Journal of Biological Dynamics
iFirst, 2012, 1–16

Numerical simulation of a susceptible–exposed–infectious space-continuous model for the spread of rabies in raccoons across a realistic landscape

Joshua P. Keller[a], Luca Gerardo-Giorda[b]* and Alessandro Veneziani[c]

[a]*Department of Biostatistics, University of Washington, Seattle, WA, USA;* [b]*BCAM – Basque Center for Applied Mathematics, Bilbao, Spain;* [c]*Department of Mathematics and Computer Science, Emory University, Atlanta, GA, USA*

(*Received 2 May 2012; final version received 4 October 2012*)

We introduce a numerical model for the spread of a lethal infectious disease in wildlife. The reference model is a Susceptible–Exposed–Infectious system where the spatial component of the dynamics is modelled by a diffusion process. The goal is to develop a model to be used for real geographical scenarios, so we do not rely upon simplifying assumptions on the shape of the region of interest. For this reason, space discretization is carried out with the finite element method on an unstructured triangulation. A diffusion term is designed to take into account landscape heterogeneities such as mountains and waterways. Numerical simulations are carried out for rabies epidemics among raccoons in New York state. A qualitative comparison of numerical results to available data from real-world epidemics is discussed.

Keywords: infectious diseases; landscape modelling; numerical approximation; finite elements

AMS Subject Classifications: 65Mxx; 65N30; 92D30; 92D40

1. Introduction

Epidemiologists have long used mathematical models to simulate the transmission of a disease throughout populations. These models allow for an understanding of the processes at work and for predicting future epidemics. The large majority of these epidemiological models have a compartmentalized design and are composed of a system of ordinary differential equations in time. Though some aspects of compartmentalized models are extremely well developed (such as the rate of infection and stability of solutions for simple cases), these models do not always account for the movement of individuals from one region to another. This can be a significant limitation, especially when dealing with wildlife diseases, as the daily movement and seasonal migration of animals can have major effects on the transmission of the disease itself. Describing the movement of individuals across a region in a mathematically rigorous and epidemiologically tractable way has been done following different approaches [14,18]. The canonical procedure is to discretize the

*Corresponding author. Email: lgerardo@bcamath.org
Author Emails: kellerjp@uw.edu; ale@mathcs.emory.edu

ISSN 1751-3758 print/ISSN 1751-3766 online

http://dx.doi.org/10.1080/17513758.2012.742578
http://www.tandfonline.com

population and the geography into geopolitical units and to consider the movement of individuals from unit to unit [18,29]. Although better than ignoring movement, this approach suffers from several limitations. The average movement of animals generally does not fit the scale and shape of geopolitical units, requiring the introduction of artifacts such as long-distance translocations [30], whose features can be difficult to interpret. Moreover, geopolitical subdivision is of little efficacy when one wants to include the geographical heterogeneities of the landscape.

In this article, we model movement as a continuous process across the geographical region of interest by means of a time-dependent diffusion process. This results in a system of partial differential equations (PDEs) of parabolic form [14,17,28]. Parabolic (diffusion) PDE's can be regarded as a macroscopic representation of Brownian motion and are therefore a natural candidate for modelling the movement of wild animals. Our aim is to describe the spread of the infection in a real two-dimensional region of interest Ω, with no simplifying assumptions on the shape of the region. Simple finite difference schemes, as used for instance by Miller Neilan and Lehnart [22], cannot incorporate realistic geographic features. Instead, we turn to unstructured discretization grids, which naturally lead to using finite elements in space, as they are well suited for these kinds of triangulations. Other researchers have used finite elements for the diffusion of age-structured populations [1,7,12,19,23], but carried out numerical tests only on simple geometries such as rectangles or one-dimensional intervals. Finite elements in an epidemiological context are considered by Kim and Park [20], but they focus only on proving stability and convergence results for the scheme and do not provide numerical results.

For illustration purposes, we focus throughout this article on the diffusion of rabies among raccoons in the state of New York in the USA. A representation of the region of interest has been retrieved by satellite maps, and the specific geographical features are included in the coefficients of the diffusive terms. Rabies, a viral encephalomyelitis, has the longest extant record of reports of any zoonotic disease in the USA. Raccoons (*Procyon lotor*) are a major carrier of the disease across North America, together with foxes, bats and skunks [26,31]. While a successful vaccine for humans and other mammals against rabies exists, and human fatalities are rare, the economic cost of widespread animal vaccination and the danger of rabies to threatened species makes understanding the spread of rabies in wildlife hosts important to epidemiologists [27,31]. It is worth noticing, however, that the results we provide can be extended to other types of infectious diseases in wildlife and to different regions of interest.

The structure of the article is as follows. In Section 2, we introduce the Susceptible–Exposed–Infectious (SEI) model with diffusion. We describe in Section 3 how we incorporate spatial heterogeneities of the landscape in the diffusion coefficients. Section 4 introduces the space and time discretization of the associated system. Finally, in Section 5, we provide a proof of concept of our numerical model by simulating the disease in New York. We first perform a sensitivity study with respect to the parameters of the model. Subsequently, in order to investigate a practical application, we simulate the spread of an epidemic of rabies in raccoons from 1990 to 1994, and compare those results qualitatively with the available public health data [25].

1.1. *Notation*

Let $\Omega \subset \mathbb{R}^2$ be a two-dimensional domain of interest with coordinates x, y. We denote by $s(x, y, t)$, $e(x, y, t)$, and $i(x, y, t)$ the densities of the *susceptible* (individuals that can be infected by the disease), *exposed* (individuals that have been exposed to the virus but do not have evident symptoms yet), and *infectious* populations, respectively, at location (x, y) and time t. The total number of SEI individuals at time t are given by

$$S(t) = \int_\Omega s(x, y, t)\,\mathrm{d}x\,\mathrm{d}y \quad E(t) = \int_\Omega e(x, y, t)\,\mathrm{d}x\,\mathrm{d}y \quad I(t) = \int_\Omega i(x, y, t)\,\mathrm{d}x\,\mathrm{d}y,$$

respectively. We introduce the auxiliary variable,

$$n_{\text{tot}}(x, y, t) = s(x, y, t) + e(x, y, t) + i(x, y, t),$$

representing the population density at location (x, y) and time t, and the total population at time t is given, therefore, by

$$N_{\text{tot}}(t) = \int_{\Omega} n_{\text{tot}}(x, y, t)\, \mathrm{d}x\, \mathrm{d}y.$$

The vector of unknowns $[s, e, i]^{\mathrm{T}}(x, y, t)$ will be denoted by $\boldsymbol{u}(x, y, t)$.

In this article, we refer to the variational formulation of the SEI model, which relies upon the definition of proper functional spaces. In particular, we denote by $L^2(\Omega)$ the space of functions f of x, y such that $\int_{\Omega}(f(x, y))^2\, \mathrm{d}x\, \mathrm{d}y < \infty$. If, in addition, $\|f\|_1^2 \equiv \int_{\Omega}(f^2 + (\partial_x f)^2 + (\partial_y f)^2)\, \mathrm{d}x\, \mathrm{d}y < \infty$, we say that the function belongs to the Sobolev space $H^1(\Omega)$. When f is also function of time, the quantity $\|f\|_1$ depends on time. If this function is such that $\int_0^T \|f\|_1^2(t)dt < \infty$, we say that $f \in L^2(0, T; H^1(\Omega))$. We use the same notation with bold font for vector functions whose components each belong to this space. Bold characters will be used generally for vectors or vector functions (according to the specific context).

2. The SEI model with diffusion

The SEI model divides the population into SEI individuals. In more general cases, a fourth group of recovered individuals could be considered (resulting in an SEIR model). For simplicity, we do not include the effect of vaccination in this work, and so the recovered class is empty in this setting, since rabies is lethal for raccoons. We point out, however, that the following arguments can be straightforwardly extended to an SEIR model.

We model the spread of an epidemic in Ω by the following nonlinear system of reaction–diffusion equations of parabolic type:

$$\begin{aligned}
\partial_t s &= \alpha(s + e) - \beta s i - \mu_0 s n_{\text{tot}} + \operatorname{div}(\boldsymbol{\nu}_s \nabla s),\\
\partial_t e &= \beta s i - \sigma e - \mu_0 e n_{\text{tot}} + \operatorname{div}(\boldsymbol{\nu}_e \nabla e),\\
\partial_t i &= \sigma e - \mu_0 i n_{\text{tot}} - \phi i + \operatorname{div}(\boldsymbol{\nu}_i \nabla i)
\end{aligned} \tag{1}$$

for all $(x, y, t) \in \Omega \times (0, T)$. Here, we assume that only susceptible and exposed individuals are able to reproduce, and we denote by α the (constant) reproduction rate. Such an assumption is quite reasonable, since on the one hand, the expected survival of a rabid raccoon is much too short to carry on a pregnancy, and on the other hand, even in the case a pregnant rabid female is able to deliver, her offspring are very unlikely to survive, as the mother is going to die shortly. In a more general epidemic, where the life expectancy of an infectious individual is longer, such an assumption can be relaxed, and the birth term can be replaced by αn_{tot}. In Equation (1), β denotes the infectiveness of a contact between a susceptible and an infectious individual, σ the reciprocal of the latency period, and ϕ the reciprocal of the life expectancy of a rabid raccoon. We assume a density-dependent mortality rate in the absence of epidemic $\mu_0 n_{\text{tot}}$. The diffusion coefficients are organized in three 2×2 tensors, $\boldsymbol{\nu}_s, \boldsymbol{\nu}_e, \boldsymbol{\nu}_i$, designed to take into account the heterogeneities of the landscape, and are thus dependent on the spatial location. The epidemiological parameters above (β, σ, ϕ) can be considered constant for a well-studied and understood disease. Concerning the parameters of the vital dynamics, the mortality rate μ_0 depends on the carrying capacity of the environment, it can vary from one region to another to account for hospitality of different areas, but it can be considered constant within the regions. It can also change in time to model

abrupt variations in the environment itself (flood, drought, fires). Finally, the reproduction rate α can be either considered constant to average births along the year or time-dependent to introduce seasonality in the process.

The vector compact representation of system (1) reads

$$\partial_t \mathbf{u} - \operatorname{div}(\boldsymbol{\nu}\nabla\mathbf{u}) = (\mathrm{A} - \mathrm{B}(\mathbf{u}))\mathbf{u} \tag{2}$$

with

$$\mathbf{u} = \begin{bmatrix} s \\ e \\ i \end{bmatrix}, \quad \boldsymbol{\nu} = \begin{bmatrix} \boldsymbol{\nu}_s & 0 & 0 \\ 0 & \boldsymbol{\nu}_e & 0 \\ 0 & 0 & \boldsymbol{\nu}_i \end{bmatrix}, \quad \mathrm{A} = \begin{bmatrix} \alpha & \alpha & 0 \\ 0 & -\sigma & 0 \\ 0 & \sigma & -\phi \end{bmatrix},$$

$$\mathrm{B}(\mathbf{u}) = \begin{bmatrix} -\mu_0 n_{\mathrm{tot}} & -\beta s & 0 \\ \beta i & -\mu_0 n_{\mathrm{tot}} & 0 \\ 0 & 0 & -\mu_0 n_{\mathrm{tot}} \end{bmatrix}.$$

This representation points out that our system is in fact a vector generalization of the well-known (scalar) normalized Fisher's equation

$$\partial_t u - \partial_{xx} u = (1 - u)u$$

that has been introduced in [10] for genetic studies. An analysis of this equation and, more specifically, of its *travelling waves* solutions (i.e. solutions of the form $u(x,t) = g(x - \gamma t)$, for $g(\cdot)$ a function and γ a scalar number to be identified) can be found in [28].

System (1) is completed by some suitable initial conditions

$$s(x,y,0) = s^0(x,y) > 0, \quad e(x,y,0) = e^0(x,y) \geq 0, \quad i(x,y,0) = i^0(x,y) \geq 0$$

for $(x,y) \in \Omega$, and by boundary conditions to be prescribed over the entire boundary $\partial\Omega$. If we assume that the region of interest is isolated, we prescribe homogeneous Neumann boundary conditions

$$\boldsymbol{n} \cdot \nabla s = \boldsymbol{n} \cdot \nabla e = \boldsymbol{n} \cdot \nabla i = 0 \quad \text{for } (x,y,t) \in \partial\Omega \times (0,T).$$

Other boundary conditions could be considered as well. For instance, a non-zero function of time in the Neumann conditions corresponds to some (known) migratory flux. To instead force $s = e = i = 0$ (homogeneous Dirichlet conditions) would assume that outside the region of interest, the habitat is so hostile that no individual survives. Alternatively, a non-homogeneous Dirichlet condition would assume that data on the actual size of the different populations at the region boundary are available. Robin conditions could also be used to have the migration dynamics depend on the population itself.

Hereafter, we consider an isolated system (homogeneous Neumann conditions).

2.1. *Variational formulation of the problem*

The finite element solution of the problem relies upon its variational formulation. Let $\boldsymbol{\nu}(x,y,t)$ be an arbitrary three-component vector in $\mathbf{H}^1(\Omega)$. We multiply Equation (2) by $\boldsymbol{\nu}$ and integrate over Ω to get

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \boldsymbol{u}\cdot\boldsymbol{\nu} + \boldsymbol{\nu}\nabla\boldsymbol{u} : \nabla\boldsymbol{\nu}\,\mathrm{d}\omega = \int_\Omega (\mathrm{A} - \mathrm{B}(\boldsymbol{u}))\boldsymbol{u}\cdot\boldsymbol{\nu}\,\mathrm{d}\omega, \tag{3}$$

where we have exploited the Green formula and the homogeneous Neumann conditions. The variational formulation of the problem we consider can be stated as: Find $\boldsymbol{u}(x,y,t) \in L^2(0,T;\mathbf{H}^1(\Omega))$

such that Equation (3) holds with the initial condition $\boldsymbol{u}(x, y, 0) = \boldsymbol{u}_0(x, y)$. Note that if Ω has a Lipschitz continuous boundary, each term of this formulation is well defined, since thanks to the Sobolev embedding theorem, the product of two functions belonging to $H^1(\Omega)$ in two dimensions is in $L^2(\Omega)$, and so $\int_\Omega (\mathrm{A} - \mathrm{B}(\boldsymbol{u}))\boldsymbol{u} \cdot \boldsymbol{v}\, \mathrm{d}\omega$ is integrable [21].

3. Modelling of landscape heterogeneities

In this work, we consider two kinds of geographical heterogeneity in the region of interest that are modelled in two different ways.

(1) *Localized heterogeneous effects*, such as the close proximity of a major waterway that locally drives the movement of the wild animals.
(2) *Extended heterogeneous regions*, where the dynamics of the infection are significantly affected by particular features. This is the case of lakes, for instance.

In the latter case, we consider a splitting of the domain of interest into homogeneous subregions. In the former one, we modify the diffusivity tensors to describe the effects of the heterogeneity on the population dynamics.

3.1. *Localized heterogeneous effects*

Some features of the region of interest can induce local modifications to the dynamics. This can be modelled by a proper dependence of the coefficients in Equation (1) on the geographical coordinates x, y. For instance, we can argue that mobility of the animals in the mountains is reduced by the presence of physical obstacles. We therefore set a reduced value for the diffusivity in these zones.

Rivers also present an obstacle, as movement across the waterway is restricted. Since their dimension is small in comparison with the region of interest, we do not introduce a bidimensional description of the rivers to avoid useless computational burdens. Instead, we resort to a monodimensional representation, where the waterways' impact on the raccoons' dynamics is represented by a modification of the diffusivity coefficient. More precisely, we reduce the mobility of raccoons across the rivers by a diffusion reduction in the direction normal to the river. In practice, to each point of the river, we associate a local tangent/normal frame of reference (Figure 1, left) with unit vectors $\boldsymbol{e_\tau}, \boldsymbol{e_n}$, respectively. In a neighbourhood Σ of width 2ε of the river the diffusivity is gradually decreased, as indicated in Figure 1, right. Denoting by $\hat{\boldsymbol{\nu}}_k$ $(k = s, e, i)$, the quantities evaluated in the normal/tangent frame of reference, respectively, we have in Σ

$$\hat{\boldsymbol{\nu}}_s = \begin{bmatrix} \nu_H & 0 \\ 0 & \nu_H - (\nu_H - \nu_L)\, \exp\left(1 - \dfrac{\epsilon^2}{\epsilon^2 - \hat{y}^2}\right) \end{bmatrix}, \tag{4}$$

where $\hat{y}$ is the distance from the river along the local normal direction, ν_H and ν_L are the high and low diffusivity values, the former being the 'regular' value and the latter being the diffusivity at the river. The function decreases the value of the diffusion coefficient $\boldsymbol{\nu}$ from ν_H to ν_L.

The tensor (4) is then transformed to the Cartesian global frame of reference, by means of the rotation matrix R, which aligns the local frame of reference to the global one. By standard vector calculus manipulations, we have that

$$\boldsymbol{\nu}_s = \mathrm{R}\hat{\boldsymbol{\nu}}_s\mathrm{R}^T.$$

Similar arguments are used for the diffusivity tensors $\boldsymbol{\nu}_e$ and $\boldsymbol{\nu}_i$.

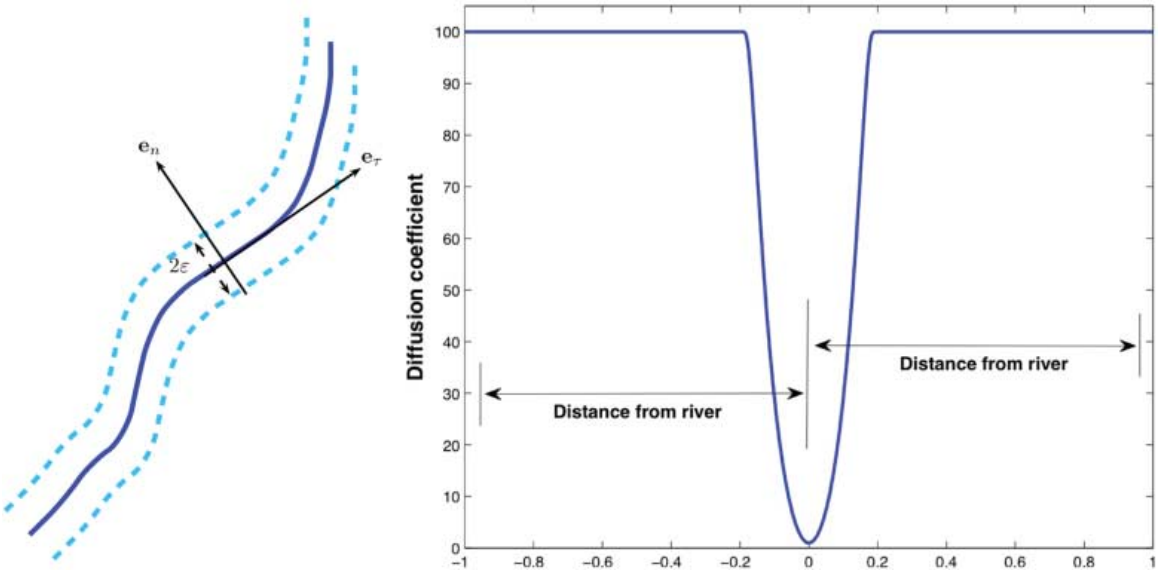

Figure 1. Left: normal/tangent reference frame along the river. Right: diffusion coefficient in the normal direction to the river, with $\nu_H = 100$, $\nu_L = 1$, and $\epsilon = 0.2$.

In particular, we have applied this to the Hudson and Mohawk rivers in our New York state, for example.

Note that in such framework, the diffusivity tensor is a function of the space coordinates only in the neighbourhood of the river, while it is (piecewise) constant in the rest of the domain. As a consequence, the divergence term in the model introduces a convection in the raccoons dynamics induced by the waterways.

3.2. *Strongly heterogeneous subregions*

In some subregions, we can postulate that the population dynamics are significantly different, due to geographical heterogeneities or other specific reasons that require modifications to Equation (1). For instance, if we assume that wild animals are subject to a strong hunting activity in a specific part of the geographical domain, we should locally correct the equations to include the effect of the predators. In the presence of large bodies of water, we can eliminate the region from the domain of interest under the assumption that no significant dynamics are occurring there. In our case, we have introduced an internal boundary to the domain of interest corresponding to lakes, with Neumann boundary conditions describing no flux of animals across the basin.

In general, when different dynamics are coupled, from the numerical standpoint, it is convenient to split the domain into subdomains where each problem is solved separately, in a globally iterative framework [33]. A more specific testing of domain decomposition techniques in the context of population dynamics was investigated in [12].

In the numerical tests of this article, we have considered only the presence of the two major Finger Lakes (Seneca and Cayuga), by excluding them from the computational domain.

4. Numerical approximation

A numerical solution of Equation (3) is obtained after a proper discretization of space and time derivatives. The finite element method (FEM) is particularly suitable for complex geometries such as the ones represented by the geographical regions we are interested in. Space discretization is

therefore performed with FEM while a classical finite difference discretization is carried out for the time discretization [9].

Let us consider a regular triangulation of the domain Ω and assume that the numerical solution denoted by $\boldsymbol{u}_h$ is piecewise linear on each element of the grid. The mesh has N nodes and h denotes a representative dimension of each element. The solution can be written as

$$s_h = \sum_{j=1}^{N} S_j(t)\varphi_j(x,y), \quad e_h = \sum_{j=1}^{N} E_j(t)\varphi_j(x,y), \quad i_h = \sum_{j=1}^{N} I_j(t)\varphi_j(x,y),$$

and $\boldsymbol{u}_h = [s_h, e_h, i_h]^{\mathrm{T}}$, where $\varphi_j(x,y)$ are the piecewise linear functions equal to 1 on the jth node of the mesh and equal to 0 in all the other nodes. The finite element formulation is obtained by plugging this representation of the solution into Equation (3) and selecting alternatively $\boldsymbol{v} = [\varphi_i, 0, 0]^{\mathrm{T}}$, $\boldsymbol{v} = [0, \varphi_i, 0]^{\mathrm{T}}$, $\boldsymbol{v} = [0, 0, \varphi_i]^{\mathrm{T}}$ for $i = 1, \ldots, N$. We obtain a $3N \times 3N$ nonlinear ordinary differential system in the form

$$\mathcal{M}\frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t} + \mathcal{A}\mathbf{U} = \mathcal{F}(\mathbf{U}), \tag{5}$$

where $\mathbf{U} = [[S_i]_{i=1,\ldots,N}, [E_i]_{i=1,\ldots,N}, [I_i]_{i=1,\ldots,N}]$ and

$$\mathcal{M} = \begin{bmatrix} M & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & M \end{bmatrix}, \quad \mathcal{A} = \begin{bmatrix} A_s & 0 & 0 \\ 0 & A_e & 0 \\ 0 & 0 & A_i \end{bmatrix}$$

with

$$M_{ij} = \int_{\Omega} \varphi_j \varphi_i \,\mathrm{d}x\,\mathrm{d}y, \tag{6}$$

and

$$[A_\tau]_{ij} = \int_{\Omega} (\nu_\tau(x)\nabla\varphi_j)\nabla\varphi_i \,\mathrm{d}x\,\mathrm{d}y, \quad \tau = s, e, i. \tag{7}$$

Finally, $\mathcal{F}(\mathbf{U}) = \int_{\Omega}(\mathrm{A} - \mathrm{B}(\boldsymbol{u}_h))\boldsymbol{u}_h \cdot \boldsymbol{v} d\omega$ when $\boldsymbol{v}$ is selected as stated above.

Let us introduce now a time step Δt and the instants $t^n = n\Delta t$, where we collocate the semi-discrete problem (5). In particular, the incremental ratio

$$\frac{\mathbf{U}^{n+1} - \mathbf{U}^n}{\Delta t}$$

is an approximation of the time derivative either in t^{n+1} or t^n, with an associated numerical error proportional to Δt. If we collocate it at t^{n+1}, we have the implicit Euler method, and we have to solve the nonlinear algebraic system

$$\mathcal{M}\mathbf{U}^{n+1} + \Delta t\mathcal{A}\mathbf{U}^{n+1} - \Delta t\mathcal{F}(\mathbf{U}^{n+1}) = \mathcal{M}\mathbf{U}^n.$$

This can be done with iterative methods, which, however, rapidly increase the computational costs, since iterations are nested in the time loop. If we collocate the derivative at t^n (explicit Euler), the system to be solved is simply ($\mathbf{U}^n$ is known by the previous time step or the initial condition)

$$\mathcal{M}\mathbf{U}^{n+1} = -\Delta t\mathcal{A}\mathbf{U}^n + \Delta t\mathcal{F}(\mathbf{U}^n) + \mathcal{M}\mathbf{U}^n.$$

The drawback of this method is that Δt is subject to restrictions from numerical stability. In this case, this restriction is proportional to h^2, so it can be significantly limiting for fine reticulations.

This is related to the explicit treatment of the (so called) stiffness term $\mathcal{A}\mathbf{U}^n$. To circumvent the drawbacks of the two approaches, we resort to a mixed implicit/explicit (IMEX) approach, where the stiffness is treated implicitly and the nonlinear term is treated explicitly. At each time step, we solve then

$$(\mathcal{M} + \Delta t\mathcal{A})\mathbf{U}^{n+1} = \Delta t\mathcal{F}(\mathbf{U}^n) + \mathcal{M}\mathbf{U}^n. \tag{8}$$

This is a linear system and the stability constraints induced by the explicit treatment of the nonlinear term are less restrictive than the ones induced by the stiffness. In our simulations, we have been in practice free to select the time step in a reasonable range of values for the application at hand. For a more extensive discussion and application of the IMEX method, see [11].

5. Numerical results

To prove the effectiveness of our proposed methodology, we apply it to the diffusion of rabies among raccoons in New York state. The simulations throughout this session are done with a self-developed code[1] running in MATLAB® 7.11.0. Due to the temporal scale of epidemic dynamics, we consider the month as the time unit measure, and in the simulations, we use a time step $\Delta t = 0.1$, roughly corresponding to 3 days. Table 1 summarizes the values of the epidemiological parameters of the model, that have either been drawn from published values and US Department of Agriculture sources (http://www.usda.org), or estimated indirectly. In particular, the birth rate α, the contact rate β, the latency period $1/\sigma$, and the infectious period $1/\phi$ are taken from the literature [2–4,6,16], while the rate of density-dependent mortality, μ_0, is estimated indirectly to produce a disease-free equilibrium within the range of reported densities for raccoon populations for the eastern US (5–17 animals/km^2; see Section 5.3.1 below for a more detailed discussion on this aspect) [32,34].

5.1. *The computational domain: main waterways, lakes, and mountains*

We derived the computational domain from a satellite image of the state of New York with the `ImageJ` [15] software. The coordinates of the computational domain will be expressed throughout this section in Longitude and Latitude. The movement across the finger lakes in west-central New York has been completely eliminated. We include the Hudson and Mohawk rivers as main waterways by forcing mesh nodes along them. Figure 2 (left) illustrates the computational domain with the included main waterways. The computational domain is then discretized by a regular triangulation with `Netgen` [24], consisting of 64,759 nodes and 127,360 triangles, plotted in Figure 2 (right).

The diffusion tensors are taken as isotropic diagonal tensors far from the Hudson and the Mohawk rivers and the mountains, with the same value along x and y given by $\nu_H = 0.01$. The mountain region is identified as the interior of the circle centred at $(-74.5, 44)$ with radius 0.7, where we assume the diffusivity $\nu_M = 5 \times 10^{-3} \times \nu_H$. These values have been tuned with a

Table 1. Coefficients of the SEI model (1).

α	Birth rate	2.67 k/f/y
β	Contact rate	1e−4 (ad)$^{-1}$
$1/\sigma$	Latency period	50 Days
$1/\phi$	Infectious period	14 Days

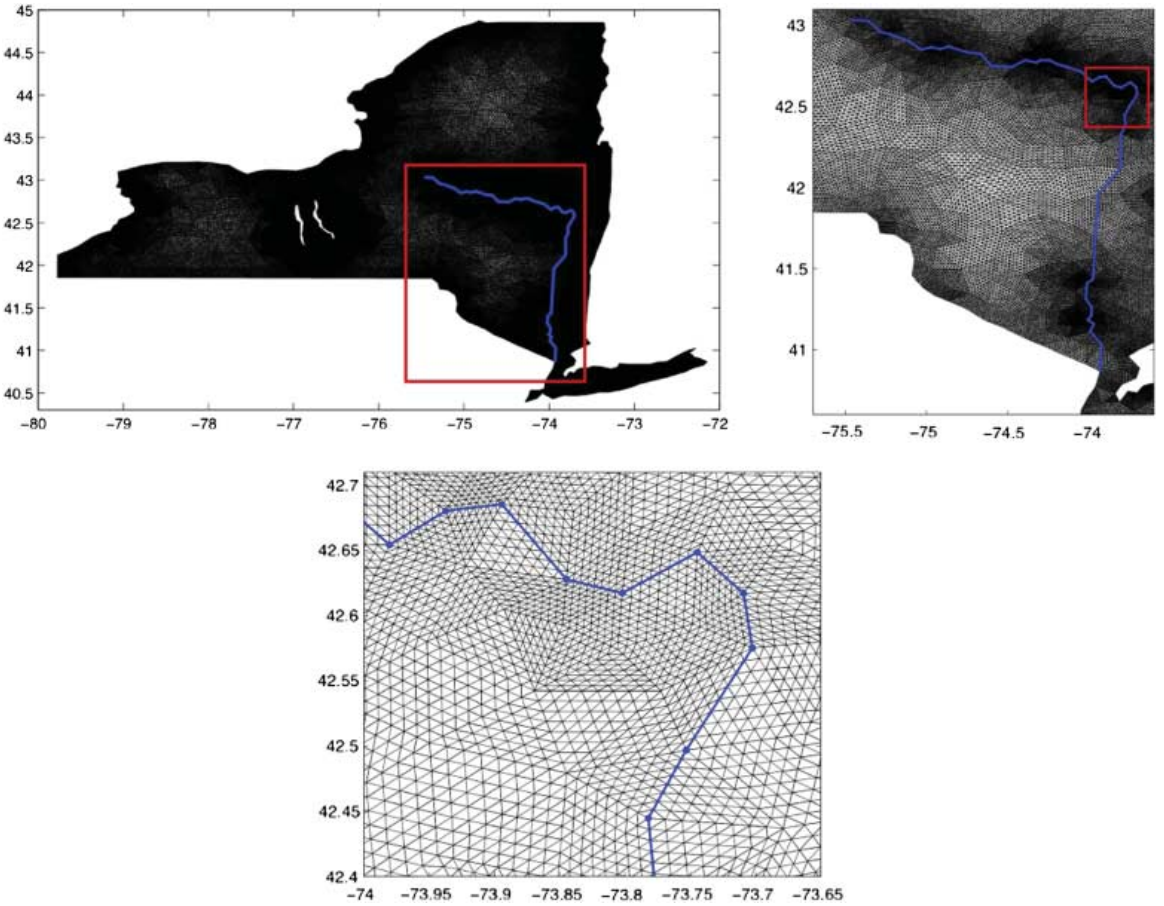

Figure 2. The computational domain including finger lakes and main waterways and its finite elements triangulation.

Figure 3. Left to right: diffusion in the x-direction, cross-diffusion, and diffusion in the y-direction ($\nu_H = 0.01$, $\nu_L = 5 \times 10^{-3} \times \nu_H$).

trial-and-error approach (with respect to real available data). Around the major rivers, we compute the tensors as described in Section 3.1. We plot in Figure 3 the resulting diffusion coefficients.

As expected, the off-diagonal terms vanish everywhere except in a narrow neighbourhood of the river, following the turns of the river itself.

The computational domain we consider obviously does not take into account all of the heterogeneities of the New York landscape. Woods and mountains in the southern region, as well as interstates and major highways could also be included. However, the treatment of such heterogeneities would not differ too much from what we proposed so far. In particular, interstates and highways can be treated in the same manner as rivers, although expecting a far less significant reduction in the normal diffusion coefficient (raccoons are not good swimmers and actually are more likely able to cross interstates).

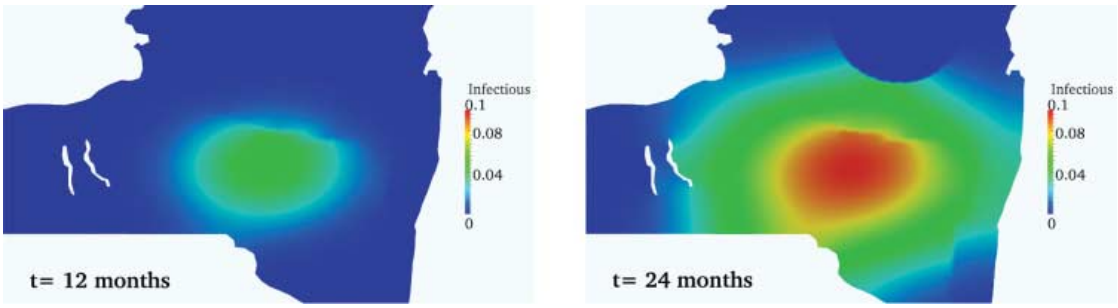

Figure 4. Infectious wave spread. Shielding effect of river, lakes and mountains on the propagation front of I_{base} at different times ($\nu_H = 10^5 \times \nu_L$). Left to right: 12, and 24 months after infection.

5.2. *Simulation of a fictitious epidemics*

To evaluate the effectiveness of our modelling of the landscape, we first simulate a fictitious epidemic starting in the centre of the state. We assume that, at the beginning of the epidemic, the raccoon population is at equilibrium at the carrying capacity. From available data in the literature, we assume that the raccoon density is 11 animals/km^2 [32]. We consider homogeneous Neumann boundary conditions and we introduce an initial density of infectious and exposed individuals, $i_0 = 0.5$ and $e_0 = 1.5$ animals per km^2, respectively, localized in a circle of radius 0.1 with centre $(-75, 42.5)$, which is near Oneonta in Otsego County. We consider an isotropic diffusion coefficient across the whole domain Ω, except along the rivers, with the following baseline values:

$$\nu_H = 10^{-2} \quad \text{km}^{-2}\,\text{year}^{-1}, \quad \nu_L = 10^{-5} \times \nu_H, \quad \epsilon = 0.08. \tag{9}$$

In the northern mountain region, identified as the interior of the circle centred in $(-74.5,44)$ with radius 0.7, we consider the diffusion coefficient $\nu_M = 5 \times 10^{-3} \times \nu_H$, and we reduce the carrying capacity to 8 animals/km^2.

We denote by $\mathbf{U}_{\text{base}} = [S_{\text{base}}, E_{\text{base}}, I_{\text{base}}]^{\text{T}}$ the solution associated with the above baseline values. We plot in Figure 4 the spread of the infectious wave: the shielding effect of the river and the influence of the presence of lakes are clearly visible.

To study the sensitivity of the numerical approximation to the modelling parameters of the major rivers, we consider the following solutions:

(1) $\mathbf{U}_{\text{noRiver}} = [S_{\text{noRiver}}, E_{\text{noRiver}}, I_{\text{noRiver}}]^{\text{T}}$, computed with a constant diffusion coefficient

$$\nu_H = \nu_L = 10^{-2}\,\text{km}^{-2}\,\text{year}^{-1}.$$

(2) $\mathbf{U}_\tau = [S_\tau, E_\tau, I_\tau]^{\text{T}}$, obtained by modifying the diffusion coefficient across the river, but keeping the baseline value of the distance ϵ unchanged:

$$\nu_H = 10^{-2}\,\text{km}^{-2}\,\text{year}^{-1}, \quad \nu_L = 10^{-\tau} \times \nu_H, \quad \epsilon = 0.08.$$

(3) $\mathbf{U}_\epsilon = [S_\epsilon, E_\epsilon, I_\epsilon]^{\text{T}}$, obtained by modifying the value of the distance ϵ, while keeping the baseline value of the diffusion across the river unchanged:

$$\nu_H = 10^{-2}\,\text{km}^{-2}\,\text{year}^{-1}, \quad \nu_L = 10^{-5} \times \nu_H, \quad \epsilon = 0.16.$$

We plot in Figure 5 the difference between the spread of an infection with uniform diffusion coefficients and the one with the landscape heterogeneities included, namely the infectious densities difference $I_{\text{noRiver}} - I_{\text{base}}$ at different times after insurgence (12, 24, and 36 months). The significant difference in the propagation speed follows from the shielding action of the river and

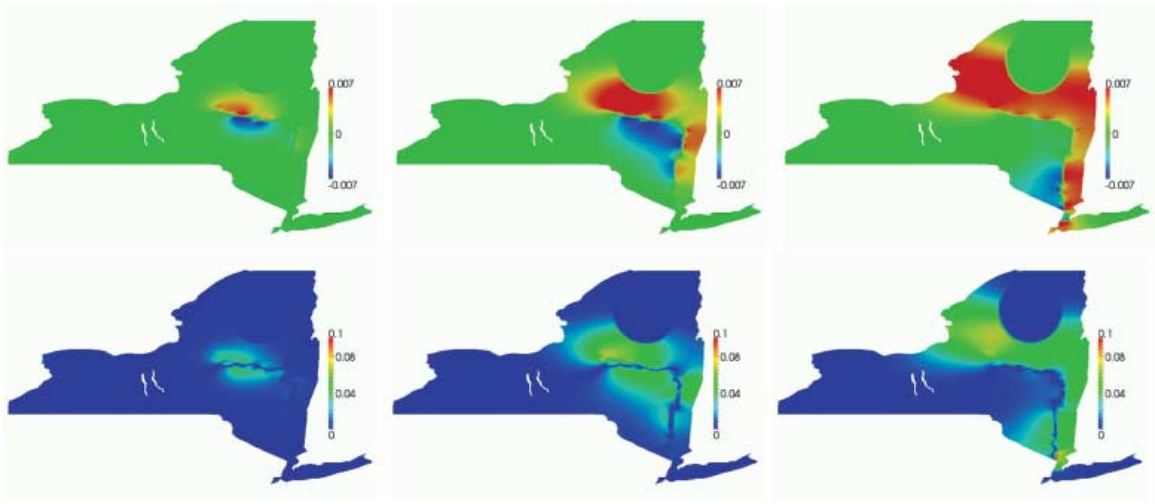

Figure 5. Effect on the propagation front of including landscape heterogeneities at (left to right) 12, 24, and 36 months. Top row: $I_{\text{noRiver}} - I_{\text{base}}$. Bottom row: relative difference $(I_{\text{noRiver}} - I_{\text{base}})/I_{\text{noRiver}}$.

Figure 6. Effect of diffusion reduction normal to the river on the propagation front at (left to right) 12, 24, and 36 months: $I_{\text{base}} - I_7$.

Figure 7. Effect of the distance ε on the propagation front at (left to right) 12, 24, and 36 months: $I_{\text{base}} - I_\epsilon$.

can be appreciated by the bottom row in Figure 5, where we plot the relative difference between I_{base} and I_{noRiver}.

In Figure 6, we highlight the sensitivity of the shielding effect of the river to the reduction factor between the normal and the tangential diffusion along the river. We plot the differences between the infectious density in the baseline solution and the infectious density in $\mathbf{U}_\tau$ for $\tau = 7$. The expected monotonicity of the shielding effect with respect to the ratio ν_H/ν_L is confirmed by the numerical simulation.

Finally, in Figure 7, we show the effect of the threshold parameter in the definition of the diffusion tensor across the river, by plotting the difference between the infectious density of the baseline solution, and the solution computed with a doubled threshold parameter ($I_{\text{base}} - I_\epsilon$). Again, as expected, the larger the threshold, the larger the shielding effect in the propagation of the infection.

These results highlight the relevance of the presence of rivers on an epidemic. The relative difference in infectious population density in the presence and absence of rivers is 10%. The numerical values of the parameters have an impact of at most 3% on the results, in the presence

of variations in the threshold. A far less significant impact is associated with changes in the ratio v_H/v_L.

5.3. *Simulation of the New York state epidemic outbreak (1990–1994)*

In order to test our method with a more realistic problem, we simulate an epidemic spread that is inspired by the raccoon rabies epizootic that spread in the state of New York from the early 1990s. Since World War II, three major regional epidemics of rabies in animals have appeared and are still ongoing. Two of them (one in Europe, one in Canada) are primarily associated with red fox (*Vulpes vulpes*), while the third one is occurring in the USA, and appears to have multiple hosts. The east coast of the USA has experienced an epidemic of rabies among raccoons (*Procyon lotor*) since the early 1970s. The epidemic began at the boundary between Virginia and West Virginia and spread northeast through Pennsylvania and Connecticut and southeast to North Carolina.

On 4 May 1990, the first case of a rabid raccoon was recorded in the state of New York, in the township of Addison, Steuben County, on the New York/Pennsylvania border. By the end of 1994, the epidemic had propagated across a great part of the state.

5.3.1. *Initial and boundary conditions*

We assume the population to be at the carrying capacity at the beginning of the epidemic. One of the major difficulties in studying epidemics in wildlife is the scarcity of available information about the actual size of the at-risk population under investigation, and the raccoon population in New York is no exception to this.

The best estimates in the literature report raccoon density in eastern USA to be in the range of 5–17 animals per km^2 [32,34]. However, according to the New York State Department of Environmental Conservation [25], raccoons are more keen to live in areas where the human presence is larger. To study the impact of the population distribution on the epidemic spread, we consider two possible initial conditions, one assuming uniform density of animals across the state, and the other one considering heterogeneous density. In the former case, we assume everywhere the reported average raccoon density of 11 animals per km^2. In the latter one, we assume the homogeneous case as baseline and we add a correction to this distribution that takes into account the human population density. As a consequence, we obtain an initial population density between 8 and 14 animals per km^2, where the highest values are associated with the most populated areas, such as New York City and its surroundings, Albany, Ithaca, and Syracuse, while the lowest values are associated with the less populated ones, such as Clinton, Franklin, and Hamilton Counties, for instance. We plot in Figure 8 (left) the heterogeneous initial condition across the whole state.

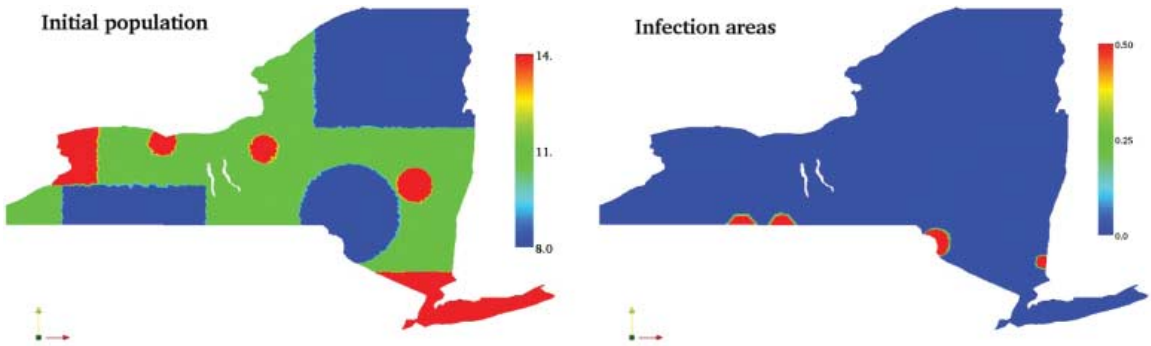

Figure 8. Initial conditions for the total raccoon population, and epidemic outbreak locations.

We also prescribe homogeneous Neumann boundary conditions on the border of the state. Note that if the assumption of an isolated environment makes perfect sense for the portion of boundary lying along water (Erie and Ontario lakes, Saint Lawrence and Hudson river, the Atlantic ocean around Long Island), it may be questionable when applied to the portion of boundary that borders Pennsylvania, Connecticut, and Vermont. However, since we do not have information about the actual movement of raccoons across those borders, we chose to consider New York state as an isolated environment. We describe the epidemic wave propagating north from Pennsylvania and west from Vermont by including source terms in the areas where the first cases have been reported.

5.3.2. *Qualitative comparison with real data*

To best reproduce the real-world data, the source terms for the epidemic were placed at the three locations highlighted in Figure 8 (right). These locations correspond to the first confirmed case of rabies in three areas of the state. Each source term was added at the month corresponding to the first confirmed case in that region. The central-western source term, centred at (−77.24,41.80), is at the location of the first confirmed case in the state, in Steuben county, thus it is introduced at $t = 0$. Also at $t = 0$, the south-central source in Sullivan county is introduced, centred at (−75.01, 41.67). Both of these first two source terms have a circular radius of 0.2. The next source terms are the western source, centred at (−77.80,41.80) with radius 0.2 and introduced at $t = 4$, as well as the south-eastern source, centred at (−73.61, 41.33) with radius 0.1 and introduced at $t = 11$. These times (4 and 11 months) correspond to the moment when the epidemic reached Cattaraugus and Westchester counties, respectively.

We use the model values from Table 1 and the coefficients of the baseline simulation of the previous section (9). To take into account the separation of Long Island from the mainland, we also reduce the diffusion across the East River, by considering a normal diffusion coefficient $\nu_{\mathrm{ER}} = 10^{-8} \times \nu_H$, still combined with a threshold parameter $\epsilon = 0.08$.

The model we simulate does not take into account the oral vaccination policies implemented in reality in Clinton, Essex, and Franklin counties beginning in the fall 1995. Thus, we limit our comparison with real data to the early stage of the epidemic outbreak (1990–1994). We run simulations both assuming a uniform raccoon population density and assuming the population distribution plotted in Figure 8. We plot in Figure 9 the spread of the infectious wave, together with the corresponding available data [25]. The epidemic waves match qualitatively, suggesting that even a moderate mesh like the one used here (the average area of each element is 1.10 km^2) is enough for the problem at hand.

A more quantitative comparison is not readily feasible, as the available data are counts of reported rabies cases, while our model provides the density of rabid raccoons in a given area. Moreover, the high uncertainty surrounding the actual total raccoon population would spoil such comparison of all meaning. On the qualitative side, the propagation of the infection wave is quite well captured, in the uniform initial population case, in the north-northeast direction, but it seems too fast in the westbound direction. In the case of heterogeneous initial population, on the other hand, the westbound spread appears too fast towards the Buffalo area and the Ontario Lake, and at the same time too slow in the area close to the state southern border. This latter behaviour is likely a consequence of the homogeneous Neumann boundary conditions along the southwest border. The epidemic wavefront is actually part of a larger epizootic moving north from Pennsylvania, and the appearance of infectious cases along the border should be taken into further account.

Both simulations feature an infection spread in the Long Island area, where no rabid cases were reported in the period under consideration. Soon after the wavefront reaches the area, the high raccoon population density triggers an outbreak even in the presence of a small density of infectious. This is particularly evident in the heterogeneous population case: the higher the

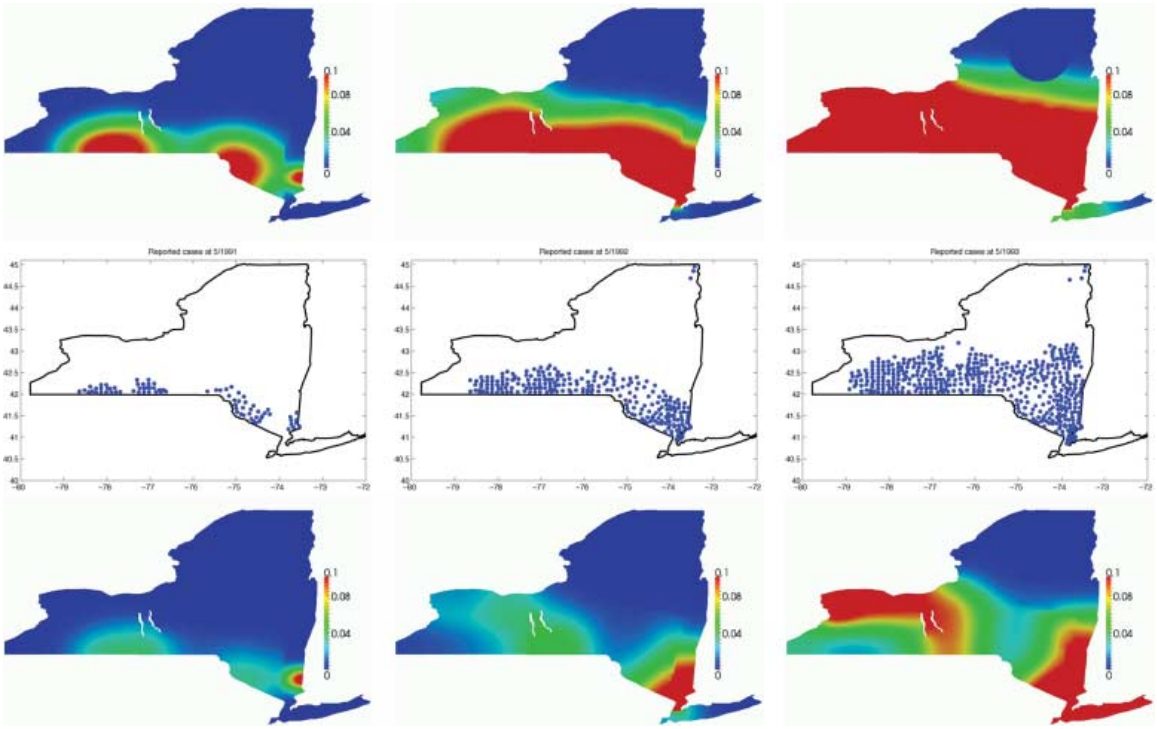

Figure 9. Propagation of a realistic rabies epidemics in New York state at different times (left to right: 1, 2, and 3 years after infection). Top: uniform initial population. Middle: reported data [25]. Bottom: heterogeneous initial population.

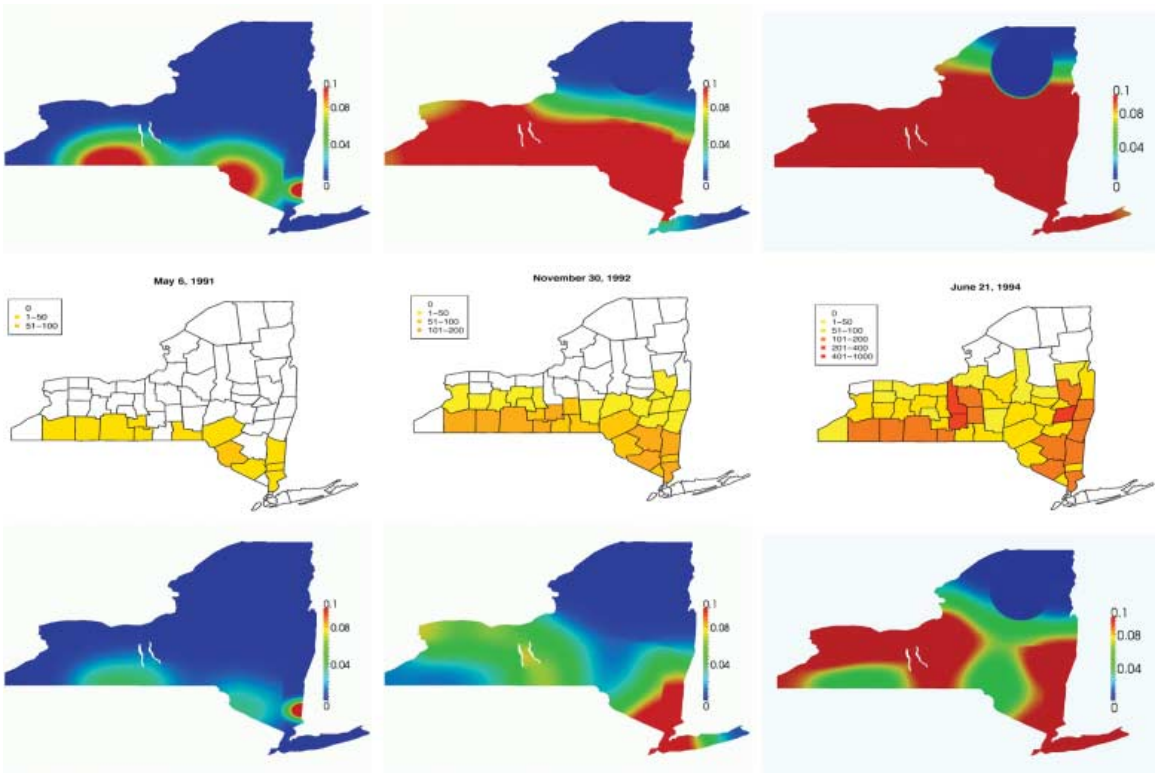

Figure 10. Propagation of a realistic rabies epidemics in New York state at different times (left to right: May 1991, November 1992, and June 1994). Top: uniform initial population. Middle: total reported data aggregated per county [25]. Bottom: heterogeneous initial population.

raccoon density, the larger the error. Since the wavefront reaches Long Island in a time consistent with the data, the problem can be addressed in a twofold manner. On the one hand, one could further reduce the normal diffusion across the East River. On the other hand, the high human population density in the area results in high levels of surveillance [13], and so infectious animals are more easily identified and removed. Thus, another possibility could be to reduce the infection rate in the area as a consequence of the high level of surveillance.

The comparison in Figure 9 does not really take into account the dynamics of the epidemic itself, as it focuses more on the propagation speed of the infection wave. To have a more accurate comparison, we considered not only the location of the reported cases, but also their quantity. We thus compared the results obtained with both the uniform and the heterogeneous population with the cumulative reported cases aggregated by county and in time, that we report in Figure 10. From this last comparison, it appears that the heterogeneous initial population, despite suffering from the problems highlighted above, performs better in identifying spatial clusters of infectious individuals that clearly emerge from reported data.

5.3.3. *Computational cost*

The discrete problem is reasonably well conditioned and its numerical solution does not present any significant difficulty. The mass and stiffness matrices are pre-assembled and the resulting linear system is associated with a $(194,277 \times 194,277)$ matrix. The simulation of 10 years of epizootic took about 3h45' on a desktop computer with 3.2 GHz Intel Core i3 dual processor and 4 GB 1333 MHz DDR3.

6. Conclusions

We present here a numerical model to simulate the spread of a lethal, infectious disease across a heterogeneous and continuous landscape. We introduced heterogeneities by including lakes, mountains, and main waterways in a finite element settings that allow the use of unstructured grids, which are better suited to represent a complex geographical domain. We provided a numerical simulation showing the effectiveness of such modelling. In addition, the method proves to be viable in reproducing the outbreak of a real epidemic of rabies among raccoons in the state of New York from 1990 to 1994. This aspect has a major potential, as it renders such modelling a solid tool, both in terms of accuracy in prediction and landscape description, and, most of all, CPU time. Moreover, such modelling requires the tuning of a small number of parameters in comparison to SEIR models based on geopolitical units [8,26], where also a full mixing matrix has to be estimated.

However, further work still needs to be done. In this work, for instance, the diffusion coefficients have been tuned empirically. Data assimilation procedures [5], that is, the parameter tuning based on a rigorous minimization of the misfit between results and available measures, is anticipated to improve the knowledge of such a crucial parameter. In a forthcoming article, we will introduce a parameter estimation procedure of Bayesian type, based on the available observed data, which is currently under investigation.

Acknowledgements

The authors thank L.A. Real (Emory University) for some fruitful discussions on the topic. The first author has been partially supported by NIEHS grant T32 ES015459. The second author has been supported by the NIH grant RO1-AI047498 to L.A. Real during his stay at the Center for Disease Ecology, Emory University, 2010–2011.

Note

1. Available under GNU General Public Licence at `http://www.mathcs.emory.edu/~ale`

References

[1] B.P. Ayati and T. Dupont, *Galerkin methods in age and space for a population model with nonlinear diffusion*, SIAM J. Numer. Anal. 40(3) (2002), pp. 1064–1076.
[2] R. Biek, J.C. Henderson, L. Waller, C.E. Rupprecht, and L.A. Real, *A high-resolution genetic signature of demographic and spatial expansion in epizootic rabies virus*, Proc. Natl. Acad. Sci. 104(19) (2007), pp. 7993–7998.
[3] W.J. Bigler, R.G. McLean, and H.A. Trevino, *Epizootiologic aspects of raccoon rabies in Florida*, Amer. J. Epidemiol. 98 (1973), pp. 326–335.
[4] T.H. Bissonnette and A.G. Csech, *Sexual photoperiodicity of raccoons on low protein diet and second litters in the same breeding season*, J. Mammalogy 19 (1938), pp. 342–348.
[5] J. Blum, F.X. Le Dimet, and I.M. Navon, *Data Assimilation for Geophysical Fluids*, XIV of Handbook of Numerical Analysis, Chapter 9, Elsevier, Amsterdam, 2008.
[6] A.B. Carey and R.G. Mclean, *The ecology of rabies – evidence of co-adaptation*, J. Appl. Ecol. 20 (1983), pp. 777–800.
[7] C. Cusulin and L. Gerardo-Giorda, *A numerical method for spatial diffusion in age-structured populations*, Numer. Methods Partial Differential Equations 26(2) (2010), pp. 253–273.
[8] S.M. Duke-Sylvester, L. Bolzoni, and L.A Real, *Strong seasonality produces spatial asynchrony in the outbreak of infectious diseases*, J. R. Soc. Interface 8(59) (2011), pp. 817–825.
[9] A. Ern and J.L. Guermond, *Theory and Practice of Finite Elements*, Springer, New York, 2004.
[10] R.A. Fisher, *The wave of advance of advantageous gene*, Ann. Eugenics 7 (1937), pp. 355–369.
[11] L. Formaggia, F. Saleri, and A. Veneziani, *Solving Numerical PDE's*, Springer-Verlag, Milan, 2012.
[12] L. Gerardo-Giorda, *Balancing waveform relaxation for age-structured populations in a multilayer environment*, J. Numer. Math. 16(4) (2008), pp. 281–306.
[13] L. Gerardo-Giorda, G. Puggioni, R.J. Rudd, L.A. Waller, and L.A. Real, *Structuring targeted-surveillance for monitoring disease emergence by mapping observational data onto ecological process*, J. R. Soc. Interface (in revision).
[14] E.E. Holmes, M.A. Lewis, J.E. Banks, and R.R. Veit, *Partial differential equations in ecology: Spatial interactions and population dynamics*, Ecology 75 (1994), pp. 17–29.
[15] IMAGEJ software. Software available at http://rsbweb.nih.gov/ij/
[16] A.C. Jackson and W. Wunner (eds.), *Rabies*, Academic Press, New York, 2007.
[17] F. John, *Partial Differential Equations*, Springer-Verlag, New York, 1982.
[18] M.J. Keeling and P. Rohani, *Modeling Infectious Disease in Humans and Animals*, Princeton University Press, Princeton, NJ, 2008.
[19] M.-Y. Kim, *Galerkin methods for a model of population dynamics with nonlinear diffusion*, Numer. Methods Partial Differential Equations 12 (1996), pp. 59–73.
[20] M.-Y. Kim and E.-J. Park, *Characteristic finite element methods for diffusion epidemic models with age-structured populations*, Comput. Math. Appl. 97 (1998), pp. 55–70.
[21] G. Leoni, *A First Course in Sobolev Spaces*, Graduate Studies in Mathematics Vol. 105, American Mathematical Society, Providence, RI, 2009.
[22] R. Miller Neilan and S. Lenhart, *Optimal vaccine distribution in a spatiotemporal epidemic model with an application to rabies and raccoons*, J. Math. Anal. Appl. 378(2) (2011), pp. 603–619.
[23] F.A. Milner, *A numerical method for a model of population dynamics with spatial diffusion*, Comp. Math Appl. 19(31) (1990), pp. 31–43.
[24] NETGEN software. Software available at http://www.hpfem.jku.at/netgen/.
[25] New York State Department of Environmental Conservation. Available at http://www.dec.ny.gov/animals/9358.html.
[26] L.A. Real and J.E. Childs, *Spatial-temporal dynamics of rabies in ecological communities*, in *Disease Ecology: Community Structure and Pathogen Dynamics*, S.K. Collinge and A.C. Ray, eds., Oxford University Press, New York, 2006, pp. 168–185.
[27] S. Ruan, *Spatial-temporal dynamics in nonlocal epidemiological models*, in *Mathematics for Life Science and Medicine*, Springer, Berlin-Heidelberg, 2007, pp. 97–122.
[28] S. Salsa, *Partial Differential Equations in action. From modeling to theory*, Springer-Verlag, Milan, 2008.
[29] D.L. Smith, B. Lucey, L.A. Waller, J.E. Childs, and L.A. Real, *Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes*, Proc. Natl. Acad. Sci. 99 (2002), pp. 3668–3672.
[30] D.L. Smith, L.A. Waller, C.A. Russell, J.E. Childs, and L.A. Real, *Assessing the role of long-distance translocation and spatial heterogeneity in the raccoon rabies epidemic in Connecticut*, Prev. Vet. Med. 71 (2005), pp. 225–240.
[31] R.T. Sterner and G.C. Smith, *Modelling wildlife rabies: Transmission, economics, and conservation*, Biol. Conservation 131(2) 2006, pp. 163–179.
[32] F.W. Stuewer, *Raccoons: Their habits and management in Michigan*, Ecol. Monogr. 13 (1943), pp. 203–257.
[33] A. Toselli and O. Widlund, *Domain Decomposition Methods*, Springer, Berlin-Heidelberg, 2004.
[34] D. Urban, *Raccoon populations, movement patterns, and predation on a managed waterfowl marsh*, J. Wildlife Manage. 34 (1970), pp. 372–382.