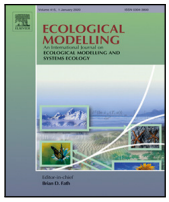
Ecological Modelling 486 (2023) 110516

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel



Parameter-varying partial differential equation to model the global change impacts on wildlife populations

Mohamad Chhaytle [a], Régis Ouvrard [a,*], Thierry Poinot [a], Lauriane Mouysset [b]

[a] Université de Poitiers, ISAE-ENSMA Poitiers, LIAS, 2 rue Pierre Brousse, TSA 41105, 86073 Poitiers cedex 9, France
[b] CNRS, Centre International de Recherche sur l'Environnement et le Développement, 45bis avenue de la Belle Gabrielle, 94736 Nogent-sur-Marne, France

ARTICLE INFO

Keywords:
Global change impacts
Parameter-varying model
Parametric estimation
Partial differential equation
Population dynamics

ABSTRACT

Accurate models of wildlife population dynamics are needed to assess the causes of global biodiversity decline. In this paper, new parameter-varying PDE models are proposed with an original methodology for parametric estimation from ecological data. The model structure makes it possible to introduce the environmental heterogeneity which can characterize global changes. Particular attention was paid to the implementation of the identification procedure: true parametric sensitivity functions in the optimization algorithm, a Galerkin method with a proper orthogonal decomposition for the PDE solution and a pre-estimate to initialize the iterative procedure. The tools are validated on simulation data and then applied to a real application modelling the impact of climate change and agricultural intensification on the population of a passerine bird.

1. Introduction

Biodiversity is currently showing alarming rates of decline (IPBES, 2019). IUCN (2022) gives percentage estimates of threatened species for the amphibian class of 41%, for reef-forming corals of 36%, for conifers of 34%, for mammals of 27%, for reptiles of 21%... The bird class is, for instance, a good indicator of global biodiversity loss (Inger et al., 2014; Stanton et al., 2018) and many surveys make it possible to measure this decline. IUCN (2022) estimates that 13% of bird species are threatened. Rosenberg et al. (2019) report population losses across much of the North American avifauna over 48 years. Similarly, Keller et al. (2020) suggest that major changes in the distribution of European breeding birds over the last 30 years are due to climate change and/or agricultural intensification. To deal with such a dramatic evolution of wildlife, predicting the trends of population dynamics requires accurate models if we are to provide management practices, at local and global scales, and implement public policies in order to halt biodiversity loss.

To address this challenge, ecologists have developed many models. Two types of models co-exist in the literature on ecology and biology conservation. On one hand, the phenomenological models, and on the other hand the mechanistic ones (Merow et al., 2011; Hefley et al., 2017), *i.e.* a black-box models class and a grey-box models class, respectively. The phenomenological models, such as the ecological niche models (Melo-Merino et al., 2020), are statistical models that link the responses of species variables to explicative variables representing, for instance, landscape or climate changes. Easy to handle, they give the spatial and temporal distribution of the considered species. Unfortunately, their long-term predictive ability is limited to the domain of calibration data and extreme behaviours are not modelled (DeAngelis and Yurek, 2016; Hefley et al., 2017). The mechanistic models are derived from ecological theories. They are based on dynamical equations such as ordinary differential equations (ODEs) which characterize the temporal dynamics with better long-term predictive ability. The mechanistic models are based on the living, but they are restricted to phenomena known and modelled by theory. The most popular are: Malthus model, Verhulst model or Lotka–Volterra model. These ODE models are typically population models and have been applied, for example, to French bird populations in Mouysset (2014) and Mouysset et al. (2016). To complete this dynamic issue, spatial considerations can be integrated into the modelling framework by considering metapopulation models (Levin, 1976; Mouysset et al., 2019; Guillet and Mouysset, 2022).

Mechanistic models based on partial differential equations (PDEs) capable of simultaneously characterizing temporal and spatial dynamics have been proposed for a long time (Fisher, 1937; Skellam, 1951). Moreover, Levin (1976) proves that the metapopulation modelling leads to Fisher's PDE by considering a continuum of patches. Since then, the models based on PDEs are widely mentioned in the literature with fixed parameters (Holmes, 1993; Lewis and Kareiva, 1993; Holmes et al., 1994; Okubo and Levin, 2001), but also with variable parameters (Shigesada et al., 1979; Okubo, 1986; Belgacem and Cosner, 1995; Cantrell and Cosner, 2001, 2003; Roques et al., 2016). However,

* Corresponding author.
E-mail address: regis.ouvrard@univ-poitiers.fr (R. Ouvrard).

https://doi.org/10.1016/j.ecolmodel.2023.110516
Received 28 April 2023; Received in revised form 5 July 2023; Accepted 26 September 2023
Available online 11 October 2023

today there are very few tools for parametric estimation of PDE models from ecological data and these models are little used in practice. Let us cite (Bateman et al., 2015).

Nevertheless, the use of the PDE models seems promising. DeAngelis and Yurek (2016) classify the PDE models among spatially explicit models and compares them to the spatially implicit models (phenomenological models). They conclude their review as follows '*we are confident that the power of spatially explicit models will be transformative in ecology over the next two decades and will lead to the marriage of basic and applied ecology*'. To confirm this conclusion, Hefley et al. (2017), for instance, show the performances of mechanistic spatio-temporal models compared to phenomenological regression-based models for forecasting spatio-temporal processes.

Some tools for the parametric estimation of PDE models from ecological data are developed. To take into account the probabilistic and mechanistic aspects, some tools are based on hierarchical Bayesian models (Wikle, 2003). For tricky numerical implementations, PDE solutions are usually approximated with finite-difference methods (Hefley et al., 2017). An example of the use of such models and implementations is given in Louvrier et al. (2020) which estimates a PDE model to explain and forecast the colonization of wolves in the South-Eastern of France.

In this paper, our main contribution is the parametric estimation of a parameter-varying PDE model. The proposed methodology is an implementation for handling the partial derivative equations and estimating the parameters of the PDE model in the challenging context of ecological applications. Particular care is taken to leverage data to collect all the information about ecological processes. The structure of the PDE model can be adapted to fit the deterministic part of the data and to represent the different effects which cause the population to evolve. The variable parameters make it possible to introduce environmental heterogeneity (in terms of agricultural intensification or global warming, for instance) to model the impacts of global changes on the dynamics of wildlife populations. We describe more broadly a methodology introduced by Ouvrard et al. (2019) to apply to the population of a passerine bird on the scale of France.

First, as mentioned above, special attention is paid to the implementation of parametric estimation algorithms. In fact, with ecological applications, the observation data are poor and sparse. The number of time samples is of the order of ten or twenty time points, sometimes less. In our algorithm, the optimization procedure, that yields the estimated parameters, is based on the true parametric sensitivity functions calculated from the model equation rather than on numerical approximations of them (Knudsen, 1994). Even though the estimation algorithm is based on the minimization of the quadratic criterion of residuals, the implementation can be extended to a hierarchical Bayesian framework and to a maximum likelihood estimation.

Second, the PDE solution is an approximate solution obtained from the Galerkin method (Polis et al., 1973) based on an orthogonal decomposition. The orthonormal basis functions used are built empirically from data by using a proper orthogonal decomposition (Newman, 1996a,b) which leads to the best approximation in the least-squares sense.

Third, like any optimization problem, an initial parameter vector must be chosen. Generally, in ecological applications, we lack prior knowledge about the PDE model parameters. Therefore, we propose a first estimate based on partial moments introduced in Trigeassou (1987) and Ouvrard and Trigeassou (2011). As with any optimization procedure, global convergence cannot be guaranteed. But we propose some practical implementations to try to obtain the global optimum.

Other contributions of this paper are the validation of the tools with realistic simulation data and the design of an accurate simulator of PDE models with varying parameters for prediction. To produce these realistic data, the simulation protocol reproduces data similar to a real application on the French population of Yellowhammers *Emberiza citrinella* and real data provided by the national French Breeding Bird Survey (Jiguet et al., 2012), by the Worldclim database (Fick and Hijmans, 2017) and by the CORINE Land Cover database (Büttner and Kosztra, 2017). The simulator allows testing of the convergence of the parametric estimation towards the true parameters. Finally, we compared the fitting performance of the proposed parameter-varying PDE model and the popular generalized linear model (GLM) model (McCullagh and Nelder, 1983) on these real data.

The paper is organized as follows. Section 2 describes the material and methodology with the parameter-varying PDE model with different effects. The proposed parametric estimation method is detailed in Section 3 for a general formulation of the PDE. The data considered, the validation of the methodology and the application to real data are given in Section 4. A discussion of the results is presented in Section 5. Section 6 concludes this paper.

2. Material and methods

2.1. A methodology to model the global change impacts on wildlife

In this paper, a methodology is proposed to model the impacts of global changes, such as land use or climate evolution, on wildlife. This procedure follows these main steps:

- Datasets and data formatting
  The approach is based on spatio-temporal data on the studied population and its environment. It requires a formatting step to produce a variable which characterizes the environmental heterogeneity. A pre-processing stage facilitates the implementation of the parametric identification step.
- PDE structure selection
  The PDE structure allows the consideration of different effects: diffusion, advection and reaction. The choice of the structure can be made on the basis of expertise or traditional criteria.
- Parametric estimation
  This step consists of an iterative procedure to estimate the PDE parameters from datasets. A POD-Galerkin method gives an approximated solution of the PDE and an optimization procedure yields an estimate of the parameters.
- Simulation for evaluation and prediction
  A numerical tool is developed to simulate PDEs with varying parameters. In the present paper, this simulator is used to evaluate the previous steps. However, its ultimate aim is to predict future population evolution.

2.2. Datasets and data formatting

The performance of the proposed methodology mainly depends on the richness of the datasets. Two datasets are necessary: data on the population studied (density or counting) and data on the phenomena that impact the population. The data, collected in space and time, must be sufficiently well sampled to describe the spatial and temporal dynamics.

Denote the population measurement dataset by $\left\{u^*(x_i, y_j, t_k)\right\}$ collected on points of coordinates (x_i, y_j) at the instants t_k. In practice, the distribution of population measurement data in time and space is not uniform.

To facilitate the implementation of the identification and simulation steps, a pre-processing tool is applied to give a population dataset uniformly sampled and denoted $\{u(i\,dx, j\,dy, k\,dt)\}_{i=0,\ldots,N_x;\ j=0,\ldots,N_y;\ k=0,\ldots,N_t}$. dx, dy and dt are, respectively, longitudinal, latitudinal and temporal steps.

The parameters of the proposed PDE models are variable to characterize the environmental heterogeneity. For instance, during a progressive degradation of a natural habitat, a population moves in the neighbourhood by a phenomenon of diffusion, even leaves an area by a phenomenon of advection, or else gradually disappears with a decreasing reaction.

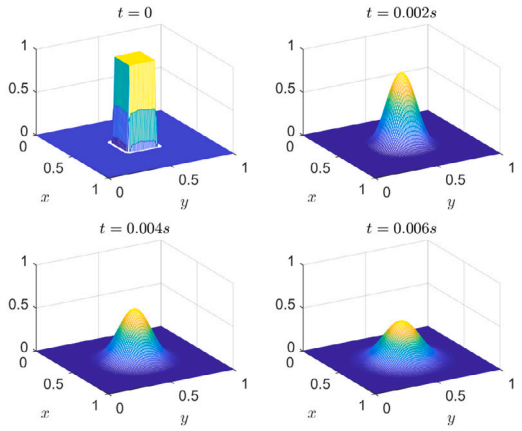


Fig. 1. Diffusion effect after a few time samples.

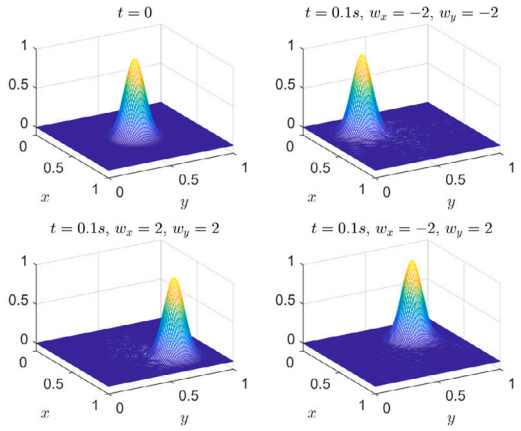


Fig. 2. Advection effect for different values of drift velocities w_x and w_y.

Therefore, the data on the phenomena that impact the population must be synthesized within an index. This latter varies in space and time to act on the different effects of the PDE and to represent a movement, a leak, a concentration, a disappearance... This index, denoted $H(x, y, t)$, is called scheduling variable in the LPV models literature (Toth, 2010).

The index $H(x, y, t)$ can be determined from an expert's assessment or by using standard tools in an ecological study like principal component analysis, generalized linear model or other niche models. Data are often available in a geographic information system (GIS) format, *e.g.* WorldClim climate data or CORINE Land Cover data. Hence, it is easy to format the index $\{H(i.dx, j.dy, k.dt)\}$ on the same population grid.

2.3. Parameter-varying PDE models

The parameter-varying PDE model is defined by

$$\begin{aligned}\frac{\partial u(x,y,t)}{\partial t} &= \frac{\partial}{\partial x}\left[D_x(H)\frac{\partial u(x,y,t)}{\partial x}\right] + \frac{\partial}{\partial y}\left[D_y(H)\frac{\partial u(x,y,t)}{\partial y}\right] \\ &\quad - w_x(H)\frac{\partial u(x,y,t)}{\partial x} - w_y(H)\frac{\partial u(x,y,t)}{\partial y} \\ &\quad + \beta_1(H)u(x,y,t) - \beta_2(H)u^2(x,y,t)\ ,\end{aligned} \tag{1}$$

where $u(x, y, t)$ is a population or a density of individuals at spatial coordinates (x, y) and time t. The varying parameters $D_x(H)$, $D_y(H)$, $w_x(H)$, $w_y(H)$, $\beta_1(H)$ and $\beta_2(H)$ are parametric functions which can be chosen to adapt with the phenomena to be modelled. The parameters contained in these parametric functions will be estimated with the identification procedure described in the following section from datasets presented in the previous section.

Such PDE models make it possible to represent both temporal and spatial dynamics. Depending on the PDE structure considered, three different effects can be modelled: diffusion, advection and reaction (Holmes et al., 1994; Okubo and Levin, 2001). The varying parameters with the scheduling variable $H(x, y, t)$ take into account the environmental heterogeneity.

Remark 1. The structure of the parameter-varying PDE model (1) can be modified to adapt to changes in the population. For the particular choice without reaction, *i.e.* for $\beta_1(H) = 0$ and $\beta_2(H) = 0$, Eq. (1) becomes non-conservative. In this case, the advection part

$$-w_x(H)\frac{\partial u(x,y,t)}{\partial x} - w_y(H)\frac{\partial u(x,y,t)}{\partial y} \tag{2}$$

must be substituted by

$$-\frac{\partial(w_x(H)u(x,y,t))}{\partial x} - \frac{\partial(w_y(H)u(x,y,t))}{\partial y} \tag{3}$$

to maintain a conservative form.

2.3.1. The diffusion effect

The diffusion effect in PDE models characterizes the Brownian random motion of an individual. For instance, consider an initial population $u(x, y, 0)$ which is zero everywhere except in a square in the centre of the study area where the population is one, as shown in the plot at the top left of Fig. 1. The simulation of the following PDE

$$\frac{\partial u(x,y,t)}{\partial t} = D_0\left(\frac{\partial^2 u(x,y,t)}{\partial x^2} + \frac{\partial^2 u(x,y,t)}{\partial y^2}\right) \tag{4}$$

with a diffusion coefficient $D_0 = 1$, presents an evolution of the population for the first time samples as drawn in Fig. 1.

2.3.2. The advection effect

The advection effect reflects the movement of the population linked, for instance, to a deterioration of habitat or a change in climatic conditions. To illustrate it, consider the simulation of the following PDE

$$\frac{\partial u(x,y,t)}{\partial t} = w_x\frac{\partial u(x,y,t)}{\partial x} + w_y\frac{\partial u(x,y,t)}{\partial y} \tag{5}$$

with an initial population $u(x, y, 0)$ given by a Gaussian centred on the middle of the study area. Fig. 2 shows the simulation for three different cases of drift velocities w_x and w_y.

2.3.3. The reaction effect

In the PDE model (1), the reaction is a logistic growth reaction. However, other reactions can be used.

To illustrate the possible reaction parts, consider the PDE model at a point of spatial coordinates (x_0, y_0) and visualize the reactions by simulating ordinary differential equations. The three most popular reactions in population dynamics are:

- the exponential growth reaction (Malthus model)

$$\frac{du(x_0,y_0,t)}{dt} = ru(x_0,y_0,t), \tag{6}$$

- the logistic growth reaction (Verhulst model)

$$\frac{du(x_0,y_0,t)}{dt} = ru(x_0,y_0,t)\left(1 - \frac{u(x_0,y_0,t)}{K}\right), \tag{7}$$

- the Allee effect

$$\begin{aligned}\frac{du(x_0,y_0,t)}{dt} &= ru(x_0,y_0,t) \\ &\quad \left(1 - \frac{u(x_0,y_0,t)}{K}\right)\left(\frac{u(x_0,y_0,t) - A}{K}\right),\end{aligned} \tag{8}$$

where r is the rate of increase of the population u, K, the carrying capacity and A, the critical density.

Figs. 3, 4 et 5 present the simulations of these three kinds of reaction.

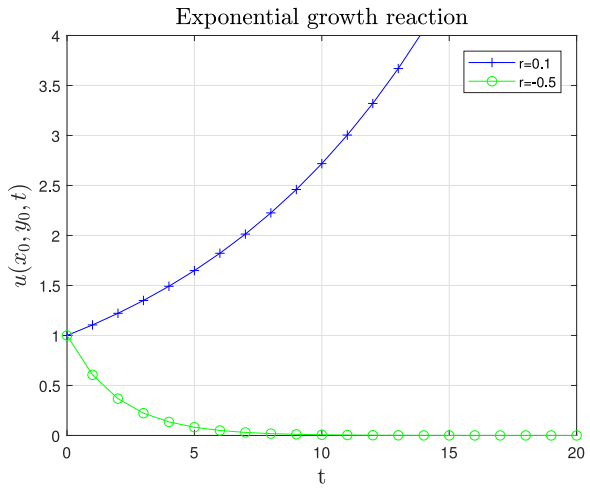


Fig. 3. Simulations of the Malthus model (6) with $r = 0.1$ and $r = -0.5$.

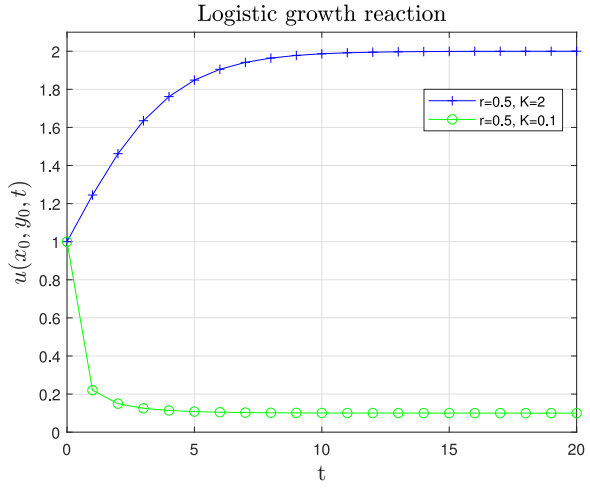


Fig. 4. Simulations of the Verhulst model (7) with $r = 0.5$, $K = 0.1$ and $K = 2$.

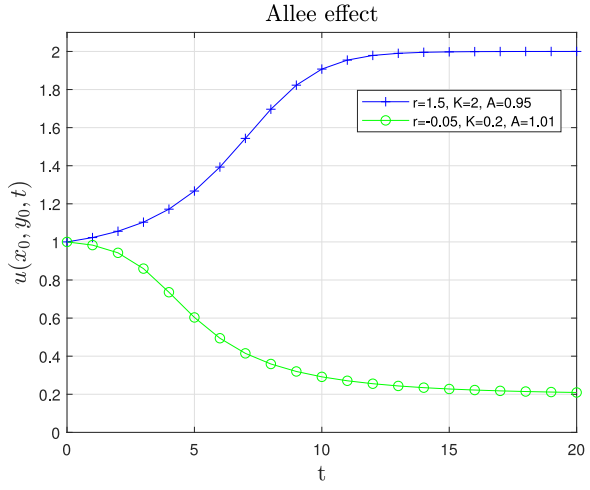


Fig. 5. Simulations of the Allee effect (8) with $r = 1.5$, $K = 2$, $A = 0.95$ and $r = -0.05$, $K = 0.2$, $A = 1.01$.

2.3.4. The contribution of the varying parameters

The proposed PDE models present varying parameters, *i.e.* D, w_x, w_y, r, K and A vary with respect to the index or the scheduling variable $H(x, y, t)$. This index allows the taking into account of environmental heterogeneity that largely improves the modelling. The value of $H(x, y, t)$ is equal to 1 for the favourable area for the considered species or 0 for the unfavourable area.

2.4. Parameter-varying PDE simulator

An accurate simulator of the parameter-varying PDE (1) is needed to predict the future evolution of the population. Moreover, the validation of the methodology proposed in this paper requires precise simulation data to show the convergence towards the true parameters. The simulator proposed in this section will also be used to validate our tools with realistic simulation data in Section 4.2.1.

The simulator is based on a finite difference method. The nonlinear part of the PDE (1) is considered via the splitting method (Havasi et al., 2001). More precisely, the principle is to split the simulation of the original nonlinear PDE into two iterative simulation steps: the simulation of a linear PDE with the effects of advection and diffusion and the simulation of a nonlinear ODE with the logistic growth reaction. Thus, the time domain is sampled with a very fine interval and, for the nth time interval denoted $[t_n, t_{n+1}]$, the simulations are carried out as follows:

1. From an initial population $u_1(x, y, t_n) = u(x, y, t_n)$, simulate the following PDE:
$$\begin{aligned}\frac{\partial u_1(x,y,t)}{\partial t} &= \frac{\partial}{\partial x}\left(D_x(H)\frac{\partial u_1(x,y,t)}{\partial x}\right)\\ &+\frac{\partial}{\partial y}\left(D_y(H)\frac{\partial u_1(x,y,t)}{\partial y}\right)\\ &-w_x(H)\frac{\partial u_1(x,y,t)}{\partial x} - w_y(H)\frac{\partial u_1(x,y,t)}{\partial y}\end{aligned} \tag{9}$$
to obtain $u_1(x, y, t_{n+1})$.
2. From an initial population $u_2(x, y, t_n) = u_1(x, y, t_{n+1})$, simulate the following ODE:
$$\frac{\partial u_2(x,y,t)}{\partial t} = \beta_1(H(x,y,t))u_2(x,y,t) - \beta_2(H(x,y,t))u_2^2(x,y,t) \tag{10}$$
to get $u(x, y, t_{n+1}) = u_2(x, y, t_{n+1})$, the solution of the PDE (1) at the instant t_{n+1}.

For the first step, Eq. (9) is simulated by applying the Crank–Nicolson scheme (Crank and Nicolson, 1996) which is an average between explicit Euler and implicit Euler schemes. In the second step, Eq. (10) is simulated by using the explicit Euler method to obtain the simulated population $u(x, y, t)$. Full details of both implementations are given in Appendices A and B.

3. Parametric estimation

3.1. Parameter-varying PDE identification in a particular context

The main issue is to estimate parameters from datasets in the difficult context of wildlife population dynamics: a small number of time samples, a spatial domain given by a map, an excitation given by the initial conditions $u(x, y, 0)$, a lack of prior knowledge about parameters and measured data not uniformly distributed in space and time which requires a kriging pre-processing.

To deal with this problem, an iterative parametric estimation procedure is proposed with a Levenberg–Marquardt algorithm which concentrates the properties of the gradient descent algorithm and the Newton algorithm. To limit the computation time, an approximate solution of the PDE based on the Galerkin method is used. To retain all information contained in the data, a proper orthogonal decomposition yields the orthonormal basis functions used in the Galerkin method. Finally, a tool based on a least-squares estimate initializes the iterative procedure and advice is given to stop the procedure.

3.2. Iterative parametric estimation procedure

Let us consider a measured population with a uniform sampling and, for simplicity, let us denote by $u(i,j,k)$ a measure of the given sampled dataset $\{u(i\ dx, j\ dy, k\ dt)\}_{i=0,\dots,N_x;\ j=0,\dots,N_y;\ k=0,\dots,N_t}$.

In the PDE model (1), the parameters can vary with respect to the scheduling variable $H(x,y,t)$ with different formulations to adapt to the phenomena impacting the population according to the studied application. In the present section, let us consider varying parameters given by

$$\begin{aligned} D_\bullet(H) &= \sum_{n=0}^{n_{D_\bullet}} D_{\bullet_n} H^n(x,y,t), \\ w_\bullet(H) &= \sum_{n=0}^{n_{w_\bullet}} w_{\bullet_n} H^n(x,y,t), \\ \beta_\bullet(H) &= \sum_{n=0}^{n_{\beta_\bullet}} \beta_{\bullet_n} H^n(x,y,t), \end{aligned} \tag{11}$$

with $n_{D_\bullet}$, $n_{w_\bullet}$ and $n_{\beta_\bullet}$, the polynomial orders, and $D_{\bullet_n}$, $w_{\bullet_n}$ and $\beta_{\bullet_n}$, the parameters to be estimated.

The parameter vector $\theta \in \mathbb{R}^{N_\theta}$ given by

$$\begin{aligned} \theta = \Big[& D_{x_0} \dots D_{x_{n_{D_x}}} \ D_{y_0} \dots D_{y_{n_{D_y}}} \ w_{x_0} \dots w_{x_{n_{w_x}}} \\ & w_{y_0} \dots w_{y_{n_{w_y}}} \ \beta_{1_0} \dots \beta_{1_{n_{\beta_1}}} \ \beta_{2_0} \dots \beta_{2_{n_{\beta_2}}} \Big]^\top \end{aligned} \tag{12}$$

is estimated with an iterative procedure by minimizing the quadratic criterion

$$J(\theta_{iter}) = \sum_{k=0}^{N_t} \sum_{j=0}^{N_y} \sum_{i=0}^{N_x} \left(u(i,j,k) - \hat{u}(i,j,k)\right)^2, \tag{13}$$

where $\hat{u}(i,j,k)$ is the model and θ_{iter} is the parameter vector at the $iter$-th iteration.

The parametric estimation procedure repeats iteratively the following two steps from an initial vector θ_0 and until convergence

- approximate a solution of (1) with $\hat{u}(i,j,k)$ based on a proper orthogonal decomposition and the Galerkin method described in Sections 3.3 and 3.4,
- compute a new parameter vector θ_{iter+1} with a Levenberg–Marquardt algorithm which minimizes the criterion (13) as presented in Section 3.5.

To deal with the problem commonly encountered in ecological applications of a lack of prior knowledge of the parameters and of initialization of the iterative procedure, an estimator based on 3D partial moments is proposed in Section 3.6 to provide the first estimation θ_0.

3.3. Proper orthogonal decomposition

The orthonormal basis functions $\{\phi_n(x,y)\}_{n=1}^N$ used in the Galerkin method (see Section 3.4) are produced with the proper orthogonal decomposition which consists of generating the following empirical covariance data matrix

$$\mathbf{K} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1M} \\ \vdots & \ddots & \vdots \\ \alpha_{M1} & \cdots & \alpha_{MM} \end{bmatrix}, \tag{14}$$

where the components $\alpha_{k\ell}$ are determined by considering the inner product between time samples as follows

$$\alpha_{k\ell} = \frac{1}{M} \int \int_D \delta(x,y,k\ dt)\delta(x,y,\ell\ dt) dx dy, \tag{15}$$

and from the so-called M zero-mean snapshots

$$\delta(x,y,k\ dt) = u(x,y,k\ dt) - \bar{u}(x,y), \tag{16}$$

with $\bar{u}(x,y) = \text{mean}_t(u(x,y,t))$, the time average of the population.

Hence, by using an eigenvalue decomposition, the basis functions are determined as a linear combination of the data snapshots

$$\phi_n(x,y) = \sum_{k=1}^{M} \mathbf{Q}_k^{(n)} \delta(x,y,k\ dt), \tag{17}$$

where $\mathbf{Q}_k^{(n)}$ is the kth component of the nth eigenvector of $\mathbf{K}$.

3.4. Galerkin method

The Galerkin method operates a time–space separation which transforms the PDE problem into a system of ODEs. Some details are given in Ouvrard et al. (2019). It consists of finding an approximate solution of the PDE defined by

$$\hat{u}(x,y,t) = \sum_{n=1}^{N} a_n(t)\phi_n(x,y) + \bar{u}(x,y), \tag{18}$$

where $\{a_n(t)\}_{n=1}^N$ is the time coefficient set to be estimated.

In the general case of the PDE (1) with the varying parameters defined by (11), the obtained system of ODEs can be written with the following matrix formulation

$$\dot{\mathbf{a}}(t) = \left(\mathbf{\Gamma}_1 + \beta_{1_0}\mathbf{I}_N - 2\mathbf{\Gamma}_2\right)\ \mathbf{a}(t) - \left(\mathbf{I}_N \otimes \mathbf{a}^\top(t)\right)\mathbf{\Lambda}\ \mathbf{a}(t) + \mathbf{b}_1, \tag{19}$$

with[1]

$$\begin{aligned} &\otimes, \text{ Kronecker product,} \\ &\mathbf{a}(t) = \begin{bmatrix} a_1(t) & \dots & a_N(t) \end{bmatrix}^\top, \\ &\mathbf{I}_N, N \times N \text{ identity matrix,} \\ &\mathbf{\Gamma}_1 = \begin{bmatrix} \int \mathcal{F}_1(\phi_1)\phi_1 & \dots & \int \mathcal{F}_1(\phi_N)\phi_1 \\ \vdots & \ddots & \vdots \\ \int \mathcal{F}_1(\phi_1)\phi_N & \dots & \int \mathcal{F}_1(\phi_N)\phi_N \end{bmatrix}, \\ &\mathbf{\Gamma}_2 = \begin{bmatrix} \int \beta_2(H)\phi_1\phi_1\bar{u} & \dots & \int \beta_2(H)\phi_N\phi_1\bar{u} \\ \vdots & \ddots & \vdots \\ \int \beta_2(H)\phi_1\phi_N\bar{u} & \dots & \int \beta_2(H)\phi_N\phi_N\bar{u} \end{bmatrix}, \\ &\mathbf{\Lambda} = \int \beta_2(H)\mathbf{\Phi} \otimes \left(\mathbf{\Phi} \otimes \mathbf{\Phi}^\top\right), \\ &\mathbf{\Phi} = \begin{bmatrix} \phi_1(x,y) & \dots & \phi_N(x,y) \end{bmatrix}^\top, \\ &\mathbf{b}_1 = \begin{bmatrix} \int \mathcal{F}_2(\bar{u})\phi_1 & \dots & \int \mathcal{F}_2(\bar{u})\phi_N \end{bmatrix}^\top, \\ &\mathcal{F}_1(\bullet) = D_x(H)\frac{\partial^2 \bullet}{\partial x^2} + D_y(H)\frac{\partial^2 \bullet}{\partial y^2} - \left(w_x(H) - \frac{\partial D_x(H)}{\partial x}\right)\frac{\partial \bullet}{\partial x} \\ &\qquad - \left(w_y(H) - \frac{\partial D_y(H)}{\partial y}\right)\frac{\partial \bullet}{\partial y} + \textstyle\sum_{v=1}^{n_{\beta_1}} \beta_{1_v} H^v(x,y,t)\bullet, \\ &\mathcal{F}_2(\bullet) = D_x(H)\frac{\partial^2 \bullet}{\partial x^2} + D_y(H)\frac{\partial^2 \bullet}{\partial y^2} - \left(w_x(H) - \frac{\partial D_x(H)}{\partial x}\right)\frac{\partial \bullet}{\partial x} \\ &\qquad - \left(w_y(H) - \frac{\partial D_y(H)}{\partial y}\right)\frac{\partial \bullet}{\partial y} + \beta_1(H)\bullet - \beta_2(H)\bullet^2. \end{aligned} \tag{20}$$

The initial conditions of the ODEs system (19) are given by

$$a_\ell(0) = \int \left(u(x,y,0) - \bar{u}(x,y)\right)\phi_\ell(x,y), \tag{21}$$

for $\ell = 1, \dots, N$.

Therefore, given parameters, the simulation of the system of ODEs (19) from the initial conditions (21) provides the approximate solution (18).

3.5. Levenberg–Marquardt algorithm

Because the considered PDE model (1) is nonlinear in parameters, the Levenberg–Marquardt algorithm is used to identify its parameters (Björck, 1996; Nocedal and Wright, 2006). This algorithm is a good compromise between the gradient and Newton algorithms. It is based on the gradient vector, defined by $\mathbf{J}'_\theta = \partial J(\theta)/\partial\theta$, which gives the direction of the steepest descent of the quadratic criterion (13) and the pseudo-Hessian matrix which approximates the step leading to the optimum, *i.e.* $\mathbf{J}''_{\theta\theta} \approx \partial^2 J(\theta)/\partial\theta^2$.

[1] The double integration $\int \int_D \bullet dx dy$ is denoted $\int \bullet$.

The considered model (1) does not have exogenous input, and the unique excitation is the initial condition $u(x, y, 0)$. Moreover, the ecological dataset often has a small number of time samples. Thus, in order to limit the approximation error of the optimization algorithm, the true parametric sensitivity functions (Knudsen, 1994) are considered with initial conditions deduced from (21).

At the *iter*-th iteration, the new parameter vector is given by

$$\boldsymbol{\theta}_{iter+1} = \boldsymbol{\theta}_{iter} - \left(\mathbf{J}''_{\theta\theta} + \mu \mathbf{I}_{N_\theta}\right)^{-1} \mathbf{J}'_{\theta}, \tag{22}$$

with μ, a tuning parameter, $\mathbf{I}_{N_\theta}$, an identity matrix with appropriate dimensions, and the gradient and the pseudo-Hessian respectively defined by

$$\begin{aligned}
\mathbf{J}'_{\theta} &= -2 \sum_{k=0}^{N_t} \sum_{j=0}^{N_y} \sum_{i=0}^{N_x} \left(u(i,j,k) - \hat{u}(i,j,k)\right) \boldsymbol{\Xi}(k) \boldsymbol{\Phi}(i,j), \\
\mathbf{J}''_{\theta\theta} &= 2 \sum_{k=0}^{N_t} \sum_{j=0}^{N_y} \sum_{i=0}^{N_x} \boldsymbol{\Xi}(k) \boldsymbol{\Phi}(i,j) \boldsymbol{\Phi}^{\top}(i,j) \boldsymbol{\Xi}^{\top}(k),
\end{aligned} \tag{23}$$

where $\hat{u}(i,j,k)$ is the approximate solution (18), $\boldsymbol{\Phi}(x,y)$ is defined in (20), and $\boldsymbol{\Xi}(t)$ is the $N_\theta \times N$ time sensitivity function matrix given by

$$\begin{aligned}
\boldsymbol{\Xi}(t) = \Big[& \boldsymbol{\sigma}^{\mathbf{a}}_{D_{x_0}}(t) \quad \cdots \quad \boldsymbol{\sigma}^{\mathbf{a}}_{D_{x_{n_{D_x}}}}(t) \quad \boldsymbol{\sigma}^{\mathbf{a}}_{D_{y_0}}(t) \\
& \cdots \quad \boldsymbol{\sigma}^{\mathbf{a}}_{D_{y_{n_{D_y}}}}(t) \quad \boldsymbol{\sigma}^{\mathbf{a}}_{w_{x_0}}(t) \quad \cdots \quad \boldsymbol{\sigma}^{\mathbf{a}}_{w_{x_{n_{w_x}}}}(t) \\
& \boldsymbol{\sigma}^{\mathbf{a}}_{w_{y_0}}(t) \quad \cdots \quad \boldsymbol{\sigma}^{\mathbf{a}}_{w_{y_{n_{w_y}}}}(t) \quad \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{1_0}}(t) \quad \cdots \\
& \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{1_{n_{\beta_1}}}}(t) \quad \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{2_0}}(t) \quad \cdots \quad \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{2_{n_{\beta_2}}}}(t) \Big]^{\top}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{D_{\bullet_n}}(t) = & \left[\sigma^{a_1}_{D_{\bullet_n}}(t) \quad \cdots \quad \sigma^{a_N}_{D_{\bullet_n}}(t)\right]^{\top}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{w_{\bullet_n}}(t) = & \left[\sigma^{a_1}_{w_{\bullet_n}}(t) \quad \cdots \quad \sigma^{a_N}_{w_{\bullet_n}}(t)\right]^{\top}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{\bullet_n}}(t) = & \left[\sigma^{a_1}_{\beta_{\bullet_n}}(t) \quad \cdots \quad \sigma^{a_N}_{\beta_{\bullet_n}}(t)\right]^{\top}, \\
\sigma^{a_m}_{\circ}(t) = & \frac{\partial a_m(t)}{\partial \circ}.
\end{aligned} \tag{24}$$

The parametric sensitivity functions are simulated by considering the following matrix formulations deduced from derivatives of (19) with respect to each parameter

$$\begin{aligned}
\dot{\boldsymbol{\sigma}}^{\mathbf{a}}_{D_{\bullet_n}}(t) = & \left(\boldsymbol{\Gamma}_1 + \beta_{1_0}\mathbf{I}_N - 2\boldsymbol{\Gamma}_2\right) \boldsymbol{\sigma}^{\mathbf{a}}_{D_{\bullet_n}}(t) \\
& + \boldsymbol{\Gamma}_{D_{\bullet_n}} \mathbf{a}(t) - \left(\mathbf{I}_N \otimes \boldsymbol{\sigma}^{\mathbf{a}}_{D_{\bullet_n}}{}^{\top}(t)\right) \boldsymbol{\Lambda}\, \mathbf{a}(t) \\
& - \left(\mathbf{I}_N \otimes \mathbf{a}^{\top}(t)\right) \boldsymbol{\Lambda}\, \boldsymbol{\sigma}^{\mathbf{a}}_{D_{\bullet_n}}(t) + \mathbf{b}_{D_{\bullet_n}}, \\
\dot{\boldsymbol{\sigma}}^{\mathbf{a}}_{w_{\bullet_n}}(t) = & \left(\boldsymbol{\Gamma}_1 + \beta_{1_0}\mathbf{I}_N - 2\boldsymbol{\Gamma}_2\right) \boldsymbol{\sigma}^{\mathbf{a}}_{w_{\bullet_n}}(t) \\
& - \boldsymbol{\Gamma}_{w_{\bullet_n}} \mathbf{a}(t) - \left(\mathbf{I}_N \otimes \boldsymbol{\sigma}^{\mathbf{a}}_{w_{\bullet_n}}{}^{\top}(t)\right) \boldsymbol{\Lambda}\, \mathbf{a}(t) \\
& - \left(\mathbf{I}_N \otimes \mathbf{a}^{\top}(t)\right) \boldsymbol{\Lambda}\, \boldsymbol{\sigma}^{\mathbf{a}}_{w_{\bullet_n}}(t) - \mathbf{b}_{w_{\bullet_n}}, \\
\dot{\boldsymbol{\sigma}}^{\mathbf{a}}_{\beta_{1_n}}(t) = & \left(\boldsymbol{\Gamma}_1 + \beta_{1_0}\mathbf{I}_N - 2\boldsymbol{\Gamma}_2\right) \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{1_n}}(t) \\
& + \boldsymbol{\Gamma}_{\beta_{1_n}} \mathbf{a}(t) - \left(\mathbf{I}_N \otimes \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{1_n}}{}^{\top}(t)\right) \boldsymbol{\Lambda}\, \mathbf{a}(t) \\
& - \left(\mathbf{I}_N \otimes \mathbf{a}^{\top}(t)\right) \boldsymbol{\Lambda}\, \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{1_n}}(t) + \mathbf{b}_{\beta_{1_n}}, \\
\dot{\boldsymbol{\sigma}}^{\mathbf{a}}_{\beta_{2_n}}(t) = & \left(\boldsymbol{\Gamma}_1 + \beta_{1_0}\mathbf{I}_N - 2\boldsymbol{\Gamma}_2\right) \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{2_n}}(t) \\
& - 2\boldsymbol{\Gamma}_{\beta_{2_n}} \mathbf{a}(t) - \left(\mathbf{I}_N \otimes \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{2_n}}{}^{\top}(t)\right) \boldsymbol{\Lambda}\, \mathbf{a}(t) \\
& - \left(\mathbf{I}_N \otimes \mathbf{a}^{\top}(t)\right) \boldsymbol{\Lambda}\, \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{2_n}}(t) \\
& - \left(\mathbf{I}_N \otimes \mathbf{a}^{\top}(t)\right) \boldsymbol{\Lambda}_{\beta_{2_n}} \mathbf{a}(t) - \mathbf{b}_{\beta_{2_n}},
\end{aligned} \tag{25}$$

for $n = 0, \ldots, n_{D_\bullet}, n_{w_\bullet}$ or $n_{\beta_\bullet}$ and where

$$\begin{aligned}
\boldsymbol{\Gamma}_{D_{\bullet_n}} &= \begin{bmatrix} \int \mathcal{F}_{D_{\bullet_n}}(\phi_1)\phi_1 & \cdots & \int \mathcal{F}_{D_{\bullet_n}}(\phi_N)\phi_1 \\ \vdots & \ddots & \vdots \\ \int \mathcal{F}_{D_{\bullet_n}}(\phi_1)\phi_N & \cdots & \int \mathcal{F}_{D_{\bullet_n}}(\phi_N)\phi_N \end{bmatrix}, \\
\boldsymbol{\Gamma}_{w_{\bullet_n}} &= \begin{bmatrix} \int \mathcal{F}_{w_{\bullet_n}}(\phi_1)\phi_1 & \cdots & \int \mathcal{F}_{w_{\bullet_n}}(\phi_N)\phi_1 \\ \vdots & \ddots & \vdots \\ \int \mathcal{F}_{w_{\bullet_n}}(\phi_1)\phi_N & \cdots & \int \mathcal{F}_{w_{\bullet_n}}(\phi_N)\phi_N \end{bmatrix}, \\
\boldsymbol{\Gamma}_{\beta_{1_0}} &= \mathbf{I}_N, \\
\boldsymbol{\Gamma}_{\beta_{1_n}} &= \begin{bmatrix} \int H^n\phi_1\phi_1 & \cdots & \int H^n\phi_N\phi_1 \\ \vdots & \ddots & \vdots \\ \int H^n\phi_1\phi_N & \cdots & \int H^n\phi_N\phi_N \end{bmatrix}, \\
\boldsymbol{\Gamma}_{\beta_{2_n}} &= \begin{bmatrix} \int H^n\phi_1\phi_1\bar{u} & \cdots & \int H^n\phi_N\phi_1\bar{u} \\ \vdots & \ddots & \vdots \\ \int H^n\phi_1\phi_N\bar{u} & \cdots & \int H^n\phi_N\phi_N\bar{u} \end{bmatrix}, \\
\boldsymbol{\Lambda}_{\beta_{2_n}} &= \int H^n \boldsymbol{\Phi} \otimes \left(\boldsymbol{\Phi} \otimes \boldsymbol{\Phi}^{\top}\right), \\
\mathbf{b}_{D_{\bullet_n}} &= \begin{bmatrix} \int \mathcal{F}_{D_{\bullet_n}}(\bar{u})\phi_1 & \cdots & \int \mathcal{F}_{D_{\bullet_n}}(\bar{u})\phi_N \end{bmatrix}^{\top}, \\
\mathbf{b}_{w_{\bullet_n}} &= \begin{bmatrix} \int \mathcal{F}_{w_{\bullet_n}}(\bar{u})\phi_1 & \cdots & \int \mathcal{F}_{w_{\bullet_n}}(\bar{u})\phi_N \end{bmatrix}^{\top}, \\
\mathbf{b}_{\beta_{1_n}} &= \begin{bmatrix} \int H^n\bar{u}\phi_1 & \cdots & \int H^n\bar{u}\phi_N \end{bmatrix}^{\top}, \\
\mathbf{b}_{\beta_{2_n}} &= \begin{bmatrix} \int H^n\bar{u}^2\phi_1 & \cdots & \int H^n\bar{u}^2\phi_N \end{bmatrix}^{\top}, \\
\mathcal{F}_{D_{\bullet_n}}(\circ) &= H^n \frac{\partial^2 \circ}{\partial \bullet^2} + \frac{\partial H^n}{\partial \bullet}\frac{\partial \circ}{\partial \bullet}, \\
\mathcal{F}_{w_{\bullet_n}}(\circ) &= H^n \frac{\partial \circ}{\partial \bullet},
\end{aligned} \tag{26}$$

and with the initial conditions

$$\begin{aligned}
\boldsymbol{\sigma}^{\mathbf{a}}_{D_{\bullet_n}}(0) &= \boldsymbol{\Gamma}_{D_{\bullet_n}}\, \mathbf{a}(0) + \mathbf{b}_{D_{\bullet_n}}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{w_{\bullet_n}}(0) &= -\boldsymbol{\Gamma}_{w_{\bullet_n}}\, \mathbf{a}(0) - \mathbf{b}_{w_{\bullet_n}}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{1_n}}(0) &= \boldsymbol{\Gamma}_{\beta_{1_n}} \mathbf{a}(0) + \mathbf{b}_{\beta_{1_n}}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{2_n}}(0) &= -2\boldsymbol{\Gamma}_{\beta_{2_n}}\, \mathbf{a}(0) - \left(\mathbf{I}_N \otimes \mathbf{a}^{\top}(0)\right) \boldsymbol{\Lambda}_{\beta_{2_n}} \mathbf{a}(0) - \mathbf{b}_{\beta_{2_n}}.
\end{aligned} \tag{27}$$

3.6. Initial estimate based on 3D partial moments

To obtain an initial estimate $\boldsymbol{\theta}_0$, a first estimation of the PDE with constant parameters, *i.e.* with $n_{D_\bullet} = 0$, $n_{w_\bullet} = 0$ and $n_{\beta_\bullet} = 0$, is performed from a least-squares (LS) method based on an extension to the three dimensions of the partial moments approach developed in Trigeassou (1987) and Ouvrard and Trigeassou (2011). Let us denote this estimation $\boldsymbol{\theta}_{LS} = \begin{bmatrix} D_{x_0} & D_{y_0} & w_{x_0} & w_{y_0} & \beta_{1_0} & \beta_{2_0} \end{bmatrix}^{\top}$.

From there, the values of the missing parameters corresponding to $n_{D_\bullet} > 0$, $n_{w_\bullet} > 0$ and $n_{\beta_\bullet} > 0$ in (12) will be chosen with the expertise of ecologists or with small values.

Let us define the 3D partial moment with orders ℓ, m and n of an arbitrary signal $f(x,y,t)$ belonging to a Hilbert space by

$$\mathcal{M}^{f}_{\ell,m,n}(X,Y,T) = \int_0^X \int_0^Y \int_0^T x^{\ell} y^{m} t^{n} f(x,y,t)\, dt\, dy\, dx, \tag{28}$$

with ℓ, m and $n \in \{\mathbb{N}, -\}$.[2]

The idea of the partial moment-based approach is to apply to the PDE (1) with $n_{D_\bullet} = 0$, $n_{w_\bullet} = 0$ and $n_{\beta_\bullet} = 0$ the following integration

$$\int_0^X \int_0^{x_1} x_2^2 \int_0^Y \int_0^{y_1} y_2^2 \int_0^T t \quad (1) \quad dt\, dy_2\, dy_1\, dx_2\, dx_1 \tag{29}$$

in such a way that all partial derivatives disappear. Thus, a formulation with only partial moments and parameters is obtained in the form of a

[2] If ℓ, m or n is the symbol –, then there is no integration with respect to the corresponding variable. For instance, $\mathcal{M}^{f}_{\ell,m,-}(X,Y,T) = \int_0^X \int_0^Y x^{\ell} y^{m} f(x,y,T)\, dy\, dx$.

linear regression. Indeed, by considering the integration by parts

$$\int_0^{\tau_0} \tau_1 \frac{\partial f}{\partial \tau} d\tau_1 = \tau_0 f(\tau_0) - \int_0^{\tau_0} f(\tau_1) d\tau_1 \tag{30}$$

or

$$\begin{aligned}\int_0^{\tau_0} \int_0^{\tau_1} \tau_2^2 \frac{\partial^2 f}{\partial \tau^2} d\tau_2 d\tau_1 &= \tau_0^2 f(\tau_0) \\ &+2\tau_0 \int_0^{\tau_0} f(\tau_1) d\tau_1 - 6 \int_0^{\tau_0} \tau_1 f(\tau_1) d\tau_1,\end{aligned} \tag{31}$$

and the Cauchy formula of repeated integrations

$$\begin{aligned}\int_0^t \int_0^{\tau_1} \cdots \int_0^{\tau_{n-1}} f(\tau_n) d\tau_n \cdots d\tau_2 d\tau_1 = \\ \frac{1}{(n-1)!} \int_0^t (t-\tau_1)^{n-1} f(\tau_1) d\tau_1,\end{aligned} \tag{32}$$

Eq. (29) can be formulated as a linear regression as follows

$$\mathcal{U}_1(X,Y,T) = \boldsymbol{\varphi}^\top(X,Y,T)\boldsymbol{\theta}_{LS}, \tag{33}$$

with

$$\begin{aligned}\boldsymbol{\varphi}(X,Y,T) = \big[&\mathcal{U}_2(X,Y,T) \quad \mathcal{U}_3(X,Y,T) \quad -\mathcal{U}_4(X,Y,T) \\ &-\mathcal{U}_5(X,Y,T) \quad \mathcal{U}_6(X,Y,T) \quad -\mathcal{U}_7(X,Y,T)\big]^\top\end{aligned} \tag{34}$$

and

$$\begin{aligned}\mathcal{U}_1 &= XYT\mathcal{M}^u_{2,2,-} - XT\mathcal{M}^u_{2,3,-} - YT\mathcal{M}^u_{3,2,-} \\ &+T\mathcal{M}^u_{3,3,-} - XY\mathcal{M}^u_{2,2,0} + X\mathcal{M}^u_{2,3,0} + Y\mathcal{M}^u_{3,2,0} - \mathcal{M}^u_{3,3,0}, \\ \mathcal{U}_2 &= X^2Y\mathcal{M}^u_{-,2,1} - X^2\mathcal{M}^u_{-,3,1} + 2XY\mathcal{M}^u_{0,2,1} \\ &-2X\mathcal{M}^u_{0,3,1} - 6Y\mathcal{M}^u_{1,2,1} + 6\mathcal{M}^u_{1,3,1}, \\ \mathcal{U}_3 &= XY^2\mathcal{M}^u_{2,-,1} - Y^2\mathcal{M}^u_{3,-,1} + 2XY\mathcal{M}^u_{2,0,1} \\ &-2Y\mathcal{M}^u_{3,0,1} - 6X\mathcal{M}^u_{2,1,1} + 6\mathcal{M}^u_{3,1,1}, \\ \mathcal{U}_4 &= 3Y\mathcal{M}^u_{2,2,1} - 3\mathcal{M}^u_{2,3,1} - 2XY\mathcal{M}^u_{1,2,1} + 2X\mathcal{M}^u_{1,3,1}, \\ \mathcal{U}_5 &= 3X\mathcal{M}^u_{2,2,1} - 3\mathcal{M}^u_{3,2,1} - 2XY\mathcal{M}^u_{2,1,1} + 2Y\mathcal{M}^u_{3,1,1}, \\ \mathcal{U}_6 &= XY\mathcal{M}^u_{2,2,1} - Y\mathcal{M}^u_{3,2,1} - X\mathcal{M}^u_{2,3,1} + \mathcal{M}^u_{3,3,1}, \\ \mathcal{U}_7 &= XY\mathcal{M}^{u^2}_{2,2,1} - Y\mathcal{M}^{u^2}_{3,2,1} - X\mathcal{M}^{u^2}_{2,3,1} + \mathcal{M}^{u^2}_{3,3,1}.\end{aligned} \tag{35}$$

Thus, the PDE with fixed parameters is estimated with the following least-squares estimator

$$\begin{aligned}\boldsymbol{\theta}_{LS} = &\left(\sum_{k=0}^{N_t} \sum_{j=0}^{N_y} \sum_{i=0}^{N_x} \boldsymbol{\varphi}(i,j,k)\boldsymbol{\varphi}^\top(i,j,k) \right)^{-1} \\ &\cdot \sum_{k=0}^{N_t} \sum_{j=0}^{N_y} \sum_{i=0}^{N_x} \boldsymbol{\varphi}(i,j,k)\mathcal{U}_1(i,j,k).\end{aligned} \tag{36}$$

Particular attention should be paid to the estimation of the diffusion coefficients D_{x_0} and D_{y_0} and to the reaction parameters β_{1_0} and β_{2_0} which must have non-negative values. Therefore, a constrained least-squares problem must be performed to implement the estimator (36) to ensure positive estimated values for D_{x_0}, D_{y_0}, β_{1_0} and β_{2_0}. Some implementations can be found in Lawson and Hanson (1974).

Note that this LS estimator is biased and may give inaccurate values due to the numerical implementation of the integrals in (28). But it is a useful first estimation to overcome the lack of prior knowledge about the parameters and the initialization problem.

3.7. Stopping test and implementation tip

There are different stopping tests (Björck, 1996; Nocedal and Wright, 2006) to stop the Levenberg–Marquardt algorithm where the optimum is considered to have been reached. For instance, it is easy to implement a stopping test based on the criterion evolution by setting *a priori* a tolerance as follows

$$J(\boldsymbol{\theta}_{iter}) - J(\boldsymbol{\theta}_{iter+1}) < \epsilon \tag{37}$$

where the tolerance ϵ is a small positive real.

Fig. 6. Spatial distribution of the squares monitored in the French Breeding Bird Survey scheme.

Due to the particular context of ecological applications with poor excitation given by the initial conditions and the difficult problem of estimating the parameters of the PDE model, the quadratic criterion (13) presents flat areas where convergence is very slow. The risk is to stop the iterative procedure prematurely. There are different solutions such as the simulated annealing approach or the particle swarm optimization to solve this problem (Rao, 2009).

We propose an easy-to-program implementation tip. When condition (37) is satisfied, before stopping the algorithm, we test a new parameter vector with randomly-drawn values close to the parameter values reached. More precisely, we test the following vector over a few iterations to find a new path in convergence

$$\boldsymbol{\theta}_{iter+1} = \boldsymbol{\theta}_{iter} + 0.005\ \boldsymbol{\xi}\ \boldsymbol{\theta}_0 \tag{38}$$

where $\boldsymbol{\xi} \in \mathbb{R}^{N_\theta}$ is a random vector with values belonging to $[-1, 1]$. In other words, we change the parameter values within a neighbourhood of the estimate reached with a random selection lower than 0.5% of the initial vector $\boldsymbol{\theta}_0$.

4. Results

The objectives of this section are to validate the methodology with realistic simulated data and to compare the varying-parameter PDE model with the popular GLM model on real data.

Let us consider the French population of Yellowhammers. This small passerine bird is in global decline and its distribution has contracted from the southwest to the northeast of France over the past 20 years (Keller et al., 2020). The main impacts on this population are agricultural intensification and global warming.

4.1. Data sets

4.1.1. Bird data set

Data on Yellowhammers is taken from the French Breeding Bird Survey described by Jiguet et al. (2012). This survey presents a spatial distribution of more than 2000 sites (squares 2×2 km shown in Fig. 6) surveyed at least once between 2002 and 2018,[3] the considered time period. In each square, an observer counts all individuals seen or heard during a fixed period of 5 min in 10 points during the breeding season.

The Yellowhammers data, plotted in Fig. 7, is the response variable.

[3] In the following, on several figures, the year 2018 has been omitted for the sake of clarity.

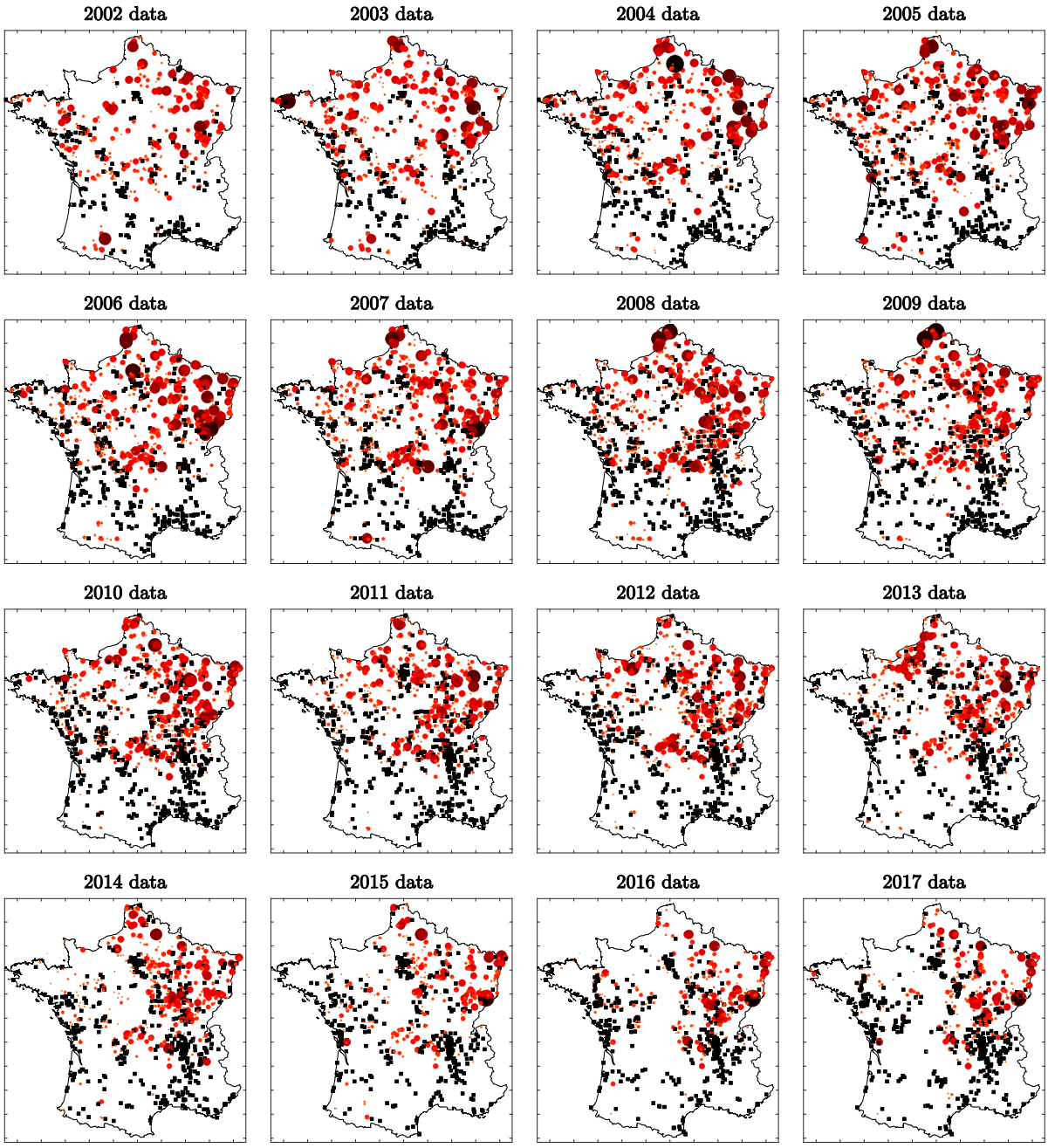


Fig. 7. Data on Yellowhammers. The absences are given by black squares and the number of individuals is shown by a red circle of varying size and darkness for numbers between 1 and 25.

4.1.2. Habitat data set

The habitat is characterized by the CORINE Land Cover database (Büttner and Kosztra, 2017) – European Union, CORINE Land Cover, 2000, 2006, 2012 and 2018. This database gives the land uses according to a 44-name nomenclature. Some examples of names are given in Table 1.

For each spatial step of the map of France, the proportion of each CORINE Land Cover code is calculated in a radius of 1.5 km, as shown on the left-hand side of Fig. 8. The right-hand side shows the map corresponding to non-irrigated arable land, *e.g.* the CORINE Land Cover code $CLC211$. The time step is six years, more precisely the data considered are from 2000, 2006, 2012 and 2018.

This gives us 44 habitat variables $CLCxxx$ (for each code in Table 1) that change every six years in time.

Table 1
Some examples of the 44 variables and names from the CORINE Land Cover database.

CORINE Land Cover code	Name
$CLC111$	Continuous urban fabric
$CLC211$	Non-irrigated arable land
$CLC221$	Vineyards
...	...

4.1.3. Bioclimatic data set

Global warming is characterized by the Worldclim database (Fick and Hijmans, 2017). This database includes 19 bioclimatic variables based on temperature and precipitation, as shown in Table 2 for some

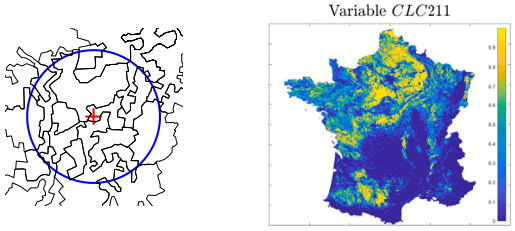


Fig. 8. Example of a map for calculating habitat variables (left). The habitat variable $CLC211$ for years 2012 to 2017 (right).

Table 2
Some examples of the 19 bioclimatic variables and names from the Worldclim database.

BIO1	Annual mean temperature
BIO2	Mean diurnal range
...	...
BIO12	Annual precipitation
BIO13	Precipitation of wettest month
...	...

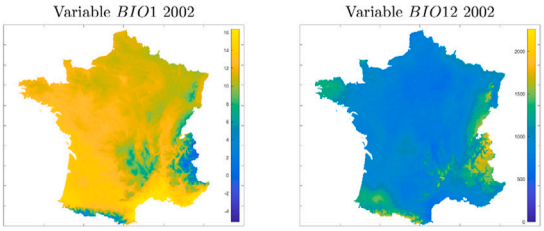


Fig. 9. Two examples of bioclimatic variables for 2002: $BIO1$, annual mean temperature, and $BIO12$, annual precipitation.

examples. Fig. 9 plotted the bioclimatic variables $BIO1$, annual mean temperature, and $BIO12$, annual precipitation.

Therefore, we have 19 bioclimatic variables $BIOxx$ (for each code in Table 2) that change every year in time.

4.2. Validation of the methodology in simulation

4.2.1. The simulation protocol

Let us consider a simulation protocol that provides data similar to the real French population of Yellowhammers shown in Fig. 7. To model such a population evolution, let us consider the true variable parameters defined in (11) as follows

$$\begin{aligned} &D_x(H) = D_y(H) = D_0 + D_1 H(x,y,t), \quad \beta_1(H) = \beta_{1_0}, \\ &\beta_2(H) = \beta_{2_0} + \beta_{2_1} H(x,y,t), \end{aligned} \tag{39}$$

with the following true parameters vector to be estimated

$$\begin{aligned} \boldsymbol{\theta}_{true} &= \begin{bmatrix} D_0 & D_1 & \beta_{1_0} & \beta_{2_0} & \beta_{2_1} \end{bmatrix}^\top \\ &= [200 \quad -150 \quad 0.01 \quad 0.1 \quad -0.089]^\top . \end{aligned} \tag{40}$$

Thereby, the effects described in Section 2.3 lead to a global decline and a contraction of the population from the southwest to the northeast:

- The diffusion effect is greater in the unfavourable areas with diffusion coefficients $D_x(H) = D_y(H)$ close to 200 than in the favourable areas where $D_x(H) = D_y(H)$ are close to 50. Indeed, if the environment is favourable, juvenile birds will disperse in the immediate vicinity of their place of birth. Otherwise, the diffusion is greater.
- The rate of growth or decline given by $\beta_1(H) = \beta_{1_0}$ is constant and the carrying capacity, which is specified by $\beta_1(H)/\beta_2(H)$, is 0.91 for $H(x,y,t) = 1$ and is 0.1 for $H(x,y,t) = 0$. Thus, globally, the population decreases with a very marked decline for the areas which become unfavourable.

To summarize, the considered true system is the following PDE

$$\begin{aligned} \frac{\partial u(x,y,t)}{\partial t} &= \frac{\partial}{\partial x}\left[\left(D_0 + D_1 H(x,y,t)\right)\frac{\partial u(x,y,t)}{\partial x}\right] \\ &+ \frac{\partial}{\partial y}\left[\left(D_0 + D_1 H(x,y,t)\right)\frac{\partial u(x,y,t)}{\partial y}\right] \\ &+ \beta_{1_0} u(x,y,t) - \left(\beta_{2_0} + \beta_{2_1} H(x,y,t)\right) u^2(x,y,t) \ . \end{aligned} \tag{41}$$

The scheduling variable $H(x,y,t)$ was calculated from the habitat and bioclimatic variables on the basis of a GLM model described in Section 4.3.1.

To obtain accurate data, the simulation is performed with a spatial step of 0.5 km and a time step 0.01 year.

As in the actual population study of the Yellowhammers, consider a small number of time samples corresponding to the years from 2002 to 2018, *i.e.* 17 time samples.

The considered scheduling variable $H(x,y,t)$ is plotted in Fig. 17 (Section 4.3.1). The latter shows that, from the initial year 2002 to 2017, the conditions for the Yellowhammers are gradually deteriorating.

Fig. 10 presents the simulated Yellowhammer population $u(x,y,t)$ obtained for a few years. The considered initial population $u(x,y,2002)$ is the kriged population taken from the French breeding bird survey for the Yellowhammer species for the year 2002, *i.e.* the map at the top left of Fig. 7. From this initial distribution, a gradual decline takes place (see Fig. 11) with a more pronounced decline in the southwest, close to extinction, and a population that remains in favourable areas notwithstanding a decline in numbers.

4.2.2. Validation of the methodology

The true system simulated in Section 4.2.1 is a reaction–diffusion PDE defined by (41) with a variable diffusion coefficient, a fixed growth and a variable carrying capacity. The true parameters are given in (40).

The goal is to estimate the parameters from a prior knowledge limited to

- a diffusion effect known to be bigger in the unfavourable areas than the favourable ones,
- and a carrying capacity smaller in the unfavourable areas than the favourable ones.

These hypotheses correspond to natural behaviours. However, it is assumed that there are no *a priori* known approximate values of the parameters.

4.2.3. First estimation with 3D partial moments

First, the least-squares method based on partial moments proposed in Section 3.6 is run to estimate initial values of the parameters D_0, β_{1_0} and β_{2_0}.

With the considered PDE structure (41), the regressor in the least-squares estimate (36) becomes

$$\begin{aligned} \boldsymbol{\varphi}(X,Y,T) = \big[&\left(\mathcal{U}_2(X,Y,T) + \mathcal{U}_3(X,Y,T)\right) \\ &\mathcal{U}_6(X,Y,T) \quad -\mathcal{U}_7(X,Y,T)\big]^\top . \end{aligned} \tag{42}$$

The considered measured dataset is $\{u(i\,dx, j\,dy, k\,dt)\}_{i=0,\ldots,N_x;\ j=0,\ldots,N_y;\ k=0,\ldots,N_t}$ with $dx = dy = 4$ km, $dt = 1$ year, $N_x = N_y = 249$ and $N_t = 16$, *i.e.* a spatial grid size 250×250 and 17 time samples.

The least-squares estimator (36) produces the following estimate

$$\boldsymbol{\theta}_{LS} = \begin{bmatrix} D_0 & \beta_{1_0} & \beta_{2_0} \end{bmatrix}^\top = [65.7 \quad -0.0248 \quad 0.0239]^\top . \tag{43}$$

Notice that the estimated parameters lead to a negative growth coefficient $r = \beta_{1_0} = -0.0248$ and a negative carrying capacity $K = \beta_{1_0}/\beta_{2_0} \approx -1.04$. The simulation of the PDE model at a point of spatial coordinates (x_0, y_0), *i.e.* the simulation of the logistic growth

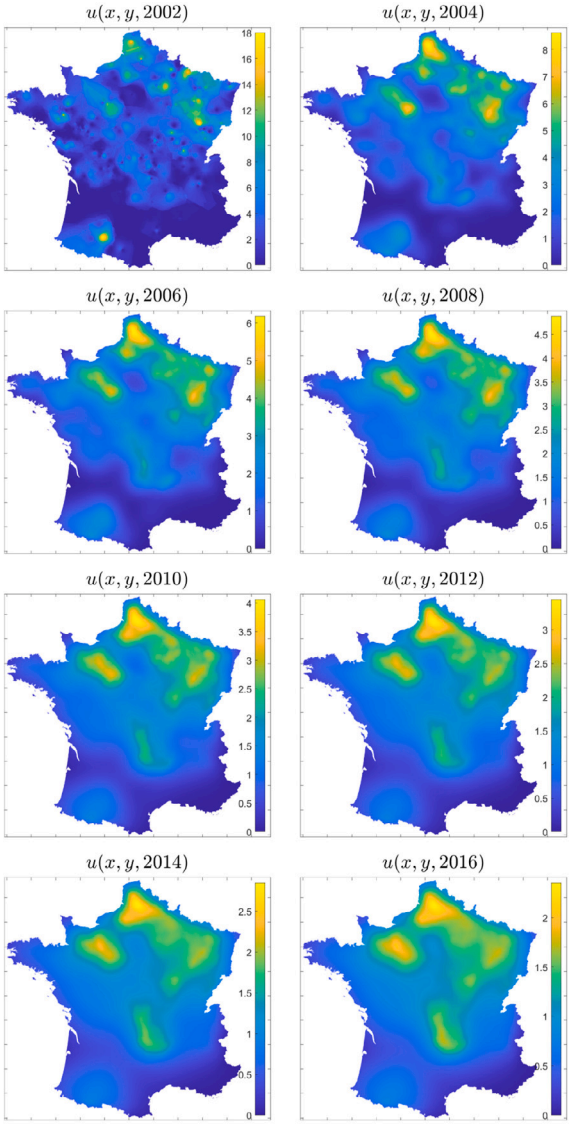


Fig. 10. The simulated Yellowhammer population for a few years (realistic simulated data example).

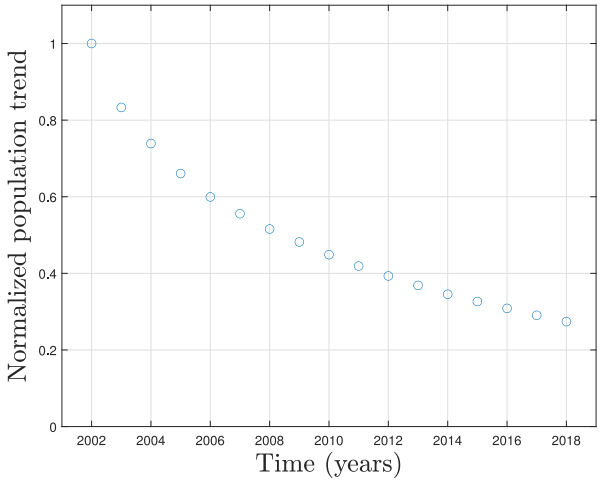


Fig. 11. The simulated Yellowhammer population trend versus time (realistic simulated data example).

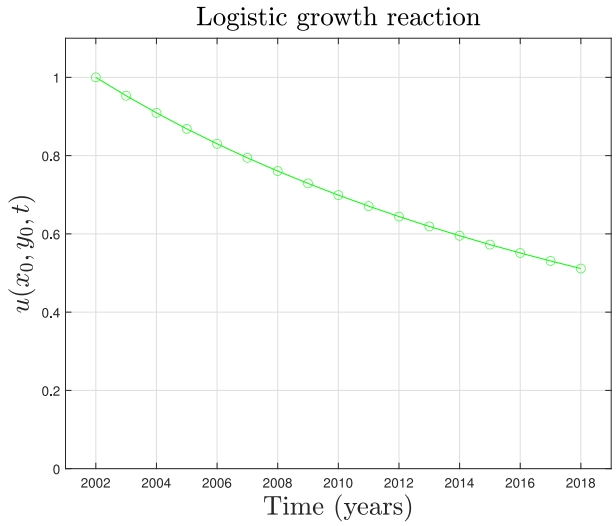


Fig. 12. Simulation of the Verhulst model (7) with $r = -0.0248$ and $K = -1.04$ (realistic simulated data example).

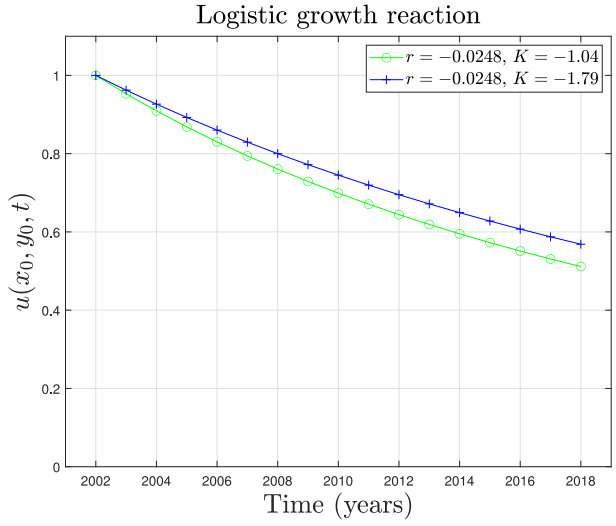


Fig. 13. Simulations of the Verhulst model (7) with $r = -0.0248$, $K = -1.04$ and $K = -1.79$ (realistic simulated data example).

reaction (7), plotted in Fig. 12 shows a declining trend which roughly approaches that of the gradual population trend presented in Fig. 11.

It is important to emphasize that the parameter-fixed PDE model estimated with the partial moment method is a very simplified modelling of the true structure, a parameter-varying PDE. Moreover, this least-squares estimate is biased. Therefore, the anomaly of the negative carrying capacity is due to the difficulties of this simplified model to fit the measured data.

4.2.4. *Estimation with the Levenberg–Marquardt algorithm*

From the first estimation (43), the initial parameter is completed with $D_1 = -40$ and $\beta_{2_1} = -0.01$. Thus, the diffusion coefficient is equal to 65.7 if $H = 0$, *i.e.* unfavourable area, and equal to 25.7 if $H = 1$, *i.e.* favourable area. The carrying capacity is equal to $K = -1.04$ or $K = -1.79$ if $H = 0$ or $H = 1$, respectively. The corresponding simulations of the logistic growth reactions, plotted in Fig. 13, show more or less significant decreases for more or less unfavourable areas.

Therefore, in the Levenberg–Marquardt algorithm (22), the initial parameter vector is given by

$$\boldsymbol{\theta}_0 = [65.7 \quad -40 \quad -0.0248 \quad 0.0239 \quad -0.01]^{\top}. \tag{44}$$

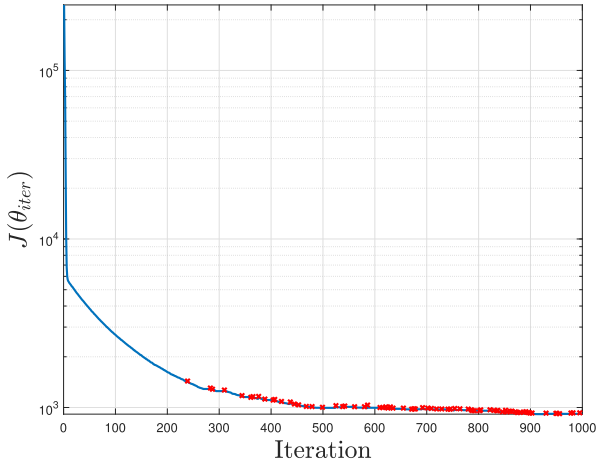


Fig. 14. Convergence of the criterion $J(\theta_{iter})$ of the Levenberg–Marquardt algorithm.

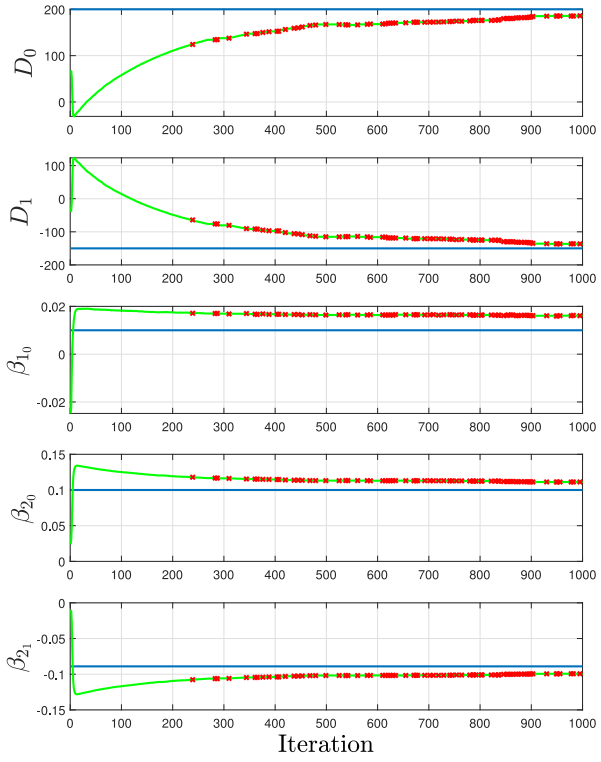


Fig. 15. Parameters convergence (green line) with the Levenberg–Marquardt algorithm, true parameter values (blue line) and implementation tip (38) (red cross) (realistic simulated data example).

From this initial parameter vector, the convergence of the Levenberg–Marquardt algorithm is represented in terms of criterion $J(\theta_{iter})$ (Eq. (13)) and parameters in Figs. 14 and 15, respectively. The red crosses correspond to the implementation tip with the random draw (38). Without this tip, the algorithm stops on a flat area which can be considered as a secondary optimum, and the random draw allows the leaving of this secondary optimum to continue the convergence.

4.3. Application to real data

4.3.1. GLM results

We estimate a GLM model (McCullagh and Nelder, 1983) to represent the response variable, *i.e.* the yellowhammer population measurements shown in Fig. 7, from the explicative variables, *i.e.* the habitat and bioclimatic data at the coordinates where Yellowhammers measurements are available in Figs. 8 and 9, respectively.

In the GLM model framework, a Poisson distribution and a log link are considered to represent the counting data of Yellowhammers. The explicative variables are averaged over a 5-year moving window. The linear predictor is composed of the intercept, linear and quadratic terms, and the selection of explicative variables is carried out by AIC criterion from the 44 habitat variables $CLCxxx$ and the 19 bioclimatic variables $BIOxx$.

At the end of the selection procedure, the estimated GLM model is composed of 79 parameters γ_i and yields to the following population

$$\hat{u}_{GLM}(x,y,t) = e^{\left(\gamma_0+\sum_{i=1}^{18}\gamma_{i+1}BIO--_{(x,y,t)}+\sum_{i=1}^{14}\gamma_{i+19}BIO--^2_{(x,y,t)}+ \sum_{i=1}^{25}\gamma_{i+33}CLC---_{(x,y,t)}+\sum_{i=1}^{21}\gamma_{i+58}CLC---^2_{(x,y,t)}\right)}. \tag{45}$$

The simulated population of Yellowhammers given by the GLM model (45) is plotted in Fig. 16 with the real data. The fitting criterion[4] between the real data and $\hat{u}_{GLM}$ at the measurement points and instants is 25%.

4.3.2. PDE results

We estimate the parameter-varying PDE model defined by

$$\begin{aligned}\frac{\partial u(x,y,t)}{\partial t} &= \frac{\partial}{\partial x}\left[(D_0+D_1H)\frac{\partial u(x,y,t)}{\partial x}\right]\\ &+\frac{\partial}{\partial y}\left[(D_0+D_1H)\frac{\partial u(x,y,t)}{\partial y}\right]\\ &-w_x(1-H)\frac{\partial u(x,y,t)}{\partial x}-w_y(1-H)\frac{\partial u(x,y,t)}{\partial y}\\ &+(\beta_{1_0}+\beta_{1_1}H)u(x,y,t)-(\beta_{2_0}+\beta_{2_1}H)u^2(x,y,t)\end{aligned} \tag{46}$$

where the scheduling variable is calculated from the GLM model (45) as follows

$$H(x,y,t) = \frac{\min(\hat{u}_{GLM}(x,y,t),5)}{5}, \tag{47}$$

i.e. $H(x,y,t)=1$ where the GLM model predicts a number of individuals greater than or equal to 5, which can be considered as an abundant population for the considered protocol and the Yellowhammer population. The considered $H(x,y,t)$ is plotted in Fig. 17.

After following the procedure described in Section 3, the approximate solution (18) of the PDE (46) yields a fitting criterion of 44%. The corresponding simulated population of Yellowhammers is plotted in Fig. 18 with the real data.

5. Discussion

Let us recall the difficult context of parametric estimation of PDE models, which is the main reason for the low use of this type of model in wildlife population dynamics.

The proposed tools are:

- A parameter-varying PDE model;
- A parameter-varying PDE simulator;
- A complete parametric estimation procedure.

The aims of this paper are to present the different tools, validate them on realistic simulated data and give some first results on real data in comparison with a GLM model.

Looking at Figs. 14 and 15, even if the convergence is slow, we notice that the estimated parameters tend towards the true parameters.

In Fig. 15, there is a difference between the estimated parameters θ_{1000} and the true parameters θ_{true} which is due to the numerical implementation and the POD-Galerkin approximate solution of the PDE. However, within the difficult context of the study of wildlife

[4] The fitting criterion is defined by $100\left(1-\frac{\|u-\hat{u}\|_2}{\|u-\text{mean}(u)\|_2}\right)$ where $\hat{u}$ is $\hat{u}_{GLM}$ or $\hat{u}$, *i.e.* Eq. (45) or (18), respectively.

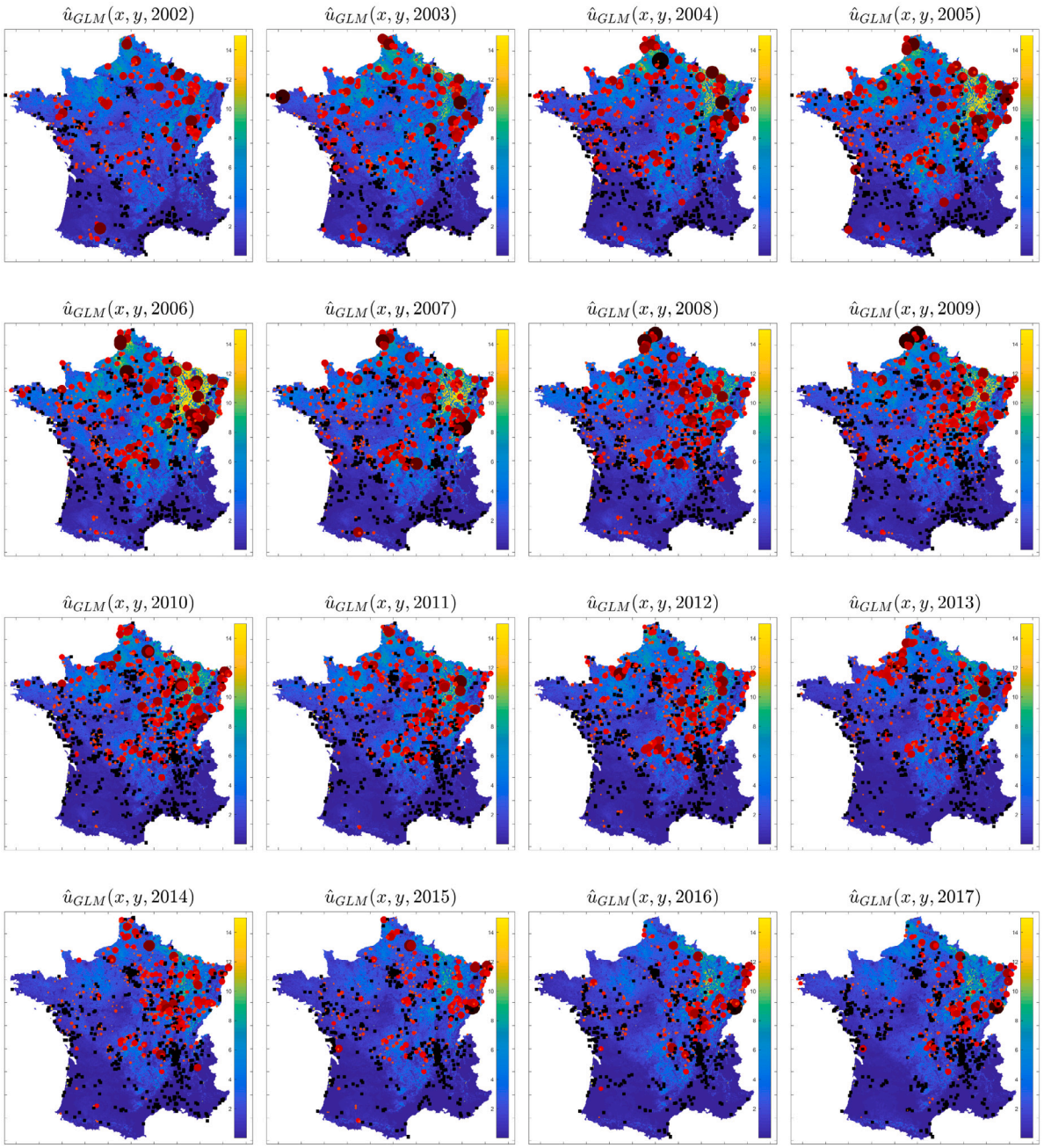


Fig. 16. Simulated population $\hat{u}_{GLM}(x, y, t)$ of Yellowhammers given by the GLM model (maps with a colour scale between blue and yellow) and real data (black squares for the absences and red circle of varying size and darkness for the number of individuals between 1 and 25).

population dynamics described in Section 3, these results are quite good and validate the methodology.

In Section 4.3, the application to real data shows that the parameter-varying PDE model gives a much better fit than the GLM model.

6. Conclusions

In this paper, a new methodology for estimating population dynamic models is proposed and validated. The considered models are particularly suitable for representing the impacts of global changes on wildlife populations. In fact, new parameter-varying PDE models are proposed and the variable parameters make it possible to introduce the environmental heterogeneity which can characterize agricultural intensification or global warming, for example.

Particular attention was paid to the implementation of the identification procedure. Parametric estimation for PDE models is a challenging problem and this difficulty is heightened in the case of ecological data. For that reason, the optimization is performed with true parametric sensitivity functions. PDE solutions are determined by a Galerkin method with basis functions given by a proper orthogonal decomposition applied to the measured data. This is an original implementation that considerably reduces the computation time of the parametric estimation procedure. For the initialization of the optimization algorithm, a pre-estimate is performed to overcome the lack of prior knowledge of the parameters.

The tools are validated for realistic simulated data. The ability of the parameter-varying PDE model to represent the impact of global change has been demonstrated for the French population of Yellowhammers.

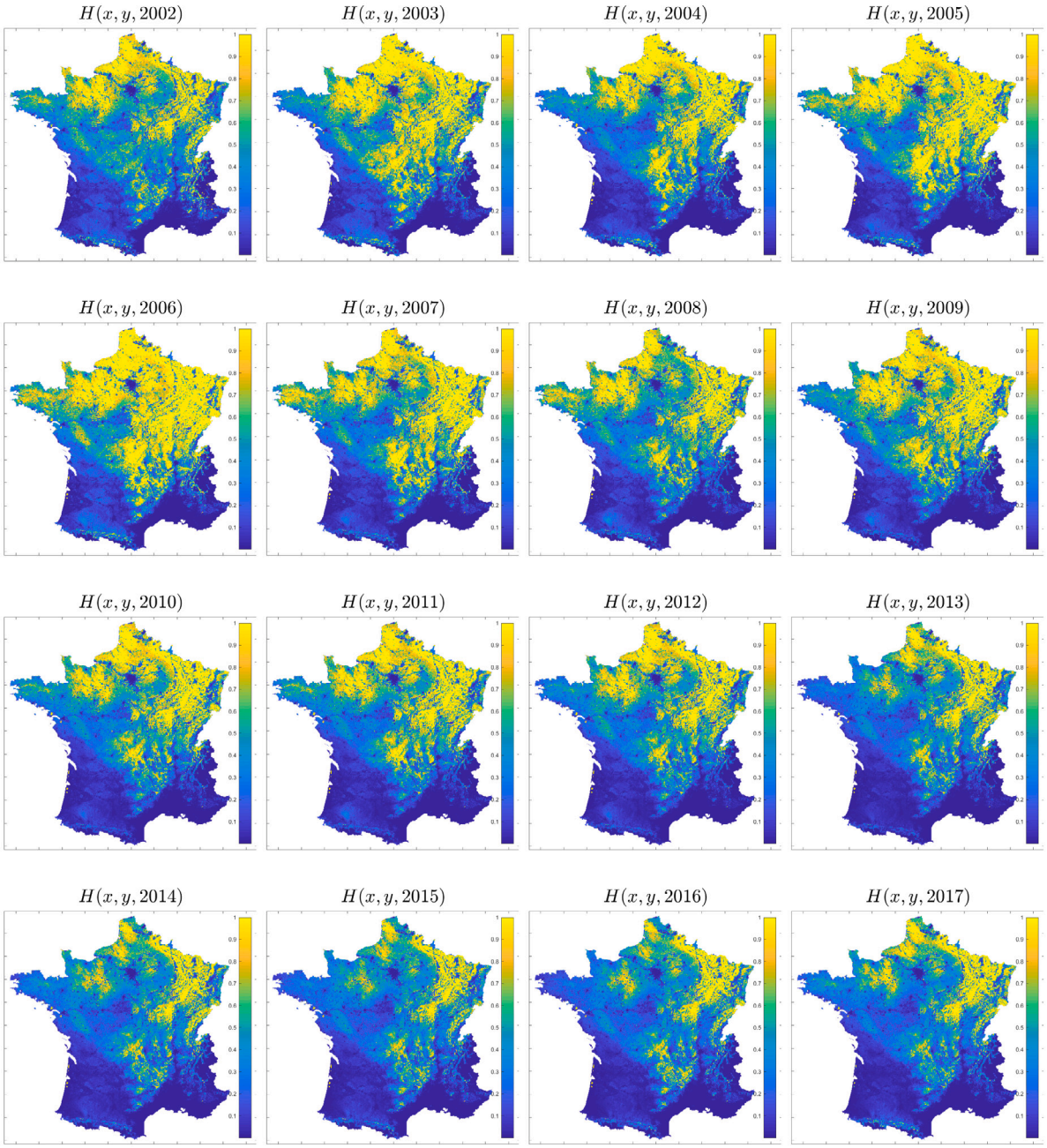


Fig. 17. The scheduling variable $H(x, y, t)$ given by (47).

CRediT authorship contribution statement

Mohamad Chhaytle, Régis Ouvrard, Thierry Poinot and Lauriane Mouysset presented the results in this paper on the basis of their respective expertise in automatic control, bioeconomic and population dynamics modelling.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is not confidential.

Acknowledgements

This work (IBIS project) benefits from the financial support of the Nouvelle-Aquitaine Region, France and the Grand Poitiers Agglomeration, France.

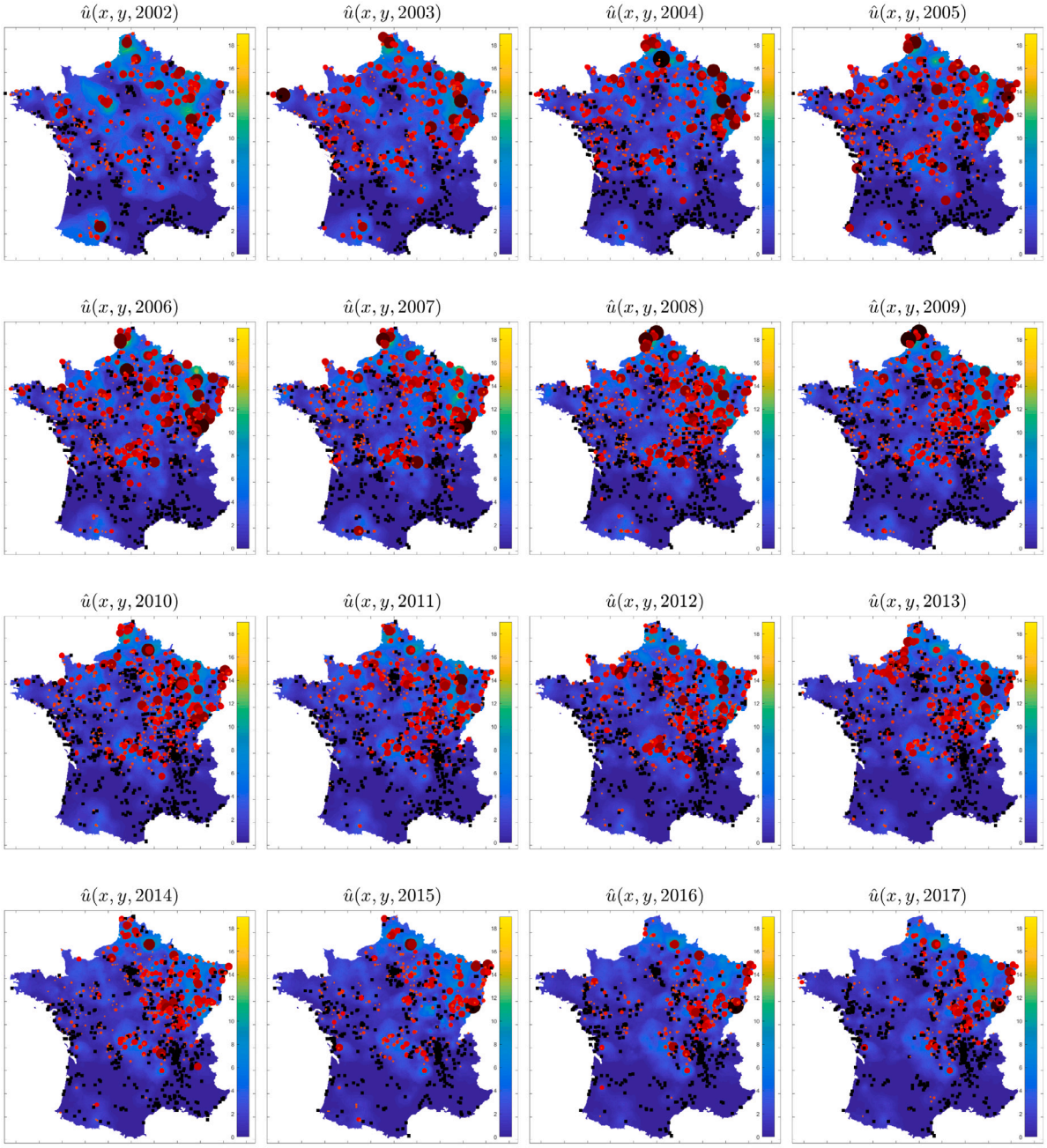


Fig. 18. Simulated population $\hat{u}(x, y, t)$ of Yellowhammers given by the POD-Galerkin approximate (46) of the PED model (46) (maps with a colour scale between blue and yellow) and real data (black squares for the absences and red circle of varying size and darkness for the number of individuals between 1 and 25).

Appendix A. Crank–Nicolson scheme for Eq. (9)

Consider the Crank–Nicolson scheme (Crank and Nicolson, 1996) to simulate Eq. (9). The interest of this method is that it gives a good compromise between the explicit and implicit Euler methods.

By adopting the discrete notation $f_{ij}^k = f(x_i, y_j, t_k)$ where $x_i = x_0 + i\ dx$, $y_j = y_0 + j\ dy$ and $t_k = k\ dt$, for any continuous function $f(x, y, t)$, the considered discretization for the time and spatial derivatives are the following

$$\begin{aligned}
\frac{\partial u(x,y,t)}{\partial t} &= \frac{u_{i,j}^{k+1} - u_{i,j}^{k}}{dt},\\
\frac{\partial^2 u(x,y,t)}{\partial x^2} &= \frac{1}{2}\left(\frac{u_{i-1,j}^{k+1} - 2u_{i,j}^{k+1} + u_{i+1,j}^{k+1}}{dx^2} + \frac{u_{i-1,j}^{k} - 2u_{i,j}^{k} + u_{i+1,j}^{k}}{dx^2}\right),\\
\frac{\partial^2 u(x,y,t)}{\partial y^2} &= \frac{1}{2}\left(\frac{u_{i,j-1}^{k+1} - 2u_{i,j}^{k+1} + u_{i,j+1}^{k+1}}{dy^2} + \frac{u_{i,j-1}^{k} - 2u_{i,j}^{k} + u_{i,j+1}^{k}}{dy^2}\right),\\
\frac{\partial u(x,y,t)}{\partial x} &= \frac{1}{2}\left(\frac{u_{i+1,j}^{k+1} - u_{i-1,j}^{k+1}}{2dx} + \frac{u_{i+1,j}^{k} - u_{i-1,j}^{k}}{2dx}\right),\\
\frac{\partial u(x,y,t)}{\partial y} &= \frac{1}{2}\left(\frac{u_{i,j+1}^{k+1} - u_{i,j-1}^{k+1}}{2dy} + \frac{u_{i,j+1}^{k} - u_{i,j-1}^{k}}{2dy}\right).
\end{aligned} \tag{A.1}$$

By replacing the derivatives by their discretizations in (9), the following discretized formulation is obtained:

$$\begin{aligned} &R_{ij}u^{k+1}_{i-1j} + Q_{ij}u^{k+1}_{ij-1} + P_{ij}u^{k+1}_{ij} + S_{ij}u^{k+1}_{ij+1} + O_{ij}u^{k+1}_{i+1j} = \\ &-R_{ij}u^{k}_{i-1j} - Q_{ij}u^{k}_{ij-1} + T_{ij}u^{k}_{ij} - S_{ij}u^{k}_{ij+1} - O_{ij}u^{k}_{i+1j} \end{aligned} \tag{A.2}$$

with

$$\begin{aligned} R_{ij} &= \frac{dt}{8dx^2}D^k_{x_{i+1j}} - \frac{dt}{8dx^2}D^k_{x_{i-1j}} - \frac{dt}{4dx}w^k_{x_{ij}} - \frac{dt}{2dx^2}D^k_{x_{ij}}, \\ Q_{ij} &= \frac{dt}{8dy^2}D^k_{y_{ij+1}} - \frac{dt}{8dy^2}D^k_{y_{ij-1}} - \frac{dt}{4dy}w^k_{y_{ij}} - \frac{dt}{2dy^2}D^k_{y_{ij}}, \\ P_{ij} &= 1 + \frac{dt}{dx^2}D^k_{x_{ij}} + \frac{dt}{dy^2}D^k_{y_{ij}}, \\ T_{ij} &= 1 - \frac{dt}{dx^2}D^k_{x_{ij}} - \frac{dt}{dy^2}D^k_{y_{ij}}, \\ S_{ij} &= -\frac{dt}{8dy^2}D^k_{y_{ij+1}} + \frac{dt}{8dy^2}D^k_{y_{ij-1}} + \frac{dt}{4dy}w^k_{y_{ij}} - \frac{dt}{2dy^2}D^k_{y_{ij}}, \\ O_{ij} &= -\frac{dt}{8dx^2}D^k_{x_{i+1j}} + \frac{dt}{8dx^2}D^k_{x_{i-1j}} + \frac{dt}{4dx}w^k_{x_{ij}} - \frac{dt}{2dx^2}D^k_{x_{ij}}. \end{aligned} \tag{A.3}$$

The applied boundary conditions are homogeneous Dirichlet boundary conditions, *i.e.* for any i and j, $u^k_{iN_y+1} = 0$, $u^k_{i0} = 0$, $u^k_{N_x+1j} = 0$ and $u^k_{0j} = 0$.

Therefore, the simulation of Eq. (9) amounts to solving the following linear equation

$$\boldsymbol{A}\boldsymbol{U}^{k+1} = \boldsymbol{B}\boldsymbol{U}^{k}, \tag{A.4}$$

where

$$\boldsymbol{U}^k = \left[u^k_{11} \dots u^k_{N_x1} \dots u^k_{1j} \dots u^k_{N_xj} \dots u^k_{1N_y} \dots u^k_{N_xN_y}\right]^\top, \tag{A.5}$$

$\boldsymbol{A}$ is an invertible matrix composed of the terms R_{ij}, Q_{ij}, P_{ij}, S_{ij} and O_{ij}, and $\boldsymbol{B}$ is an matrix composed of the terms R_{ij}, Q_{ij}, T_{ij}, S_{ij} and O_{ij} with, for instance, for $N_x = N_y = 3$, the following formulation

$$\boldsymbol{A} = \left[\begin{array}{ccc|ccc|ccc} P_{1,1} & S_{1,1} & 0 & O_{1,1} & 0 & 0 & 0 & 0 & 0 \\ Q_{2,1} & P_{2,1} & S_{2,1} & 0 & O_{2,1} & 0 & 0 & 0 & 0 \\ 0 & Q_{3,1} & P_{3,1} & 0 & 0 & O_{3,1} & 0 & 0 & 0 \\ \hline R_{1,2} & 0 & 0 & P_{1,2} & S_{1,2} & 0 & O_{1,2} & 0 & 0 \\ 0 & R_{2,2} & 0 & Q_{2,2} & P_{2,2} & S_{2,2} & 0 & O_{2,2} & 0 \\ 0 & 0 & R_{3,2} & 0 & Q_{3,2} & P_{3,2} & 0 & 0 & O_{3,2} \\ \hline 0 & 0 & 0 & R_{1,3} & 0 & 0 & P_{1,3} & S_{1,3} & 0 \\ 0 & 0 & 0 & 0 & R_{2,3} & 0 & Q_{2,3} & P_{2,3} & S_{2,3} \\ 0 & 0 & 0 & 0 & 0 & R_{3,3} & 0 & Q_{3,3} & P_{3,3} \end{array}\right],$$

$$\boldsymbol{B} = \left[\begin{array}{ccc|ccc|ccc} T_{1,1} & -S_{1,1} & 0 & -O_{1,1} & 0 & 0 & 0 & 0 & 0 \\ -Q_{2,1} & T_{2,1} & -S_{2,1} & 0 & -O_{2,1} & 0 & 0 & 0 & 0 \\ 0 & -Q_{3,1} & T_{3,1} & 0 & 0 & -O_{3,1} & 0 & 0 & 0 \\ \hline -R_{1,2} & 0 & 0 & T_{1,2} & -S_{1,2} & 0 & -O_{1,2} & 0 & 0 \\ 0 & -R_{2,2} & 0 & -Q_{2,2} & T_{2,2} & -S_{2,2} & 0 & -O_{2,2} & 0 \\ 0 & 0 & -R_{3,2} & 0 & -Q_{3,2} & T_{3,2} & 0 & 0 & -O_{3,2} \\ \hline 0 & 0 & 0 & -R_{1,3} & 0 & 0 & T_{1,3} & -S_{1,3} & 0 \\ 0 & 0 & 0 & 0 & R_{2,3} & 0 & -Q_{2,3} & T_{2,3} & -S_{2,3} \\ 0 & 0 & 0 & 0 & 0 & R_{3,3} & 0 & -Q_{3,3} & T_{3,3} \end{array}\right].$$

Solving the linear system (A.4) requires inverting the large matrix $\boldsymbol{A}$ of size $\mathbb{R}^{N_xN_y \times N_xN_y}$. It is preferable to use a fast and precise implementation such as the biconjugate gradients stabilized method (Barrett et al., 1994) which allows the consideration of a fine spatial mesh, *i.e.* small values for dx and dy.

Appendix B. Explicit Euler method for Eq. (10)

The implementation to simulate the ODE (10) is carried out with the explicit Euler method and the following discretized formulation

$$u^{k+1}_{ij} = u^k_{ij} + dt\left(\beta^k_{1_{ij}}u^k_{ij} - \beta^k_{2_{ij}}(u^k_{ij})^2\right). \tag{B.1}$$

References

Barrett, R., Berry, M., Chan, T.F., Demmel, J., Donato, J.M., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., der Vorst, H.V., 1994. Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, second ed. SIAM, Philadelphia, PA.

Bateman, A.W., Lewis, M.A., Gall, G., Manser, M.B., Clutton-Brock, T.H., 2015. Territoriality and home-range dynamics in meerkats, *suricata suricatta*: a mechanistic modelling approach. J. Anim. Ecol. 84, 260–271.

Belgacem, F., Cosner, C., 1995. The effects of dispersal along environmental gradients on the dynamics of populations in heterogeneous environment. Can. Appl. Math. Q. 3 (4), 379–397.

Björck, A., 1996. Numerical Methods for Least Squares Problems. SIAM.

Büttner, G., Kosztra, B., 2017. CLC2018 Technical Guidelines. Tech. Rep., European Environment Agency.

Cantrell, R.S., Cosner, C., 2001. Spatial heterogeneity and critical patch size: area effects via diffusion in closed environments. J. Theoret. Biol. 209, 161–171.

Cantrell, R.S., Cosner, C., 2003. Spatial Ecology Via Reaction–Diffusion Equations. Wiley, Chichester.

Crank, J., Nicolson, P., 1996. A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. Adv. Comput. Math. 6, 207–226.

DeAngelis, D.L., Yurek, S., 2016. Spatially explicit modeling in ecology: A review. Ecosystems.

Fick, S., Hijmans, R., 2017. Worldclim 2: new 1km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37 (12), 4302–4315.

Fisher, R.A., 1937. The wave of advance of advantageous genes. Ann. Eugen. 7 (4), 355–369, URL http://dx.doi.org/10.1111/j.1469-1809.1937.tb02153.x.

Guillet, T., Mouysset, L., 2022. Productive versus environmental objectives of agricultural policies dealing with climate change: a French case study. Front. Environ. Sci..

Havasi, A., Bartholy, J., Faragó, I., 2001. Splitting method and its application in air pollution modeling. Quart. J. Hung. Meteorol. Serv. 105 (1), 39–58.

Hefley, T., Hooten, M., Russell, R., Walsh, D., Powell, J., 2017. When mechanism matters: Bayesian forecasting using models of ecological diffusion. Ecol. Lett. 20, 640–650.

Holmes, E.E., 1993. Are diffusion models too simple? A comparison with telegraph models of invasion. Amer. Nat. 142–5, 779–795.

Holmes, E.E., Lewis, M.A., Banks, J.E., Veit, R.R., 1994. Partial differential equations in ecology: spatial interactions and population dynamics. Ecology 75 (1), 17–29.

Inger, R., Gregory, R., Duffy, J.P., Stott, I., Vorisek, P., Gaston, K.J., 2014. Common European birds are declining rapidly while less abundant species' numbers are rising. Ecol. Lett. URL http://dx.doi.org/10.1111/ele.12387.

IPBES, 2019. Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. IPBES Secretariat, Bonn, Germany, URL http://dx.doi.org/10.5281/zenodo.3831673.

IUCN, 2022. The IUCN red list of threatened species. Version 2022-2. URL https://www.iucnredlist.org. Accessed on 3 april 2023.

Jiguet, F., Devictor, V., Julliard, R., Couvet, D., 2012. French citizens monitoring ordinary birds provide tools for conservation and ecological sciences. Acta Oecol. 44, 58–66.

Keller, V., Herrando, S., Voříšek, P., Franch, M., Kipson, M., Milanesi, P., Martí, D., Anton, M., Klvaňová, A., Kalyakin, M., Bauer, H., Foppen, R., 2020. European Breeding Bird Atlas 2: Distribution, Abundance and Change. European Bird Census Council & Lynx Edicions, Barcelona.

Knudsen, M., 1994. A sensitivity approach for estimation of physical parameters. In: 10th IFAC Symposium on System Identification, Vol. 2. pp. 231–237.

Lawson, C.L., Hanson, R.J., 1974. Solving Least-Squares Problems. Prentice Hall, Upper Saddle River, NJ.

Levin, S.A., 1976. Population dynamic models in heterogeneous environments. Annu. Rev. Ecol. Syst. 7, 287–310.

Lewis, M.A., Kareiva, P., 1993. Allee dynamics and the spread of invading organisms. Theor. Popul. Biol. 43–2, 141–158, URL http://dx.doi.org/10.1006/tpbi.1993.1007.

Louvrier, J., Papaïx, J., Duchamp, C., Gimenez, O., 2020. A mechanistic-statistical species distribution model to explain and forecast wolf (*canis lupus*) colonization in South-Eastern France. Spat. Statist. 36, URL http://dx.doi.org/10.1016/j.spasta.2020.100428.

McCullagh, P., Nelder, J., 1983. Generalized Linear Models, second ed. London New York.

Melo-Merino, S.M., Reyes-Bonilla, H., Lira-Noriega, A., 2020. Ecological niche models and species distribution models in marine environments: A literature review and spatial analysis of evidence. Ecol. Model. 415, 108837.

Merow, C., LaFleur, N., Silander, J.A., Wilson, A.M., Rubega, M., 2011. Developing dynamic mechanistic species distribution models: Predicting bird-mediated spread of invasive plants across northeastern north America. Amer. Nat. 178 (1), 30–43.

Mouysset, L., 2014. Agricultural public policy: Green or sustainable? Ecol. Econom. 102, 15–23.

Mouysset, L., Miglianico, M., Makowski, D., Jiguet, F., Doyen, L., 2016. Selection of dynamic models for bird populations in farmlands. Environ. Model. Assess. 21 (3), 407–418.

Mouysset, L., Rais Assaa, C., Ay, J., Jiguet, F., Lorrilière, R., Doyen, L., 2019. Bioeconomic impacts of agroforestry policies in France. Land Use Policy 85, 239–248.

Newman, A., 1996a. Model Reduction Via the Karhunen–Loeve Expansion, Part I: An Exposition. Technical Report, T.R. 96-32, University of Maryland, ISR. MD, USA.

Newman, A., 1996b. Model Reduction Via the Karhunen–Loeve Expansion, Part II: Some Elementary Examples. Technical Report, T.R. 96-32, University of Maryland, ISR. MD, USA.

Nocedal, J., Wright, S.J., 2006. Numerical Optimization, second ed. Springer.

Okubo, A., 1986. Dynamical aspects of animal grouping: Swarms, schools, flocks, and herds. Adv. Biophys. 22, 1–94, URL http://dx.doi.org/10.1016/0065-227X(86)90003-1.

Okubo, A., Levin, S.A., 2001. Diffusion and Ecological Problems: Modern Perspectives, second ed. Springer.

Ouvrard, R., Mercère, G., Poinot, T., Jiguet, F., Mouysset, L., 2019. Dynamic models for bird population - A parameter-varying partial differential equation identification approach. Control Eng. Pract. 91, URL http://dx.doi.org/10.1016/j.conengprac.2019.07.009.

Ouvrard, R., Trigeassou, J.C., 2011. On embedded FIR filter models for identifying continuous-time and discrete-time transfer functions: the RPM approach. Internat. J. Control 84 (3), 616–632.

Polis, M., Goodson, R., Wozny, M., 1973. On parameter identification for distributed systems using Galerkin's criterion. Automatica 9, 53–64.

Rao, S., 2009. Engineering Optimization: Theory and Practice, fourth ed. Wiley.

Roques, L., Walker, E., Franck, P., Soubeyrand, S., Klein, E., 2016. Using genetic data to estimate diffusion rates in heterogeneous landscapes. J. Math. Biol. 73 (2), 397–422.

Rosenberg, K.V., Dokter, A.M., Blancher, P.J., Sauer, J.R., Smith, A.C., Smith, P.A., Stanton, J.C., Panjabi, A., Helft, L., Parr, M., Marra, P.P., 2019. Decline of the North American avifauna. Science 366, 120–124.

Shigesada, N., Kawasaki, K., Teramoto, E., 1979. Spatial segregation of interacting species. J. Theoret. Biol. 79–1, 83–99, URL http://dx.doi.org/10.1016/0022-5193(79)90258-3.

Skellam, J., 1951. Random dispersal in theoretical populations. Biometrika 38, 196–218, URL http://www.jstor.org/stable/2332328.

Stanton, R.L., Morrissey, C.A., Clark, R.G., 2018. Analysis of trends and agricultural drivers of farmland bird declines in North America: A review. Agric. Ecosys. Environ. 254, 244–254, URL http://dx.doi.org/10.1016/j.agee.2017.11.028.

Toth, R., 2010. Modeling and Identification of Linear Parameter-Varying Systems. In: Lecture Notes in Control and Information Sciences, Springer.

Trigeassou, J.C., 1987. Contribution à l'Extension de la Méthode des Moments en Automatique. Application à l'Identification des Systèmes Linéaires (Thèse d'etat). Université de Poitiers, France.

Wikle, C., 2003. Hierarchical bayesian models for predicting the spread of ecological processes. Ecology 84, 1382–1394.