Control Engineering Practice 91 (2019) 104091

Contents lists available at ScienceDirect

Control Engineering Practice

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

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

ARTICLE INFO

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

ABSTRACT

In order to study the global decline of biodiversity, accurate models of animal population dynamics are required. In this paper, this challenging problem of biodiversity decline analysis is tackled by modelling the dynamics of bird populations. More specifically, a new data-driven modelling of bird population dynamics is suggested which resorts to a parameter-varying partial differential equation (PDE) model, the Galerkin method and the proper orthogonal decomposition. The parameter-varying formulation allows us to introduce in models prior information such as temperatures or landscape patterns. Hence, such models can be used to study the impact on biodiversity of the global warming or the agricultural intensification, for instance. In order 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. These 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.

1. Introduction

The global decline of biodiversity is an undeniable fact today (Hallmann et al., 2017; IPBES, 2018a; WWF, 2018). As mentioned in the assessment reports of IPBES (2018b, 2018c, 2018d), in Americas, Central Asia or Europe, such a biodiversity loss is mainly due to the land degradation and, for instance, to an unsustainable agricultural. These landscape modifications have a strong impact on animal population dynamics. As a direct consequence of this correlation between land degradation and animal population dynamics, it is nowadays essential to have accurate models of animal population dynamics in order to analyse the biodiversity decline as well as to predict future trends and influence public policies. Different population dynamic models have been developed in literature such as in Chamberlain, Fuller, Bunce, Duckworth, and Shrubb (2000), Mouysset, Miglianico, Makowski, Jiguet, and Doyen (2016) and Siriwardena et al. (1998). 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, new numerical tools allow us to model population dynamics with partial differential equations (PDEs) in order to characterize temporal and spatial dynamics simultaneously. To deal with this objective, automatic control community and more specifically system identification teams should investigate models based on PDEs.

Birds are known as the best taxon to measure the global biodiversity decline (Inger et al., 2014; Stanton, Morrissey, & Clark, 2018). The developments introduced in this paper are thus 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 in terms of time and space are needed. In Mouysset (2012), Mouysset, Doyen, and Jiguet (2012) and Mouysset et al. (2016), dynamic models for bird populations in France have been estimated from large sets 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 tools which allow researchers to analyse the direct impacts of economic scenarios of the Common Agricultural Policy from European Union. The dynamic models estimated in

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

https://doi.org/10.1016/j.conengprac.2019.07.009
Received 21 November 2018; Received in revised form 5 July 2019; Accepted 14 July 2019
Available online xxxx
0967-0661/

Mouysset (2012) and Mouysset et al. (2012, 2016) are based on the well-known logistic equation which is often used to model a population growth in ecology (Okubo & Levin, 2001; Wu, 1996). More specifically, these models are deduced from a nonlinear ODE, the Verhulst's model, which characterizes the temporal dynamic. In the present paper, PDE models are introduced to improve the modelling of the population by adding spatial variables. Notice that, even if no effective tools for data-driven parameter estimation have been developed in the literature applied to ecology so far (to the authors' knowledge), PDEs have been used to represent population dynamics for many years (Holmes, Lewis, Banks, & Veit, 1994; Levin, 1976; Okubo, 1986; Okubo & Levin, 2001; Shigesada, Kawasaki, & Teramoto, 1979; Skellam, 1951). Okubo and Levin (2001) survey 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 the models to evaluate, for instance, the environmental footprint of agricultural intensification, linear parameter-varying (LPV) PDEs are suggested. Although LPV models have been widely studied for two decades (Toth, 2010), the LPV PDE model applications are rarely considered in literature. The following references can be cited (Belforte, Dabbene, & Gay, 2005; Farah, Mercère, Ouvrard, & Poinot, 2016; Pham, Mercère, Ouvrard, & Poinot, 2018; Schorsch, Gilson, Laurain, & Garnier, 2013; Schorsch, Laurain, Gilson, & Garnier, 2014). In the sequel, the term parameter-varying is favoured because the considered PDE models can be linear or nonlinear.

System identification consists in estimating the parameters of a dynamic 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. 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, Devictor, Julliard, and Couvet (2012). The sampling time is limited to one data by breeding season, *i.e.* one sample per year at each location, during thirteen years. This sampling limitation enforces us to develop new identification methods to estimate PDE parameters. More specifically, an iterative approach based on time–space separation (Li & Qi, 2010) is suggested. In order to bypass some difficulties encountered when PDEs come into play, this iterative approach consists in combining the Galerkin method (Polis, Goodson, & Wozny, 1973) and the proper orthogonal decomposition (Newman, 1996a, 1996b), respectively. This iterative procedure requires an initialization which is given by an estimator based on 3D partial moments, an extension of partial moments described in Ouvrard and Trigeassou (2011). These new 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 the approach by specialists of system identification. As mentioned in perspectives in the conclusion section, a future work will involve 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 POD-Galerkin method is presented in Section 3 for a general formulation of the PDE. The identification procedure is introduced in Section 4. Section 5 is dedicated to the validation of the identification approach with simulated data and a simplified PDE. An application to the population of European Stonechat *Saxicola torquatus* is shown in Section 6 with parameters varying with the landscape pattern in France. Section 7 concludes this paper.

2. Parameter-varying PDE model

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

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

where x and y are spatial coordinates, t is time, $u(x,y,t)$ is a population, $H(x,y)$ is the so-called scheduling variable (Toth, 2010) while 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 stand for the dispersal rate (Holmes et al., 1994). 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 us to take into account, for example, the land use to analyse agricultural changes to animal populations or the temperature to study effects of global warming.

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

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

3. POD-Galerkin method

The parameter estimation of the PDE (1) requires to use an optimization algorithm, *i.e.* an iterative procedure which minimizes a cost function based on the measured data and the PDE simulation. But the simulation of a parameter-varying PDE such as (1) is a difficult issue. Therefore, it is preferable to calculate an approximate solution of the PDE by using the so-called POD-Galerkin method (Newman, 1996a, 1996b) and the time–space separation. More specifically, a Galerkin method is considered and 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 & Mikusinski, 2005) proves that any arbitrary signal f belonging to a Hilbert space like $\mathcal{L}_2(D)$ (the space of Lebesgue-measurable square-integrable functions on a domain D, *i.e.* $\int_D |f(x)|^2 dx < \infty$), 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 $\mathcal{L}_2(D)$ has been selected. More specifically, for signals like $u(x, y, t)$ belonging to the Hilbert space $\mathcal{L}_2(D)$, for $D = \{x \in [0, X] \text{ and } y \in [0, Y]\}$, with the inner product

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

the signal $u(x, y, t)$ satisfies

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

where $a_n(t) = \langle u, \phi_n \rangle$ 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 & Qi, 2010).

3.2. Proper orthogonal decomposition

In this paper, in order to benefit from available data-sets, it was decided to use 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 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, 1996b). 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 by considering the inner product between time samples as follows[1]

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

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.

Remark 1. For the sake of conciseness, in the sequel, the double integration $\int \int_D \bullet dx dy$ introduced with the inner product will be denoted $\int \bullet$ and, for instance, Eq. (7) is now denoted

$$\alpha_{k\ell} = \frac{1}{M} \int \delta(x, y, k\ dt) \delta(x, y, \ell\ dt). \tag{9}$$

[1] dt is the time sampling period.

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,

$$\mathbf{K} = \boldsymbol{Q} \boldsymbol{R} \boldsymbol{Q}^{\top}, \tag{10}$$

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), it is guaranteed that

$$\sum_{n=1}^{N} a_n(t) \phi_n(x, y) \tag{11}$$

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

$$\phi_n(x, y) = \sum_{k=1}^{M} \boldsymbol{Q}_k^{(n)} \delta(x, y, k\ dt), \tag{12}$$

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

3.3. Galerkin method

As previously mentioned, the Galerkin method operates a time–space separation which transforms the PDE problem into a system of ODEs. Some details of the calculations are given in Appendix. It consists in finding an approximate solution of the PDE defined by

$$\hat{u}(x, y, t) = \sum_{n=1}^{N} a_n(t) \phi_n(x, y) + \bar{u}(x, y), \tag{13}$$

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 (13) into (1) leads to

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

By applying the inner product $\langle$Eq. (14), $\phi_\ell\rangle$, the orthonormal property of $\{\phi_n(x,y)\}_{n=1}^N$ yields the following system of ODEs for $\ell = 1,\dots,N$

$$\begin{aligned}\dot{a}_\ell(t) &= \sum_{n=1}^{N} a_n(t) \int \left\{ D_x(H)\frac{\partial^2\phi_n}{\partial x^2} + D_y(H)\frac{\partial^2\phi_n}{\partial y^2} \right.\\ &\left. - w_x(H)\frac{\partial\phi_n}{\partial x} - w_y(H)\frac{\partial\phi_n}{\partial y} + \beta_1(H)\phi_n \right\}\phi_\ell\\ &-2\sum_{n=1}^{N} a_n(t)\int \beta_2(H)\phi_n\phi_\ell\bar{u}\\ &-\sum_{n=1}^{N} a_n^2(t)\int \beta_2(H)\phi_n^2\phi_\ell\\ &-2\sum_{n=1}^{N-1}\sum_{m=n+1}^{N} a_n(t)a_m(t)\int \beta_2(H)\phi_n\phi_m\phi_\ell\\ &+\int\left\{ D_x(H)\frac{\partial^2\bar{u}}{\partial x^2} + D_y(H)\frac{\partial^2\bar{u}}{\partial y^2} - w_x(H)\frac{\partial\bar{u}}{\partial x}\right.\\ &\left. - w_y(H)\frac{\partial\bar{u}}{\partial y} + \beta_1(H)\bar{u} - \beta_2(H)\bar{u}^2\right\}\phi_\ell.\end{aligned} \tag{15}$$

For an easier implementation, this system of ODEs can be rewritten with a matrix formulation

$$\begin{aligned}\dot{\mathbf{a}}(t) = &\left(\boldsymbol{\Gamma}_1 + \beta_{1_0}\mathbf{I}_N - 2\boldsymbol{\Gamma}_2\right)\mathbf{a}(t)\\ &-\left(\mathbf{I}_N\otimes\mathbf{a}^\top(t)\right)\boldsymbol{\Lambda}\,\mathbf{a}(t) + \mathbf{b}_1,\end{aligned} \tag{16}$$

with

$$\begin{aligned}&\otimes,\ \text{Kronecker product},\\ &\mathbf{a}(t) = \begin{bmatrix}a_1(t) & \dots & a_N(t)\end{bmatrix}^\top,\\ &\mathbf{I}_N,\ N\times N\ \text{identity matrix},\\ &\boldsymbol{\Gamma}_1 = \begin{bmatrix}\int \mathcal{F}_1(\phi_1)\phi_1 & \dots & \int \mathcal{F}_1(\phi_N)\phi_1\\ \vdots & \ddots & \vdots\\ \int \mathcal{F}_1(\phi_1)\phi_N & \dots & \int \mathcal{F}_1(\phi_N)\phi_N\end{bmatrix},\\ &\boldsymbol{\Gamma}_2 = \\ &\begin{bmatrix}\int \beta_2(H)\phi_1\phi_1\bar{u} & \dots & \int \beta_2(H)\phi_N\phi_1\bar{u}\\ \vdots & \ddots & \vdots\\ \int \beta_2(H)\phi_1\phi_N\bar{u} & \dots & \int \beta_2(H)\phi_N\phi_N\bar{u}\end{bmatrix},\\ &\boldsymbol{\Lambda} = \int \beta_2(H)\boldsymbol{\Phi}\otimes\left(\boldsymbol{\Phi}\otimes\boldsymbol{\Phi}^\top\right),\\ &\boldsymbol{\Phi} = \begin{bmatrix}\phi_1(x,y) & \dots & \phi_N(x,y)\end{bmatrix}^\top,\\ &\mathbf{b}_1 = \begin{bmatrix}\int \mathcal{F}_2(\bar{u})\phi_1 & \dots & \int \mathcal{F}_2(\bar{u})\phi_N\end{bmatrix}^\top,\\ &\mathcal{F}_1(\bullet) = D_x(H)\frac{\partial^2\bullet}{\partial x^2} + D_y(H)\frac{\partial^2\bullet}{\partial y^2} - w_x(H)\frac{\partial\bullet}{\partial x}\\ &\qquad -w_y(H)\frac{\partial\bullet}{\partial y} + \sum_{\upsilon=1}^{n_{\beta_1}}\beta_{1_\upsilon}H^\upsilon(x,y)\bullet,\\ &\mathcal{F}_2(\bullet) = D_x(H)\frac{\partial^2\bullet}{\partial x^2} + D_y(H)\frac{\partial^2\bullet}{\partial y^2} - w_x(H)\frac{\partial\bullet}{\partial x}\\ &\qquad -w_y(H)\frac{\partial\bullet}{\partial y} + \beta_1(H)\bullet - \beta_2(H)\bullet^2.\end{aligned} \tag{17}$$

The initial conditions of the ODEs system (16) is obtained by making the initial signal error orthogonal to the basis functions $\phi_\ell(x,y)$, *i.e.*

$$\langle u(\cdot,\cdot,0) - \hat{u}(\cdot,\cdot,0), \phi_\ell\rangle_{\ell=1,\dots,N} = 0, \tag{18}$$

or

$$a_\ell(0) = \int \left(u(x,y,0) - \bar{u}(x,y)\right)\phi_\ell(x,y), \tag{19}$$

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

Therefore, given parameters, the simulation of the system of ODEs (16) from the initial conditions (19) provides the approximate solution (13). 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

$$\begin{aligned}\boldsymbol{\theta} = \Big[&D_{x_0}\dots D_{x_{n_{D_x}}}\ D_{y_0}\dots D_{y_{n_{D_y}}}\ w_{x_0}\dots w_{x_{n_{w_x}}}\\ &w_{y_0}\dots w_{y_{n_{w_y}}}\ \beta_{1_0}\dots\beta_{1_{n_{\beta_1}}}\ \beta_{2_0}\dots\beta_{2_{n_{\beta_2}}}\Big]^\top\end{aligned} \tag{20}$$

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

$$J(\boldsymbol{\theta}_{iter}) = \sum_{k=0}^{N_t}\sum_{j=0}^{N_y}\sum_{i=0}^{N_x}\left(u(i,j,k) - \hat{u}(i,j,k)\right)^2, \tag{21}$$

where $\hat{u}(i,j,k)$ is the approximate solution (13) of the PDE model (1), and $\boldsymbol{\theta}_{iter}$ is the parameter vector at the *iter*-th iteration. This procedure repeats iteratively the following two steps from an initial vector $\boldsymbol{\theta}_0$ and until convergence

- approximate a solution $\hat{u}(i,j,k)$ by the Galerkin method, Eq. (13),
- compute a new parameter vector $\boldsymbol{\theta}_{iter+1}$ with a Levenberg-Marquardt algorithm which minimizes the criterion (21).

4.2. 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 & Wright, 2006). This model does 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. 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 (19).

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 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 (13), $\boldsymbol{\Phi}(x,y)$ is defined in (17), and $\boldsymbol{\Xi}(t)$ is the $N_\theta\times N$ time sensitivity function matrix given by

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

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

The parametric sensitivity functions are simulated by considering the following matrix formulations deduced from derivatives of (16) with respect to each parameter

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

for $n = 0, \dots, n_{D_\bullet}, n_{w_\bullet}$ 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 & \dots & \int \mathcal{F}_{D_{\bullet_n}}(\phi_N)\phi_1 \\ \vdots & \ddots & \vdots \\ \int \mathcal{F}_{D_{\bullet_n}}(\phi_1)\phi_N & \dots & \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 & \dots & \int \mathcal{F}_{w_{\bullet_n}}(\phi_N)\phi_1 \\ \vdots & \ddots & \vdots \\ \int \mathcal{F}_{w_{\bullet_n}}(\phi_1)\phi_N & \dots & \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 & \dots & \int H^n\phi_N\phi_1 \\ \vdots & \ddots & \vdots \\ \int H^n\phi_1\phi_N & \dots & \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} & \dots & \int H^n\phi_N\phi_1\bar{u} \\ \vdots & \ddots & \vdots \\ \int H^n\phi_1\phi_N\bar{u} & \dots & \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 & \dots & \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 & \dots & \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 & \dots & \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 & \dots & \int H^n\bar{u}^2\phi_N \end{bmatrix}^{\top}, \\
&\mathcal{F}_{D_{\bullet_n}}(\circ) = H^n(x,y)\frac{\partial^2 \circ}{\partial \bullet^2}, \quad \mathcal{F}_{w_{\bullet_n}}(\circ) = H^n(x,y)\frac{\partial \circ}{\partial \bullet},
\end{aligned}
\tag{26}
$$

and with the initial conditions

$$
\begin{aligned}
\boldsymbol{\sigma}^{\mathbf{a}}_{D_{\bullet_n}}(0) &= \boldsymbol{\Gamma}_{D_{\bullet_n}}\ \mathbf{a}(0) + \mathbf{b}_{D_{\bullet_n}}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{w_{\bullet_n}}(0) &= \boldsymbol{\Gamma}_{w_{\bullet_n}}\ \mathbf{a}(0) + \mathbf{b}_{w_{\bullet_n}}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{1_n}}(0) &= \boldsymbol{\Gamma}_{\beta_{1_n}}\mathbf{a}(0) + \mathbf{b}_{\beta_{1_n}}, \\
\boldsymbol{\sigma}^{\mathbf{a}}_{\beta_{2_n}}(0) &= -2\boldsymbol{\Gamma}_2\ \mathbf{a}(0) \\
&\quad - \left(\mathbf{I}_N \otimes \mathbf{a}^{\top}(0)\right)\boldsymbol{\Lambda}_{\beta_{2_n}}\mathbf{a}(0) - \mathbf{b}_{\beta_{2_n}}.
\end{aligned}
\tag{27}
$$

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 initialization of the iterative identification procedure given in (22). For the PDE model, the choice of the initialization is particularly crucial to tend toward the global optimum of the criterion. It can be done from prior knowledge. In ecological applications, such a prior knowledge is often missing. Therefore, a least-squares method is suggested. It is based on partial moments developed in Trigeassou (1987) and Ouvrard and Trigeassou (2011) for systems represented by ODEs and extended to the PDE models thereafter.

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 $n \in \{\mathbb{N},\ -\}$.[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

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

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

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

and the Cauchy formula of repeated integrations

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

Eq. (29) can be reformulated with 3D partial moments as follows

$$
\begin{aligned}
\mathcal{U}_1(X,Y,T) &= D_{x_0}\mathcal{U}_2(X,Y,T) + D_{y_0}\mathcal{U}_3(X,Y,T) \\
&\quad -w_{x_0}\ \mathcal{U}_4(X,Y,T) - w_{y_0}\ \mathcal{U}_5(X,Y,T) \\
&\quad +\beta_{1_0}\ \mathcal{U}_6(X,Y,T) + \beta_{2_0}\ \mathcal{U}_7(X,Y,T),
\end{aligned}
\tag{33}
$$

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

where

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

Hence, (33) can be rewritten in a linear regression form

$$\mathcal{U}_1(X,Y,T) = \boldsymbol{\varphi}^\top(X,Y,T)\boldsymbol{\theta}_{LS}, \tag{34}$$

with

$$
\begin{aligned}
\boldsymbol{\varphi}(X,Y,T) = \big[&\mathcal{U}_2(X,Y,T) \quad \mathcal{U}_3(X,Y,T) \\
&-\mathcal{U}_4(X,Y,T) \quad -\mathcal{U}_5(X,Y,T) \\
&\mathcal{U}_6(X,Y,T) \quad \mathcal{U}_7(X,Y,T)\big]^\top,
\end{aligned} \tag{35}
$$

and a least-squares estimate of parameters is given by

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

In that way, the vector $\boldsymbol{\theta}_{LS}$ gives 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 (20), small values for the missing parameters corresponding to $n_{D_\bullet} > 0$, $n_{w_\bullet} > 0$ and $n_{\beta_\bullet} > 0$ are considered.

5. Simulation

The goal of this simulated test is to validate the proposed 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 5 presents such an application with real data.

5.1. Simulated system

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

$$
\begin{aligned}
&\frac{\partial u(x,y,t)}{\partial t} = \\
&\quad D(H)\left(\frac{\partial^2 u(x,y,t)}{\partial x^2} + \frac{\partial^2 u(x,y,t)}{\partial y^2}\right) \\
&\quad +\beta_1(H)u(x,y,t) - \beta_2(H)u^2(x,y,t)\ ,
\end{aligned} \tag{37}
$$

with

$$
\begin{aligned}
D(H) &= D_0 + D_1 H(x,y), \\
\beta_\bullet(H) &= \beta_{\bullet_0} + \beta_{\bullet_1} H(x,y).
\end{aligned} \tag{38}
$$

For the identification procedure, the developments are those presented in the previous section without parameters $w_\bullet$ and with $n_{D_\bullet} = 1$ and $n_{\beta_\bullet} = 1$.

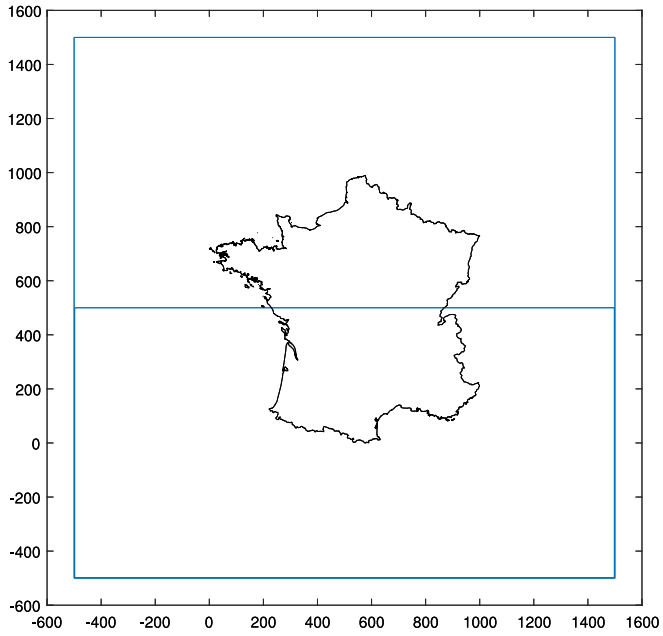

Fig. 1. Simulation domain (blue square), data domain (France map) and scheduling variable (blue rectangles).

5.2. Simulation protocol

An ecological application often presents some specificities: a small number of time samples, a spatial domain given by a map and a kriging preprocessing of data. It is proposed to test and to validate the tools in these specific conditions with 21 time samples, the France map and noise-free data since the kriging preprocessing smooths the noise.

The data-set is provided by a simulation of the PDE (37) 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 Fig. 1. In order to limit the effect of the boundary conditions, the simulation domain is a wide square (blue square in Fig. 1) and, to build the data-set, only the data that belongs to the France map are considered. 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.

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 initial conditions are defined as follows

$$u(x,y,0) = \sum_{n=1}^{10} e^{-((x-x_n)^2+(y-y_n)^2)/2\sigma_n^2}, \tag{39}$$

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

By analogy with the ecological application of Section 5, 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. Fig. 2 presents the population $u(x,y,t)$ for some time samples.

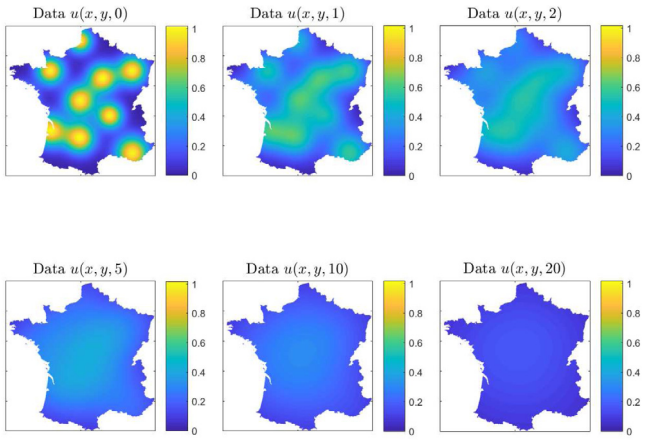

Fig. 2. Data $u(x,y,t)$ for $t = 0, 1, 2, 5, 10$ and 20.

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

Notice that, hereafter, the estimation data-set presented in Fig. 2 will be used for the estimation step. For the estimated model validation, another data-set is generated in the same way but with different values of (x_n, y_n) and σ_n^2 for the initial condition (39).

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. (13). For M, in the proposed context of a small number of time samples, it is preferable to consider all the 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 Fig. 3. By considering increasing values of N and by comparing Levenberg–Marquardt algorithm results in terms of fitting defined by

$$FIT = 100\left(1 - \frac{\|u - \hat{u}\|_2}{\|u - \text{mean}(u)\|_2}\right), \tag{40}$$

Table 1 presents the fittings obtained for the estimation and validation data-sets. 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 behaviour. The first seven orthonormal basis functions $\phi_n(x,y)$ are presented in Fig. 4.

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, two PDE models defined by (37) are estimated: one with parameters varying 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 $N = 7$.

Table 2 shows the obtained fittings and Fig. 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 Fig. 5 presents large amplitudes between top and bottom rectangles.

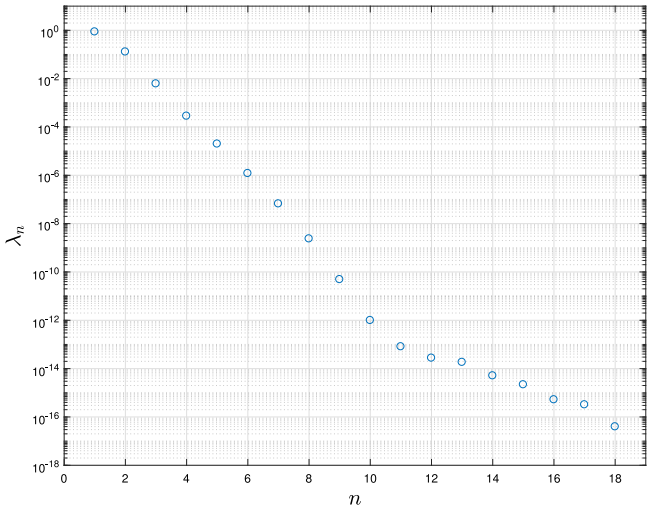

Fig. 3. Normalized eigenvalue λ_n in a decreasing order.

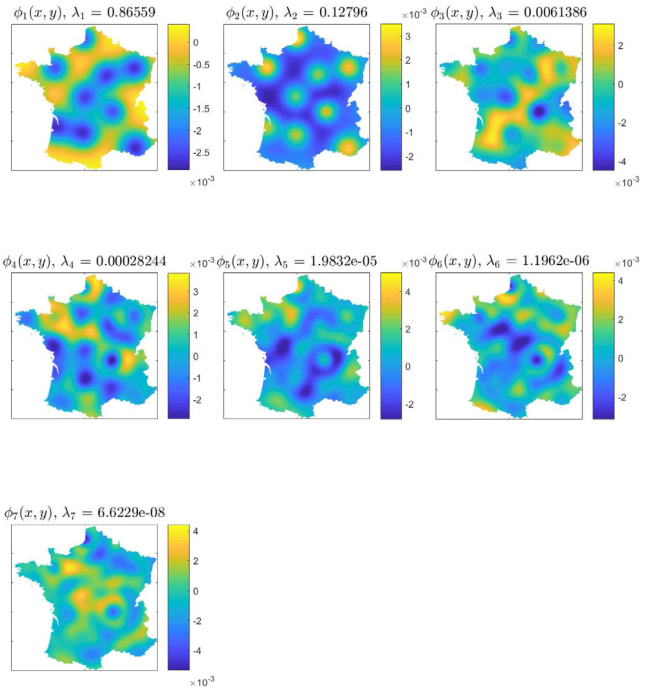

Fig. 4. The seven considered basis functions $\phi_n(x,y)$ with the corresponding eigenvalue λ_n.

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

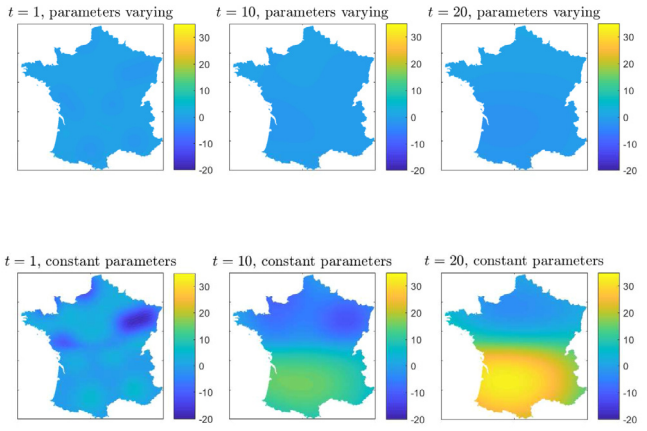

Fig. 5. Spatial error (%) defined by $(u(x,y,t)-\hat{u}(x,y,t))/mean_{x,y}(u(x,y,t))$ for any x, any y and $t = 1, 10, 20$ with the validation data-set.

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

algorithm, and more specifically in the context of a small number of time samples. In other words, the numerical implementation of integrals in (16)–(17) and (25)–(26) causes some approximations. How do these approximations affect the convergence?

In order to observe the effect of the numerical implementation of integrals, a new data-set is used. Consider the so-called POD-Galerkin data-set where u is generated with the POD-Galerkin solution (13) 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. The design parameter N is 7. Two different sampling times are used in the simulation of the ODE system (16) and two different snapshot numbers M in Eq. (6) are considered: $dt = 1$ 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. It can be concluded 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. Fitting in terms of trend

As mentioned in the introduction, in ecological applications, the commonly used tools are based on ODE models which describe the global trend of the population in time domain (Mouysset, 2012; Mouysset et al., 2012, 2016; Okubo & Levin, 2001; Wu, 1996). Even if our PDE models represent the population in both spatial domain and time domain, the fitting in terms of global trend is now considered. The normalized temporal trend is defined by

$$f_t(t) = \sum_{j=0}^{N_y}\sum_{i=0}^{N_x} f(i.dx, j.dy, t) / \sum_{j=0}^{N_y}\sum_{i=0}^{N_x} f(i.dx, j.dy, 0). \quad (41)$$

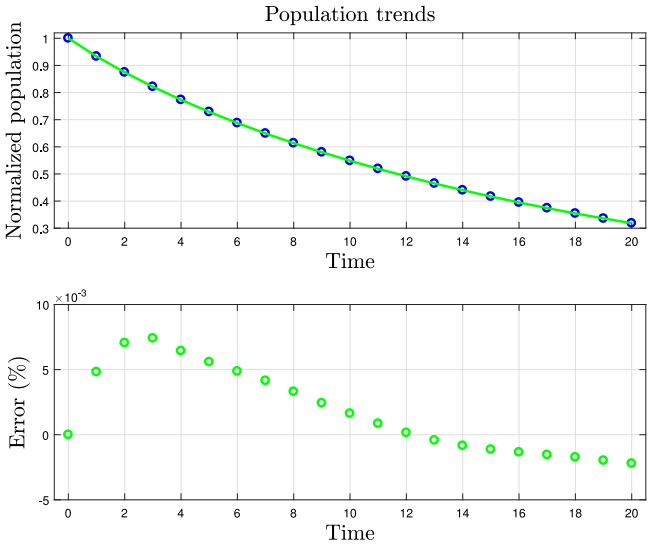

Fig. 6. Trends calculated with the true data (blue circles) and the estimated parameter-varying PDE model simulation (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

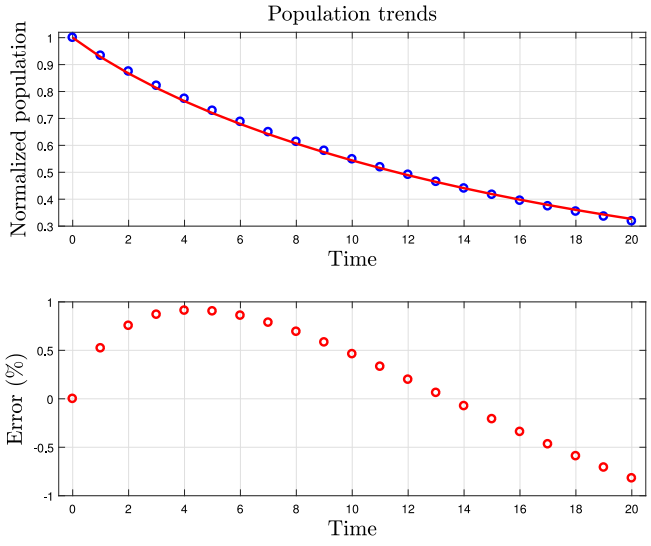

Fig. 7. Trends calculated with the true data (blue circles) and the estimated constant-parameter PDE model simulation (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The considered data-set is the one presented in Section 5.2 and the design parameters are $M = 21$ and $N = 7$, respectively. Fig. 6 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, in order to evaluate the benefit of the parameters varying, Fig. 7 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.

5.7. Validation with a richer scheduling variable

Here, the goal is to test the proposed tool with a more complicated scheduling variable than the one used in the previous sub-sections. Consider the simulation protocol introduced in Section 5.2 with the simulation domain and the scheduling variable presented in Fig. 8.

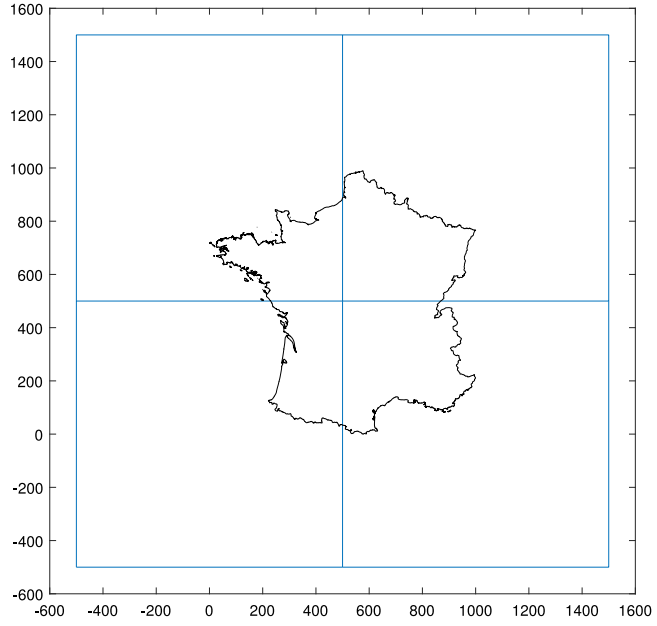

Fig. 8. Simulation domain (outer blue square), data domain (France map) and scheduling variable (inner blue squares) for the global implementation validation.

Table 4
Relative errors (%) after convergence of the Levenberg–Marquardt algorithm.

	Relative error
D_0	0.28
D_1	0.01
β_{1_0}	0.89
β_{1_1}	6.54
β_{2_0}	11.34
β_{2_1}	5.73

Here, the four inner squares specify the value of the scheduling variable: $H(x, y) = 1$ for the coordinates inside the right top square, $H(x, y) = 0.5$ inside the right bottom one, $H(x, y) = 0.25$ inside the left bottom one and $H(x, y) = 0$ inside the left top one. With this new scheduling variable, the selected true parameters are $D_0 = 2000$, $D_1 = -250$, $\beta_{1_0} = -0.04$, $\beta_{1_1} = 0.007$, $\beta_{2_0} = -0.015$ and $\beta_{2_1} = -0.03$.

In order to test the convergence, the Levenberg–Marquardt algorithm is initialized with a parameter vector $\boldsymbol{\theta}_0$ at 80% of the true values. The design parameters are $M = 21$ and $N = 6$, respectively. The convergence is shown in Fig. 9.

After convergence, the estimated parameters are close to true values and the obtained fitting is 99.94 %. Table 4 shows the obtained relative errors on the parameters.

It is important to note that 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 residual error cannot be null because, in the present simulation example, u is given by a finite element method which is a more accurate way to simulate the PDE (37) than the Galerkin method used to calculate $\hat{u}$. This explains why the relative error is larger than the values shown in Table 3. Nevertheless, this error, with a maximum of 11.34 %, is acceptable taking into account the difficult context of the PDE model identification with only 21 time samples.

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

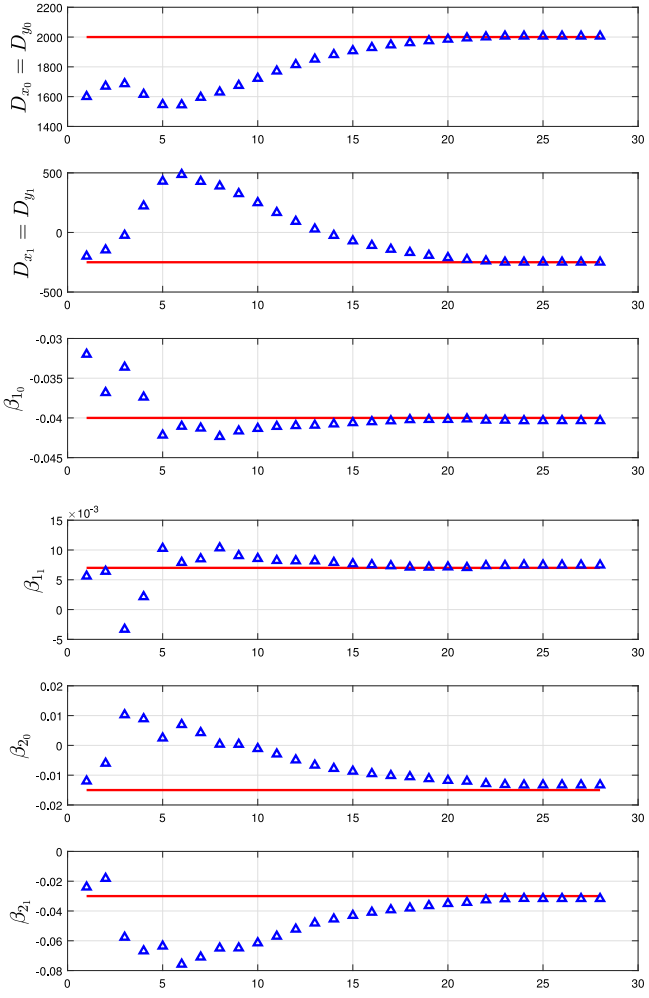

Fig. 9. Convergence of the Levenberg–Marquardt algorithm: true parameters (red line) and estimated parameters with the finite-element data-set (blue triangles).

monitoring based on birdwatchers which count spring birds annually during 5 min exactly on 10 fix points within a randomly selected square. Fig. 10 shows the spatial distribution of the monitored squares. Therefore, this database provides the information related to the bird abundances across the whole country. Abundance values for each species are available for the period 2002–2014. In this paper, only one specie is considered, the European Stonechat *Saxicola torquatus* drawn at the bottom right of Fig. 11. A spatial interpolation of these abundance data is performed 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 Fig. 11.

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

$$H(x, y) = -\sum_{i=1}^{R} p_i \ln p_i, \tag{42}$$

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

Fig. 10. Spatial distribution of the squares monitored in continental France by the Breeding Bird Survey scheme.

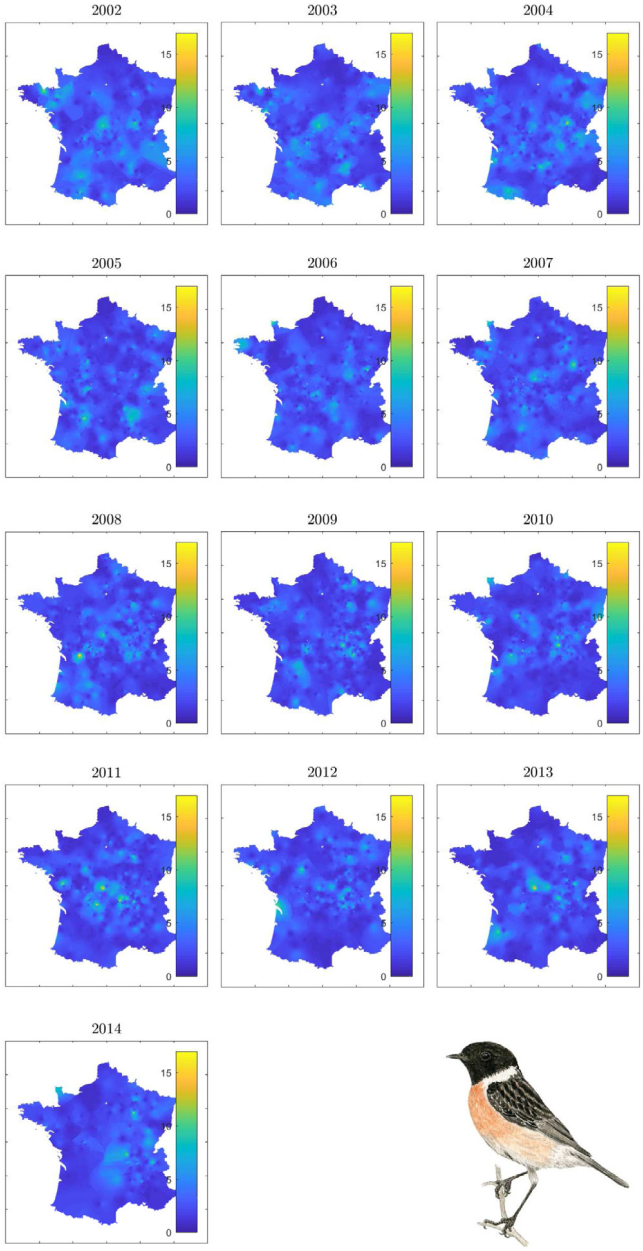

Fig. 11. European Stonechat *Saxicola torquatus* ©Katia LIPOVOI-LPO (bottom rigth) and kriged abundance of this specie in France for 2002–2014.

Fig. 12. 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).

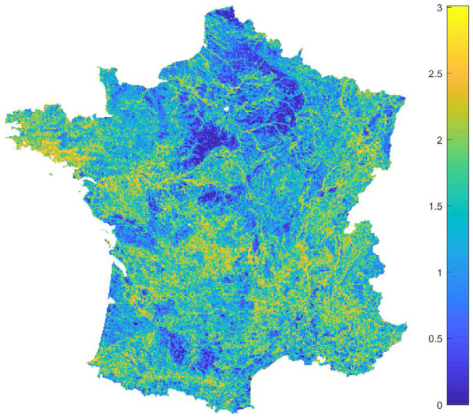

Fig. 13. Land use index H from CORINE Land Cover 2006.

with p_i, the proportion of the ith land use and R, the number of land uses on C (radius 1.5 km) deduced from the CORINE Land Cover. Fig. 12 shows two examples of computation. Thereby, the scheduling variable is small if the landscape diversity is small (Fig. 12 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 (Fig. 12 left). The considered land use index has been calculated from CORINE Land Cover 2006 and is plotted in Fig. 13.

6.3. Considered model

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

$$\begin{aligned}\frac{\partial u(x,y,t)}{\partial t} = \\ D(H)\left(\frac{\partial^2 u(x,y,t)}{\partial x^2} + \frac{\partial^2 u(x,y,t)}{\partial y^2}\right) \\ -w_x(H)\frac{\partial u(x,y,t)}{\partial x} - w_y(H)\frac{\partial u(x,y,t)}{\partial y} \\ +\beta_1(H)u(x,y,t) - \beta_2(H)u^2(x,y,t)\ ,\end{aligned} \tag{43}$$

where

$$\begin{aligned}D(H) &= D_0 + D_1 H(x,y), \\ w_\bullet(H) &= w_{\bullet_0} + w_{\bullet_1} H(x,y), \\ \beta_\bullet(H) &= \beta_{\bullet_0} + \beta_{\bullet_1} H(x,y).\end{aligned} \tag{44}$$

Therefore, the parameter vector to be estimated is

$$\boldsymbol{\theta} = \left[D_0\ D_1\ w_{x_0}\ w_{x_1}\ w_{y_0}\ w_{y_1}\ \beta_{1_0}\ \beta_{1_1}\ \beta_{2_0}\ \beta_{2_1}\right]^\top. \tag{45}$$

6.4. Identification results

From the 3D partial moments and the least-squares estimate (36), the following estimated parameters are obtained: $D_0 = 113.638$, $w_{x_0} = 4.718$, $w_{y_0} = 2.823$, $\beta_{1_0} = -0.124$ and $\beta_{2_0} = -0.036$. Then small values

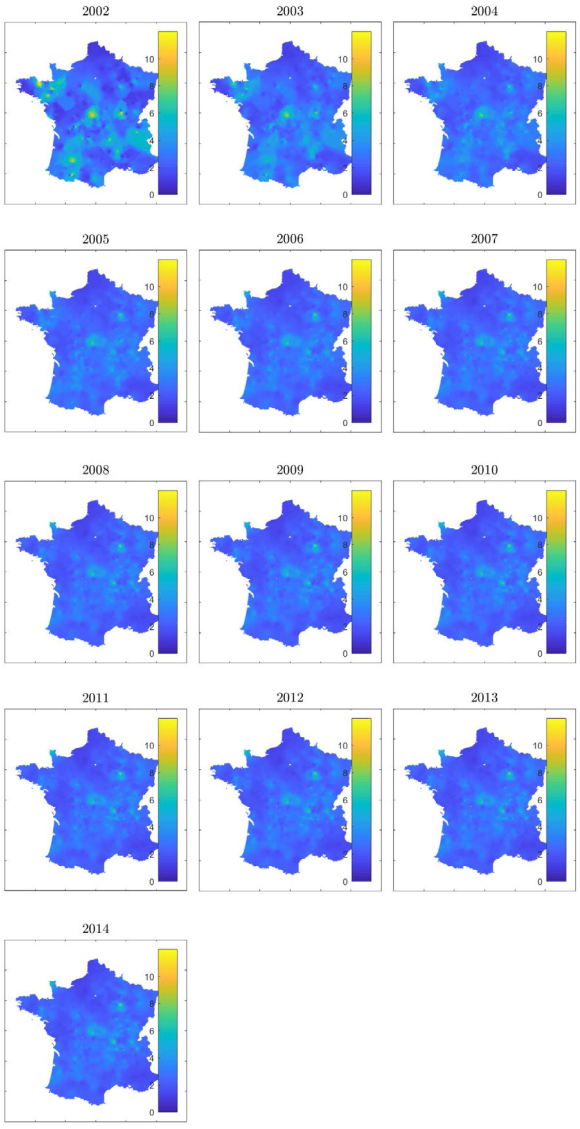

Fig. 14. Estimated abundance given by the parameter-varying PDE model (43) with parameters (47).

for the missing parameters are selected in the initialization vector as follows

$$\boldsymbol{\theta}_0 = [113.638 \quad -0.1 \quad 4.718 \quad -0.1 \quad 2.823 \\ -0.1 \quad -0.124 \quad 0.01 \quad -0.036 \quad 0.001]^\top . \tag{46}$$

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

$$\boldsymbol{\theta}_{22} = [37.617 \quad 67.858 \quad -35.243 \quad 54.480 \quad 19.641 \\ -14.391 \quad 0.697 \quad -0.824 \quad 0.228 \quad -0.255]^\top . \tag{47}$$

The simulation of this estimated model is shown in Fig. 14.

6.5. Fitting in terms of trend

Once again, even if our PDE model represents the population in both spatial domain and time domain, the fitting in terms of global trend is considered by using the normalized temporal trend defined by (41). More specifically, the goal is to compare the estimated PDE model with the ODE logistic model (Mouysset et al., 2016) defined by

$$u_t((k+1)\ dt) = u_t(k\ dt) + u_t(k\ dt) R_s \left(1 - \frac{u_t(k\ dt)}{K}\right), \tag{48}$$

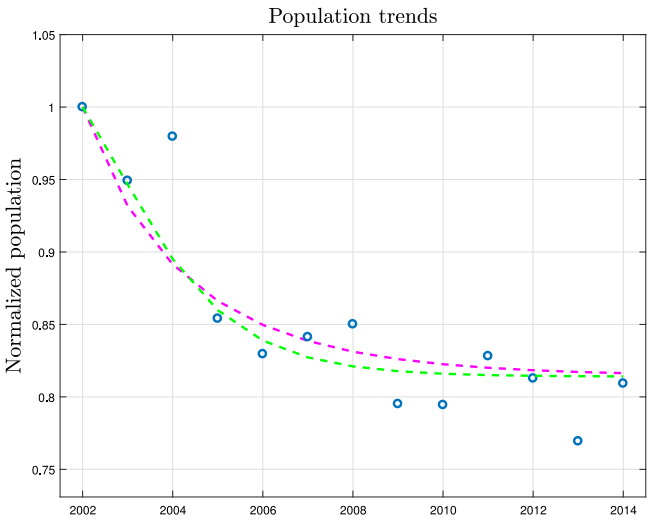

Fig. 15. Measured population trend $u_t(t)$ (blue circles), ODE logistic model simulation (magenta dashed line) and estimated PDE model trend simulated with (47) and calculated from (41) (green dashed line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

where $u_t(t)$ is the data-set with $t \in [2002,\ 2014]$ calculated from (41) and plotted (blue circles) in Fig. 15, R_s and K are parameters estimated by least-squares from $u_t(t)$. As shown in Fig. 15, 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) .

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 consists in applying tools to a community of birds specialized of agricultural area, considering a Shannon index in the parameter-varying formulation which characterize the agricultural practice and developing tools to help take biodiversity goals into public policies into account for Common Agricultural Policy from European Union, for instance.

Declaration of competing interest

None declared.

Appendix. Galerkin method calculations

The Galerkin method (Section 3.3) aims at determining the time coefficients $\{a_n(t)\}_{n=1}^{N}$ of (13) from the known function basis $\{\phi_n(x,y)\}_{n=1}^{N}$. With the substitution of the approximate solution (13) in (1), in the left-hand-side of the corresponding equation, the time derivative becomes

$$\begin{aligned}\frac{\partial \hat{u}(x,y,t)}{\partial t} &= \sum_{n=1}^{N} \frac{\partial a_n(t)}{\partial t}\phi_n(x,y),\\ &= \sum_{n=1}^{N} \dot{a}_n(t)\phi_n(x,y),\end{aligned} \tag{A.1}$$

while, on the right-hand-side, the spatial derivatives become

$$\begin{aligned}\frac{\partial^2 \hat{u}(x,y,t)}{\partial x^2} &= \sum_{n=1}^{N} a_n(t)\frac{\partial^2 \phi_n(x,y)}{\partial x^2} + \frac{\partial^2 \bar{u}(x,y)}{\partial x^2},\\ \frac{\partial^2 \hat{u}(x,y,t)}{\partial y^2} &= \sum_{n=1}^{N} a_n(t)\frac{\partial^2 \phi_n(x,y)}{\partial y^2} + \frac{\partial^2 \bar{u}(x,y)}{\partial y^2},\\ \frac{\partial \hat{u}(x,y,t)}{\partial x} &= \sum_{n=1}^{N} a_n(t)\frac{\partial \phi_n(x,y)}{\partial x} + \frac{\partial \bar{u}(x,y)}{\partial x},\\ \frac{\partial \hat{u}(x,y,t)}{\partial y} &= \sum_{n=1}^{N} a_n(t)\frac{\partial \phi_n(x,y)}{\partial y} + \frac{\partial \bar{u}(x,y)}{\partial y}.\end{aligned} \tag{A.2}$$

Lastly, the quadratic term in (1) yields

$$\begin{aligned}\hat{u}^2(x,y,t) &= \left(\sum_{n=1}^{N} a_n(t)\phi_n(x,y)\right)^2\\ &+2\sum_{n=1}^{N} a_n(t)\phi_n(x,y)\bar{u}(x,y) + \bar{u}^2(x,y)\ ,\end{aligned} \tag{A.3}$$

and

$$\begin{aligned}\left(\sum_{n=1}^{N} a_n(t)\phi_n(x,y)\right)^2 &= \sum_{n=1}^{N} a_n^2(t)\phi_n^2(x,y)\\ &+2\sum_{n=1}^{N-1}\sum_{m=n+1}^{N} a_n(t)a_m(t)\phi_n(x,y)\phi_m(x,y).\end{aligned} \tag{A.4}$$

Therefore, the substitution of (13) in (1) leads to Eq. (14).

By applying the inner product $\langle \text{Eq. (14)}, \phi_\ell \rangle = \int \text{Eq. (14)}\ \phi_\ell$, the orthonormal property of $\{\phi_n(x,y)\}_{n=1}^{N}$ given by

$$\langle \phi_n, \phi_\ell \rangle = \int \phi_n \phi_\ell = \begin{cases} 1 \text{ if } n = \ell \\ 0 \text{ if } n \neq \ell \end{cases} \tag{A.5}$$

leads to the following simplification on the left-hand-side

$$\int \sum_{n=1}^{N} \dot{a}_n(t)\phi_n\phi_\ell = \dot{a}_\ell(t) \tag{A.6}$$

and to Eq. (15).

Remind that

$$\beta_1(H) = \sum_{\upsilon=0}^{n_{\beta_1}} \beta_{1_\upsilon} H^\upsilon(x,y). \tag{A.7}$$

Therefore, on the right-hand-side of (15), the following simplification can be done

$$\begin{aligned}\sum_{n=1}^{N} a_n(t)\int \beta_1(H)\phi_n\phi_\ell &= \beta_{1_0} a_\ell(t)\\ &+\sum_{n=1}^{N} a_n(t)\int \sum_{\upsilon=1}^{n_{\beta_1}} \beta_{1_\upsilon} H^\upsilon(x,y)\phi_n\phi_\ell.\end{aligned} \tag{A.8}$$

Hence, (15) can be rewritten as follows

$$\begin{aligned}\dot{a}_\ell(t) &= \sum_{n=1}^{N} a_n(t)\int \left\{ D_x(H)\frac{\partial^2 \phi_n}{\partial x^2} + D_y(H)\frac{\partial^2 \phi_n}{\partial y^2}\right.\\ &\left. -w_x(H)\frac{\partial \phi_n}{\partial x} - w_y(H)\frac{\partial \phi_n}{\partial y} + \sum_{\upsilon=1}^{n_{\beta_1}} \beta_{1_\upsilon} H^\upsilon(x,y)\phi_n \right\}\phi_\ell\\ &\beta_{1_0} a_\ell(t) - 2\sum_{n=1}^{N} a_n(t)\int \beta_2(H)\phi_n\phi_\ell\bar{u}\\ &-\sum_{n=1}^{N} a_n^2(t)\int \beta_2(H)\phi_n^2\phi_\ell\\ &-2\sum_{n=1}^{N-1}\sum_{m=n+1}^{N} a_n(t)a_m(t)\int \beta_2(H)\phi_n\phi_m\phi_\ell\\ &+\int \left\{ D_x(H)\frac{\partial^2 \bar{u}}{\partial x^2} + D_y(H)\frac{\partial^2 \bar{u}}{\partial y^2} - w_x(H)\frac{\partial \bar{u}}{\partial x}\right.\\ &\left. -w_y(H)\frac{\partial \bar{u}}{\partial y} + \beta_1(H)\bar{u} - \beta_2(H)\bar{u}^2 \right\}\phi_\ell\end{aligned} \tag{A.9}$$

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

By considering the vector $\mathbf{a}(t)$, the matrix formulation (16) is obtained where

- $\mathbf{\Gamma}_1$, $\mathbf{\Gamma}_2$ and $\mathbf{b}_1$ are directly deduced from (A.9),
- the simplification (A.8) leads to the term $\beta_{1_0}\mathbf{I}_N\mathbf{a}(t)$ on the right-hand-side of (16) and to the definition of $\mathcal{F}_1(\bullet)$ which presents the particular sum $\sum_{\upsilon=1}^{n_{\beta_1}} \beta_{1_\upsilon} H^\upsilon(x,y)\bullet$ in (17),
- all terms $\int \bullet\phi_n\phi_m\phi_\ell$ due to the square non-linearity in the PDE definition are included in the term $\left(\mathbf{I}_N \otimes \mathbf{a}^\mathsf{T}(t)\right)\mathbf{\Lambda}\ \mathbf{a}(t)$.

References

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

Björck, A. (1996). *Numerical methods for least squares problems*. SIAM.

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. *International Journal of 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., et al. (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.), (2018a). *IPBES-6 plenary*. Medellin, Colombia.

IPBES (2018b). In 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, & L. Willemen (Eds.), *Summary for policymakers of the assessment report on land degradation and restoration of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. Bonn, Germany: IPBES secretariat.

IPBES (2018c). In 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, & M. Christie (Eds.), *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*. Bonn, Germany: IPBES secretariat.

IPBES (2018d). In J. Rice, C. S. Seixas, M. E. Zaccagnini, M. Bedoya-Gaitán, N. Valderrama, C. B. Anderson, T. K. M.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, & J. S. Farinaci (Eds.), *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*. Bonn, Germany: IPBES secretariat.

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

Knudsen, M. (1994). A sensitivity approach for estimation of physical parameters. In *10th IFAC symposium on system identification, Vol. 2* (pp. 231–237).

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*, 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* (2nd ed.). Upper Saddle River, N.J.: Prentice-Hall.

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*. ISR. MD, USA: University of Maryland.

Newman, A. (1996b). *Model reduction via the Karhunen-Loeve expansion, Part II: Some elementary examples. Technical report, T.R. 96-32*. ISR. MD, USA: University of Maryland.

Nocedal, J., & Wright, S. (2006). *Numerical optimization* (2nd ed.). Springer.

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* (2nd ed.). Springer.

Ouvrard, R., & Trigeassou, J. C. (2011). On embedded FIR filter models for identifying continuous-time and discrete-time transfer functions: the RPM approach. *International Journal of 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). *Lecture notes in control and information sciences*, *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* (Thése d'etat), France: Université de Poitiers.

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*. New York: Springer.

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