Dynamic models for bird population – A parameter-varying partial differential equation identification approach

Régis Ouvrard[a,*], Guillaume Mercère[a], Thierry Poinot[a], Frédéric Jiguet[b], Lauriane Mouysset[c]

[a] *Université de Poitiers, Laboratoire d'Informatique et d'Automatique pour les Systèmes, 2 rue Pierre Brousse, TSA 41105, 86073 Poitiers cedex 9, France*
[b] *Sorbonne Universités, Muséum National d'Histoire Naturelle, Centre d'Écologie et des Sciences de la Conservation, 43 rue Buffon, 75005 Paris, France*
[c] *Université de Bordeaux, Groupe de Recherche en Économie Théorique et Appliquée, avenue Léon Duguit, 33608 Pessac cedex, France*

Abstract

In order to study the global decline of biodiversity, accurate models of animal population dynamics are required. In this paper, we propose a data-driven modelling procedure of a parameter-varying partial differential equation (PDE) model based on the Galerkin method and the proper orthogonal decomposition, respectively. The parameter-varying formulation allows us to introduce in models information such as temperature or landscape pattern in studies about the impact on biodiversity of the global warming or the agricultural intensification, respectively. To deal with the specific conditions of ecological applications, a specific attention is paid to the initialization as well as the implementation of an iterative identification procedure based on 3D partial moments and a Levenberg-Marquardt algorithm. The tools are tested on the data-driven modelling of the population of European Stonechat *Saxicola torquatus*, a European common bird species, by using data from the national French Breeding Bird Survey and the CORINE Land Cover. The perspectives are to model specific community of birds in order to evaluate the effect of global changes in population trends and to develop tools to help decision makers take into account biodiversity goals into public policies.

Keywords: Biodiversity, Ecology, Galerkin method, Parameter-varying, Partial differential equation, Population dynamics, Proper orthogonal decomposition, System identification

1. Introduction

The global decline of biodiversity is an undeniable fact today (Hallmann et al., 2017; IPBES, 2018a; WWF, 2018). Birds are known as the best taxon to measure this global decline (Inger et al., 2014; Stanton et al., 2018). In order to analyse this decline, as well as to predict future trends and influence public policies, it is essential to have accurate models of population dynamics. Moreover, as mentioned in assessment reports of IPBES (2018b,c,d), in Americas, Central Asia or Europe, the biodiversity loss is mainly due to the land degradation, and for instance, to an unsustainable agricultural. Thus, in order to study the correlation between agricultural intensification and biodiversity decline, different population dynamic models have been developed in literature such as in (Siriwardena et al., 1998; Chamberlain et al., 2000; Mouysset et al., 2016). All these models, based on curve fitting approaches or ordinary differential equations (ODEs), describe overall population trends in an area or a country. At best, these models characterize the temporal dynamics of each patch which composes the considered area. Nowadays we consider that the numerical tools allow us to model population dynamics with partial differential equations (PDEs) in order to characterize simultaneously temporal and spatial dynamics. To deal with this objective, automatic control community and more specifically system identification should investigate models based on PDEs.

The developments presented in this paper are motivated by the necessary improvements of the population dynamic models of birds in order to understand past dynamics related to global changes such as agricultural changes. The main perspective is to develop tools to help decision makers take into account biodiversity goals into public policies. Thus accurate models

*Corresponding author

in terms of time and space are needed. In (Mouysset, 2012; Mouysset et al., 2012, 2016), dynamic models for French bird populations have been estimated from abundances of a large set of common bird species and accurate agricultural patterns. These models have been used to evaluate different agricultural policy scenarios which promote targets of biodiversity up to 2050. This is an example of tool which allows to analyse the direct impacts of economic scenarios of the Common Agricultural Policy from European Union. The dynamic models estimated in (Mouysset, 2012; Mouysset et al., 2012, 2016) are based on the well-known logistic equation which is often used to model a population growth in ecology (Wu, 1996; Okubo and Levin, 2001). More specifically, these models are deduced from a nonlinear ODE, the Verhulst's model, which characterizes the temporal dynamic. In the present paper, we propose to consider PDE models to improve the modelling of the population by adding spatial variables. Notice that, even if no effective tool to the parameter estimation from data-set has been developed in the literature applied to ecology to our knowledge, the PDEs are considered to represent population dynamics for many years (Skellam, 1951; Levin, 1976; Shigesada et al., 1979; Okubo, 1986; Holmes et al., 1994; Okubo and Levin, 2001). Okubo and Levin (2001) surveys a wide variety of PDEs to model dynamic behaviours of populations. It is a way to describe spatial variations of a population: for instance, the diffusion equation can be seen as an extension of the random walk model of an individual, the advection equation can represent the effect of the non-homogeneities of environment producing an oriented movement of animals, the logistic growth can be used to introduce a temporal dynamic to characterize increase or decrease of a population.

In order to introduce the landscape pattern in our models to evaluate, for instance, the environmental footprint of agricultural intensification, we suggest considering linear parameter-varying (LPV) PDEs. Although LPV models have been widely studied for two decades (Toth, 2010), the LPV PDE model applications are rarely considered in literature. Let us cite (Belforte et al., 2005; Schorsch et al., 2013, 2014; Farah et al., 2016; Pham et al., 2018). In the sequel, we prefer to use the term parameter-varying since the considered PDE models can be linear or nonlinear.

System identification consists in estimating the parameters of a dynamical model by adjusting these parameters to minimize the discrepancy between measured data and system behaviour. Today, the system identification theory is widely developed for linear time-invariant systems (Ljung, 1999), *i.e.* for systems whose behaviour can be described by ODEs. The system identification for distributed-parameter systems, *i.e.* for systems whose behaviour can be described by PDEs, is more complicated. Measurements and spatial distribution of the sensors are critical points to achieve accurate results (Uciński, 2005). In this paper, the considered data-sets are given by the national French breeding bird survey described in (Jiguet et al., 2012). The sampling time is limited to one data by breeding season, *i.e.* one sample per year at each location, during thirteen years. The sampling limitations enforce to develop new identification methods to estimate PDE parameters. Therefore we propose an iterative approach based on time-space separation (Li and Qi, 2010), and more precisely the Galerkin method (Polis et al., 1973) and the proper orthogonal decomposition (Newman, 1996a,b). This iterative procedure requires an initialisation which is given by an estimator based on 3D partial moments, an extension of partial moments described in (Ouvrard and Trigeassou, 2011). Our tools are validated in simulation before being used on real data from the national French breeding bird survey.

The present paper is a preliminary work which aims at validating our approach by specialists of system identification. As mentioned in perspectives in the conclusion section, a future work will to consider specific communities of birds to evaluate the effect of global changes in population declines.

The paper is organized as follows. Section 2 describes the parameter-varying PDE model while the identification procedure is presented in Section 3 for a general formulation of the PDE. Section 4 is dedicated to the validation of the identification approach in simulation with a simplified PDE. An application to the population of European Stonechat *Saxicola torquatus* is shown in Section 5 with parameters varying with the landscape pattern in France. Section 6 concludes this paper.

2. Parameter-varying PDE model

The goal is to present our identification procedure for a general formulation of the parameter-varying PDE model, *i.e.* with diffusion, advection and logistic growth, and with parameter variations given by polynomial functions. Thus, we consider the following

parameter-varying PDE model

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

where x and y are spatial coordinates, t is time, $u(x,y,t)$ is a population, $H(x,y)$ is the scheduling variable and the varying parameters are given by

$$\begin{aligned} D_\bullet(H) &= \sum_{n=0}^{n_{D_\bullet}} D_{\bullet_n} H^n(x,y), \\ w_\bullet(H) &= \sum_{n=0}^{n_{w_\bullet}} w_{\bullet_n} H^n(x,y), \\ \beta_\bullet(H) &= \sum_{n=0}^{n_{\beta_\bullet}} \beta_{\bullet_n} H^n(x,y), \end{aligned} \tag{2}$$

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 diffusion coefficients D_x and D_y represent the Brownian random motion. The advection term relies on w_x and w_y, the drift velocities, which characterize oriented motions due to an external stimuli or an attractive area, for instance. β_1 and β_2 are parameters which yield to a logistic growth of the population. β_1 is the intrinsic growth rate and β_2/β_1 represents the carrying capacity. The scheduling variable $H(x,y)$ allows to take into account, for example, the land use to analyse agricultural changes to animal populations or again the temperature to study effects of global warming.

Notice that most of the PDE models described in (Fisher, 1937; Levin, 1976; Holmes et al., 1994; Wu, 1996; Okubo and Levin, 2001) are special cases of the parameter-varying PDE model (1).

The main issue is to estimate the parameters of polynomials (2) from a data-set.

3. POD-Galerkin method

The so-called POD-Galerkin method allows the time-space separation to provide an approximate solution of the parameter-varying PDE model (1). More specifically, we consider a Galerkin method based on function basis which is built with a proper orthogonal decomposition (POD).

3.1. Time-space separation basic idea

The Hilbert space theory (Debnath and Mikusinski, 2005) proves that any arbitrary signal f belonging to a Hilbert space like $\mathcal{L}_2(\mathcal{D})$ (the space of Lebesgue-measurable square-integrable functions on a domain $\mathcal{D}$, *i.e.* $\int_{\mathcal{D}} |f(x)|^2 dx < \infty$), with $\langle \bullet, \bullet \rangle$ its inner product, can be decomposed into a Fourier expansion of the form

$$f = \sum_{n=1}^{\infty} \langle f, \phi_n \rangle \phi_n, \tag{3}$$

once an orthonormal basis $\{\phi_n(x)\}_{n=1}^{\infty}$ of $\mathcal{L}_2(\mathcal{D})$ has been selected. More specifically, for signals like $u(x,y,t)$ belonging to the Hilbert space $\mathcal{L}_2(\mathcal{D})$, for $\mathcal{D} = \{x \in [0,X] \text{ and } y \in [0,Y]\}$, with the inner product

$$\langle f(x,y,t), g(x,y,t) \rangle = \int \int_{\mathcal{D}} f(x,y,t) g(x,y,t) dx dy, \tag{4}$$

we have

$$u(x,y,t) = \sum_{n=1}^{\infty} a_n(t) \phi_n(x,y), \tag{5}$$

where $a_n(t) = \langle u(x,y,t), \phi_n(x,y) \rangle$ are called the Fourier coefficients of this expansion. Eq. (5) is the basic idea to approximate the spatio-temporal signals involved in the population dynamics via the well-known time-space separation approach (Li and Qi, 2010).

For simplicity, in the sequel, the double integration $\int \int_{\mathcal{D}} \bullet dxdy$ introduced with the inner product will be denoted $\int \bullet$.

3.2. Proper orthogonal decomposition

In this paper, in order to benefit from available data-sets, we suggest resorting to the proper orthogonal decomposition to generate data-based orthonormal basis functions $\{\phi_n(x,y)\}_{n=1}^{\infty}$. As explained hereafter, such a technique:

- does not require specific prior knowledge on the way the system is excited (which is the case, *e.g.*, when basis functions like cosine and sine functions are introduced),
- helps the user to select the truncation order by looking at the magnitude of singular or eigenvalues generated by the POD.

Also called Karhunen-Loeve decomposition or principal component analysis (PCA) (Liang et al., 2002), the POD mainly consists in generating an empirical covariance data matrix, then determining the orthonormal

basis $\{\phi_n(x,y)\}_{n=1}^N$ by extracting the main vectors spanning its column space by using an eigenvalue or singular value decomposition (Newman, 1996a,b). The empirical sampled covariance matrix used by the POD method is more precisely defined by

$$\mathbf{K} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1M} \\ \vdots & \ddots & \vdots \\ \alpha_{M1} & \cdots & \alpha_{MM} \end{bmatrix}, \tag{6}$$

where the components $\alpha_{k\ell}$ are determined as follows[1]

$$\alpha_{k\ell} = \frac{1}{M} \int \delta(x,y,k.dt)\delta(x,y,\ell.dt), \tag{7}$$

and from the so-called M zero-mean snapshots

$$\delta(x,y,k.dt) = u(x,y,k.dt) - \bar{u}(x,y), \tag{8}$$

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

Once the matrix $\mathbf{K}$ is generated, the determination of the basis functions $\{\phi_n(x,y)\}_{n=1}^\infty$ is performed by computing the eigenvalue decomposition of the real symmetric matrix $\mathbf{K}$, feature which guarantees that the eigenvalues are all real and the eigenvectors orthogonal to each other. More specifically,

$$\mathbf{K} = \boldsymbol{Q}\boldsymbol{R}\boldsymbol{Q}^\top, \tag{9}$$

with $\boldsymbol{Q} \in \mathbb{R}^{M\times M}$ with $\boldsymbol{Q}\boldsymbol{Q}^\top = \mathbf{I}_M$, a $M \times M$ identity matrix, while $\boldsymbol{R} \in \mathbb{R}^{M\times M}$ contains the eigenvalues of $\mathbf{K}$ assumed to be ordered in a decreasing order. By using the direct link between the POD and the PCA, it is clear that the eigenvectors corresponding to the largest eigenvalues are the main modes of the covariance matrix $\mathbf{K}$. Said differently, by selecting the basis functions $\{\phi_n(x,y)\}_{n=1}^N$ from the first N eigenvectors of $\mathbf{K}$ with the largest eigenvalues (*i.e.*, the largest energy among all the available components), we can guarantee that

$$\sum_{n=1}^N a_n(t)\phi_n(x,y) \tag{10}$$

is the best approximation of $u(x,y,t)$ (in the least-squares sense) given $\{\phi_n(x,y)\}_{n=1}^N$ (Liang et al., 2002). Finally, the basis function set $\{\phi_n(x,y)\}_{n=1}^N$ is determined as a linear combination of the data snapshots

$$\phi_n(x,y) = \sum_{k=1}^M \boldsymbol{Q}_k^{(n)} \delta(x,y,k.dt), \tag{11}$$

where $\boldsymbol{Q}_k^{(n)}$ is the k-th line of the n-th eigenvector of $\mathbf{K}$. These computations are performed one time as an initial step of the iterative identification procedure.

[1] dt is the time sampling period.

3.3. Galerkin method

The Galerkin method operates a time-space separation as mentioned in Section 3.1 which transforms the PDE problem into a system of ODEs. It consists in finding an approximate solution of the PDE defined by

$$\hat{u}(x,y,t) = \sum_{n=1}^N a_n(t)\phi_n(x,y) + \bar{u}(x,y), \tag{12}$$

where $\{\phi_n(x,y)\}_{n=1}^N$ is the known function basis determined by POD, and $\{a_n(t)\}_{n=1}^N$ is the time coefficient set to be estimated.

The substitution of (12) in (1) leads to

$$\begin{aligned}
&\sum_{n=1}^N \dot{a}_n(t)\phi_n(x,y) = \\
&\sum_{n=1}^N a_n(t)\Big(D_x(H)\frac{\partial^2\phi_n(x,y)}{\partial x^2} + D_y(H)\frac{\partial^2\phi_n(x,y)}{\partial y^2} \\
&-w_x(H)\frac{\partial\phi_n(x,y)}{\partial x} - w_y(H)\frac{\partial\phi_n(x,y)}{\partial y} \\
&+\beta_1(H)\phi_n(x,y)\Big) \\
&-2\beta_2(H)\sum_{n=1}^N a_n(t)\phi_n(x,y)\bar{u}(x,y) \\
&-\beta_2(H)\sum_{n=1}^N a_n^2(t)\phi_n^2(x,y) \\
&-2\beta_2(H)\sum_{n=1}^{N-1}\sum_{m=n+1}^N a_n(t)a_m(t)\phi_n(x,y)\phi_m(x,y) \\
&+D_x(H)\frac{\partial^2\bar{u}(x,y)}{\partial x^2} + D_y(H)\frac{\partial^2\bar{u}(x,y)}{\partial y^2} \\
&-w_x(H)\frac{\partial\bar{u}(x,y))}{\partial x} - w_y(H)\frac{\partial\bar{u}(x,y)}{\partial y} \\
&+\beta_1(H)\bar{u}(x,y) - \beta_2(H)\bar{u}^2(x,y)\,.
\end{aligned} \tag{13}$$

By applying the inner product $\langle Eq.(13), \phi_\ell(x,y)\rangle$, the orthonormal property of $\{\phi_n(x,y)\}_{n=1}^N$ yields the follow-

ing system of ODEs

$$\begin{aligned}
\dot{a}_\ell(t) &= \sum_{n=1}^{N} a_n(t) \int \left\{ D_x(H)\frac{\partial^2\phi_n}{\partial x^2} + D_y(H)\frac{\partial^2\phi_n}{\partial y^2} \right. \\
&\left. -w_x(H)\frac{\partial\phi_n}{\partial x} - w_y(H)\frac{\partial\phi_n}{\partial y} + \beta_1(H)\phi_n \right\} \phi_\ell \\
&-2\sum_{n=1}^{N} a_n(t) \int \beta_2(H)\phi_n\phi_\ell\bar{u} \\
&-\sum_{n=1}^{N} a_n^2(t) \int \beta_2(H)\phi_n^2\phi_\ell \\
&-2\sum_{n=1}^{N-1}\sum_{m=n+1}^{N} a_n(t)a_m(t) \int \beta_2(H)\phi_n\phi_m\phi_\ell \\
&+\int \left\{ D_x(H)\frac{\partial^2\bar{u}}{\partial x^2} + D_y(H)\frac{\partial^2\bar{u}}{\partial y^2} - w_x(H)\frac{\partial\bar{u}}{\partial x} \right. \\
&\left. -w_y(H)\frac{\partial\bar{u}}{\partial y} + \beta_1(H)\bar{u} - \beta_2(H)\bar{u}^2 \right\} \phi_\ell
\end{aligned} \tag{14}$$

for $\ell = 1, \ldots, N$. For an easier implementation, this system of ODEs can be rewritten with a matrix formulation

$$\begin{aligned}
\dot{\mathbf{a}}(t) = (\boldsymbol{\Gamma}_1 + \beta_{1_0}\mathbf{I}_N - 2\boldsymbol{\Gamma}_2)\ \mathbf{a}(t) \\
- (\mathbf{I}_N \otimes \mathbf{a}^\top(t))\, \boldsymbol{\Lambda}\, \mathbf{a}(t) + \mathbf{b}_1,
\end{aligned} \tag{15}$$

with

$$\begin{aligned}
&\otimes, \text{ Kronecker product,} \\
&\mathbf{a}(t) = \begin{bmatrix} a_1(t) & \ldots & a_N(t) \end{bmatrix}^\top, \\
&\mathbf{I}_N, N \times N \text{ identity matrix,} \\
&\boldsymbol{\Gamma}_1 = \begin{bmatrix} \int \mathcal{F}_1(\phi_1)\phi_1 & \ldots & \int \mathcal{F}_1(\phi_N)\phi_1 \\ \vdots & \ddots & \vdots \\ \int \mathcal{F}_1(\phi_1)\phi_N & \ldots & \int \mathcal{F}_1(\phi_N)\phi_N \end{bmatrix}, \\
&\boldsymbol{\Gamma}_2 = \\
&\begin{bmatrix} \int \beta_2(H)\phi_1\phi_1\bar{u} & \ldots & \int \beta_2(H)\phi_N\phi_1\bar{u} \\ \vdots & \ddots & \vdots \\ \int \beta_2(H)\phi_1\phi_N\bar{u} & \ldots & \int \beta_2(H)\phi_N\phi_N\bar{u} \end{bmatrix}, \\
&\boldsymbol{\Lambda} = \int \beta_2(H)\boldsymbol{\Phi} \otimes (\boldsymbol{\Phi} \otimes \boldsymbol{\Phi}^\top), \\
&\boldsymbol{\Phi} = \begin{bmatrix} \phi_1(x,y) & \ldots & \phi_N(x,y) \end{bmatrix}^\top, \\
&\mathbf{b}_1 = \begin{bmatrix} \int \mathcal{F}_2(\bar{u})\phi_1 & \ldots & \int \mathcal{F}_2(\bar{u})\phi_N \end{bmatrix}^\top, \\
&\mathcal{F}_1(\bullet) = D_x(H)\tfrac{\partial^2\bullet}{\partial x^2} + D_y(H)\tfrac{\partial^2\bullet}{\partial y^2} - w_x(H)\tfrac{\partial\bullet}{\partial x} \\
&\qquad -w_y(H)\tfrac{\partial\bullet}{\partial y} + \textstyle\sum_{\upsilon=1}^{n_{\beta_1}} \beta_{1_\upsilon} H^\upsilon(x,y)\bullet, \\
&\mathcal{F}_2(\bullet) = D_x(H)\tfrac{\partial^2\bullet}{\partial x^2} + D_y(H)\tfrac{\partial^2\bullet}{\partial y^2} - w_x(H)\tfrac{\partial\bullet}{\partial x} \\
&\qquad -w_y(H)\tfrac{\partial\bullet}{\partial y} + \beta_1(H)\bullet - \beta_2(H)\bullet^2 .
\end{aligned} \tag{16}$$

The initial conditions of the ODEs system (15) is obtained by making the initial signal error orthogonal to the basis functions $\phi_\ell(x, y)$, *i.e.* :

$$\langle u(x,y,0) - \hat{u}(x,y,0), \phi_\ell(x,y)\rangle_{\ell=1,\ldots,N} = 0 \tag{17}$$

or

$$a_\ell(0) = \int \left(u(x,y,0) - \bar{u}(x,y)\right)\phi_\ell(x,y) \tag{18}$$

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

Therefore, given parameters, the simulation of the system of ODEs (15) from the initial conditions (18) provides the approximate solution (12). The latter can be used in the following iterative identification procedure.

4. Iterative identification procedure

4.1. Parameter estimation

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

$$\begin{aligned}
\boldsymbol{\theta} = \Big[& D_{x_0} \ldots D_{x_{n_{D_x}}}\ D_{y_0} \ldots D_{y_{n_{D_y}}}\ w_{x_0} \ldots w_{x_{n_{w_x}}} \\
& w_{y_0} \ldots w_{y_{n_{w_y}}}\ \beta_{1_0} \ldots \beta_{1_{n_{\beta_1}}}\ \beta_{2_0} \ldots \beta_{2_{n_{\beta_2}}} \Big]^\top
\end{aligned} \tag{19}$$

is estimated from a measured population[2] $\{u(i,j,k)\}_{i=0,\ldots,N_x;\ j=0,\ldots,N_y;\ k=0,\ldots,N_t}$ with an iterative procedure by minimizing the quadratic criterion

$$J(\boldsymbol{\theta}_{it}) = \sum_{k=0}^{N_t}\sum_{j=0}^{N_y}\sum_{i=0}^{N_x} \left(u(i,j,k) - \hat{u}(i,j,k)\right)^2, \tag{20}$$

where $\hat{u}(i,j,k)$ is the approximate solution (12) of the PDE model (1), and $\boldsymbol{\theta}_{it}$ is the parameter vector at the it-th iteration. This procedure repeats iteratively the following two steps from an initial vector $\boldsymbol{\theta}_0$ and until the convergence:

- to approximate a solution $\hat{u}(i,j,k)$ by the Galerkin method, Eq. (12),
- to calculate a new parameter vector $\boldsymbol{\theta}_{it+1}$ with a Levenberg-Marquardt algorithm which minimizes the criterion (20).

4.2. Levenberg-Marquardt algorithm

Some precautions are required in the implementation of the optimization algorithm. The considered PDE models (1) do not have exogenous input, and the unique excitation is the initial condition $u(x, y, 0)$. Moreover, the ecological data-set often has a small number of time samples. Therefore, we propose a Levenberg-Marquardt algorithm based on simulation of the true sensitivity functions with initial conditions deduced from (18). This reduces the approximations.

[2]For simplicity, $v(i,j,k)$ denotes sampled data $v(i.dx, j.dy, k.dt)$ with corresponding sampling distances dx, dy and dt, respectively.

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

$$\boldsymbol{\theta}_{it+1} = \boldsymbol{\theta}_{it} - \left(\mathbf{J}''_{\theta\theta} + \mu \mathbf{I}_{N_\theta}\right)^{-1} \mathbf{J}'_{\theta}, \tag{21}$$

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

$$\begin{aligned}
\mathbf{J}'_{\theta} &= -2 \sum_{k=0}^{N_t} \sum_{j=0}^{N_y} \sum_{i=0}^{N_x} \left(u(i,j,k) - \hat{u}(i,j,k)\right) \boldsymbol{\Xi}(k) \boldsymbol{\Phi}(i,j), \\
\mathbf{J}''_{\theta\theta} &= 2 \sum_{k=0}^{N_t} \sum_{j=0}^{N_y} \sum_{i=0}^{N_x} \boldsymbol{\Xi}(k) \boldsymbol{\Phi}(i,j) \boldsymbol{\Phi}^\top(i,j) \boldsymbol{\Xi}^\top(k),
\end{aligned} \tag{22}$$

where $\hat{u}(i,j,k)$ is the approximate solution (12), the $\boldsymbol{\Phi}(x,y)$ is defined in (16), and $\boldsymbol{\Xi}(t)$ is the time sensitivity function matrix given by

$$\begin{aligned}
\boldsymbol{\Xi}(t) = \Big[& \boldsymbol{\sigma}^{\mathbf{a}}_{D_{x_0}}(t) \quad \cdots \quad \boldsymbol{\sigma}^{\mathbf{a}}_{D_{x_{n_{D_x}}}}(t) \quad \boldsymbol{\sigma}^{\mathbf{a}}_{D_{y_0}}(t) \\
& \cdots \quad \boldsymbol{\sigma}^{\mathbf{a}}_{D_{y_{n_{D_y}}}}(t) \quad \boldsymbol{\sigma}^{\mathbf{a}}_{w_{x_0}}(t) \quad \cdots \quad \boldsymbol{\sigma}^{\mathbf{a}}_{w_{x_{n_{w_x}}}}(t) \\
& \boldsymbol{\sigma}^{\mathbf{a}}_{w_{y_0}}(t) \quad \cdots \quad \boldsymbol{\sigma}^{\mathbf{a}}_{w_{y_{n_{w_y}}}}(t) \quad \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{1_0}}(t) \quad \cdots \\
& \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{1_{n_{\beta_1}}}}(t) \quad \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{2_0}}(t) \quad \cdots \quad \boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{2_{n_{\beta_2}}}}(t) \Big]^\top, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{D_{\bullet_n}}(t) = & \left[\sigma^{a_1}_{D_{\bullet_n}}(t) \quad \ldots \quad \sigma^{a_N}_{D_{\bullet_n}}(t) \right]^\top, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{w_{\bullet_n}}(t) = & \left[\sigma^{a_1}_{w_{\bullet_n}}(t) \quad \ldots \quad \sigma^{a_N}_{w_{\bullet_n}}(t) \right]^\top, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{\bullet_n}}(t) = & \left[\sigma^{a_1}_{\beta_{\bullet_n}}(t) \quad \ldots \quad \sigma^{a_N}_{\beta_{\bullet_n}}(t) \right]^\top, \\
\sigma^{a_m}_{\circ}(t) = & \frac{\partial a_m(t)}{\partial \circ}.
\end{aligned} \tag{23}$$

The sensitivity functions are simulated by considering the following matrix formulations deduced from (15)

$$\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{24}$$

for $n = 0, \cdots, n_{D_\bullet}, n_{w_\bullet}$ 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 & \ldots & \int \mathcal{F}_{D_{\bullet_n}}(\phi_N)\phi_1 \\ \vdots & \ddots & \vdots \\ \int \mathcal{F}_{D_{\bullet_n}}(\phi_1)\phi_N & \ldots & \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 & \ldots & \int \mathcal{F}_{w_{\bullet_n}}(\phi_N)\phi_1 \\ \vdots & \ddots & \vdots \\ \int \mathcal{F}_{w_{\bullet_n}}(\phi_1)\phi_N & \ldots & \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 & \ldots & \int H^n\phi_N\phi_1 \\ \vdots & \ddots & \vdots \\ \int H^n\phi_1\phi_N & \ldots & \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} & \ldots & \int H^n\phi_N\phi_1\bar{u} \\ \vdots & \ddots & \vdots \\ \int H^n\phi_1\phi_N\bar{u} & \ldots & \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}} = & \left[\int \mathcal{F}_{D_{\bullet_n}}(\bar{u})\phi_1 \quad \ldots \quad \int \mathcal{F}_{D_{\bullet_n}}(\bar{u})\phi_N \right]^\top, \\
\mathbf{b}_{w_{\bullet_n}} = & \left[\int \mathcal{F}_{w_{\bullet_n}}(\bar{u})\phi_1 \quad \ldots \quad \int \mathcal{F}_{w_{\bullet_n}}(\bar{u})\phi_N \right]^\top, \\
\mathbf{b}_{\beta_{1_n}} = & \left[\int H^n\bar{u}\phi_1 \quad \ldots \quad \int H^n\bar{u}\phi_N \right]^\top, \\
\mathbf{b}_{\beta_{2_n}} = & \left[\int H^n\bar{u}^2\phi_1 \quad \ldots \quad \int H^n\bar{u}^2\phi_N \right]^\top, \\
\mathcal{F}_{D_{\bullet_n}}(\circ) = & H^n(x,y)\frac{\partial^2 \circ}{\partial \bullet^2}, \; \mathcal{F}_{w_{\bullet_n}}(\circ) = H^n(x,y)\frac{\partial \circ}{\partial \bullet},
\end{aligned} \tag{25}$$

and with the initial conditions

$$\begin{aligned}
\boldsymbol{\sigma}^{\mathbf{a}}_{D_{\bullet_n}}(0) &= \boldsymbol{\Gamma}_{D_{\bullet_n}}\, \mathbf{a}(0) + \mathbf{b}_{D_{\bullet_n}}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{w_{\bullet_n}}(0) &= \boldsymbol{\Gamma}_{w_{\bullet_n}}\, \mathbf{a}(0) + \mathbf{b}_{w_{\bullet_n}}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{1_n}}(0) &= \boldsymbol{\Gamma}_{\beta_{1_n}} \mathbf{a}(0) + \mathbf{b}_{\beta_{1_n}}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{2_n}}(0) &= -2\boldsymbol{\Gamma}_2\, \mathbf{a}(0) \\
& \quad - \left(\mathbf{I}_N \otimes \mathbf{a}^\top(0)\right) \boldsymbol{\Lambda}_{\beta_{2_n}} \mathbf{a}(0) - \mathbf{b}_{\beta_{2_n}}.
\end{aligned} \tag{26}$$

4.3. Initial estimate based on 3D partial moments

The aim of this section is to obtain a initial estimate $\boldsymbol{\theta}_0$ of the parameters for the initialisation of the iterative identification procedure given in (21). For the PDE model, the choice of the initialisation is particularly crucial to tend toward the global optimum of the criterion. It can be done from prior knowledge. Here, we propose another way based on partial moments developed in (Trigeassou, 1987; Ouvrard and Trigeassou, 2011) for systems represented by ODEs and extended to the PDE models thereafter.

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

$$\mathcal{M}^{f}_{\ell,m,n}(X,Y,T)=\int_0^X\int_0^Y\int_0^T x^\ell y^m t^n f(x,y,t)\partial t\partial y\partial x, \tag{27}$$

with ℓ, m and $n\in\{\mathbb{N},\ -\}$ [3].

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

$$\int_0^X\int_0^{x_1} x_2^2\int_0^Y\int_0^{y_1} y_2^2\int_0^T t\quad(1)\quad \partial t\partial y_2\partial y_1\partial x_2\partial x_1 \tag{28}$$

in such a way that all partial derivatives disappear. Thus we obtain a formulation with only partial moments and parameters. Indeed by considering the integration by parts

$$\int_0^{\tau_0}\tau_1\frac{\partial f}{\partial\tau}d\tau_1=\tau_0 f(\tau_0)-\int_0^{\tau_0} f(\tau_1)d\tau_1 \tag{29}$$

or again

$$\begin{aligned}\int_0^{\tau_0}\int_0^{\tau_1}\tau_2^2\frac{\partial^2 f}{\partial\tau^2}d\tau_2 d\tau_1&=\tau_0^2 f(\tau_0)\\ &+2\tau_0\int_0^{\tau_0} f(\tau_1)d\tau_1-6\int_0^{\tau_0}\tau_1 f(\tau_1)d\tau_1,\end{aligned} \tag{30}$$

and the Cauchy formula of repeated integrations

$$\begin{aligned}\int_0^t\int_0^{\tau_1}\cdots\int_0^{\tau_{n-1}} f(\tau_n)d\tau_n\cdots d\tau_2 d\tau_1&=\\ \frac{1}{(n-1)!}\int_0^t(t-\tau_1)^{n-1}f(\tau_1)d\tau_1,&\end{aligned} \tag{31}$$

Eq. (28) can be reformulated with 3D partial moments as follows

$$\begin{aligned}\mathcal{U}_1(X,Y,T)&=D_{x_0}\mathcal{U}_2(X,Y,T)+D_{y_0}\mathcal{U}_3(X,Y,T)\\ &-w_{x_0}\,\mathcal{U}_4(X,Y,T)-w_{y_0}\,\mathcal{U}_5(X,Y,T)\\ &+\beta_{1_0}\,\mathcal{U}_6(X,Y,T)+\beta_{2_0}\,\mathcal{U}_7(X,Y,T),\end{aligned} \tag{32}$$

where

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

[3] If ℓ, m or n is the symbol $-$, then there is no integration with respect of the corresponding variable.

Hence, (32) can be rewritten in a linear regression form

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

with

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

and a least-squares estimate of parameters is given by

$$\begin{aligned}\boldsymbol{\theta}_{LS}=&\left(\sum_{k=0}^{N_t}\sum_{j=0}^{N_y}\sum_{i=0}^{N_x}\boldsymbol{\varphi}(i,j,k)\boldsymbol{\varphi}^\top(i,j,k)\right)^{-1}\\ &\cdot\sum_{k=0}^{N_t}\sum_{j=0}^{N_y}\sum_{i=0}^{N_x}\boldsymbol{\varphi}(i,j,k)\mathcal{U}_1(i,j,k).\end{aligned} \tag{35}$$

In that way, with $\boldsymbol{\theta}_{LS}$ we have the initial parameters of a PDE model with $n_{D_\bullet}=0$, $n_{w_\bullet}=0$ and $n_{\beta_\bullet}=0$. To complete the initial vector (19), we consider small values for the missing parameters corresponding to $n_{D_\bullet}>0$, $n_{w_\bullet}>0$ and $n_{\beta_\bullet}>0$.

5. Simulation

The goal of this simulated test is to validate our tools for applications in ecology, more specifically for the study of bird population evolution based on data collected with the monitoring program presented in Jiguet et al. (2012). Section 6 presents such an application with real data.

5.1. Simulated system

Let us consider a population described by a parameter-varying formulation of the Fisher PDE (Fisher, 1937) given by

$$\begin{aligned}\frac{\partial u(x,y,t)}{\partial t}=&\\ &D(H)\left(\frac{\partial^2 u(x,y,t)}{\partial x^2}+\frac{\partial^2 u(x,y,t)}{\partial y^2}\right)\\ &+\beta_1(H)u(x,y,t)-\beta_2(H)u^2(x,y,t)\ ,\end{aligned} \tag{36}$$

with

$$\begin{aligned}D(H)&=D_0+D_1H(x,y),\\ \beta_\bullet(H)&=\beta_{\bullet_0}+\beta_{\bullet_1}H(x,y).\end{aligned} \tag{37}$$

For the identification procedure, we keep all developments of the previous section without parameters $w_\bullet$ and with $n_{D_\bullet}=1$ and $n_{\beta_\bullet}=1$.

5.2. Simulation protocol

An ecological application, like the one introduced in Section 6, presents some specificities: a small number of time samples, a spatial domain given by a map and a kriging preprocessing of data. We propose to test and to validate our tools in these specific conditions with 21 time samples, the France map and we consider noise-free data since the kriging preprocessing smooths the noise.

The data-set is provided by a simulation of (36) using a finite element method. A triangular mesh with a maximum mesh edge length of 20 is generated for the spatial domain. The geometry of the spatial domain is presented in Figure 1. To limit the effect of the boundary conditions, the simulation domain is a wide square (blue square in Figure 1) and, to build the data-set, we consider only the data that belongs to the France map. In other words, all grid points outside the France map are simulated with the finite element method, but they are ignored in the final data-set. The rectangles specify the value of the scheduling variable: $H(x, y) = 0$ for the coordinates inside the bottom rectangle, and $H(x, y) = 1$ inside the top one.

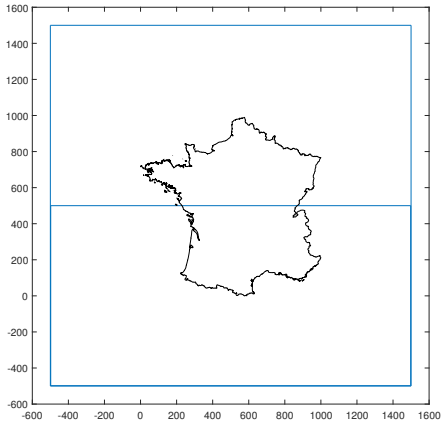

Figure 1: Simulation domain (blue square), data domain (France map) and scheduling variable (blue rectangles).

Let us select the true parameters $D_0 = 2000$, $D_1 = -200$, $\beta_{1_0} = -0.04$, $\beta_{1_1} = 0.006$, $\beta_{2_0} = -0.01$ and $\beta_{2_1} = -0.02$, and the sampling distances $dt = 1$, $dx = 2$ and $dy = 2$, respectively. The units are omitted in this simulation application. Neumann boundary conditions are specified and the initial condition is defined as follows

$$u(x, y, 0) = \sum_{n=1}^{10} e^{-\left((x-x_n)^2+(y-y_n)^2\right)/2\sigma_n^2}, \qquad (38)$$

i.e. a sum of Gaussian centered on $(x_n, y_n) \in [100, 900]$ with variance $\sigma_n^2 \in [3000, 6000]$.

By analogy with the ecological application of Section 6, the scheduling variable $H(x, y)$ has been chosen to represent the agricultural intensification: the bottom rectangle is an area unfavourable to biodiversity and the top rectangle is favourable. Therefore, the diffusion linked to the parameter D is larger in the bottom rectangle than the top one, the decreasing rate (β_1) is faster and the carrying capacity (β_2/β_1) is smaller. In other words, the decreasing trend of the bird population is larger in the bottom rectangle than in the top rectangle.

The data-set size is $N_x = 500$, $N_y = 500$ and $N_t = 20$, respectively. Figure 2 presents the population $u(x, y, t)$ for some time samples.

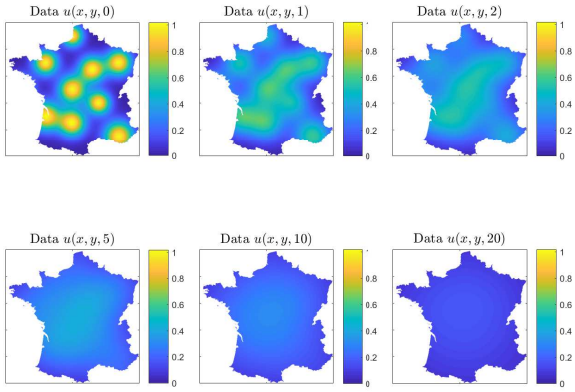

Figure 2: Data $u(x, y, t)$ for $t = 0, 1, 2, 5, 10$ and 20.

Notice that, hereafter, we will consider the estimation data-set presented in Figure 2 and a validation data-set simulated in the same way but with different values of (x_n, y_n) and σ_n^2 for the initial condition (38).

5.3. Selection of the design parameters

The iterative identification procedure introduced in Section 4 requires to choose two design parameters: M, the number of snapshots in Eq. (6) and N, the number of basis functions $\phi_n(x, y)$ in Eq. (12). For M, in our context of a small number of time samples, it is preferable to consider a maximum of snapshots, *i.e.* $M = 21$ for the above simulation protocol. The design parameter N must be large enough to consider the largest eigenvalues λ_n and to represent correctly the system behaviour, and a limited value gives a parsimonious model. The POD yields the normalized eigenvalues λ_n shown in Figure 3. Let us carry out the Levenberg-Marquardt algorithm for

increasing values of N and compare results in terms of fitting defined by

$$FIT = 100\left(1 - \frac{\|u - \hat{u}\|_2}{\|u - \text{mean}(u)\|_2}\right). \quad (39)$$

Table 1 presents the fittings obtained for the estimation and validation data-sets. We can consider that a number $N = 7$ gives a sufficiently good fitting and a parsimonious model. In other words, the eigenvalues from $\lambda_1 = 0.866$ to $\lambda_7 = 6.6e-8$ are sufficient to represent the system. The first seven orthonormal basis functions $\phi_n(x, y)$ are presented in Figure 4.

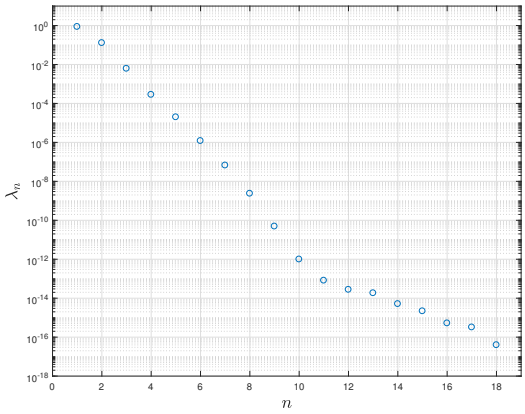

Figure 3: Normalized eigenvalue λ_n in a decreasing order.

Table 1: *FIT* (%) *vs N*

N	Results with estimation data-set	Results with validation data-set
3	61.80	54.83
4	83.31	80.98
5	94.13	93.41
6	99.31	99.24
7	99.44	99.43

5.4. Comparison with a constant-parameter PDE model

Here, the goal is to evaluate the benefit of the parameter-varying PDE model compared to a constant-parameter PDE model. By considering data-sets presented in Section 5.2, let us estimate two PDE models defined by (36): one with parameters varying, *i.e.* with $n_D = 1$, $n_{\beta_1} = 1$, $n_{\beta_2} = 1$, and one with constant parameters, *i.e.* with $n_D = 0$, $n_{\beta_1} = 0$, $n_{\beta_2} = 0$, respectively. The design parameters are $M = 21$ and $N = 7$.

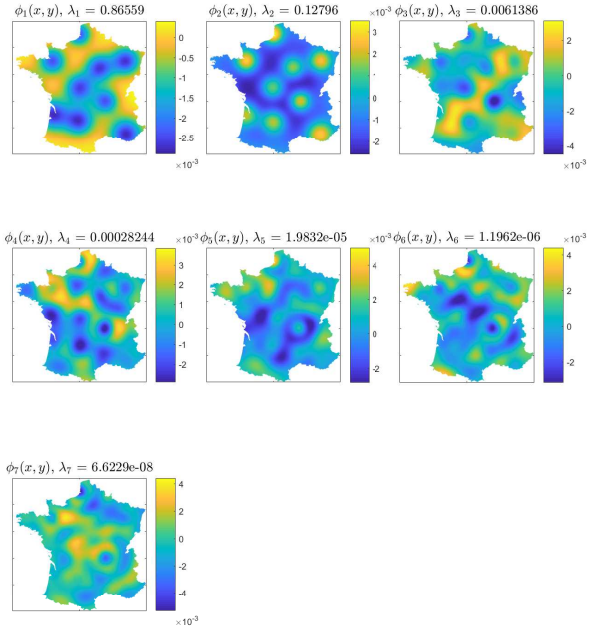

Figure 4: The seven considered basis functions $\phi_n(x, y)$ with the corresponding eigenvalue λ_n.

Table 2 shows the obtained fittings and Figure 5 presents the spatial error at some time samples. It is clear that the parameter-varying PDE model gives a better representation of the system than the constant-parameter PDE model. The spatial error of the constant-parameter PDE model in Figure 5 presents large amplitudes between top and bottom rectangles.

Table 2: FIT (%) for PDE models with parameters varying and constant parameters

PDE model with	Estimation data	Validation data
parameters varying	99.44	99.43
constant parameters	93.53	92.79

5.5. Numerical implementation of integrals

Here, the goal is to evaluate the impact of the numerical implementation of integrals in the convergence of the Levenberg-Marquardt algorithm, and more specifically in the context of a small number of time samples. In other words, the numerical implementation of integrals in (15-16) and (24-25) causes some approximations. How do these approximations affect the convergence?

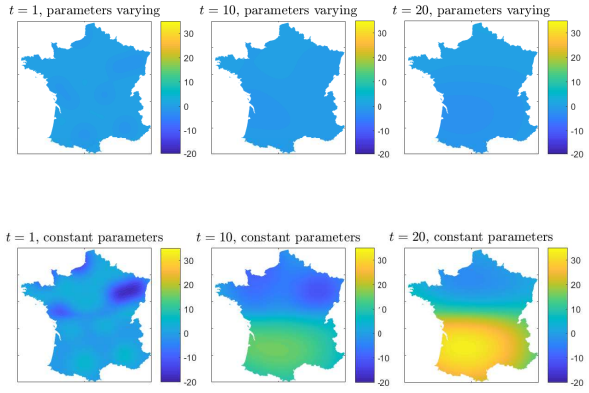

Figure 5: Spatial error (%) defined by $(u(x,y,t) - \hat{u}(x,y,t))/mean_{x,y}(u(x,y,t))$ for any x, any y and $t = 1, 10, 20$ with the validation data-set.

In order to observe the effect of the numerical implementation of integrals, we propose a new data-set. Let us consider the so-called POD-Galerkin data-set where u is generated with the POD-Galerkin solution (12) with the true parameters and the basis functions set $\{\phi_n(x,y)\}_{n=1}^N$ obtained from the data-set simulated with the finite element method of Section 5.2. Thus, the POD-Galerkin data-set must lead to a zero error on the estimated parameters, *i.e.* the Levenberg-Marquardt algorithm converges towards the exact true parameters.

All integrals are implemented with the trapezoidal rule. Moreover the design parameter N is 7, and we consider two different sampling times in simulation of the ODE system (15) and two different snapshot numbers M in Eq. (6): $dt = 1$ and $M = 21$, for the first case, and $dt = 0.5$ and $M = 41$, for the second one.

The relative errors between true and estimated parameters are given in Table 3. We can conclude that the estimation error tends towards zero for a decreasing sampling time. For applications with a small number of time samples such as $dt = 1$ and $M = 21$, the approximation due to the numerical implementation of integrals is acceptable, less than 1 %, even if it is not negligible.

5.6. Global implementation validation

Let us study the convergence of the Levenberg-Marquardt algorithm to evaluate the impact of the global implementation. The estimation data-set presented in Section 5.2 is simulated with a finite element method. This method is an accurate way to generate a population u. The Galerkin method (12) which gives $\hat{u}$ is less accurate, it is an approximate solution. Therefore, the quadratic criterion (20) cannot tend towards zero, there exists a residual error due to the approximated population $\hat{u}$. How does this residual error affect the estimation?

Table 3: Relative errors (%) after convergence of the Levenberg-Marquardt algorithm with the POD-Galerkin data-set

	$dt = 1$ and $M = 21$	$dt = 0.5$ and $M = 41$
D_0	0.04	0.001
D_1	0.68	0.018
β_{1_0}	0.02	0.001
β_{1_1}	0.26	0.015
β_{2_0}	0.74	0.029
β_{2_1}	0.72	0.032

In order to test the convergence, let us initialize the Levenberg-Marquardt algorithm with a parameter vector $\boldsymbol{\theta}_0$ at 80% of the true values. The convergence is shown in Figure 6 for the following data-sets :

- the estimation data-set simulated with the finite element method of Section 5.2 called finite-element data-set,
- the POD-Galerkin data-set presented in Section 5.5.

In both cases, after convergence, the estimated parameters are closed to true values.

Table 4 shows the relative errors obtained for the finite-element data-set with $dt = 1$ and $M = 21$. In this case, the estimation error is due to both the numerical implementation of integrals and the residual error between the simulations of u and $\hat{u}$. This explains why the relative error is larger (see Table 3). Nevertheless, this error, between 0.32 % and 8.88 %, is acceptable taking into account the difficult context of the PDE model identification with 21 time samples.

Table 4: Relative errors (%) with the finite-element data-set

	$dt = 1$ and $M = 21$
D_0	0.93
D_1	2.69
β_{1_0}	0.32
β_{1_1}	2.29
β_{2_0}	8.88
β_{2_1}	1.98

5.7. Fitting in terms of trend

As mentioned in the introduction, in ecological applications, the commonly used tools are based on ODE

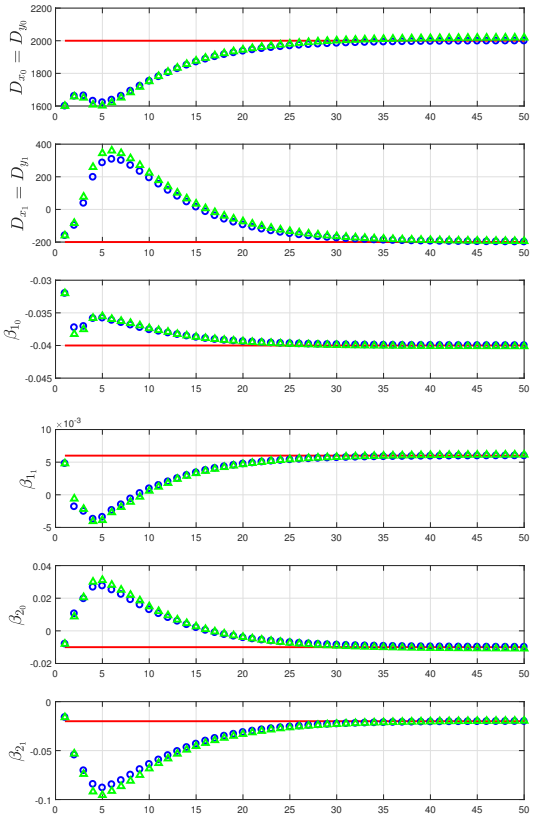

Figure 6: Convergence of the Levenberg-Marquardt algorithm: true parameters (red line), estimated parameters with the POD-Galerkin data-set (blue circles) and estimated parameters with the finite-element data-set (green triangles).

models which describe the global trend of the population in time domain (Wu, 1996; Okubo and Levin, 2001; Mouysset, 2012; Mouysset et al., 2012, 2016). Even if our PDE models represent the population in both spatial domain and time domain, let us see the fitting in terms of global trend. The normalized temporal trend is defined by

$$f_t(t) = \sum_{j=0}^{N_y}\sum_{i=0}^{N_x} f(i.dx, j.dy, t) / \sum_{j=0}^{N_y}\sum_{i=0}^{N_x} f(i.dx, j.dy, 0). \tag{40}$$

Figure 7 presents the true temporal trend of the population with the one obtained by simulating the parameter-varying PDE model estimated with the Levenberg-Marquardt algorithm. Notice that the estimated PDE model represents almost perfectly the global trend. Once again to evaluate the benefit of the parameters varying, Figure 8 presents the same curves but with the constant-parameter PDE model. In this case, the error is more significant which justifies the use of a parameter-varying formulation.

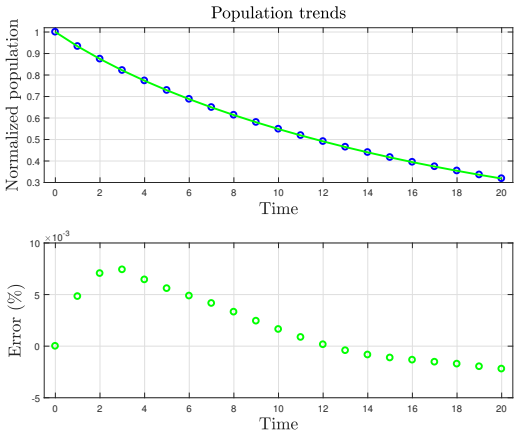

Figure 7: Trends calculated with the true data (blue circles) and the estimated parameter-varying PDE model simulation (green).

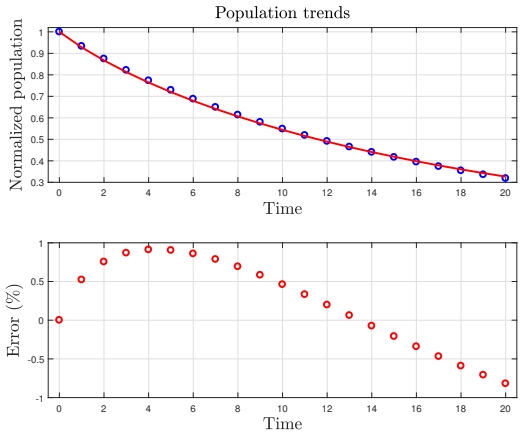

Figure 8: Trends calculated with the true data (blue circles) and the estimated constant-parameter PDE model simulation (red).

6. European Stonechat *Saxicola torquatus* population

6.1. Bird data

The considered bird data is given by the national French Breeding Bird Survey (Jiguet et al., 2012). This is a volunteer-based standardized monitoring based on birdwatchers which count spring birds annually during 5 minutes exactly on 10 fix points within a randomly selected square. Figure 9 shows the spatial distribution

of the monitored squares. Therefore, this database provides the informations related to the bird abundances across the whole country. Abundance values for each species are available for the period 2002-2014. In this paper, we consider only one specie, the European Stonechat *Saxicola torquatus* drawn at the bottom right of Figure 10. We perform a spatial interpolation of these abundance data to obtain relative abundance values for each possible 2×2 km square in the country (*i.e.* about 136 000 squares) using kriging models based on spatial autocorrelation and the exponential function (Doxa et al., 2010; Mouysset et al., 2012). The considered kriged data-set for 2002-2014 is presented in Figure 10.

Figure 9: Spatial distribution of the squares monitored in continental France by the Breeding Bird Survey scheme.

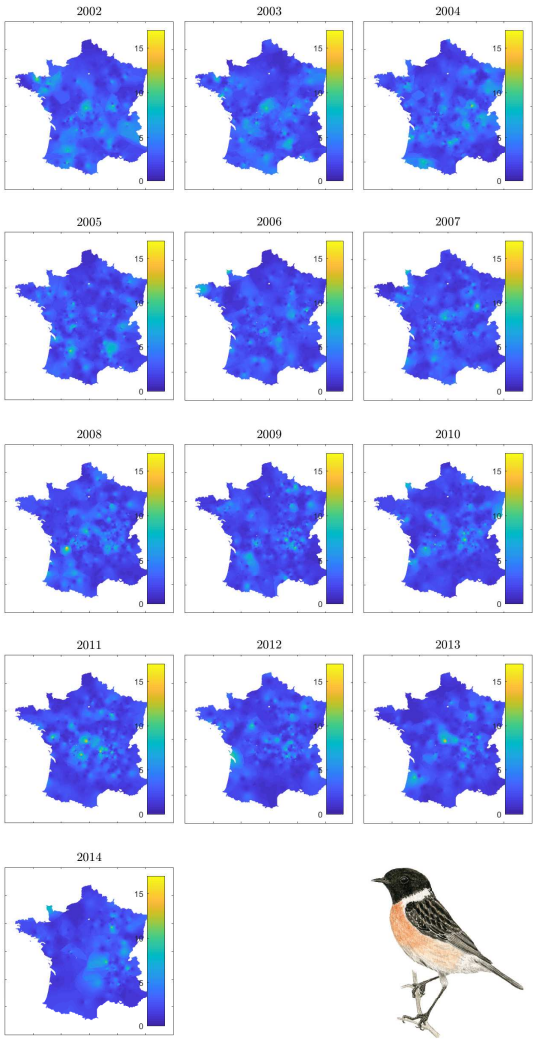

Figure 10: European Stonechat *Saxicola torquatus* ©Katia LIPOVOI-LPO (bottom rigth) and kriged abundance of this specie in France for 2002-2014.

6.2. *Land use index*

In order to introduce the landscape pattern given by the CORINE Land Cover[4] in the parameter-varying PDE model, the scheduling variable $H(x, y)$ stands for a land use diversity index inspired by the ecological Shannon index (Hill, 1973) defined on a disc C of centre (x, y) by

$$H(x, y) = -\sum_{i=1}^{R} p_i \ln p_i, \tag{41}$$

with p_i, the proportion of the i-th land use and R, the number of land uses on C (radius 1.5 km) deduced from the CORINE Land Cover. Figure 11 shows two examples of computation. Thereby, the scheduling variable is small if the landscape diversity is small (Figure 11 right), for instance in area with an intensive agriculture or a wide urban density. On the contrary, the scheduling variable is large for region with varied biotopes (Figure 11 left). The considered land use index has been calculated from CORINE Land Cover 2006 and is plotted in Figure 12.

[4] `https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-corine`

Figure 11: Examples of computation of the land use index. CORINE Land Cover frontiers (black), disc C (blue) and centre (red cross) with $H(456, 400) = 2.71$ (left) and $H(458, 676) = 0.27$ (right).

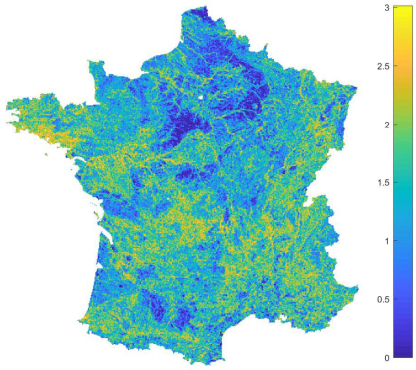

Figure 12: Land use index H from CORINE Land Cover 2006.

6.3. Considered model

The parameter-varying PDE model has been chosen with the following structure

$$\begin{aligned}\frac{\partial u(x,y,t)}{\partial t} = \\ D(H)\left(\frac{\partial^2 u(x,y,t)}{\partial x^2} + \frac{\partial^2 u(x,y,t)}{\partial y^2}\right) \\ -w_x(H)\frac{\partial u(x,y,t)}{\partial x} - w_y(H)\frac{\partial u(x,y,t)}{\partial y} \\ +\beta_1(H)u(x,y,t) - \beta_2(H)u^2(x,y,t)\,,\end{aligned} \tag{42}$$

where

$$\begin{aligned} D(H) &= D_0 + D_1 H(x,y), \\ w_\bullet(H) &= w_{\bullet_0} + w_{\bullet_1} H(x,y), \\ \beta_\bullet(H) &= \beta_{\bullet_0} + \beta_{\bullet_1} H(x,y). \end{aligned} \tag{43}$$

Therefore, the parameter vector to be estimated is

$$\boldsymbol{\theta} = \left[D_0\ D_1\ w_{x_0}\ w_{x_1}\ w_{y_0}\ w_{y_1}\ \beta_{1_0}\ \beta_{1_1}\ \beta_{2_0}\ \beta_{2_1}\right]^\top. \tag{44}$$

6.4. Identification results

From the 3D partial moments and the least-squares estimate (35), we obtain the following estimated parameters: $D_0 = 113.638$, $w_{x_0} = 4.718$, $w_{y_0} = 2.823$, $\beta_{1_0} = -0.124$ and $\beta_{2_0} = -0.036$. Then we select small values for the missing parameters in the initialization vector as follows

$$\begin{aligned}\boldsymbol{\theta}_0 = [113.638 \quad -0.1 \quad 4.718 \quad -0.1 \quad 2.823 \\ -0.1 \quad -0.124 \quad 0.01 \quad -0.036 \quad 0.001]^\top.\end{aligned} \tag{45}$$

From this initialization, the Levenberg-Marquardt algorithm yields the estimated parameter vector

$$\begin{aligned}\boldsymbol{\theta}_{22} = [37.617 \quad 67.858 \quad -35.243 \quad 54.480 \quad 19.641 \\ -14.391 \quad 0.697 \quad -0.824 \quad 0.228 \quad -0.255]^\top.\end{aligned} \tag{46}$$

The simulation of this estimated model is shown in Figure 13.

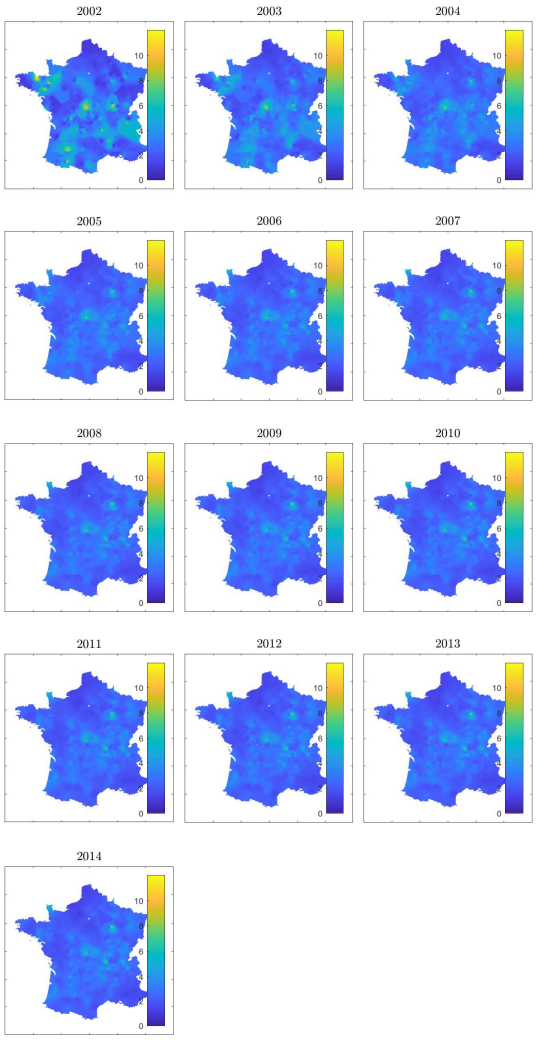

Figure 13: Estimated abundance given by the parameter-varying PDE model (42) with parameters (46).

6.5. *Fitting in terms of trend*

Once again, even if our PDE model represents the population in both spatial domain and time domain, let us see the fitting in terms of global trend by using the normalized temporal trend defined by (40). More specifically, we want to compare our estimated PDE model with the ODE logistic model (Mouysset et al. (2016)) defined by

$$u_t((k+1).dt) = u_t(k.dt) + u_t(k.dt)R_s\left(1 - \frac{u_t(k.dt)}{K}\right), \tag{47}$$

where $u_t(t)$ is the data-set with $t \in [2002,\ 2014]$ calculated from (40) and plotted (blue circles) in Figure 14, R_s and K are parameters estimated by least-squares from $u_t(t)$. As shown in Figure 14, the population trend given by the estimated PDE model (green dashed curve) approximates the global population trend like the ODE logistic model (magenta dashed curve) .

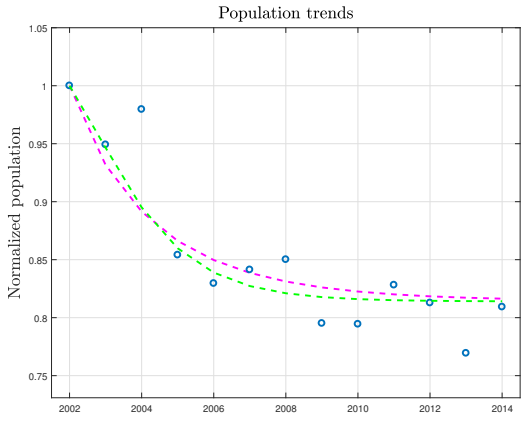

Figure 14: Measured population trend $u_t(t)$ (blue circles), ODE logistic model simulation (magenta dashed line) and estimated PDE model trend simulated with (46) and calculated from (40) (green dashed line).

7. Conclusion

In this paper, a specific attention has been paid to the estimation of parameters of a parameter-varying partial differential equation model for the bird population dynamic. In order to reach this goal, a new identification method based on the Galerkin method and the proper orthogonal decomposition has been presented. This method is implemented with a Levenberg-Marquardt algorithm initialized by a least-squares estimate based on 3D partial moments.

The proposed tools have been validated in simulation in the following conditions:

- a parametric estimation of nonlinear PDE models with parameters varying,
- considered systems without exogenous input and with a unique excitation provided by initial conditions,
- and a small number of time samples.

Finally, the algorithms are applied to European Stonechat population with data collected in the national French Breeding Bird Survey. The scheduling variable of the parameter-varying formulation is the landscape pattern given by the CORINE Land Cover data.

The future works will be to apply tools to a community of birds specialized of agricultural area, to consider a Shannon index in the parameter-varying formulation which characterize the agricultural practice and to develop tools to help take biodiversity goals into public policies into account for Common Agricultural Policy from European Union, for instance.

References

Belforte, G., Dabbene, F., Gay, P., 2005. LPV approximation of distributed parameter systems in environmental modelling. Environmental Modelling and Software 20, 1063–1070.

Chamberlain, D., Fuller, R., Bunce, R., Duckworth, J., Shrubb, M., 2000. Changes in the abundance of farmland birds in relation to the timing of agricultural intensification in England and Wales. Journal of Applied Ecology 37, 771–788.

Debnath, L., Mikusinski, P., 2005. Introduction to Hilbert spaces with applications. Academic Press.

Doxa, A., Bas, Y., Paracchini, M. L., Pointereau, P., Terres, J. M., Jiguet, F., 2010. Low-intensity agriculture increases farmland bird abundances in France. Journal of Applied Ecology 47, 1348–1356.

Farah, M., Mercère, G., Ouvrard, R., Poinot, T., 2016. Combining least-squares and gradient-based algorithms for the identification of a co-current flow heat exchanger. Int. J. Control, 1–13.

Fisher, R. A., 1937. The wave of advance of advantageous genes. Annals of Eugenics 7(4), 355–369.

Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., Goulson, D., de Kroon, H., 2017. More than 75 percent decline over 27 years in total flying insect biomass in protected areas. PLoS ONE 12(10).

Hill, M. O., 1973. Diversity and evenness: a unifying notation and its consequences. Ecology 54, 427–432.

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. Ecology Letters.

IPBES (Ed.), March 2018a. IPBES-6 Plenary. Medellin, Colombia.

IPBES, 2018b. Summary for policymakers of the assessment report on land degradation and restoration of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. R. Scholes, L. Montanarella, A. Brainich, N. Barger, B. ten Brink, M. Cantele, B. Erasmus, J. Fisher, T. Gardner, T. G. Holland, F. Kohler, J. S. Kotiaho, G. Von Maltitz, G. Nangendo, R. Pandit, J. Parrotta, M. D. Potts, S. Prince, M. Sankaran and L. Willemen (Eds.). IPBES secretariat, Bonn, Germany.

IPBES, 2018c. Summary for policymakers of the regional assessment report on biodiversity and ecosystem services for Europe and Central Asia of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. M. Fischer, M. Rounsevell, A. Torre-Marin Rando, A. Mader, A. Church, M. Elbakidze, V. Elias, T. Hahn, P.A. Harrison, J. Hauck, B. Martín-López, I. Ring, C. Sandström, I. Sousa Pinto, P. Visconti, N.E. Zimmermann and M. Christie (Eds.), IPBES secretariat, Bonn, Germany.

IPBES, 2018d. Summary for policymakers of the regional assessment report on biodiversity and ecosystem services for the Americas of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. J. Rice, C.S. Seixas, M.E. Zaccagnini, M. Bedoya-Gaitán, N. Valderrama, C.B. Anderson, M.T.K. Arroyo, M. Bustamante, J. Cavender-Bares, A. Diaz-de-Leon, S. Fennessy, J. R. García Márquez, K. Garcia, E.H. Helmer, B. Herrera, B. Klatt, J.P. Ometo, V. Rodríguez Osuna, F.R. Scarano, S. Schill and J. S. Farinaci (Eds.), IPBES secretariat, Bonn, Germany.

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

Levin, S., 1976. Population dynamic models in heterogeneous environments. Annual review of ecology and systematics 7, 287–310.

Li, H., Qi, C., 2010. Modeling of distributed parameter systems for applications - A synthesized review from time-space separation. Journal of Process Control 20 (2010) 20, 891–901.

Liang, Y., Lee, H., Lim, S., Lin, W., Lee, K., Wu, C., 2002. Proper orthogonal decomposition and its applications. Part I: theory. Journal of Sound and Vibration 252, 527–544.

Ljung, L., 1999. System identification - Theory for the User. Prentice-Hall, Upper Saddle River, N.J., 2nd edition.

Mouysset, L., 2012. Les politiques publiques au défi de la biodiversité : modèles et scénarios bio-économiques pour une agriculture durable. Ph.D. thesis, Muséum National d'Histoire Naturelle.

Mouysset, L., Doyen, L., Jiguet, F., 2012. Different policy scenarios to promote various targets of biodiversity. Ecological Indicators 14, 209–221.

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

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.

Okubo, A., 1986. Dynamical aspects of animal grouping: Swarms, schools, flocks, and herds. Advances in Biophysics 22, 1–94.

Okubo, A., Levin, S., 2001. Diffusion and ecological problems: modern perspectives, Second Edition. Springer.

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

Pham, D., Mercère, G., Ouvrard, R., Poinot, T., 2018. Heat equation parameter estimation based on the POD-Galerkin approach. In: 18th IFAC Symposium on System Identification.

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

Schorsch, J., Gilson, M., Laurain, V., Garnier, H., 2013. Identification of LPV partial differential equation models. In: 52nd IEEE Conference on Decision and Control. Florence, Italy.

Schorsch, J., Laurain, V., Gilson, M., Garnier, H., 2014. Refined IV-based method for LPV partial differential equation model identification. In: 13th European Control Conference. Strasbourg, France.

Shigesada, N., Kawasaki, K., Teramoto, E., 1979. Spatial segregation of interacting species. Journal of Theoretical Biology 79-1, 83–99.

Siriwardena, G., Baillie, S., Buckland, S., Fewster, R., Marchant, J., Wilson, J., 1998. Trends in the abundance of farmland birds: a quantitative comparison of smoothed Common Birds Census indices. Journal of Applied Ecology 35, 24–43.

Skellam, J. G., 1951. Random dispersal in theoretical populations. Biometrika 38, 196–218.

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. Agriculture, Ecosystems and Environment 254, 244–254.

Toth, R., 2010. Modeling and identification of linear parameter-varying systems. 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.

Uciński, D., 2005. Optimal measurement methods for distributed parameter system identification. CRC Press.

Wu, J., 1996. Theory and applications of partial functional differential equations. Vol. 119. Springer, New York.

WWF, 2018. Living planet report - 2018: Aiming higher. Grooten, M. and Almond, R.E.A. (Eds.), WWF, Gland, Switzerland.