spatially concentrated around the area that includes points A and C in Fig 4, which are located in the municipalities of Landerd and Heusden, respectively (see Fig 2b). Fig 5 shows the approximate standard error maps associated with Fig 4. As expected in a Poisson setting, larger variances are found in areas with higher incidence rates.

From the previous ST-PCLM estimates, we can also visualize the disaggregated (weekly) temporal evolution of the Q fever disease at specifics spatial coordinates of the fine grid. Fig 6 shows the down-scaled smoothed temporal incidence (per week) at three specific locations A, B, and C, in the study area. We observe that the temporal evolution of the incidence at point B
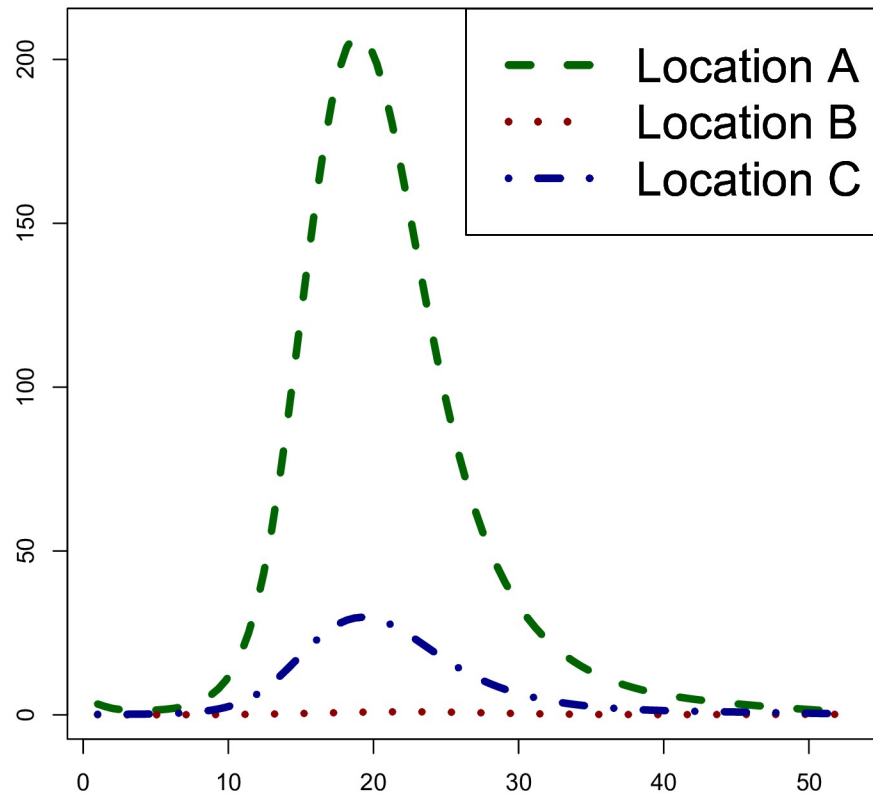
**Fig 6. Weekly temporal evolution of Q fever incidence in three specific points (A, B, and C), spatially presented in Fig 4 on the high-resolution map at week 19.**

https://doi.org/10.1371/journal.pone.0263711.g006

is constant and almost zero, whereas, at points A and C, the temporally smoothed incidence present a unimodal behaviour, where the peak is reached around week 19 (May). This is consistent with the summaries of the incidences given in Figs 1 and 2.

## Simulation studies

In this Section, we present one of two simulation studies that have been carried out to examine the predictive performance of the model under different scenarios. In both of them, we have done spatial and temporal disaggregation, but the focus of each simulation exercise is different. The first aims to check the performance of the ST-PCLM approach under different degrees of spatial dependence. In the second, we will study the behaviour of the model under different levels of temporal disaggregation (results from this second study are available in the S1 File).

**Simulation study 1.**   Data are generated using the fine grid in Fig 3a as the spatial region of study, and the 53 weeks in a year as the disaggregated temporal scale. The simulation is conducted as follows:

1. The fine-resolution incidence vector is constructed based on the smoothed Q fever incidences obtained in the previous section. We denote these smoothed incidences as $\text{inc}(\boldsymbol{u}_k)$, where $\boldsymbol{u}_k$, $k = 1, \ldots, K$, with $K = 4371 \times 53 = 258163$, represents the spatio-temporal coordinates at fine resolution. The different levels of spatial dependence are achieved by changing the values of the variance components that control the spatial term in the model, $\tau_1^2$ and $\tau_2^2$ (the variance component for the temporal term, $\tau_3^2$, will remain fixed at the optimal value

obtained in the fit). The values for the optimal variance components in the fit were: $\tau_1^2 = 218.7$, $\tau_2^2 = 127.7$, and $\tau_3^2 = 92.4$. Based on these results we set 3 different scenarios:

- Scenario 1: Variances for the spatial component are those in the fit.

- Scenario 2: Variances for the spatial component are 100 times larger than those in the fit.

- Scenario 3: Variance for the spatial component are 1000 times smaller than those in the fit.

2. Calculation of the aggregated expected number of cases (over municipalities and months), $\mu_{kl}$, $k = 1, \ldots, 72$, $l = 1, \ldots, 12$:

$$\boldsymbol{\mu} = (\mathbf{C}_t \otimes \mathbf{C}_s)\mathrm{inc}(\boldsymbol{u}_k).$$

3. 100 realizations of the number of cases in each municipality and month are generated through a random drawing from a Poisson distribution with mean parameter $\boldsymbol{\mu}$.

Fig 7 shows the incidences, $\mathrm{inc}_g(\boldsymbol{u}_k)$, corresponding to week 19 (which is the one with the highest number of cases) used in the three scenarios of simulation study 1. They reflect the three different levels of smoothness obtained by increasing or decreasing the variance components that control the spatial effect in the model.

For all realizations $l = 1, \ldots, 100$, the predicted incidences $\mathrm{inc}_{\mathrm{P}_g}^{(l)}(\boldsymbol{u}_k)$ obtained from the ST-PCLM approach for each scenario $g$, with $g = 1, 2, 3$, were compared to the smoothed incidences $\mathrm{inc}_g(\boldsymbol{u}_k)$. To evaluate the performance of the model we computed, for each scenario, the mean absolute error (MAE), the root mean squared error (RMSE):

$$\mathrm{MAE}_g^{(l)} \quad = \frac{1}{K}\sum_{k=1}^{K}|\mathrm{inc}_{\mathrm{P}_g}^{(l)}(\boldsymbol{u}_k) - \mathrm{inc}_g(\boldsymbol{u}_k)|,$$

$$\mathrm{RMSE}_g^{(l)} \quad = \sqrt{\frac{1}{K}\sum_{k=1}^{K}(\mathrm{inc}_{\mathrm{P}_g}^{(l)}(\boldsymbol{u}_k) - \mathrm{inc}_g(\boldsymbol{u}_k))^2},$$

the correlation between the observed and predicted incidence, and the percent of grid cells with true incidence falling within 95% prediction intervals of the model. These metrics were averaged over 100 simulated data sets and they are summarized in Table 1.

The model performance criteria reported in Table 1 show good model performance in terms of prediction accuracy in all scenarios. All criteria showed that slightly worse predictions where obtained in scenario 2, this is mainly due to the fact that, for consistency, we have used
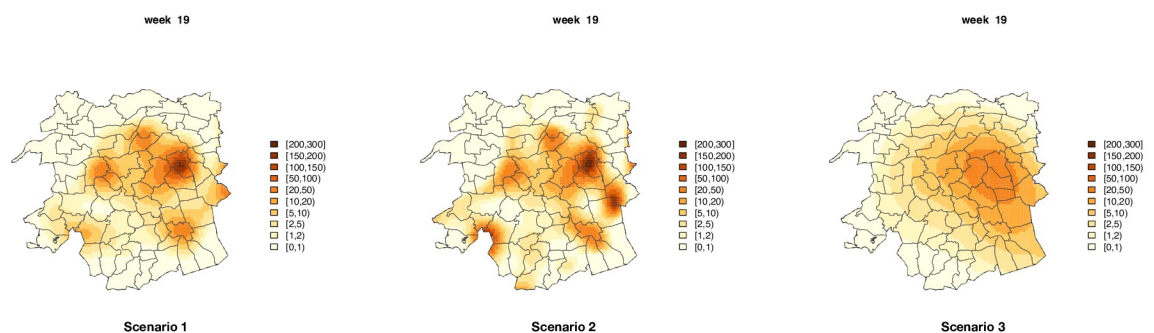


**Fig 7. Plots of incidences (for one of the weeks) used in each simulated scenario.**

**Table 1. Performance comparison of the ST-PCLM approach in three different types of scenarios of simulation study 1.** Correlation coefficient average (avg) of the 100 replicates, average coverage % and mean absolute errors (MAE) and root mean squared error (RMSE) are also shown.

| Scenarios | Correlation | Coverage | MAE | RMSE |
|-----------|-------------|----------|-----|------|
| 1 | 0.9829 | 92.06% | 0.00159 | 0.00996 |
| 2 | 0.9566 | 89.50% | 0.00255 | 0.01619 |
| 3 | 0.9941 | 96.44% | 0.00086 | 0.00434 |

the same size of B-spline basis for all scenarios (those used in the data analysis), but in the case of rapidly changing spatial patters (as is the case in scenario 2), a larger basis would be necessary to correctly capture the spatial effect, or adaptive P-splines [43] could be used; however, this approach is beyond of the scope of this paper.

In Fig 8 we provide further insight on the 95% coverage achieved in scenario 1 (similar plots for the other scenarios can be found in the S1 File). The lowest coverage over time (averaged over all pixels) was 85%, and was found at the extremes of the time interval (corresponding to the periods of lower incidence). The coverage for each pixel at the fine-grid resolution (averaged over time), shows that areas with lower coverage (60%–80%) correspond to locations with the highest number of cases. However, the coverage for these pixels is not constant over time, reaching almost 100% in weeks with low incidence rates. In scenario 2, the results are similar, although the coverage is lower (especially for a small number of pixels where the averaged coverage is between 20% and 30%), this is mainly due to the poor fitting for points with a sudden high incidence pick. The coverage in scenario 3 (the smoothest spatial trend) is nearly constant over the grid, ranging from 90% to 100%.

Taking into account the fact that we are predicting over more than 4000 grid cells and 53 weeks using aggregated data from 72 municipalities and 12 months, these results indicate a good predictive performance of the model.
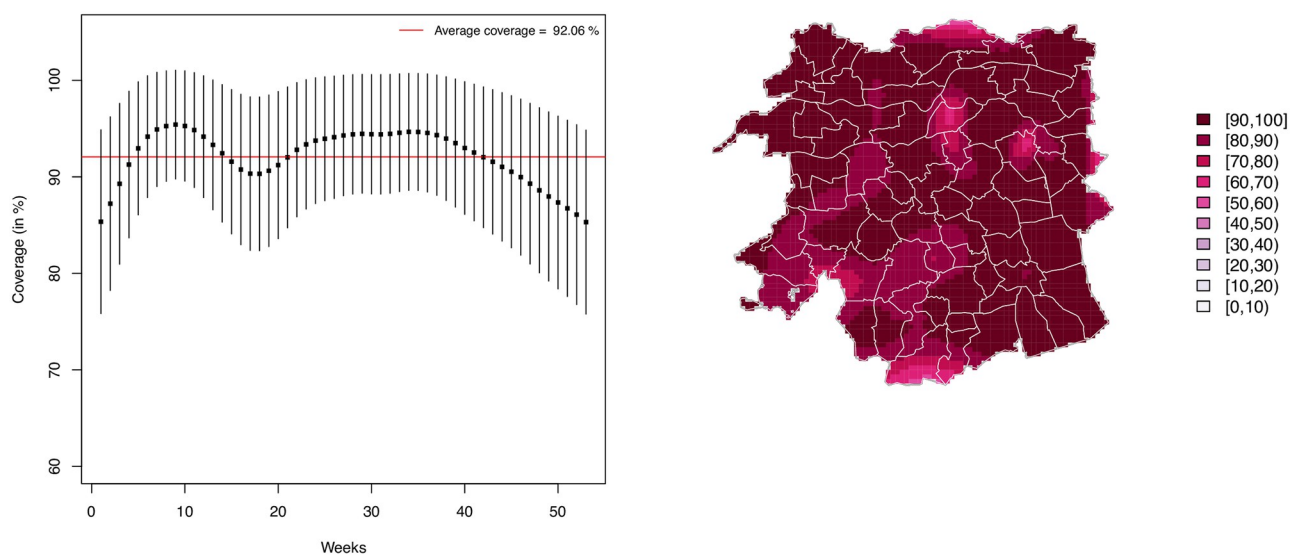


**Fig 8. Percentage of cells in the grid with true incidence falling within 95% prediction in scenario 1.** On the left, averaged coverage per week over all cells in the grid, and on the right, averaged coverage per cell over all weeks.

## Discussion

We have presented a novel model for the disaggregation of grouped data in both space and time, based on the spatio-temporal penalized composite link model approach. This framework was used to model Q fever counts (recorded in municipalities and months) to obtain Q fever incidence estimates over the fine grid and weeks. The model allows to obtain detailed trends in disease incidence, mortality risks, or any other vital rates at a desirable fine spatio-temporal resolution. Therefore, the resulting ST-PCLM outcomes can be displayed as a dynamic map. It also allows to include population information at a fine resolution in the estimation process. The flexibility of the model is provided by the use of B-splines, along with a penalty on the regression coefficients, and the link between the areas and the fine-resolution grid is achieved through the composition matrix. Our proposal can also address situations in which population information is recorded over small spatial units that are nested in coarser ones (for example, from municipalities to census tracts). In that case, the centroids of the small units can be used to represent the fine spatial scale.

Furthermore, the model allows the incorporation of covariates of interest (such as, for example, socio-economic, demographic, and environmental factors) in the ST-PCLM formulation to improve the estimation of the latent trend. They can be included at the aggregated level or at the fine-scale level (or at both levels simultaneously). The inclusion of covariates at the fine-scale level is immediate by adding them as columns in the design matrix $\mathbf{X}$ in (6) (if a linear relationship is assumed), or adding columns in $\mathbf{Z}$ if a non-linear relationship is expected. Details on how to include explanatory variables measured at the aggregated level can be found in [20].

It is important to acknowledge the use of GLAM methods in conjunction with the SAP algorithm, to avoid storage problems and to speed up computations. However, we are aware that the disaggregation of grouped data into a very detailed resolution could lead to increased computational load and storage problems. The sparsity of the marginal composition matrices can be exploited to deal with these issues (see, for example, [44]).

We have conducted two simulation studies to evaluate the prediction accuracy of the ST-PCLM. The first study aimed to check the performance of the model under different degrees of spatial dependence, and the second to asses the effect of different levels of temporal disaggregation. The overall performance of the model was good, although, as expected, it was affected by the number of units available at the aggregated level.

In the ST-PCLM formulation, we assume the spatial unit boundaries remain fixed over time. But it may happen that some of these boundaries change over the years. This issue is known in the statistical literature as the *spatio-temporal misalignment problem*. As future work, we plan to extend the ST-PCLM approach to handle this problem, where marginal composition matrices will play an important role. Furthermore, we can exploit the unique relationship between the penalties associated with basis and the covariance structure they yield to explore the use of other common spatio-temporal covariance matrices.

The implementation of our proposal and all data analyses were carried out using the statistical software R [45].

## Supporting information

**S1 File.**
(PDF)

## Author Contributions

**Conceptualization:** Dae-Jin Lee, María Durbán, Jan Van de Kassteele.

**Data curation:** Diego Ayma.

**Formal analysis:** Dae-Jin Lee, Diego Ayma.

**Investigation:** Jan Van de Kassteele.

**Methodology:** Dae-Jin Lee, María Durbán.

**Resources:** Jan Van de Kassteele.

**Software:** Dae-Jin Lee, María Durbán, Diego Ayma.

**Supervision:** Dae-Jin Lee, María Durbán.

**Validation:** Dae-Jin Lee, María Durbán.

**Visualization:** Dae-Jin Lee.

**Writing – original draft:** Dae-Jin Lee, María Durbán, Diego Ayma.

**Writing – review & editing:** Dae-Jin Lee, María Durbán.

## References

1. Waller LA, Gotway CA. Applied Spatial Statistics for Public Health Data. John Wiley & Sons, New York; 2004.

2. MacNab YC, Dean CB. Autoregressive spatial smoothing and temporal spline smoothing for mapping rates. Biometrics. 2001; 57:949–956. https://doi.org/10.1111/j.0006-341X.2001.00949.x PMID: 11550949

3. Ugarte MD, Goicoa T, Militino AF. Spatio-temporal modeling of mortality risks using penalized splines. Environmetrics. 2010; 21(3-4):270–289. https://doi.org/10.1002/env.1011

4. Waller LA, Carlin BP, Xia H, Gelfand AE. Hierarchical Spatio-Temporal Mapping of Disease Rates. Journal of the American Statistical Association. 1997; 92(438):607–617. https://doi.org/10.1080/01621459.1997.10474012

5. Knorr-Held L. Bayesian modelling of inseparable space-time variation in disease risk. Statistics in Medicine. 2000; 19(17-18):2555–2567. https://doi.org/10.1002/1097-0258(20000915/30)19:17/18%3C2555::AID-SIM587%3E3.0.CO;2-%23 PMID: 10960871

6. Martínez-Beneito MA, López-Quilez A, Botella-Rocamora P. An autoregressive approach to spatio-temporal disease mapping. Statistics in Medicine. 2008; 27(15):2874–2889. https://doi.org/10.1002/sim.3103 PMID: 17979141

7. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society Series B: Statistical Methodology. 2009; 71(2):319–392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

8. Schrödle B, Held L. Spatio-temporal disease mapping using INLA. Environmetrics. 2011; 22(6):725–734. https://doi.org/10.1002/env.1065

9. Ugarte MD, Adin A, Goicoa T, Militino AF. On fitting spatio-temporal disease mapping models using approximate Bayesian inference. Statistical Methods in Medical Research. 2014; 23(6):507–530. https://doi.org/10.1177/0962280214527528 PMID: 24713158

10. Bauer C, Wakefield J, Rue H, Self S, Feng Z, Wang Y. Bayesian penalized spline models for the analysis of spatio-temporal count data. Statist Med. 2016; 35(11):1848–1865. https://doi.org/10.1002/sim.6785 PMID: 26530705

11. Gelfand AE, Diggle PJ, Fuentes M, Guttorp P. Handbook of Spatial Statistics. CRC Press, Boca Raton, Florida; 2010.

12. Eilers PCH. Ill-posed problems with counts, the composite link model and penalized likelihood. Statistical Modelling. 2007; 7(3):239–254. https://doi.org/10.1177/1471082X0700700302

13. Lee D.-J., Durban M. P-spline ANOVA-type interaction models for spatio-temporal smoothing. Statistical Modelling. 2011; 11(1):49–69. https://doi.org/10.1177/1471082X1001100104

14. Lee D.-J., Durbán M. Smooth-CAR mixed models for spatial count data. Computational Statistics and Data Analysis. 2009; 53(8):2968–2979. https://doi.org/10.1016/j.csda.2008.07.025

15. Rodríguez-Álvarez MX, Durbán M, Lee D.-J., Eilers PHC. Spatio-temporal adaptive penalized splines with application to Neuroscience; 2016. Available from: http://arxiv.org/abs/1610.06860.

16. Eilers PHC, Marx BD. Flexible smoothing with B-splines and penalties. Statistical Science. 1996; 11 (2):89–121. https://doi.org/10.1214/ss/1038425655

17. Berke O. Exploratory disease mapping: Kriging the spatial risk function from regional count data. International Journal of Health Geographics. 2004; 3(18):–. https://doi.org/10.1186/1476-072X-3-18 PMID: 15333131

18. Goovaerts P. Geostatistical analysis of disease data: Accounting for spatial support and population density in the isopleth mapping of cancer mortality risk using area-to-point Poisson kriging. International Journal of Health Geographics. 2006; 5(52). https://doi.org/10.1186/1476-072X-5-52 PMID: 17137504

19. Diggle PJ, Moraga P, Rowlingson B, Taylor BM. Spatial and spatio-temporal log-gaussian Cox processes: Extending the geostatistical paradigm. Statistical Science. 2013; 28(4):542–563. https://doi.org/10.1214/13-STS441

20. Ayma D, Durbán M, Lee D.-J., Eilers PHC. Penalized composite link models for aggregated spatial count data: A mixed model approach. Spatial Statistics. 2016; 17:179–198. https://doi.org/10.1016/j.spasta.2016.06.007

21. Allcroft DJ, Glasbey CA. A latent Gaussian Markov random-field model for spatiotemporal rainfall disaggregation. Journal of the Royal Statistical Society Series C: Applied Statistics. 2003; 52(4):487–498. https://doi.org/10.1111/1467-9876.00419

22. Prairie J, Rajagopalan B, Lall U, Fulp T. A stochastic nonparametric technique for space-time disaggregation of streamflows. Water Resources Research. 2007; 43(3). https://doi.org/10.1029/2005WR004721

23. Segond ML, Neokleous N, Makropoulos C, Onof C, Maksimovic C. Simulation and spatio-temporal disaggregation of multi-site rainfall data for urban drainage applications. Hydrological Sciences Journal. 2007; 52(5):917–935. https://doi.org/10.1623/hysj.52.5.917

24. Schleiss M, Berne A. Stochastic space-time disaggregation of rainfall into dsd fields. Journal of Hydrometeorology. 2012; 13(6):1954–1969. https://doi.org/10.1175/JHM-D-12-013.1

25. Bindhu VM, Narasimhan B. Development of a spatio-temporal disaggregation method (DisNDVI) for generating a time series of fine resolution NDVI images. ISPRS Journal of Photogrammetry and Remote Sensing. 2015; 101:57–68. https://doi.org/10.1016/j.isprsjprs.2014.12.005

26. Ballas D, Rossiter D, Thomas B, Clarke G, Dorling D. Geography Matters: Simulating the local impacts of national social policies. Joseph Rowntree Foundation; 2005. 3. Available from: http://www.dannydorling.org/wp-content/files/dannydorling_publication_id1233.pdf.

27. Brown GL. Q Fever. British Medical Journal. 1973; 2(5857):43–45. https://doi.org/10.1136/bmj.2.5857.43 PMID: 4735198

28. Roest HIJ, Tilburg JJHC, Van Der Hoek W, Vellema P, Van Zijderveld FG, Klaassen CHW, et al. The Q fever epidemic in the Netherlands: History, onset, response and reflection. Epidemiology and Infection. 2011; 139(1):1–12. https://doi.org/10.1017/S0950268810002268 PMID: 20920383

29. van der Hoek W, Dijkstra F, Schimmer B, Schneeberger PM, Vellema P, Wijkmans C, et al. Q fever in the Netherlands: An update on the epidemiology and control measures. Eurosurveillance. 2010; 15(12): pii:19520. https://doi.org/10.2807/ese.15.12.19520-en PMID: 20350500

30. Thompson R, Baker RJ. Composite Link Functions in Generalized Linear Models. Applied Statistics. 1981; 30(2):125. https://doi.org/10.2307/2346381

31. Eilers PHC, Currie ID, Durbán M. Fast and compact smoothing on large multidimensional grids. Computational Statistics & Data Analysis. 2006; 50(1):61–76. https://doi.org/10.1016/j.csda.2004.07.008

32. Breslow NE, Clayton DG. Approximate Inference in Generalized Linear Mixed Models. Journal of the American Statistical Association. 1993; 88(421):9. https://doi.org/10.2307/2290687

33. Patterson HD, Thompson R. Recovery of inter-block information when block sizes are unequal. Biometrika. 1971; 58(3):545–554. https://doi.org/10.1093/biomet/58.3.545

34. Lin X, Zhang D. Inference in generalized additive mixed models by using smoothing splines. Journal of the Royal Statistical Society Series B: Statistical Methodology. 1999; 61(2):381–400. https://doi.org/10.1111/1467-9868.00183

35. Rodríguez-Álvarez MX, Lee D.-J., Kneib T, Durbán M, Eilers PHC. Fast smoothing parameter separation in multidimensional generalized P-splines: the SAP algorithm. Statistics and Computing. 2015; 25 (5):941–957. https://doi.org/10.1007/s11222-014-9464-2

36. Currie ID, Durbán M, Eilers PHC. Generalized linear array models with applications to multidimensional smoothing. J R Statist Soc B. 2006; 68(2):259–280. https://doi.org/10.1111/j.1467-9868.2006.00543.x

37. Harville DA. Maximum likelihood approaches to variance component estimation and to related problems. Journal of the American Statistical Association. 1977; 72(358):320–338. https://doi.org/10.2307/2286798

38. Schall R. Estimation in generalized linear models with random effects. Biometrika. 1991; 78(4):719–727. https://doi.org/10.1093/biomet/78.4.719

39. Harville DA. Matrix Algebra From a Statistician's Perspective. Springer-Verlag, New York; 1997.

40. Hastie TJ, Tibshirani RJ. Generalized Additive Models. Chapman & Hall, London; 1990.

41. van der Hoek W, Van De Kassteele JVD, Bom B, DE Bruin AD, Dijkstra F, Schimmer B, et al. Smooth incidence maps give valuable insight into Q fever outbreaks in the Netherlands. Geospatial Health. 2013; 7(1):127–134. https://doi.org/10.4081/gh.2012.111

42. De Koeijer AA, Hagenaars TJ, Van Leuken JPG, Swart AN, Boender GJ. Spatial transmission risk during the 2007- 2010 Q fever epidemic in the Netherlands: Analysis of the farm-to-farm and farm-toresident transmission. PLoS ONE. 2020; 15(2):e0227491. https://doi.org/10.1371/journal.pone.0227491

43. Rodríguez-Álvarez MX, Durban M, Lee D.-J., Eilers PHC. On the estimation of variance parameters in non-standard generalised linear mixed models: application to penalised smoothing. Statistics and Computing. 2019; 29(3):483–500. https://doi.org/10.1007/s11222-018-9818-2

44. Bates D, Martin M. Package 'Matrix' Sparse and Dense Matrix Classes and Methods Author; 2015. Available from: https://cran.r-project.org/package=Matrix.

45. Team R Development Core. A Language and Environment for Statistical Computing; 2018. Available from: http://www.r-project.org.