Research

Impact of Environmental Factors on the Distribution Patterns of Nephropathia Epidemica Cases in Western Europe

Diana Erazo,[1] Maria Fernanda Vincenti-Gonzalez,[1] Guillaume Ghisbain,[1,2] Mirko Faber,[3] Chantal Reusken,[4] Virginie Sauvage,[5] William Wint,[6] Herwig Leirs,[7] Simon Dellicour,[1,8]* and Katrien Tersago[7,9]*

[1]Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, Brussels, Belgium
[2]Laboratory of Zoology, Research Institute for Biosciences, University of Mons, Mons, Belgium
[3]Department for Infectious Disease Epidemiology, Robert Koch-Institute, Berlin, Germany
[4]Department Virology, Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, the Netherlands
[5]Université Paris Cité, Unité Environnement et Risques Infectieux, Centre National de Référence des Hantavirus, Institut Pasteur, Paris, France
[6]Department of Biology, Environmental Research Group Oxford Ltd, Oxford, UK
[7]Evolutionary Ecology Group, University of Antwerp, Antwerp, Belgium
[8]Laboratory for Clinical and Epidemiological Virology, Rega Institute, Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium
[9]Epidemiology Unit, Scientific Institute of Public Health, Brussels, Belgium

BACKGROUND: Environmental factors, such as fluctuations of climatic conditions and land cover, play a pivotal role in driving infectious disease epidemics, particularly those originating from wildlife reservoirs. *Orthohantavirus puumalaense*, hosted by bank voles in Europe, is the causative agent of a form of hemorrhagic fever and renal syndrome called nephropathia epidemica. Despite two decades of consistent presence in western Europe, nephropathia epidemica outbreaks still pose challenges due to localized periodic occurrences and a lack of understanding of its environmental drivers.

OBJECTIVE: Our study aims to bridge this gap by investigating the specific ecological and climatic factors influencing nephropathia epidemica outbreaks in western Europe.

METHODS: We compiled monthly, serologically confirmed nephropathia epidemica case data obtained from public health authorities in Belgium, France, Germany, and the Netherlands for the period 2004–2012. Cases were georeferenced to the finest available administrative unit. We selected 28 covariates, including climatic variables, land cover, tree species distributions, and human population, and implemented a Bayesian spatiotemporal model using integrated nested Laplace approximation (INLA) with zero-inflated Poisson distribution, including fixed effects and spatial, temporal, and nonstructured random effects.

RESULTS: We identified key triggers for nephropathia epidemica outbreaks, particularly climate-mediated changes in all seasons up to 2 years before, favoring tree mast impacting bank vole abundance. Our findings revealed that while land-cover factors mostly determine hotspot locations, climatic fluctuation patterns rather tend to modulate outbreak intensity.

DISCUSSION: Crucially, our model allows for the generation of yearly maps showcasing nephropathia epidemica incidence and risk factors, aiding in public health preparedness against climate change–induced disease emergence. This work represents a significant step toward developing targeted forecasting tools for *Orthohantavirus puumalaense* outbreaks, offering valuable insights for epidemic control strategies. https://doi.org/10.1289/EHP15457

Introduction

Environmental factors can act as major drivers of infectious disease epidemics, typically epidemics of zoonotic infections originating in wildlife animal reservoirs.[1,2] Environmental factors, such as climatic, meteorological, and land-use factors, are acknowledged as major drivers of changes in infectious disease transmission dynamics.[3,4] In that context, epidemiological models explicitly considering environmental factors are key for understanding the spatiotemporal variation of infectious disease incidence.[5–10] Previous work has shown the potential and applicability of spatiotemporal environment-based early warning models for highly impactful vectorborne diseases, including malaria, dengue, and West Nile fever.[11,12] Yet, due to the multifactorial complexity by which climate and land cover affect host and/or vector ecology, pathogen ecology, and human infection patterns, such models can be challenging to implement.[13]

Orthohantaviruses (*Hantaviridae* family, *Bunyanvirales* order) constitute a prime example of (re)emerging zoonotic viruses whose risk of local circulation are impacted by climatic and land-use factors.[14,15] Rodents, moles, shrews, and bats host these well-studied RNA viruses, and, though primarily rodentborne orthohantaviruses have been found to be pathogenic to humans.[16] In Europe, *Orthohantavirus puumalaense* (PUUV), hosted by the bank vole (*Clethrionomys glareolus*), is the causative agent of a usually mild form of hemorrhagic fever and renal syndrome (HFRS), called nephropathia epidemica (NE). Human infection occurs during recreational activities or outdoor work through inhalation of infectious virus particles originating from the secretions, e.g., saliva, urine, and feces of infected bank voles.[17–19] NE is marked by abrupt onset of fever and headache, followed by gastrointestinal problems, vomiting and abdominal pain often accompanied by visual disturbances.[20] Half of NE cases also show renal symptoms although acute renal failure is only detected in severe cases.[16] NE is the most common orthohantavirus syndrome in Europe.[21] There have been more than 1,000 cases every year in Fennoscandia (Norway, Sweden, and Finland) since the 1990s,[16] but in the past 20 years, NE incidence has doubled in western Europe compared to the previous decades.[22,23] Particularly, localized epidemics every 2 to 4 years have been observed in countries such as Belgium, France, and Germany.[22,23]

Although risk factors related to bank vole ecology, virus ecology, and human behavior are all pivotal in understanding NE incidence patterns, a sudden increase in local bank vole abundance is one of the main drivers of NE epidemics when PUUV is circulating in a region.[22,24,25] Fluctuations in bank vole density through time are the result of climate-mediated seasonal changes in resources, varying energetic requirements, and bank vole self-regulatory density-dependent processes (reviewed by Andreassen et al.[26]). In addition, bank vole population cycles are also modulated by top-down processes, where predation by small carnivores limits the rodent population.[27] Multiple mechanisms have been described to formally explain climate-mediated fluctuations in bank vole density.[28,29] For instance, warm and dry summers are associated with massive tree seed production in the autumn the following year, affecting bank vole reproductive behavior and survival. An overview of how multiple climatic variables can be associated with massive tree seed production (masting) of tree species relevant to bank vole nutrition and how they might influence bank vole behavior has previously been provided.[9,30] Furthermore, in temperate western Europe, mild winters are assumed to be beneficial for bank vole survival and breeding.[31–33] Whereas, in Fennoscandia, deep snow cover during extended periods is hypothesized to positively affect bank vole survival and reproduction.[34,35] Spring/autumn temperature and precipitation affect the length and amount of primary production of biomass, which allows extended reproductive periods in bank voles.[34,36–38] Besides effects on bank vole abundance, precipitation and temperature can also alter PUUV transmission patterns through their effect on PUUV survival in the environment—colder and more humid conditions being known to improve PUUV longevity[18,39]—and levels of human–bank vole contact during outdoor activities.[28]

The peak of NE in Fennoscandia typically occurs between October and March, with additional cases often reported in winter with abundant vole populations.[40] In contrast, western Europe generally presents NE cases in rodent peak years in summer and a minor peak in early winter.[21] However, there is a complex seasonal pattern in the region, and a key question arises: do local patterns of climatic fluctuations account for broader PUUV epidemiological trends across western European countries? Evidence indicates that the timing of NE peaks varies by region. For instance, in the Ardennes in France and Belgium, cases could peak from March to September.[41] In Germany, the peak usually extends from late spring to early summer of a mast year.[30] In this study, we aimed to identify general patterns of NE cases across western Europe by developing a Bayesian spatiotemporal model and create an early warning system for NE in the region.

Methodology

Data

Nephropathia epidemica cases. We obtained numbers of monthly registered NE cases from epidemiological records maintained by the health institutes in the four countries considered in the present study: Belgium, France, Germany, and the Netherlands. Details on the NE data available from the different countries, including data source, time period, and spatial resolution associated with the NE data analyzed in our study, are provided in Table 1. The data sources considered are the following: Sciensano, the National Public Health Institute of Belgium, based on data from the network of sentinel laboratories; the National Hantavirus Reference Center (France) and Unit of Biology of Emerging Viral Infections of the Pasteur Institute (Paris, France); the Department for Infectious Disease Epidemiology of the Robert Koch Institute (Berlin, Germany); and the Center for Infectious Disease Control of the National Institute for Public Health and the Environment (Bilthoven, the Netherlands). The length of the time series varied among countries between 2000 and 2012 (Figure 1). This period was selected because NE case data were collected and processed as part of the EDENext project (Biology and Control of Vector-borne Infections in Europe; EU grant FP7-261504 EDENext) between 2011 and 2014.

Reported NE cases were serologically confirmed human cases for PUUV early antibody response (IgM), late antibody response (IgG), or both. Whenever feasible, we used the date of disease onset. If unavailable, for France, we relied on the date of sampling or sample reception, and for other countries, we relied on the date of diagnosis or report date. We georeferenced NE cases preferentially with the suspected location of PUUV infection, but if that was unavailable, we used the place of residence. The spatial resolution of case data corresponds to the smallest administrative units available in each country, as detailed below, with spatial boundaries obtained from Eurostat.

Cartographic data and covariate selection. We merged national administrative polygon maps, representing the unit of NE case data resolution for each country, into a single polygon map that was used throughout the study (Figure 2). The administrative polygons reflected the data formats provided by each health institute: For Belgium, data were at the LAU2 level (589 municipalities); for France, data were at the LAU1 level (3,785 cantons); for Germany, data were at the NUTS 3 level (412 districts); and for the Netherlands, data were at the LAU2 level (418 municipalities). We extracted values of climatic and land-cover variables (hereafter referred to as "covariates") for each administrative polygon using the "exactextractr" R package (version 0.9.1; R Development Core Team). A first selection of covariates was based on a literature review on variables affecting either human–PUUV contact, bank vole ecology, PUUV and/or tree ecology in the temperate region, followed by an assessment of data availability and relevance on the spatial and temporal scale of the present study (Table S1). These variables encompassed key climatic factors such as maximum land surface temperature, total precipitation, frost days, and snow days, which were previously hypothesized to be significant due to their influence on the ecologies of bank voles, trees, and viruses.[9,28] Additionally, we included two proxies for vegetation growth as follows: annual net primary production (NPP) and length of the

Table 1. Overview of available nephropathia epidemica (NE) case data. For each country involved, we report the data source, time period, considered spatial resolution, i.e., administrative polygons, associated with the NE data analyzed in our study, total number of NE cases, and median, minimum, and maximum population.

Country	Time period	Spatial resolution[a] (units)	Total NE cases	Median population (minimum–maximum)
Belgium[b]	2000–2010	LAU2 (589 municipalities)	1,862	11,184 (80–483,505)
France[c]	2003–2012	LAU1 (3,785 cantons)	753	10,104 (100–440,204)
Germany[d]	2001–2012	NUTS 3 (412 districts)	6,979	142,315 (33,807–3,501,872)
The Netherlands[e]	2008–2012	LAU2 (418 municipalities)	52	25,055 (955–779,810)

[a]According to national structures (Eurostat).
[b]Data from the network of sentinel laboratories, provided by Sciensano, the National Public Health Institute of Belgium.
[c]Data from the National Hantavirus Reference Center (France) and Unit of Biology of Emerging Viral Infections of the Pasteur Institute (Paris, France).
[d]Data from the Department for Infectious Disease Epidemiology of the Robert Koch Institute (Berlin, Germany).
[e]Data from the Center for Infectious Disease Control of the National Institute for Public Health and the Environment (Bilthoven, The Netherlands).
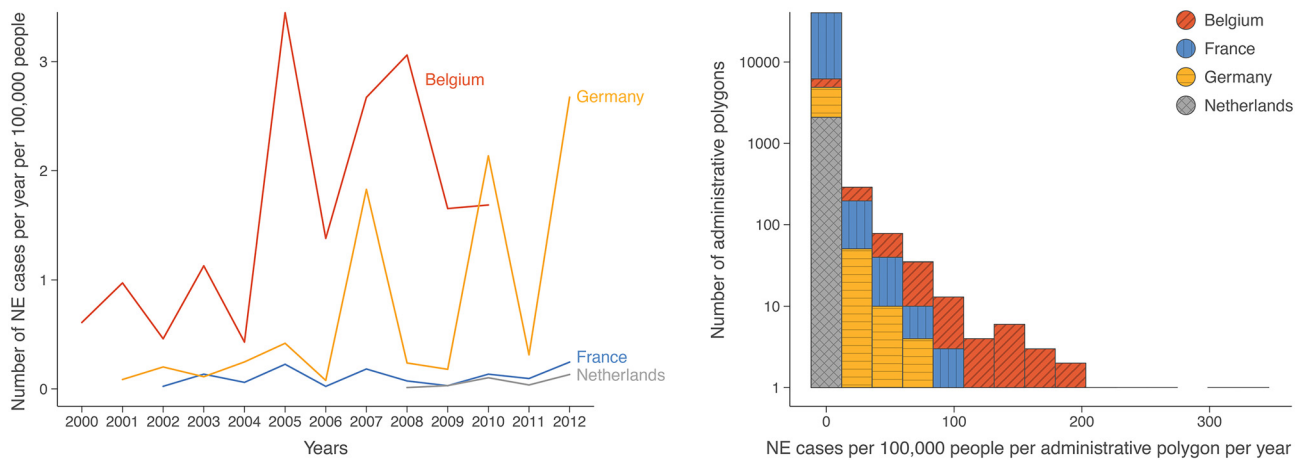
Figure 1. Number of monthly nephropatia epidemica (NE) cases per 100,000 people in western Europe and distribution of NE cases per administrative polygon per year. The first graph displays the annual incidence of nephropathia epidemica (NE) cases per 100,000 people from 2000 to 2012 for Belgium (red diagonal lines), France (blue vertical lines), Germany (yellow horizontal lines), and the Netherlands (gray crossed lines). This visualization provides insight into the varying levels of NE cases over the specified time period across these four countries. The second graph presents a histogram reporting the distribution of NE cases per year per 100,000 people from 2000 to 2012 in the aforementioned countries. This histogram offers a detailed perspective on the frequency and distribution of NE cases within administrative polygons across Belgium, France, Germany, and the Netherlands during the specified timeframe. Corresponding numeric data are available in the supplemental material (Excel figure1a.csv and figure1b.csv).

greening season (LGS), derived from data obtained through the Moderate-Resolution Imaging Spectroradiometer (MODIS) satellite imagery. Primary productivity is the rate at which light energy is converted into plant biomass, and NPP is the total amount of converted energy, accounting for plant respiration cost, that is used as a yearly measure of bank vole resource availability.[42,43] LGS is a metric of vegetation phenology that measures the duration of the availability of green plant material in the environment. Green plant material is considered a trigger for vole reproduction; hence, the period between greening-up and vegetation senescence can be used as a measure of the bank vole reproductive period.[44,45] The MODIS Land Cover Dynamics algorithm identifies phenophase transition dates based on logistic functions fit to the time series of the enhanced vegetation index (EVI). By subtracting the greening-up date from the senescence date, the length of the greening period (only based on major greening-up events in spring and greening-down events later in the same year) was calculated and averaged per polygon.[46] For climatic variables, data from previous years (lag) from NE case occurrence were used, with lag zero referring to the same year, lag 1 referring to the year before, and so forth (Table S1).

To this initial set of covariates, we also included land-cover variables associated with the bank vole habitat, the bank vole–human contact interface, and the proportion of tree species sensitive to massive tree seed production (Table S1).[47] We constructed bank vole habitat layers based on the 300-m resolution GlobCover land-cover map 2009 (ESA). Specifically, we conducted an expert-based categorization of land-cover classes as follows: *a*) preferred habitat (broadleaved forest, mixed forest),[48] *b*) suitable habitat (coniferous forest, green urban areas, transitional woodland-shrub, agro-forestry areas, fruit tree and berry plantations, agricultural land with significant areas of natural vegetation, moors and heathland, and inland marshes),[49,50] and *c*) unsuitable habitat for bank voles (all others). Both the proportion of preferred and suitable land-cover categories were quantified per polygon and subsequently considered as distinct variables in our model (see below). As for the bank vole–human contact interface layers, they were built upon the bank vole habitat layers: For both
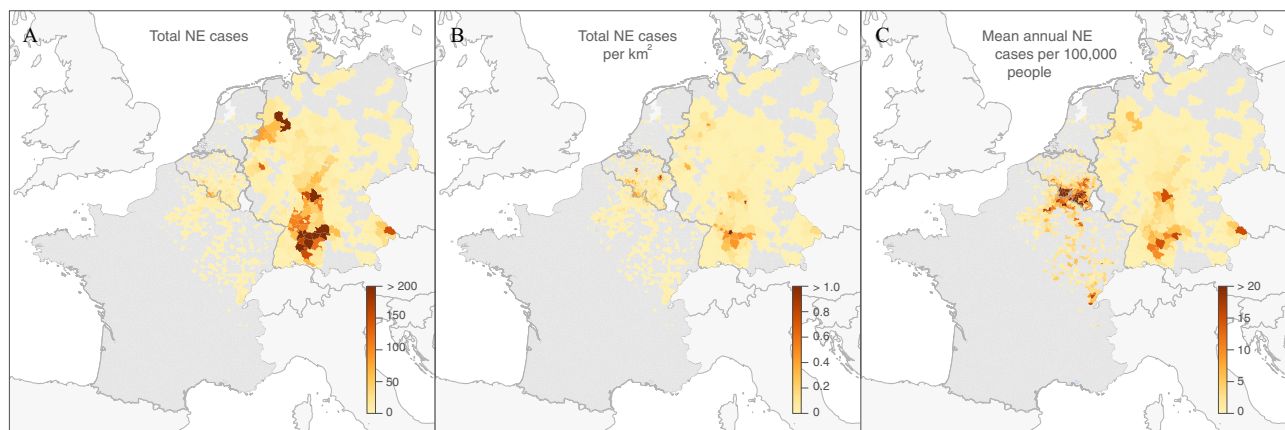

Figure 2. Distribution of nephropathia epidemica (NE) cases between 2000 and 2012 in western Europe. Dark gray background areas indicate the geographical regions under study, while light gray background areas represent other territorial land and white background areas signify the sea. The reported data include the total number of NE cases (A) per kilometer squared (B) and mean annual cases per 100,000 people (C), based on the administrative units outlined in Table 1. The legend color spectrum ranges from gray to dark red, with gray indicating zero cases and dark red denoting more than 200 NE cases (A), 1 NE case per kilometer squared (B) and 20 NE cases per year per 100,000 people (C). Administrative boundaries were obtained from the GADM database of Global Administrative Areas (http://gadm.org). Corresponding numeric data are available in the supplemental material (Excel Figure 2.csv).

preferred and suitable habitat, the proportion of vole habitat within a 300-m and 1-km radius of urban areas was calculated per kilometer squared per polygon. The resulting 300-m- and 1-km-radius-based maps were then considered as distinct variables. Therefore, 300 m reflects the typical movement range within bank vole home ranges and 1 km captures potential longer-distance movements.[51–55]

Finally, we calculated the proportion of each polygon area per kilometer squared covered by *Fagus sylvatica*, *Quercus robur*, or *Quercus petraea* per polygon based on the tree species occurrence maps (2000) provided by the European Forest Data Center (EFDAC). These three tree species are heavily consumed by various granivorous rodent species across Europe, including bank voles.[56]

Variable selection. Overall, we gathered 49 initial variables that were classified in six main groups: *a*) precipitation, *b*) temperature, *c*) frosting/snowing, *d*) vegetation phenology, *e*) land cover, and *f*) human population (Table S1). Each variable group, except human population, was evaluated following the covariate selection procedure implemented by Shrestha et al.[57] The initial variable selection therefore consisted of assessing multicollinearity by computing the Pearson correlation among variables in the group followed by the variable inflation factor (VIF) computed for the linear regression model using the "car" R package (version 3.1-2). First, one of the variables was discarded; when a pair of variables was associated with a correlation coefficient above 0.8,[58] we randomly discarded one of the variables in that pair. In addition, variables presenting a VIF above 10 were considered redundant because VIF measures the feasibility of a predictor variable to predict a response variable from a linear regression.[57,59] After this procedure, a total of 34 statistically independent variables was ultimately retained in the final dataset.

To avoid model overfitting, we implemented a model selection based on the deviance information criterion (DIC)[60] using the "ggregplot" R package. This procedure consists of removing variables one by one and observing how model fit varies according to the resulting model's DIC. The number of variables is reduced iteratively until the simplest model is obtained where removing any variables increases the DIC by greater than a set threshold value. Finally, 28 variables out of the subset of 34 variables were selected for training the predictive model (Table 2; Figure S1).

Model

Spatiotemporal predictive model. To implement our model, we used an integrated nested Laplace approximation (INLA),[61] which has been suggested as a computationally efficient approach for posterior distribution calculations in large-scale geostatistical models.[58,59,62,63] In practice, we used the "R-INLA" R package (version 23.04.24)[64] to implement a Bayesian spatiotemporal model of NE cases in western Europe. INLA has three main advantages: *a*) It deals with spatial correlation among observations, meaning that closer observations tend to present similar NE cases compared to distant ones.[58] Specifically, information from neighboring points was used to smooth the outlier values, which usually are a few points and provide robust estimations from sparse data.[59] *b*) INLA provides a hierarchical structure to estimate covariate effects, spatial covariance structure, and missing data prediction.[59] *c*) INLA includes both fixed and random effects. The fixed effects on NE incidence were here determined by the covariates and the random ones by the spatial variation. In this sense, the model investigates the relationship between NE cases, covariates, and spatial dependence.[65]

Model definition: zero-inflated Poisson model. Let Y_i be a random variable following a Poisson distribution with μ_i rate of

Table 2. Mean coefficient estimates and 95% Bayesian credible interval (BCI) for the environmental and sociodemographic variables considered in the models for a study of nephropathia epidemica outbreaks in Western Europe (2004–2012). Regression coefficients for covariates are presented as a risk ratio (RR), which represents the change in prevalence for a unit change in that covariate given that all other variables are kept constant.

Covariate	Mean	SD	Q2.5%	Q9.75%	RR	95% BCI
Human population (number of inhabitants per polygon)	**0.23**	**0.02**	**0.18**	**0.27**	**1.25**	**1.20–1.31**[a]
Area (km^2)	**0.28**	**0.04**	**0.21**	**0.35**	**1.32**	**1.23–1.42**[a]
Fagus sylvatica (% per km^2 per polygon area)	**0.07**	**0.03**	**0.01**	**0.14**	**1.08**	**1.01–1.15**[a]
Bank vole–human interface 1 (%)[b]	0.04	0.03	−0.02	0.11	1.04	0.98–1.11
Quercus petraea (% per km^2 per polygon area)	**0.08**	**0.03**	**0.01**	**0.14**	**1.08**	**1.01–1.15**[a]
Quercus robur (% per km^2 per polygon area)	0.04	0.04	−0.03	0.12	1.04	0.97–1.12
Bank vole preferred habitat (% per polygon)	**0.35**	**0.06**	**0.22**	**0.47**	**1.41**	**1.25–1.59**[a]
Bank vole suitable habitat (% per polygon)	−0.04	0.08	−0.20	0.11	0.96	0.82–1.12
Length greening season (lag 1) (number of days)	0.06	0.04	−0.03	0.14	1.06	0.97–1.15
Length greening season (lag 2) (number of days)	−0.01	0.04	−0.09	0.07	0.99	0.91–1.07
Length greening season (no lag) (number of days)	0.06	0.05	−0.03	0.15	1.06	0.97–1.16
Annual net primary productivity (lag 1) (kg × C/m^2)	−0.21	0.11	−0.41	0.00	0.81	0.66–1.00
Autumn snow days (lag 1) (number of days)	**−0.14**	**0.05**	**−0.24**	**−0.04**	**0.87**	**0.78–0.96**[a]
Spring snow days (lag 1) (number of days)	0.07	0.06	−0.05	0.20	1.07	0.95–1.22
Winter snow days (lag 1) (number of days)	−0.08	0.07	−0.22	0.05	0.92	0.81–1.05
Spring frost days (no lag) (number of days)	−0.02	0.04	−0.11	0.06	0.98	0.90–1.07
Autumn precipitation (lag 1) (mm)	**−0.25**	**0.10**	**−0.45**	**−0.06**	**0.78**	**0.64–0.94**[a]
Spring precipitation (lag 1) (mm)	−0.08	0.09	−0.26	0.11	0.92	0.77–1.11
Summer precipitation (lag 1) (mm)	0.15	0.08	−0.01	0.31	1.17	0.99–1.36
Summer precipitation (lag 2) (mm)	0.12	0.08	−0.04	0.28	1.13	0.96–1.32
Summer precipitation (lag 3) (mm)	0.08	0.08	−0.07	0.24	1.09	0.93–1.27
Spring precipitation (no lag) (mm)	−0.09	0.11	−0.30	0.12	0.91	0.74–1.12
Summer maximum temperature (lag 1) (°C)	−0.04	0.36	−0.74	0.67	0.96	0.48–1.95
Winter maximum temperature (lag 1) (°C)	**−1.65**	**0.31**	**−2.26**	**−1.04**	**0.19**	**0.10–0.35**[a]
Summer maximum temperature (lag 2) (°C)	**−0.73**	**0.35**	**−1.42**	**−0.04**	**0.48**	**0.24–0.96**[a]
Winter maximum temperature (lag 2) (°C)	0.03	0.29	−0.54	0.60	1.03	0.59–1.82
Summer maximum temperature (lag 3) (°C)	−0.19	0.31	−0.80	0.42	0.83	0.45–1.52
Spring maximum temperature (no lag) (°C)	**1.25**	**0.33**	**0.59**	**1.90**	**3.48**	**1.81–6.69**[a]
Spring land surface temperature (no lag) (°C)	0.01	0.11	−0.20	0.22	1.01	0.82–1.25

Note: LGS, length of greening season; LST, land surface temperature; NPP, annual net primary productivity; Q, quantile; SD, standard deviation.
[a]Significant variables are highlighted in bold and shown in Figure S1.
[b]Percentage within a 300-m radius of urban areas per kilometer squared per polygon.

NE cases with realizations y_i for $i = 1, \ldots, 5{,}124$ administrative units.

$$\text{NE}_{\text{cases}} = Y_i \sim \text{Poisson}\,(\mu_i),$$

$$\log\,(\mu_i) = \text{offset} + \text{intercept} + \text{fixed covariates},$$

$$\log\,(\mu_i) = -1 + \text{intercept} + \sum_{j=1}^{n} \beta_j \text{covariate value}_i,$$

where β_j, covariate value$_i$, and n represent the regression coefficient of variable j, covariate value for administrative unit i, and the total number of covariates, respectively. Given the large proportion of zeros in the dataset (Figure 1), we used a zero-inflated Poisson (ZIP) distribution to fit the Bayesian spatiotemporal model of NE cases in western Europe:

$$Y_i \sim \text{ZIP}\,(\mu_i \pi_i),$$

where Y_i is modeled as a combination of a Bernoulli distribution with $1 - \pi_i$ probability that creates zeros and ones and a Poisson distribution that creates counts. NE cases ($Y_i > 0$), therefore, are assumed to be generated from a Poisson distribution with probability $1 - \pi_i$. These assumptions are represented as follows:

$$Y_i \sim \text{Poisson}\,(\mu_i W_i),$$

$$W_i \sim \text{Bernoulli}\,(\pi_i),$$

$$P(Y_i = y_i | \beta_j, \gamma) = \begin{cases} \pi_i + (1 - \pi_i)e^{-\mu_{it}} & \text{for } y_i = 0 \\ (1 - \pi_i) f_{\text{Poisson}}(Y | \beta_j) & \text{for } y_i > 0 \end{cases},$$

where γ represents the intercept. This distribution assumes that an observation can be zero, even if there were cases, which is convenient when there is underreporting and the surveillance system cannot detect all of the "true" positive cases.[66] Missing cases in the database were designated as "nonavailable" and were not imputed; instead, they were handled directly within the INLA, which allows for missing values while still generating predictions.

Mesh construction. For the observed number of cases, we assumed a spatially continuous variable modeled with Gaussian random fields. Particularly, we used the stochastic partial differential equation (SPDE) approach from the "R-INLA" R package (version 23.04.24) to fit a spatial model and estimate the number of cases for the unsampled units.[67] The finite element method is an approximate solution to the SPDE, and here its representation is simplified and defined on the triangulation of domain D, which is subdivided into a triangulate mesh.[68] The way to construct this triangulate mesh is by first placing the vertices of the triangle at the sample locations and then adding additional vertices in the area for spatial predictions.

Following Zuur et al.,[69] the mesh range was based on the histogram of distances between sampling locations, determining the distance at which spatial dependency diminishes. This range was initially set at ~ 500 km (Figure S2A). The maximum edge length for the inner mesh was defined as one-fifth of this range (100 km) to accurately capture spatial correlations within this region. For the outer area, the maximum edge length was set equal to the range (500 km) to account for correlations near the boundaries of the inner region. Three mesh configurations with ranges of 250 km, 500 km, and 750 km (Figure S2B–D) were evaluated to assess their impact on model prediction performance. The results of these tests are presented in Table 3 for comparison.

The σ_i term is the nonspatial random effect, and the v_i term is the spatial correlated random effect, where the SPDE approach is used to estimate it, which is represented as follows:

$$\log\,(\mu_i) = -1 + \text{intercept} + \sum_{j=1}^{n} \beta_j \text{covariate value}_i + \sigma_i + v_i.$$

Model implementation and performance evaluation. We implemented four competing models assuming a zero-inflated Poisson distribution. Model 1, considering only fixed effects; model 2, identical to model 1 plus nonspatial random effect (σ_i); model 3, based on model 2 plus spatial random effect (v_i); and model 4, including all of the previous effects plus year random effect (ϑ_{it}), represented as follows:

$$\log\,(\mu_{it}) = -1 + \text{intercept} + \sum_{j=1}^{n} \beta_j \text{covariate value}_i + \sigma_i + v_i + \vartheta_{it},$$

$$Y_{it} \sim \text{ZIP}\,(\mu_{it} \pi_{it}).$$

In model 4, Y_{it} is a random variable following a Poisson distribution of NE cases, and μ_{it} is the rate of NE cases, both with realizations y_{it} for $i = 1, \ldots, 5{,}124$ administrative units for $t = 1, \ldots, 9$ years. The same database was used across all models.

The predictive performance of the models was assessed through the deviance information criterion (DIC) and the Watanabe–Akaike information criterion (WAIC).[60] The DIC is a parameter based on the posterior mean deviance, which measures the fit to the data, and the effective number of parameters, which measures model complexity. Therefore, the lower the DIC, the better the model predictive performance. Similar to DIC, WAIC estimates the effective number of parameters to adjust for overfitting. To assess the overall predictive performance of the models, we calculated three comprehensive metrics: accuracy, precision, and recall. These metrics were derived by withholding 10% of the data during model training and comparing the observed values with the predictions. Incorporating these metrics offers a more robust evaluation of the models' reliability and effectiveness as early warning systems. Mean DIC, WAIC, accuracy, precision, and recall were computed for 10 replicates per model and

Table 3. Implemented INLA models with the corresponding estimates for the study of nephropathia epidemica outbreaks in Western Europe (2004–2012).

Model	DIC	WAIC	Precision	Recall	Accuracy
Fixed effects	22,414 (22,296; 22,531)	57,700 (56,168; 59,231)	0.17 (0.16, 0.18)	0.76 (0.73, 0.79)	0.81 (0.80, 0.82)
Random effects	16,226 (16,200; 16,253)	17,367 (17,202; 17,531)	0.52 (0.50, 0.55)	0.54 (0.51, 0.57)	0.95 (0.95, 0.96)
Spatial (250)	14,895 (14,877; 14,913)	15,992 (15,898; 16,086)	0.49 (0.45, 0.53)	0.61 (0.60, 0.62)	0.95 (0.95, 0.95)
Spatial (500)	15,001 (14,983; 15,019)	16,068 (15,976; 16,160)	0.48 (0.45, 0.52)	0.61 (0.60, 0.62)	0.95 (0.95, 0.95)
Spatial (750)	15,090 (15,067; 15,113)	16,105 (16,012; 16,199)	0.47 (0.43, 0.51)	0.61 (0.60, 0.62)	0.95 (0.95, 0.95)
Spatiotemporal (250)	13,890 (13,838; 13,941)	13,731 (13,468; 13,994)	0.55 (0.54, 0.56)	0.62 (0.61, 0.63)	0.96 (0.96, 0.96)
Spatiotemporal (500)	14,020 (14,006; 14,035)	13,873 (13,821; 13,924)	0.56 (0.54, 0.57)	0.61 (0.60, 0.62)	0.96 (0.96, 0.96)
Spatiotemporal (750)	14,121 (14,076; 14,166)	14,061 (13,989; 14,134)	0.54 (0.53, 0.56)	0.61 (0.60, 0.62)	0.96 (0.96, 0.96)

Note: The deviance information criterion (DIC) and Watanabe–Akaike information criterion (WAIC) were calculated for integrated nested Laplace approximation (INLA) model comparison. Precision, recall, and accuracy were calculated to evaluate the predictive performance of each model. We report the mean estimate (as well as the 95% confidence interval) for ten replicates. Three mesh configurations with ranges of 250 km, 500 km, and 750 km (Figure S2B–D) were evaluated to assess their impact on the model prediction performance of the spatial and spatiotemporal models.

reported in Table 3 with their corresponding 95% confidence intervals.

We further estimated the regression coefficients for each model covariate. Considering INLA is a Bayesian approach, we reported the 95% Bayesian credible interval (BCI) of the regression coefficients. We also exponentiated the covariate regression coefficients to calculate the more interpretable risk ratio (RR) as suggested by Moraga et al.[70] and also implemented by Shrestha et al.[57] The RR represents the ratio of NE case number when the covariate increases in one unit to the NE case number when the covariate is fixed.[57] The association was considered significant when the 95% BCI interval did not include zero, and the RR was considered significant when the 95% BCI interval did not include one.

We fitted a smoothing curve between the predicted posterior mean of NE cases from the generalized linear model and the different covariates to evaluate their association using the "gam" smoothing function from the "ggplot2" R package (version 3.4.2) for visualization purposes (Figure 3). All analyses, figures, and maps were done in R (version 4.2.3).

Results

As the complexity of the model increased, so did its performance. Model 4, which incorporated both nonspatial and spatial random effects, along with year and a range of 250 km, yielded the smallest DIC and WAIC [DIC = 13,890 (13,838; 13,941); WAIC = 13,731 (13,468; 13,994)]. Following closely were the variations of model 4 with a 500-km range [DIC = 14,020 (14,006; 14,035); WAIC = 13,873 (13,821; 13,924)] and a 750-km range [DIC = 14,121 (14,076; 14,166); WAIC = 14,061 (13,989; 14,134)]. Among models, model 3 (both nonspatial and spatial random effects) with a 250-km range followed model 4 in performance [DIC = 14,895 (14,877; 14,913); WAIC = 15,992 (15,898; 16,086)]. Model 2 (nonspatial random effects) [DIC = 16,226 (16,200; 16,253); WAIC = 17,367 (17,202; 17,531)] and model 1 exhibited the largest DICs [DIC = 22,414 (22,296; 22,531); WAIC = 57,700 (56,168; 59,231)] (Table 3). Model 4 also demonstrated the strongest predictive performance, with the highest accuracy (0.96) and recall (0.62) and the second-highest precision (0.55) among all models tested (Table 3). These results,

combined with the lowest DIC and WAIC values, supported the selection of the fourth model for further analysis.

Out of 28 covariates considered for the final models, 10 covariates were significantly associated with NE case number (based on 95% BCI) (Table 2; Figure S1). Spring maximum temperature with no lag showed the highest association with NE case number (RR = 3.48; 95% BCI: 1.81, 6.69) followed by bank vole preferred habitat (RR = 1.41; 95% BCI: 1.25, 1.59), polygon area (RR = 1.32, 95% BCI: 1.23, 1.42), human population (RR = 1.25; 95% BCI: 1.2, 1.31), the area proportion covered by *F. sylvatica* (RR = 1.08; 95% BCI: 1.01, 1.15), and *Quercus petraea* (RR = 1.08; 95% BCI: 1.01, 1.15). On the other hand, winter maximum temperature (lag 1) was the most negatively associated with NE cases (RR = 0.19; 95% BCI: 0.10, 0.35) followed by summer maximum temperature (lag 2; RR = 0.48; 95% BCI: 0.24, 0.96), autumn precipitation (lag 1; RR = 0.78; 95% BCI: 0.64, 0.94) and autumn snow days (lag 1; RR = 0.87; 95% BCI: 0.78, 0.96).

The model results and the maps of predicted incidence per 100,000 people showed spatial and temporal heterogeneity in NE cases in western Europe (Figure S3). During the 2004–2012 period, two NE important hotspots were detected: the southwestern/central Germany hotspot (Figure 4B) and southern Belgium–northeastern France hotspot (Figure 4B). In general, the presence of NE cases was observed in these hotspots every year, but their intensity varied temporally (Figure S3). The hotspot in Germany was notably intense in 2007, 2010, and 2012, and the hotspot in the Belgium–France region had the most cases in 2005, 2008, and 2012. Except for northeastern France, the rest of the country presented a near to zero incidence in the study period, similarly to the Netherlands. Locations in the Netherlands and southeastern Germany showed low uncertainty in estimating the number of cases as well as low incidence of NE cases, regardless of the absence of Netherlands NE data before 2008 (Figure 4A,B). In general, the hotspots in southwestern/central Germany and southern Belgium–northeastern France presented the lowest uncertainty (standard deviation below 0.001) (Figure 4C). These predictions represent cumulative estimates over the 2004–2012 period and were generated using model 4.

The relationship profiles between NE predicted cases and relevant covariate values indicated how ecological conditions affected
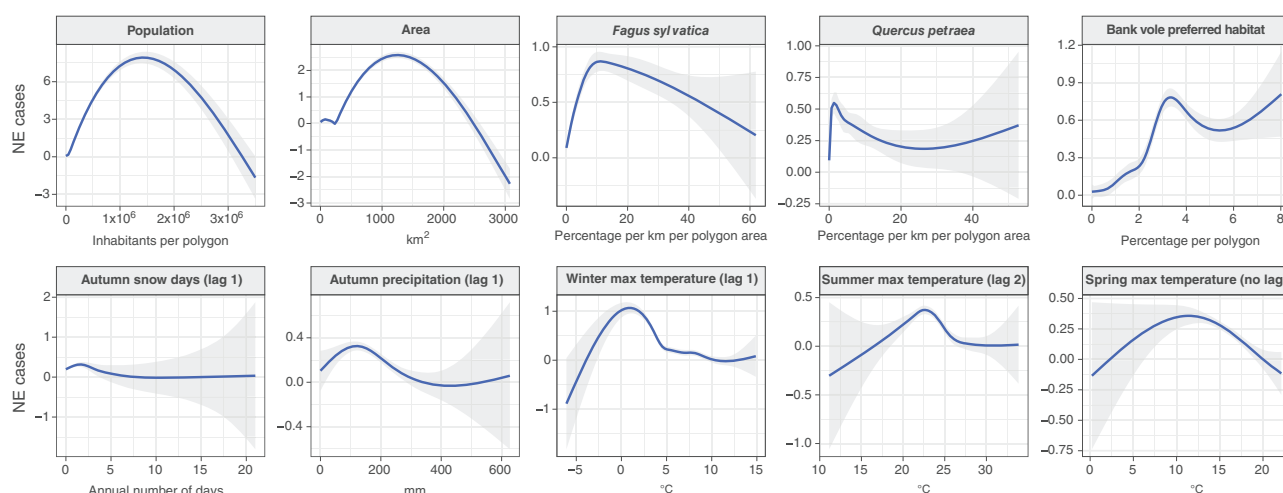

Figure 3. Estimated relationship between the predicted posterior mean of nephropathia epidemica (NE) cases and the environmental covariates associated with a significant regression coefficient. These covariates include human population density; polygon area; percentage of *Fagus sylvatica*, *Quercus petraea*, and bank vole preferred habitat; autumn snow days the previous year; autumn precipitation the previous year; winter maximum temperature the previous year; summer maximum temperature 2 years prior; and spring maximum temperature in the present year. The curve was fitted using the "gam" smoothing function available in the "ggplot2" R package (version 3.4.2) for visualization purposes. The blue line represents the mean value, while the shaded region around each curve signifies the 95% confidence interval. Corresponding numeric data are available in the supplemental material (Excel Figure 3.csv). Note: SD, standard deviation.
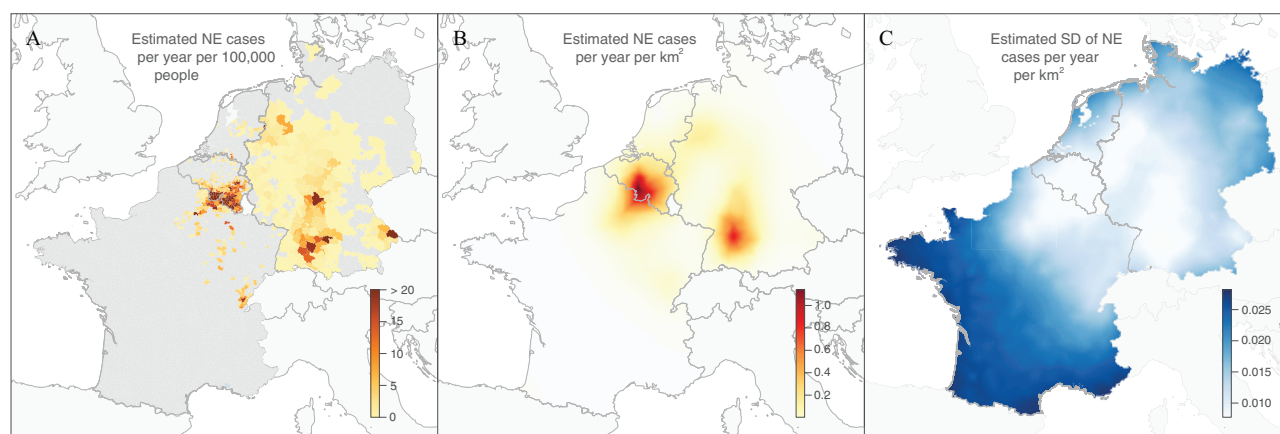
Figure 4. Cumulative estimated number of nephropathia epidemica (NE) cases per year per 100,000 people (A), estimated number of NE cases per year per kilometer squared in western Europe between 2004 and 2012 (B), and associated standard deviation (C) based on model 4. In panels A and B, the legend color spectrum ranges from light gray to dark red, where light gray represents 0 estimated number of NE cases per year per 100,000 people (A) and kilometer squared (B), and dark red indicates more than 20 NE cases per year per 100,000 people (A) and per year per kilometer squared (B). In panel C, the legend color gradient extends from white to dark blue, with white representing a standard deviation of 0 and dark blue representing a standard deviation of 0.025. The spatial unit for panel A is the polygon (administrative units), whereas for panels B and C it depends on the constructed mesh (with a defined pixel size of 450 km^2). Administrative boundaries were obtained from the GADM database of Global Administrative Areas (http://gadm.org). Corresponding numeric data are available in the supplemental material (Excel Figures 4a.csv, 4b.csv, and 4c.csv). Note: max, maximum.

NE local circulation (Figure 3). As human population increases per polygon, there is an increased uncertainty in estimating the number of cases and a rise in the predicted NE cases per polygon until 1.5 million people, when the number of predicted cases steadily decreased. There was a linear increase in the predicted NE cases up to around 10% coverage of *F. sylvatica* and 2% for *Q. petraea*, which then decreased with higher coverage and uncertainty, showing the lowest NE cases per polygon at 0% and 60% *F. sylvatica* coverage and 0% *Q. petraea* coverage. Similar to the previous association, NE cases were predicted to incrementally increase with percent coverage of bank vole preferred habitat up to around 3.5%, followed by a decrease and an increase in the NE predicted cases per polygon with high uncertainty. A few days of snow in autumn was associated with an increase of NE cases per polygon the following year. A decrease in NE cases per polygon was associated with an increased precipitation in autumn 1 year before the outbreak (between 100 mm and 300 mm of precipitation). Conversely, temperatures in winter above 0°C were expected to decrease the predicted NE cases per polygon for the following year, whereas summer temperatures above and/or below 22.5°C were expected to decrease the predicted NE cases 2 years later, though these estimates had high uncertainty. Finally, increases in spring maximum temperature above 12.5°C were negatively associated with predicted NE cases.

Discussion

We found that the percentage per polygon of bank vole preferred habitat (broad-leaved forest, mixed forest) (Figure S1D) was the land-cover variable that best predicted NE cases in western Europe. The presence of bank vole preferred habitat in a polygon implied an increased NE risk. Host habitat preference, used as a proxy for vole density, is a significant driver of PUUV local outbreaks, consistent with findings in multiple pathogen transmission systems.[71,72] Furthermore, optimal proportions of beech (*Fagus sylvatica*) and oak (*Quercus petraea*) underscore the importance of masting (high seed production) events in NE outbreaks, as suggested by several studies.[9,25,30,32,73,74] These masting events, which cause explosive fluctuations in bank vole abundance and PUUV prevalence, likely explain the rise in PUUV spillover events to humans.[75–77] Piechotowski et al.[78] and Faber et al.[79] found a clear link between NE occurrence and beech-dominated broadleaf forests during outbreaks in Germany.

The optimal proportions of *F. sylvatica* and *Q. petraea*, as well as human population are likely expected: low tree coverage suggests unsuitable habitat for reservoir animals, while very high coverage indicates predominantly wooded areas with limited human presence, reducing contact and disease occurrence. Similarly, very high population density points to highly urbanized areas with unsuitable vole habitats, whereas low population density indicates fewer people. An optimal mix of suitable vole habitat and human settlements—such as wooded areas interspersed with communities—may facilitate ideal conditions for contact between infected bank voles and humans. Although the exact reasons for the dramatic increases in NE cases in 2007 in Germany and 2008 in Belgium remain unclear,[79,80] our study confirms that climatic fluctuation patterns and the presence of beech forests are likely significant in the spatial and temporal occurrence patterns observed in western Europe. While land-cover factors act as the determinants of hotspot locations, climatic fluctuation patterns constitute drivers of peak years.

Spring maximum temperature in the year of NE occurrence, winter maximum temperature, number of snow days in autumn and autumn precipitation the year before, and summer maximum temperature 2 years before are the most informative climatic variables in our models. Since four out of five climatic variables correspond to at least 1 year before the outbreak, risk maps could be produced a year prior to NE occurrence. In particular, high summer temperatures are related to high NE incidence in Belgium,[9] and July maximum temperature 2 years before has been proposed as a good predictor of bank vole abundance and therefore NE cases in Germany.[29] Moreover, a mild winter could enhance the survival and reproduction of the bank vole.[31–33] Summer temperatures in the preceding year of a masting event are pivotal for the initiation of floral buds in the trees,[81,82] a prerequisite for tree mast. Therefore, in some years, this system follows a 3-year trophic cascade, where tree masting links local climatic fluctuations to the number of NE cases, with events occurring at 1-year intervals[9] (see Figure 5 for an illustration of the trophic cascade).

Bank vole populations can remain high even in years without a mast event; therefore, further investigation is essential to understand NE epidemics under these circumstances. Notably, our findings suggest that autumn conditions in the year prior to an
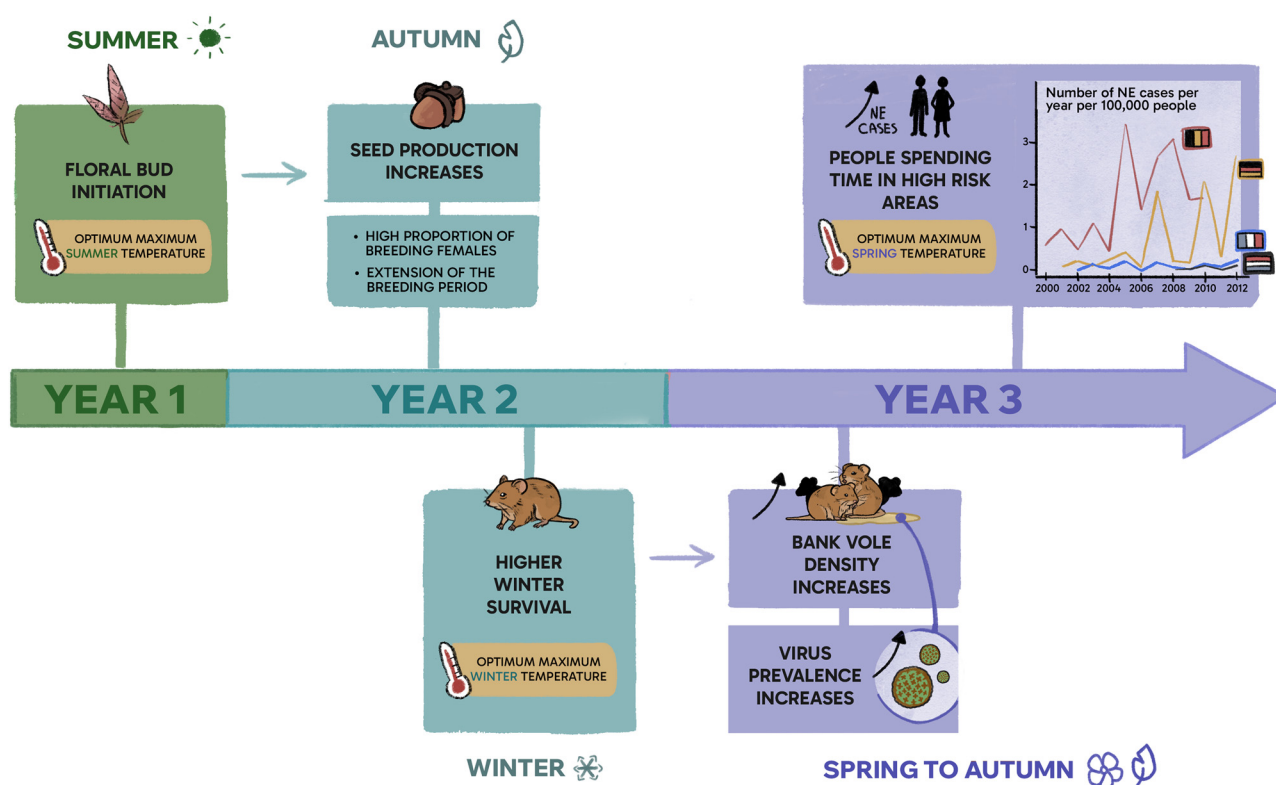
Figure 5. Trophic cascade linking climatic fluctuation, tree masting, bank vole population dynamics, and nephropathia epidemica (NE) outbreaks in western Europe. Summer temperatures are crucial for initiating floral buds in trees,[81,82] which is essential for tree masting in the following autumn. During mast years, abundant seed production significantly influences bank vole populations by prolonging the breeding season, increasing the proportion of breeding females, and improving winter survival.[31–33] These combined events result in higher bank vole population densities that can persist into the following spring. Bank vole population increases are closely associated with increments in *Orthohantavirus puumalaense* (PUUV) prevalence and NE outbreaks in western Europe.[9,29] Therefore, this process illustrates a 3-year trophic cascade, with tree masting serving as a key intermediary connecting climatic fluctuation patterns to NE outbreaks.[9] Figure created using Adobe Photoshop 2025 (Adobe Inc.).

outbreak may support an increase in NE cases. This association is not easily explained by tree seeding patterns,[9] as seed fall occurs during this season. A possible explanation could involve the availability of green biomass and vole body condition. During the normal reproductive season, herbaceous plants make up a significant part of the bank vole diet. Typically, green foliage declines in early autumn, though favorable conditions can extend its availability.[37]

In several orthohantavirus systems, climatic fluctuation cascade effects related to rodent food resources have been described as an explanation for the occurrence of local epizootics or epidemics.[13,14,83] A novelty of our work lies in the result that similar climatic fluctuation mechanisms seem to operate over vast geographical areas and can predict epidemic dynamics, despite the noise induced by complex land-cover patterns and socioeconomic factors.[84,85] Such complexity is challenging to parameterize over large regions and is, hence, likely to contribute to the relatively low sensitivity (high standard deviation in some regions of France). We acknowledge the limitation of using data from different administrative levels (LAU2, LAU1, NUTS3) without harmonization, which introduces heterogeneity in spatial resolution and impacts key variables like population density. To address this, we included both population and polygon area as covariates, which proved relevant for model predictions. However, larger polygons may average dense and sparse preferred habitats, masking localized effects and influencing observed relationships between habitat percentage and NE cases. Future studies should consider harmonizing data for improved consistency or smaller polygons for NE data incidence, if available.

Additionally, a large proportion of the observed NE cases was assigned to the location of residence instead of a confirmed PUUV infection source. Even though it is likely that, on average, a strong spatial correlation exists between residence and infection source (NE associations with living near forests and peridomestic activities), it remains a strong assumption that unconfirmed PUUV infection sources correspond to residence location. Forest leisure or work activities are not necessarily conducted near the place of residence.[86,87] We applied lagged effects to climatic variables influencing masting events, which significantly boost bank vole populations and to net primary productivity (NPP) to account for temporal variations in food resource availability. On the contrary, for key masting tree species such as *F. sylvatica* and *Quercus* spp., we relied on static layers, as available broad-scale land-cover datasets (e.g., CORINE and ESA CCI) lack the resolution to capture their dynamics, and adapting these layers for fine-scale modeling was beyond the scope of this study. These limitations highlight the challenges of modeling temporal dynamics without high-resolution data. Future work could incorporate more detailed spatial and temporal datasets to better understand rodent population responses to habitat and food resource changes.

While environmental conditions can forecast heightened risk of nephropathia epidemica (NE), the response of vole populations and tree seeding patterns to these cues may be hindered by physiological or behavioral limitations. In contrast to wood mice, which switch between crop fields and forests based on resource availability and respond differently to masting events, bank voles are less opportunistic and remain in their habitats. Bank voles depend on masting events, which drive population peaks by

providing increased food availability and supporting winter breeding,[36] followed by strong density-dependent regulation.[76] Therefore, the density-dependent population regulation of bank voles and potential tree resource depletion make it highly improbable for outbreak years to occur consecutively.[9,76] Consequently, even if local climatic conditions are ideal for triggering PUUV outbreaks following a significant outbreak year, such as those in Germany in 2007 or Belgium in 2008, it is unlikely that outbreaks of comparable magnitude will occur the following year, as shown in Figure S3. These constraints should be considered when developing PUUV control strategies by using regional thresholds of NE incidence in the previous year.

Local climatic fluctuations and land-cover factors have the potential to impact the epidemics of multiple pathogens across ecoregions in Europe. Our study presents an approach that could be effectively applied to distinct zoonoses and geographical areas. This was previously hindered by methodological constraints but is now overcome with the capabilities of R-INLA, which includes handling zero-inflated data while considering spatial–temporal correlations.[69] Exploiting the recent developments of Bayesian spatiotemporal modeling (i.e., R-INLA) using epidemiological and environmental data allows disentangling epidemiological drivers and comparison between ecoregions as well as predicting the risk of hotspots.

In contrast to prior research,[73,74,88,89] which focus on small regions or a single country and, in most cases, either emphasize time or space exclusively, our study is a pioneering effort aiming to understand the spatiotemporal dynamics of NE incidence across a wide expanse of the western Europe cases at a high resolution. While Kazasidis et al.[73] attempted to develop a PUUV early warning system in Germany using machine learning, their model showcased a limited predictive performance when predicting widespread high outbreak risk, likely due to limitations in capturing spatiotemporal dependencies. Here we used R-INLA, a method specifically designed to account for spatiotemporal correlations, providing a more performant framework for predicting PUUV outbreak risk. Expert knowledge on host habitat preference, as well as beech and oak presence, were fundamental to determine the location of risk hotspots, but seasonal climate data were required to accurately establish hotspot intensity throughout time and subsequent outbreaks. Although our model currently relies on historical data, it demonstrates the strong potential to produce yearly maps for NE incidence—a key milestone for implementing targeted risk prevention strategies within the study region (Figure S3). With additional input data aiming to improve sensitivity, and following an extended validation period, this modeling framework is promising as a forecasting tool, allowing targeted interventions and preparedness. Fine-scale warning tools are particularly relevant in the context of climate change, which is projected to increase PUUV outbreak frequency by 3- to 4-fold by the end of this century.[29]

Acknowledgments

We are grateful to two anonymous reviewers for their constructive and helpful comments, as well as to Markus Neteler, Els Ducheyne, Luigi Sedda, and David Rogers for their support in the initial phase of the project. We also acknowledge all persons involved in the data collection conducted by *a)* the Belgian network of sentinel laboratories and the service Epidemiology of Infectious Diseases at Sciensano, the National Public Health Institute of Belgium; *b)* Robert Koch Institute, Department for Infectious Disease Epidemiology in Germany; *c)* the Virology Department, Centre for Infectious Disease Control, National Institute for Public Health and the Environment in The Netherlands; and *d)* the Centre National de Référence des Hantavirus, Unité de Environnement et Risques Infectieux at Institut Pasteur in France.

References

1. Kilpatrick AM, Randolph SE. 2012. Drivers, dynamics, and control of emerging vector-borne zoonotic diseases. Lancet 380(9857):1946–1955, PMID: 23200503, https://doi.org/10.1016/S0140-6736(12)61151-9.

2. Estrada-Peña A, Ostfeld RS, Peterson AT, Poulin R, De La Fuente J. 2014. Effects of environmental change on zoonotic disease risk: an ecological primer. Trends Parasitol 30(4):205–214, PMID: 24636356, https://doi.org/10.1016/j.pt.2014.02.003.

3. Altizer S, Dobson A, Hosseini P, Hudson P, Pascual M, Rohani P. 2006. Seasonality and the dynamics of infectious diseases. Ecol Lett 9(4):467–484, PMID: 16623732, https://doi.org/10.1111/j.1461-0248.2005.00879.x.

4. Gottdenker NL, Streicker DG, Faust CL, Carroll CR. 2014. Anthropogenic land use change and infectious diseases: a review of the evidence. EcoHealth 11(4):619–632, PMID: 24854248, https://doi.org/10.1007/s10393-014-0941-z.

5. Altizer S, Ostfeld RS, Johnson PTJ, Kutz S, Harvell CD. 2013. Climate change and infectious diseases: from evidence to a predictive framework. Science 341(6145):514–519, PMID: 23908230, https://doi.org/10.1126/science.1239401.

6. Lafferty KD. 2009. The ecology of climate change and infectious diseases. Ecology 90(4):888–900, PMID: 19449681, https://doi.org/10.1890/08-0079.1.

7. Morand S, Owers KA, Waret-Szkuta A, McIntyre KM, Baylis M. 2013. Climate variability and outbreaks of infectious diseases in Europe. Sci Rep 3(1):1774, PMID: 23639950, https://doi.org/10.1038/srep01774.

8. Semenza JC, Sudre B, Oni T, Suk JE, Giesecke J. 2013. Linking environmental drivers to infectious diseases: the European environment and epidemiology network. PLoS Negl Trop Dis 7(7):e2323, PMID: 23936561, https://doi.org/10.1371/journal.pntd.0002323.

9. Tersago K, Verhagen R, Servais A, Heyman P, Ducoffre G, Leirs H. 2009. Hantavirus disease (nephropathia epidemica) in Belgium: effects of tree seed production and climate. Epidemiol Infect 137(2):250–256, PMID: 18606026, https://doi.org/10.1017/S0950268808000940.

10. Metcalf CJE, Walter KS, Wesolowski A, Buckee CO, Shevliakova E, Tatem AJ, et al. 2017. Identifying climate drivers of infectious disease dynamics: recent advances and challenges ahead. Proc R Soc B Biol B 284(1860):20170901, https://doi.org/10.1098/rspb.2017.0901.

11. Stewart-Ibarra AM, Rollock L, Best S, Brown T, Diaz AR, Dunbar W, et al. 2022. Co-learning during the co-creation of a dengue early warning system for the health sector in Barbados. BMJ Glob Health 7(1):e007842, PMID: 34992079, https://doi.org/10.1136/bmjgh-2021-007842.

12. Finch E, Lotto Batista M, Alcayna T, Lee SA, Fletcher IK, Lowe R. 2024. Early warning systems for vector-borne diseases: engagement, methods and implementation. In: *Planetary Health Approaches to Understand and Control Vector-Borne Diseases*. Wageningen, the Netherlands: Wageningen Academic, 347–386.

13. Mills JN, Gage KL, Khan AS. 2010. Potential influence of climate change on vector-borne and zoonotic diseases: a review and proposed research plan. Environ Health Perspect 118(11):1507–1514, PMID: 20576580, https://doi.org/10.1289/ehp.0901389.

14. Dearing MD, Dizney L. 2010. Ecology of hantavirus in a changing world. Ann NY Acad Sci 1195(1):99–112, PMID: 20536819, https://doi.org/10.1111/j.1749-6632.2010.05452.x.

15. Kuhn JH, Schmaljohn CS. 2023. A brief history of bunyaviral family *Hantaviridae*. Diseases 11(1):38, PMID: 36975587, https://doi.org/10.3390/diseases11010038.

16. Vaheri A, Strandin T, Hepojoki J, Sironen T, Henttonen H, Mäkelä S, et al. 2013. Uncovering the mysteries of hantavirus infections. Nat Rev Microbiol 11(8):539–550, PMID: 24020072, https://doi.org/10.1038/nrmicro3066.

17. Yanagihara R, Amyx HL, Gajdusek DC. 1985. Experimental infection with Puumala virus, the etiologic agent of nephropathia epidemica, in bank voles (*Clethrionomys glareolus*). J Virol 55(1):34–38, PMID: 2861296, https://doi.org/10.1128/JVI.55.1.34-38.1985.

18. Kallio ER, Klingström J, Gustafsson E, Manni T, Vaheri A, Henttonen H, et al. 2006. Prolonged survival of Puumala hantavirus outside the host: evidence for indirect transmission via the environment. J Gen Virol 87(pt 8):2127–2134, PMID: 16847107, https://doi.org/10.1099/vir.0.81643-0.

19. Hardestam J, Karlsson M, Falk KI, Olsson G, Klingström J, Lundkvist Å. 2008. Puumala hantavirus excretion kinetics in bank voles (*Myodes glareolus*). Emerg Infect Dis 14(8):1209–1215, PMID: 18680643, https://doi.org/10.3201/eid1408.080221.

20. Settergren B. 2000. Clinical aspects of nephropathia epidemica (Puumala virus infection) in Europe: a review. Scand J Infect Dis 32(2):125–132, PMID: 10826895, https://doi.org/10.1080/003655400750045204.

21. Vapalahti O, Mustonen J, Lundkvist Å, Henttonen H, Plyusnin A, Vaheri A. 2003. Hantavirus infections in Europe. Lancet Infect Dis 3(10):653–661, PMID: 14522264, https://doi.org/10.1016/s1473-3099(03)00774-6.

22. Heyman P, Thoma BR, Marié JL, Cochez C, Essbauer SS. 2012. In search for factors that drive hantavirus epidemics. Front Physiol 3:237, PMID: 22934002, https://doi.org/10.3389/fphys.2012.00237.

23. European Centre for Disease Prevention and Control. 2024. Surveillance Atlas of Infectious Diseases. https://atlas.ecdc.europa.eu/public/ [accessed 1 April 2024].

24. Reusken C, Heyman P. 2013. Factors driving hantavirus emergence in Europe. Curr Opin Virol 3(1):92–99, PMID: 23384818, https://doi.org/10.1016/j.coviro.2013.01.002.

25. Reil D, Imholt C, Eccard JA, Jacob J. 2015. Beech fructification and bank vole population dynamics - combined analyses of promoters of human Puumala virus infections in Germany. PLoS One 10(7):e0134124, PMID: 26214509, https://doi.org/10.1371/journal.pone.0134124.

26. Andreassen HP, Sundell J, Ecke F, Halle S, Haapakoski M, Henttonen H, et al. 2021. Population cycles and outbreaks of small rodents: ten essential questions we still need to solve. Oecologia 195(3):601–622, PMID: 33369695, https://doi.org/10.1007/s00442-020-04810-w.

27. Hansson L, Henttonen H. 1985. Gradients in density variations of small rodents: the importance of latitude and snow cover. Oecologia 67(3):394–402, PMID: 28311574, https://doi.org/10.1007/BF00384946.

28. Olsson GE, Leirs H, Henttonen H. 2010. Hantaviruses and their hosts in Europe: reservoirs here and there, but not everywhere? Vector Borne Zoonotic Dis 10(6):549–561, PMID: 20795916, https://doi.org/10.1089/vbz.2009.0138.

29. Imholt C, Reil D, Eccard JA, Jacob D, Hempelmann N, Jacob J. 2015. Quantifying the past and future impact of climate on outbreak patterns of bank voles (*Myodes glareolus*). Pest Manag Sci 71(2):166–172, PMID: 24889216, https://doi.org/10.1002/ps.3838.

30. Cunze S, Kochmann J, Kuhn T, Frank R, Dörge DD, Klimpel S. 2018. Spatial and temporal patterns of human puumala virus (PUUV) infections in Germany. PeerJ 6:e4255, PMID: 29404206, https://doi.org/10.7717/peerj.4255.

31. Clement J, Maes P, Van Ranst M. 2006. Hantaviruses in the old and new world. Perspect Med Virol 16:161–177, https://doi.org/10.1016/S0168-7069(06)16008-5.

32. Clement J, Vercauteren J, Verstraeten WW, Ducoffre G, Barrios JM, Vandamme A-M, et al. 2009. Relating increasing hantavirus incidences to the changing climate: the mast connection. Int J Health Geogr 8:1, PMID: 19149870, https://doi.org/10.1186/1476-072X-8-1.

33. Piovesan G, Adams JM. 2001. Masting behaviour in beech: linking reproduction and climatic variation. Can J Bot 79(9):1039–1047.

34. Smyth M. 1966. Winter breeding in woodland mice, *Apodemus sylvaticus*, and voles, *Clethrionomys glareolus* and *Microtus agrestis*, near Oxford. J Anim Ecol 35(3):471–485, https://doi.org/10.2307/2486.

35. Ylonen H, Viitala J. 1985. Social organization of an enclosed winter population of the bank vole Clethrionomys glareolus. Ann Zool Fennici 22(3):353–358.

36. Jensen TS. 1982. Seed production and outbreaks of non-cyclic rodent populations in deciduous forests. Oecologia 54(2):184–192, PMID: 28311427, https://doi.org/10.1007/BF00378391.

37. Pucek Z, Jędrzejewski W, Jędrzejewska B, Pucek M. 1993. Rodent population dynamics in a primeval deciduous Forest (Białowieża National Park) in relation to weather, seed crop, and predation. Acta Theriol 38:199–232, https://doi.org/10.4098/AT.arch.93-18.

38. Crespin L, Verhagen R, Stenseth NC, Yoccoz NG, Prévot-Julliard A, Lebreton J. 2002. Survival in fluctuating bank vole populations: seasonal and yearly variations. Oikos 98(3):467–479, https://doi.org/10.1034/j.1600-0706.2002.980311.x.

39. Sipari S, Khalil H, Magnusson M, Evander M, Hörnfeldt B, Ecke F. 2022. Climate change accelerates winter transmission of a zoonotic pathogen. Ambio 51(3):508–517, PMID: 34228253, https://doi.org/10.1007/s13280-021-01594-y.

40. Olsson GE, Dalerum F, Hörnfeldt B, Elgh F, Palo TR, Juto P, et al. 2003. Human hantavirus infections, Sweden. Emerg Infect Dis 9(11):1395–1401, PMID: 14718081, https://doi.org/10.3201/eid0911.030275.

41. Sauvage F, Penalba C, Vuillaume P, Boue F, Coudrier D, Pontier D, et al. 2002. Infection in humans and in the reservoir host, Ardennes region, France. Emerg Infect Dis 8(12):1509–1511, PMID: 12498675, https://doi.org/10.3201/eid0812.010518.

42. Huston MA, Wolverton S. 2009. The global distribution of net primary production: resolving the paradox. Ecol Monogr 79(3):343–377, https://doi.org/10.1890/08-0588.1.

43. Loehman RA, Elias J, Douglass RJ, Kuenzi AJ, Mills JN, Wagoner K. 2012. Prediction of *Peromyscus maniculatus* (deer mouse) population dynamics in Montana, USA, using satellite-driven vegetation productivity and weather data. J Wildl Dis 48(2):348–360, PMID: 22493110, https://doi.org/10.7589/0090-3558-48.2.348.

44. Barrios JM, Verstraeten WW, Maes P, Clement J, Aerts J-M, Haredasht SA, et al. 2010. Satellite derived Forest phenology and its relation with nephropathia epidemica in Belgium. Int J Environ Res Public Health 7(6):2486–2500, PMID: 20644685, https://doi.org/10.3390/ijerph7062486.

45. Barrios JM, Verstraeten WW, Maes P, Aerts JM, Farifteh J, Coppin P. 2013. Relating land cover and spatial distribution of nephropathia epidemica and Lyme borreliosis in Belgium. Int J Environ Health Res 23(2):132–154, PMID: 22894742, https://doi.org/10.1080/09603123.2012.708918.

46. Ganguly S, Friedl MA, Tan B, Zhang X, Verma M. 2010. Land surface phenology from MODIS: characterization of the collection 5 global land cover dynamics product. Remote Sens Environ 114(8):1805–1816, https://doi.org/10.1016/j.rse.2010.04.005.

47. Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, et al. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol Evol 24(3):127–135, PMID: 19185386, https://doi.org/10.1016/j.tree.2008.10.008.

48. Bujalska G. 1990. Social system of the bank vole, *Clethrionomys glareolus*. In: *Social Systems and Population Cycles in Voles*. Tamarin RH, Ostfeld RS, Pugh SR, Bujalska G, eds. Basel, Switzerland: Birkhäuser, 155–167.

49. Spitzenberger F. 2000. *Clethrionomys glareolus*. In: *The Atlas of European Mammals*, vol. 75. Cambridge, MA: Academic Press.

50. Viro P, Niethammer J. 1982. *Clethrionomys glareolus* (Schreber, 1780). In: *Handbuch Der Säugetiere Europas. ötelmaus*. Niethammer J, Krapp F, eds. Leipzig, Germany: Akademische Verlagsgesellschaft.

51. Alain B, Gilles P, Yannick D. 2006. Factors driving small rodents assemblages from field boundaries in agricultural landscapes of Western France. Landscape Ecol 21(3):449–461, https://doi.org/10.1007/s10980-005-4118-6.

52. Gliwicz J, Ims RA. 2000. Dispersal in the bank vole. Polish J Ecol 48:51–61.

53. Kozakiewicz M, Chołuj A, Kozakiewicz A. 2007. Long-distance movements of individuals in a free-living bank vole population: an important element of male breeding strategy. Acta Theriol 52(4):339–348, https://doi.org/10.1007/BF03194231.

54. Mazurkiewicz M. 1994. Factors influencing the distribution of the bank vole in Forest habitats. Acta Theriol 39(2):113–126, https://doi.org/10.4098/AT.arch.94-16.

55. Van Apeldoorn RC, Oostenbrink WT, Van Winden A, Van Der Zee FF. 1992. Effects of habitat fragmentation on the bank vole, *Clethrionomys glareolus*, in an agricultural landscape. Oikos 65(2):265–274, https://doi.org/10.2307/3545018.

56. Birkedal M, Fischer A, Karlsson M, Löf M, Madsen P. 2009. Rodent impact on establishment of direct-seeded *Fagus sylvatica*, *Quercus robur* and *Quercus petraea* on forest land. Scand J For Res 24(4):298–307, https://doi.org/10.1080/02827580903055125.

57. Shrestha H, McCulloch K, Hedtke SM, Grant WN. 2022. Geospatial modeling of pre-intervention nodule prevalence of *Onchocerca volvulus* in Ethiopia as an aid to onchocerciasis elimination. PLoS Negl Trop Dis 16(7):e0010620, PMID: 35849615, https://doi.org/10.1371/journal.pntd.0010620.

58. Moraga P, Cano J, Baggaley RF, Gyapong JO, Njenga SM, Nikolay B, et al. 2015. Modelling the distribution and transmission intensity of lymphatic filariasis in sub-Saharan Africa prior to scaling up interventions: integrated use of geostatistical and mathematical modelling. Parasit Vectors 8(1):560, PMID: 26496983, https://doi.org/10.1186/s13071-015-1166-x.

59. Kang SY, Battle KE, Gibson HS, Ratsimbasoa A, Randrianarivelojosia M, Ramboarina S, et al. 2018. Spatio-temporal mapping of Madagascar's malaria indicator survey results to assess Plasmodium falciparum endemicity trends between 2011 and 2016. BMC Med 16(1):71, PMID: 29788968.

60. Blangiardo M, Cameletti M. 2015. *Spatial and Spatio-temporal Bayesian Models with R-INLA*. 1st ed. Hoboken, NJ: Wiley.

61. Rue H, Martino S, Chopin N. 2009. Approximate Bayesian inference for latent gaussian models by using integrated nested Laplace approximations. J R Stat Soc Ser B Stat Methodol 71(2):319–392, https://doi.org/10.1111/j.1467-9868.2008.00700.x.

62. Karagiannis-Voules DA, Scholte RGC, Guimarães LH, Utzinger J, Vounatsou P. 2013. Bayesian geostatistical modeling of Leishmaniasis incidence in Brazil. PLoS Negl Trop Dis 7(5):e2213, PMID: 23675541, https://doi.org/10.1371/journal.pntd.0002213.

63. Osgood-Zimmerman A, Millear AI, Stubbs RW, Shields C, Pickering BV, Earl L, et al. 2018. Mapping child growth failure in Africa between 2000 and 2015. Nature 555(7694):41–47, PMID: 29493591, https://doi.org/10.1038/nature25760.

64. Martins TG, Simpson D, Lindgren F, Rue H. 2013. Bayesian computing with INLA: new features. Comput Stat Data Anal 67:68–83, https://doi.org/10.1016/j.csda.2013.04.014.

65. Moraga P. 2019. *Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny*. 1st ed. Boca Raton, FL: Chapman and Hall/CRC.

66. Nyandwi E, Osei FB, Veldkamp T, Amer S. 2020. Modeling schistosomiasis spatial risk dynamics over time in Rwanda using zero-inflated Poisson

regression. Sci Rep 10(1):19276, PMID: 33159143, https://doi.org/10.1038/s41598-020-76288-8.

67. Lindgren F, Rue H, Lindström J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J R Stat Soc Ser B Stat Methodol 73(4):423–498, https://doi.org/10.1111/j.1467-9868.2011.00777.x.

68. Cameletti M, Lindgren F, Simpson D, Rue H. 2013. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. AStA Adv Stat Anal 97(2):109–131, https://doi.org/10.1007/s10182-012-0196-3.

69. Zuur AF, Ieno EN, Saveliev AA. 2018. *GAM and Zero-Inflated Models*. Newburgh, United Kingdom: Highland Statistics Ltd.

70. Moraga P, Dean C, Inoue J, Morawiecki P, Noureen SR, Wang F. 2021. Bayesian spatial modelling of geostatistical data using INLA and SPDE methods: a case study predicting malaria risk in Mozambique. Spat Spatiotemporal Epidemiol 39:100440, PMID: 34774255, https://doi.org/10.1016/j.sste.2021.100440.

71. Leach CB, Webb CT, Cross PC. 2016. When environmentally persistent pathogens transform good habitat into ecological traps. R Soc Open Sci 3(3):160051, PMID: 27069672, https://doi.org/10.1098/rsos.160051.

72. Dougherty ER, Seidel DP, Blackburn JK, Turner WC, Getz WM. 2022. A framework for integrating inferred movement behavior into disease risk models. Mov Ecol 10(1):31, PMID: 35871637, https://doi.org/10.1186/s40462-022-00331-8.

73. Kazasidis O, Geduhn A, Jacob J. 2024. High-resolution early warning system for human Puumala hantavirus infection risk in Germany. Sci Rep 14(1):9602, PMID: 38671000, https://doi.org/10.1038/s41598-024-60144-0.

74. Swart A, Bekker DL, Maas M, de Vries A, Pijnacker R, Reusken CBEM, et al. 2017. Modelling human Puumala hantavirus infection in relation to bank vole abundance and masting intensity in the Netherlands. Infect Ecol Epidemiol 7(1):1287986, PMID: 28567209, https://doi.org/10.1080/20008686.2017.1287986.

75. Sauvage F, Langlais M, Pontier D. 2007. Predicting the emergence of human hantavirus disease using a combination of viral dynamics and rodent demographic patterns. Epidemiol Infect 135(1):46–56, PMID: 16753079, https://doi.org/10.1017/S0950268806006595.

76. Tersago K, Verhagen R, Vapalahti O, Heyman P, Ducoffre G, Leirs H. 2011. Hantavirus outbreak in Western Europe: reservoir host infection dynamics related to human disease patterns. Epidemiol Infect 139(3):381–390, PMID: 20450527, https://doi.org/10.1017/S0950268810000956.

77. Tersago K, Schreurs A, Linard C, Verhagen R, Van Dongen S, Leirs H. 2008. Population, environmental, and community effects on local bank vole (*Myodes glareolus*) Puumala virus infection in an area with low human incidence. Vector Borne Zoonotic Dis 8(2):235–244, PMID: 18370592, https://doi.org/10.1089/vbz.2007.0160.

78. Piechotowski I, Brockmann SO, Schwarz C, Winter CH, Ranft U, Pfaff G. 2008. Emergence of hantavirus in South Germany: rodents, climate and human

infections. Parasitol Res 103(suppl 1):131–137, PMID: 19030895, https://doi.org/10.1007/s00436-008-1055-8.

79. Faber M, Wollny T, Schlegel M, Wanka KM, Thiel J, Frank C, et al. 2013. Puumala virus outbreak in Western Thuringia, Germany, 2010: epidemiology and strain identification. Zoonoses Public Health 60(8):549–554, PMID: 23398736, https://doi.org/10.1111/zph.12037.

80. Ettinger J, Hofmann J, Enders M, Tewald F, Oehme RM, Rosenfeld UM, et al. 2012. Multiple synchronous outbreaks of Puumala virus, Germany, 2010. Emerg Infect Dis 18(9):1461–1464, PMID: 22932394, https://doi.org/10.3201/eid1809.111447.

81. Gurnell J. 1993. Tree seed production and food conditions for rodents in an oak wood in Southern England. Forestry 66(3):291–315, https://doi.org/10.1093/forestry/66.3.291.

82. Schauber EM, Kelly D, Turchin P, Simon C, Lee WG, Allen RB, et al. 2002. Masting by eighteen New Zealand plant species: the role of temperature as a synchronizing cue. Ecology 83(5):1214–1225.

83. Klempa B. 2009. Hantaviruses and climate change. Clin Microbiol Infect 15(6):518–523, PMID: 19604276, https://doi.org/10.1111/j.1469-0691.2009.02848.x.

84. Linard C, Lamarque P, Heyman P, Ducoffre G, Luyasu V, Tersago K, et al. 2007. Determinants of the geographic distribution of Puumala virus and Lyme borreliosis infections in Belgium. Int J Health Geogr 6(1):15, PMID: 17474974, https://doi.org/10.1186/1476-072X-6-15.

85. Linard C, Tersago K, Leirs H, Lambin EF. 2007. Environmental conditions and Puumala virus transmission in Belgium. Int J Health Geogr 6(1):55, PMID: 18078526, https://doi.org/10.1186/1476-072X-6-55.

86. Van Loock F, Thomas I, Clement J, Ghoos S, Colson P. 1999. A case-control study after a hantavirus infection outbreak in the south of Belgium: who is at risk? Clin Infect Dis 28(4):834–839, PMID: 10825047, https://doi.org/10.1086/515196.

87. Abu Sin M, Stark K, van Treeck U, Dieckmann H, Uphoff H, Hautmann W, et al. 2007. Risk factors for hantavirus infection in Germany, 2005. Emerg Infect Dis 13(9):1364–1366, PMID: 18252110, https://doi.org/10.3201/eid1309.070552.

88. Haredasht SA, Taylor CJ, Maes P, Verstraeten WW, Clement J, Barrios M, et al. 2013. Model-based prediction of nephropathia epidemica outbreaks based on climatological and vegetation data and bank vole population dynamics. Zoonoses Public Health 60(7):461–477, PMID: 23176630, https://doi.org/10.1111/zph.12021.

89. Vanwambeke SO, Zeimes CB, Drewes S, Ulrich RG, Reil D, Jacob J. 2019. Spatial dynamics of a zoonotic orthohantavirus disease through heterogenous data on rodents, rodent infections, and human disease. Sci Rep 9(1):2329, PMID: 30787344, https://doi.org/10.1038/s41598-019-38802-5.