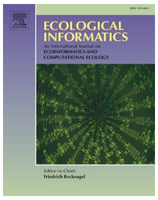
Ecological Informatics 77 (2023) 102232

Contents lists available at ScienceDirect

Ecological Informatics

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



Estimation and analysis of insect population dynamics parameters *via* physiologically based models and hybrid genetic algorithm MCMC methods

Luca Rossini [a,b,*], Octavio A. Bruzzone [c], Stefano Speranza [b], Ines Delfino [d]

[a] *Service d'Automatique et d'Analyse des Systèmes, Université Libre de Bruxelles, Av. F.D. Roosvelt 50, CP 165/55, 1050 Brussels, Belgium*
[b] *Dipartimento di Scienze Agrarie e Forestali, Università degli Studi della Tuscia, Via San Camillo de Lellis snc, 01100 Viterbo, Italy*
[c] *IFAB, CONICET/INTA, Modesta Victoria 4450, San Carlos de Bariloche, Rio Negro, Argentina*
[d] *Dipartimento di Scienze Ecologiche e Biologiche, Università degli Studi della Tuscia, Via San Camillo de Lellis snc, 01100 Viterbo, Italy*

ARTICLE INFO

Keywords:
Parameter estimation
Field monitoring
Least square method
Metropolis hasting algorithm
Insect pest populations

ABSTRACT

Decision support systems are gaining importance in several fields of agriculture, forest, and ecological systems management. Their predictive potential, entrusted to mathematical models, is of fundamental importance to set up opportune strategies to control pests and adversities that may occur and that may seriously compromise the natural equilibria. Among the others, population dynamics is one of the crucial challenges in the field. Despite the scientific community in recent years providing valuable models that faithfully represent terrestrial arthropods populations, such as insects, one of the main concerns is still represented by the parameter estimation. Parameters, in fact, characterise the species and their estimation are often entrusted to dedicated laboratory experiments that require specific equipment and highly qualified personnel. In this study we propose a novel method to estimate the model parameters directly from field data, where experimental activities are less expensive and less time consuming. In this study we propose a combination of least squares methods via genetic algorithms to preliminary evaluate the best parameter values and Markov Chain Monte Carlo approach to obtain their distribution. The algorithm has been tested in the special case of *Drosophila suzukii*, to quantify part of the parameters of an almost validated model in two steps: *i)* a first pseudo-validation using perturbed numerical solutions, and *ii)* a validation using real field data. The results highlighted the potentialities of the algorithm in estimating model parameters and opened several perspectives for further improvements from both the computational and experimental point of view.

1. Introduction

The quantitative interpretation of biological phenomena is a complex process that requires a highly multidisciplinary approach. Models in ecology have been considered a powerful tool to better understand the processes and make decisions from 40 years ago (Conway, 1977), but the advent of personal computers ulteriorly endorsed the increasing interest of the scientific community in mathematically describing biological processes. These efforts have led to the formulation of several models (Orlandini et al., 2018; Sinclair and Seligman, 1996) and of novel numerical methods to support their application and validation with field data (Buffoni et al., 1990; Ratkowsky, 1993; Ratkowsky and Reddy, 2017).

Population dynamics is one of the most recurrent topics of modelling, since it has been applied to humans, animals, plants, and organisms of any kind, including insects. Insects, and more in general terrestrial arthropods, are well described by the so-called physiologically based models (Bellagamba et al., 1987; Cappio Borlino et al., 1991, 1990; di Cola and Gilioli, 1996). These models have the advantage to faithfully describe the insects' biology by considering their ectothermic nature, which makes the stage development highly dependent on environmental conditions (Baumgärtner and Severini, 1987; Gilioli and Pasquali, 2007; Gutierrez et al., 2017; Ponti et al., 2021; Rossini et al., 2022b, 2021a; Severini et al., 1990). Physiologically based models spurred the interest of the scientific community, given their suitable implementation in Decision Support Systems (DSSs) (Lessio and Alma, 2021). DSSs aim to simulate the evolution of pest outbreaks supporting the framework of Precision Agriculture (PA) (Rossi et al., 2010), according to which inputs

[*] Corresponding author at: Service d'Automatique et d'Analyse des Systèmes, Université Libre de Bruxelles, Av. F.D. Roosvelt 50, CP 165/55, 1050 Brussels, Belgium.
E-mail address: luca.rossini@unitus.it (L. Rossini).

https://doi.org/10.1016/j.ecoinf.2023.102232
Received 12 February 2023; Received in revised form 19 May 2023; Accepted 23 July 2023
Available online 26 July 2023

in cultivated fields should be limited as much as possible and provided only where and when they are really needed (Sadovski, 2020). The decision-making process and the formulation of opportune strategies to control pest species infesting cultivations can be strongly endorsed by DSSs, with a subsequent economically and environmentally-friendly management of farms.

Even though a great effort has been made in developing pest population models, several questions are still open (Petrovskii et al., 2012). For instance, most existing models have a versatile structure that makes them suitable to describe the biology of different species after the estimation of a set of parameters for each case of study (Damos et al., 2018; Rossini et al., 2019a, 2019b). Conducting laboratory experiments under controlled conditions is a fundamental step of the model application, although highly time and economic resources consuming. For insects, the normal procedure is to rear populations of a given species under study in climatic chambers where temperature, humidity, photoperiod, and diet are controlled (Deevey, 1947; Govindan and Hutchison, 2020; Harcourt, 1969). This operation allows us to obtain quantitative information on how development, fertility, and mortality (to cite some examples) vary depending on the above-mentioned environmental variables (Naranjo and Ellsworth, 2017). Once we have this quantitative information we can estimate the parameters of specific mathematical functions used in the physiologically based models (Damos et al., 2018), carry out simulations, and validate the outputs through field data (Bellocchi et al., 2011; Ikemoto and Kiritani, 2019; Orlandini et al., 2018; Rossini et al., 2019b). Field data usually consists of a time series containing the daily average values of the environmental variables (e.g., temperature, relative humidity) recorded in the insects' living environment and the count of individuals being in a particular stage (Rossini et al., 2021b, 2020b, 2020c). The counting process is a step of monitoring, usually carried out through traps (e.g., pheromone-based, food-based or chromotropic, depending on the species) or by employing different sampling techniques such as visual inspections or plant shaking (Preti et al., 2021).

Field surveys are affected by high costs in terms of resources (e.g., manpower, materials) and time (Petrovskaya et al., 2012), but they are more convenient than laboratory experiments (Padmanabha and Streif, 2019). Accordingly, it may be reasonable to ask if we can invert the usual process of application of physiologically based models to estimate the biological parameters characterising the species. In other words, if a model is supposed to faithfully represent the biology of a given species, we may use datasets from field surveys to estimate its parameters. A first issue that we identify is that field data are affected by a higher variability than those obtained in the laboratory environment, where the conditions are strictly controlled (Wang and Ma, 2022), with a subsequent possible reduction of the reliability of the estimations. Additionally, the time range between two consecutives samplings is wider in field than in laboratory trials (e.g., one week versus one day, respectively) with a subsequent lower availability of data series for model parameters' estimation. Hence, an algorithm that may consider all these issues would be necessary.

The main challenge in modelling population dynamics is the nonlinear nature of the models, which complicates the research of an optimal and biologically meaningful combination of parameters (i.e.: all the parameters within a range of values which are biologically meaningful) (Quinn, 2017; Shi et al., 2017). Some authors proposed to solve the problem using Markov Chain Monte Carlo families (MCMC) algorithms, that calculate the distribution of the parameters in a fully Bayesian Framework (e.g., Bruzzone and Utgés, 2022; Dorazio, 2016; Gillespie and Golightly, 2010; Heydari et al., 2014; Lanzarone et al., 2017). Other valuable approaches are based on a reversible jump MCMC that helps to automatically find, among a list of candidates, the best-explaining models (e.g., Bruzzone et al., 2018). Despite powerful, pure MCMC approaches are however slow and often require lengthy calculations, making the process slow and tedious.

An alternative method for model parameters estimation is the use of gradient-based numerical optimization to find the optimal values (e.g., Broyden-Fletcher-Goldfarb-Shanno algorithms - BFGS), as in the works of Chau et al. (2014), Forouzanfar and Reynolds (2014), Kegl and Kovač Kralj (2020). These algorithms resulted faster than MCMC but were more prone to get stuck in local minima (Alain and Bengio, 2014; Zhang et al., 2022).

The availability of computational tools is currently endorsing the use of Genetic Algorithms (GA), useful methods to optimize the exploration of the parameters' space. Applications of these methods in ecological contexts are always more frequent (Durgabai et al., 2018), mostly because GA can be combined with different algorithms, such as the Least Squares (LS) method (Song et al., 2012), the AutoRegressive Integrated Moving Average (ARIMA) method (Rathod et al., 2017), or artificial neural networks (Shang and Zhu, 2018).

This study aims to test the combination of two methodologies to cope with the physiologically based model parameter estimation. More specifically, we developed and tested an algorithm that explores the possible combinations of model parameters and provides their best estimate using field monitoring instead of laboratory datasets. We faced the problem of model parameter estimation through a hybrid method that involves the strengths of GA, LS, and MCMC. Particularly, the method is based on *i)* direct or iterative estimation via a genetic algorithm using LS as an estimation method of the error to find an optimal value of the parameters, and *ii)* MCMC approach to find the statistical distribution of parameters. Least squares produce results similar to the Maximum Likelihood Estimation (MLE) under the hypothesis that the statistical distribution of the errors is Gaussian, while the best single estimation via MCMC is usually the MLE estimation of the parameters.

Although these two methodologies are based on different mathematical backgrounds, their joint use can be of great help in the model parameters estimation from field data. This study proposes a workflow that can be applied to most physiologically based models existing in the literature. The basic idea is to use the LS method to preliminary explore the space of the parameters via genetic algorithms, providing a preliminary optimization of the parameter values. Then, the MCMC algorithm will tune the parameter estimation, providing the statistical distribution and, accordingly, the uncertainty associated with every single parameter. For the sake of exposition, we will apply the algorithm to the model of Rossini et al. (2022a, 2021a) in two steps: *i)* a first step concerning a theoretical test on perturbed numerical series, and *ii)* a second step concerning a test using field data. The methodology introduced with this work is totally general and can be extended to any physiologically based model describing insects' stage development.

2. Materials and methods

According to the objectives of the study, this section presents the theoretical background briefly focusing on the eco-physiologically based model used as a case-of-study and highlighting the features that it has in common with other existing models. Subsequently, we introduce the algorithm for parameter estimation and how it has been tested with theoretical and experimental data.

2.1. The physiologically-based model

Although the theoretical workflow introduced with this study aims to be as general as possible, focusing on a specific eco-physiologically based model may simplify the exposition. These types of models that describe populations of terrestrial arthropods are composed of two parts (Severini and Gilioli, 2002). The first part, commonly identified as “phenological models” (Chuine and Régnière, 2017; Rebaudo and Rabhi, 2018), mathematically represents the effect of the environmental parameters on stage development, fertility, and mortality. The second part, commonly identified as “population dynamics” (Severini et al., 1990), mathematically represents the variation over time of the individuals between the different life stages (Bellagamba et al., 1987;

Gutierrez et al., 1984; Rossini et al., 2020d). According to this general vision, physiologically based models take as input a series of parameters required by phenological models, that characterise the species, and a series of environmental parameters' values directly measured (usually on a daily basis) in the insects' living environment.

Over the years, several authors have developed physiologically based models with these features (Ainseba et al., 2011; Cappio Borlino et al., 1990; de Roos et al., 1992; Diekmann et al., 2020; Gutierrez et al., 1994; Holst and Ruggle, 1997; Nance et al., 2018; Otero et al., 2006; Sharov, 1996; Vansickle, 1977; Voulgaris et al., 2013), but for the sake of this study, we will focus on the model of Rossini et al. (2022b, 2021a) without any loss of generality. Referring the most interested readers to the cited literature, let us report only the essential information of the model that is helpful to understand the rationale behind the present study. The life cycle of an insect can be schematized with a series of chained stages, each one being identified by a label i, directly corresponding to the stages biologically defined by entomologists (i.e., egg, larval or nymphal stages, pupa, adult). Each life stage i is associated with a state variable $x_i(t)$, accounting for the corresponding number of individuals over time, resulting in an overall compartmental structure where each life stage is a discrete compartment. The overall flux of individuals coming in and out of the life stages is described by a system of Ordinary Differential Equations (ODEs). The main feature of the general physiologically based model introduced by Rossini et al. (2021a) is the need for a case of study to be particularised. For this reason, we need to focus on a species before presenting its final mathematical formulation.

2.2. The target species, *Drosophila suzukii*

For this study we have chosen the case of the spotted wing drosophila *Drosophila suzukii* (Matsumura), an insect infesting soft fruit cultivations worldwide. The life cycle of this pest species is composed of an egg stage $x_e(t)$, three larval instars $x_{L1}(t)$, $x_{L2}(t)$ and $x_{L3}(t)$, a pupa stage $x_P(t)$, an adult male stage $x_{Am}(t)$, and an adult female stage. Adult females, as already indicated in Rossini et al. (2021a), are in turn divided into non-mated, $x_{Af1}(t)$, and mated, $x_{Af2}(t)$, substages. Considering this subdivision, the resulting physiologically based model is composed of 8 equations, graphically summarised in Fig. 1.

At this point, we need to define the development, $G_i(t)$, mortality, $M_i(t)$, and fertility, $\beta_i(t)$, rate functions involved. The current knowledge of both *D. suzukii* and phenological models allows us to consider only temperature, T, as the main environmental driving variable.

The development rate function considered in this study is the Briére development rate function (Briere et al., 1999), mathematically defined as:

$$G_i[T(t)] = a\,T(t)(T(t) - T_L)\,(T_M - T(t))^{1/m} \tag{1}$$

In eq. (1) a and m are empirical parameters with no biological meaning, while T_L and T_M are the lower and upper temperature thresholds below and above which the development is theoretically not possible, respectively. Based on the dataset available in the current literature (Rossini et al., 2020a; Ryan et al., 2016; Tochen et al., 2014; Wang et al., 2018; Winkler et al., 2021), the following subdivision was considered for the stages i covered by each set of parameters of the function (1): $i = eL$, the set of parameters describing the development from egg to the third larval instar (included), $i = P$, the set of parameters describing the development rate of pupae, and $i = A$, the set of parameters describing adult survival. This subdivision is respected because these literature values can be considered as references to assess the validity of the model hereafter presented.

The temperature-dependent mortality rate is expressed by the (Kim and Lee, 2003) equation, subsequently revised by Son and Lewis (Son and Lewis, 2005):

$$M_i[T(t)] = 1 - \left[k\,exp\left(1 + \frac{T_{MAX} - T(t)}{\rho_T} - exp\left(\frac{T_{MAX} - T(t)}{\rho_T}\right)\right)\right] \tag{2}$$

where k and ρ_T (°C) are empirical parameters, and T_{MAX} (°C) is the temperature where the mortality is lower, namely the abscissa of the minimum of the function (2).

Based on the literature data (Tochen et al., 2014), we consider a single set of parameters for the function (2) to describe the egg to adult stages. According to the theory presented by Rossini et al. (2021a), a further step is needed to correctly represent adult mortality. The latter can be expressed as a combination of additive terms describing the different types of mortality (e.g., survival rate, temperature-dependent mortality, insecticide action, etc.). In this study, we consider two contributions in the adult mortality: the survival rate $G_A(t)$, mathematically expressed by eq. (1), and the mortality, expressed by eq. (2). Mathematically, the mortality of the three adult stages is expressed by:

$$M_A[T(t)] = G_A[T(t)] + M[T(t)] \tag{3}$$

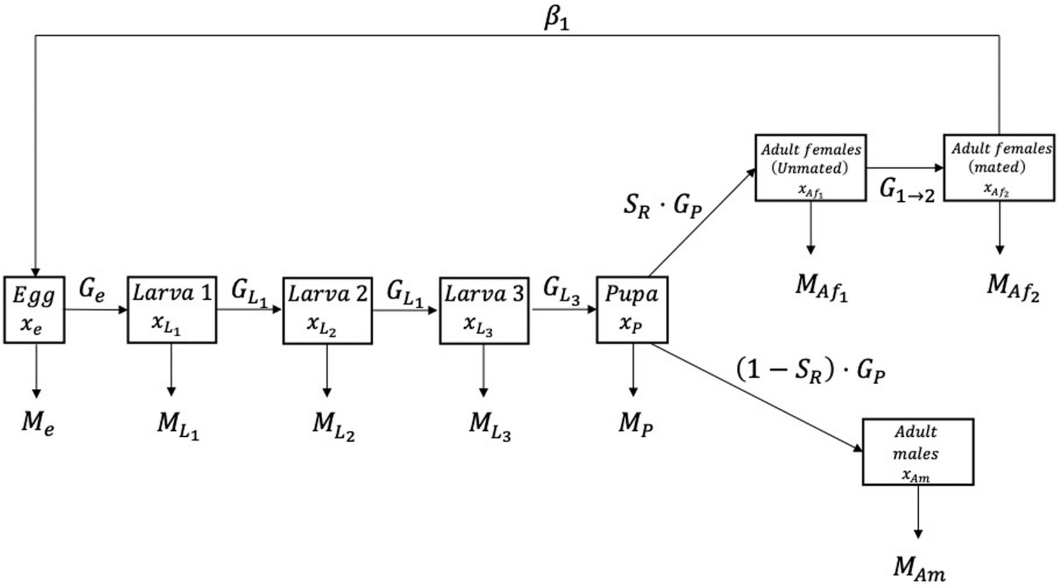


Fig. 1. General scheme of the compartmental eco-physiologically based ODE framework considered in this study. This scheme is a representation of the biological life cycle of *Drosophila suzukii*, considering its sex-division as well.

It is worth reminding that the current state-of-the-art provides quantitative information only about the temperature-dependent mortality. However, wild populations are subject to other sources of mortality (that depends on the life stage) such as natural enemies (e.g., entomopathogenic bacteria, viruses, and fungi, or predators and parasitoids) or a sudden food shortage (that depends on the presence and on the health status of the host plant). Despite relevant, the present study cannot currently consider these factors, implicitly assuming that their effect is limited and included in the uncertainty of the parameters.

Fertility is entrusted to the Gaussian-like function published by (Ryan et al., 2016):

$$\beta[T(t)] = \alpha\left[\frac{\gamma+1}{\pi\,\lambda^{2\gamma+2}}\left(\lambda^2 - \left([T(t)-\tau]^2+\delta^2\right)\right)^{\gamma}\right] \tag{4}$$

where α, γ, λ and δ are empirical parameters, and τ is the optimal temperature (°C) for egg production.

Table 1 summarises the overall model parameters considered for the *Drosophila suzukii* case of study.

To complete the model description, let us consider that in a cultivated field there is a high probability for males and females to mate (Rossini et al., 2022b, 2021a), that is a transition rate from non-mated to mated females of $G_{1\to 2}(t) = 1 - M[T(t)]$. Putting together all the assumptions made, and considering a sex ratio $S_R = 0.5$ (1:1, males: females) (Emiljanowicz et al., 2014) we obtain the final version of the physiologically based model to test the algorithm introduced in the following sections:

$$\frac{d}{dt}x_e(t) = \beta(t)x_{Amf}(t) - G_{eL}(t)x_e(t) - M(t)x_e(t)$$

$$\frac{d}{dt}x_{L1}(t) = G_{eL}(t)x_e(t) - G_{eL}(t)x_{L1}(t) - M(t)x_{L1}(t)$$

$$\frac{d}{dt}x_{L2}(t) = G_{eL}(t)x_{L1}(t) - G_{eL}(t)x_{L2}(t) - M(t)x_{L2}(t)$$

$$\frac{d}{dt}x_{L3}(t) = G_{eL}(t)x_{L2}(t) - G_{eL}(t)x_{L3}(t) - M(t)x_{L3}(t)$$

$$\frac{d}{dt}x_P(t) = G_{eL}(t)x_{L3}(t) - G_P(t)x_P(t) - M(t)x_P(t)$$

$$\frac{d}{dt}x_{Am}(t) = (1-S_R)G_P(t)x_P(t) - G_A(t)x_{Am}(t) - M(t)x_{Am}(t)$$

$$\frac{d}{dt}x_{Af1}(t) = S_RG_P(t)x_P(t) - x_{Af1}(t)$$

$$\frac{d}{dt}x_{Af2}(t) = (1-G_A(t))x_{Af1}(t) - M(t)x_{Af1}(t) - M(t)x_{Af2}(t) - G_A(t)x_{Af2}(t) \tag{5}$$

For the sake of exposition, the explicit dependence of the eqs. (1)–(3) on temperature T has been omitted in the ODE system (4), however it can be exploited by considering, for instance, the following notation: $G_i[T(t)]$, $M_i[T(t)]$, and $\beta[T(t)]$.

2.3. The hybrid MCMC algorithm

After the introduction of the physiologically based model, let us detail the hybrid MCMC algorithm, the main objective of this study. It is worth pointing out that what follows in this section can be adapted to any model having the features described in Section 2.1. The algorithm can be divided into two macro steps: the first one is based on a Least Squares approach for finding the optimal combination of parameters, and the second one on a Metropolis-Hastings-like algorithm to sample the a posteriori distribution of parameters. A schematic representation of the logical steps is detailed in Fig. 2, while we hereafter report the

Table 1
Model parameters considered in this study for the specific case of *Drosophila suzukii*. a^{eL}, m^{eL}, a^P, m^P, a^A, m^A, k, α, γ, λ, δ are adimensional parameters, while temperatures ${T_L}^{eL}$, ${T_M}^{eL}$, ${T_L}^P$, ${T_M}^P$, ${T_L}^A$, ${T_M}^A$, T_{MAX}, ρ_T, and τ are reported in °C. These values are considered as theoretical references to assess the performance of the algorithm.

Model function and life stage	Parameter value	Reference of the dataset of provenance	Perturbed numerical solution		Field data	
			Best value from Least-Squares macro step	Final values after the Metropolis-Hastings like step (mean ± SD)	Best value from Least-Squares macro step	Final values after the Metropolis-Hastings like step (mean ± SD)
Development rate function (1) Egg to pupa	$a^{eL} = 1.59 \cdot 10^{-4}$	Tochen et al. (2014)	$1.23 \cdot 10^{-4}$	$(1.1 \pm 0.3) \cdot 10^{-4}$	$3.02 \cdot 10^{-4}$	$(1.8 \pm 0.9) \cdot 10^{-4}$
	${T_L}^{eL} = 2.09$		1.01	$(9 \pm 5) \cdot 10^{-1}$	0.98	1.2 ± 0.8
	${T_M}^{eL} = 32.08$		28.00	35 ± 7	27.00	37 ± 9
	$m^{eL} = 4.0$		2.51	3 ± 1	2.50	4 ± 2
Development rate function (1) Pupa to adult	$a^P = 2.36 \cdot 10^{-4}$	Tochen et al. (2014)	$1.00 \cdot 10^{-4}$	$(1.1 \pm 0.5) \cdot 10^{-4}$	$1.16 \cdot 10^{-5}$	$(4 \pm 2) \cdot 10^{-5}$
	${T_L}^P = 4.0$		3.87	5 ± 2	3.43	4 ± 2
	${T_M}^P = 33.16$		30.00	39 ± 9	35.00	31 ± 7
	$m^P = 4.0$		4.01	5 ± 2	2.85	3 ± 2
Development rate function (1) Adult survival	$a^A = 6.84 \cdot 10^{-5}$	Tochen et al. (2014)	$3.00 \cdot 10^{-5}$	$(4 \pm 1) \cdot 10^{-5}$	$3.22 \cdot 10^{-5}$	$(4 \pm 2) \cdot 10^{-5}$
	${T_L}^A = -3.0$		−0.99	-1.0 ± 0.6	−1.12	-1 ± 2
	${T_M}^A = 30.03$		28.00	34 ± 9	28.00	20 ± 10
	$m^A = 2.5$		1.50	3 ± 1	2.56	3 ± 1
Mortality rate function (3)	$k = 1.0$ $T_{MAX} = 23.42$ $\rho_T = -5.54$	Ryan et al. (2016)	–	–	–	–
Fertility rate function (4)	$\alpha = 659.06$ $\gamma = 88.53$ $\lambda = 52.32$ $\delta = 6.06$ $\tau = 22.87$	Ryan et al. (2016)	–	–	–	–

Parameter values of the function (1) provided by the best iteration from the least-squares genetic algorithm and by the Metropolis-Hastings like step (mean ± standard deviation), respectively, in the case of the perturbed numerical solution and of the field data. The uncertainties estimated by the LS genetic algorithm were too low to be considered reliable and were not reported.

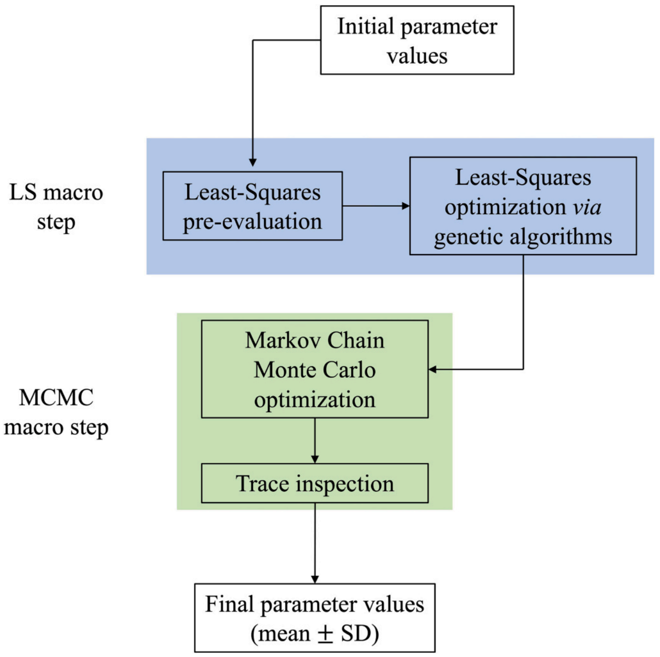

Initial parameter values
LS macro step
Least-Squares pre-evaluation
Least-Squares optimization *via* genetic algorithms
MCMC macro step
Markov Chain Monte Carlo optimization
Trace inspection
Final parameter values (mean ± SD)

Fig. 2. Graphical representation of the logical steps carried out by the hybrid MCMC algorithm.

details. For a clearer exposition and without loss of generality, we will centre the description of the algorithm and the practical example considered as a case of study, to the estimation of the parameters of the function (1) for the egg-pupa, pupa-adult, and adult life stages, that is the three sets of a, T_L, T_M and m. This choice is motivated mainly by the amount of field data available and to not complicate too much the presentation of the case of study.

2.3.1. *The iterative least-squares fit macro step via genetic algorithm*

The Least Squares macro step starts with a combination of initial values for the parameters to estimate, that is the set of initial conditions for the process. These initial parameter values may be based on biological assumptions, empirical measurements, or assigned in an arbitrary manner as well. The process starts by considering that each initial value assigned to the parameters to be estimated at the end of the algorithm has a Gaussian distribution. Accordingly, each initial value is the mean of a Gaussian distribution, μ, while the associated variance is defined as $\sigma^2 = (z\mu)^2$, with $z \in [0,1]$, that is a given portion of the mean value μ. It is worth pointing out that the choice of the Gaussian distribution is not obliged, and that in general any kind of distribution can be considered for each parameter. Additionally, it is possible to choose more refined values for σ^2 based on the biological information available. We have hereafter consider $z = 0.4$ for the LS macro step of the process.

The algorithm starts by choosing a random value for each parameter from its specific Gaussian distribution. The random values are generated by using the *random.normal()* function from the Python 3.6.8 library *numpy*, version 1.19.5. These random values are subsequently taken as input for a classic LS minimization process: the ODE system (5) is numerically solved using the *odeint* function from the Python 3.6.8 library *scipy*, version 1.5.4, and then the solution is compared with field data. In *odeint* we selected the Runge-Kutta 4–5 method as an option.

To be solved, the ODE system (5) needs the array of daily temperature values associated with the field monitoring. The dataset composed of daily average temperatures and trap catches is absorbed by using the Python 3.6.8 library *pandas*, version 1.1.5.

The LS algorithm is operated by the Python 3.6.8 library *lmfit*, version 1.0.3, through the functions *Minimizer()*, *Parameters()*, and *report_fit()*. The numerical solution of the ODE system (5), previously calculated using the random parameters as input, is then compared point-by-point with the field monitoring data. This part of the process applies the Levenberg-Marquardt algorithm, minimising the sum of the residuals between the numerical values of the ODE system (5) and the field data. At the end of every single LS iteration, the set of best-fitting parameter values (vector of parameters), their standard errors (standard deviations vector), and fitting information in the form of the sum of the residuals (pseudo-χ^2 function) are stored in a database consisting in a dedicated Python dictionary. The results of each iteration are saved in a separate database entry. This first step is repeated for an arbitrary number of iterations, in our case set to $n_{LSF} = 192$ based on the number of cores of the machine used for the calculation.

The entries of the database are subsequently ordered from the smaller to the higher sum of the residual values. This operation ends the first step of the LS algorithm and leads to the second part, hereafter defined as the *genetic algorithm*. This is in turn based on an iterative optimization, and its purpose is to partially optimize the results of the estimation and to better explore the space of the parameters. This part of the process is analogous to the previously described step: the first quarter of the best fit parameter values stored in the database is considered as input for the process.

From each combination of values belonging to the first quarter of values stored in the Python dictionary, the GA generates four random combinations of initial values considered as input for the LS procedure previously described. Each best fit value stored in a single row of the dictionary is considered as an expected value of a Gaussian distribution, μ, while the associated variance is still considered as $\sigma^2 = (z\mu)^2$. Each new combination of best fit values calculated during the iterations of the genetic algorithm is again stored in the database together with all the previously estimated sets of values.

At the end of a genetic algorithm cycle, the rows of the database are again ordered according to the values of the sum of the residual. The genetic algorithm can be repeated an arbitrary number of times, but for simplicity we considered a $n_{GAC} = 2$, that leads to a total number of $n_{TGA} = 2^{n_{GAC}+1} \cdot n_{LSF} = 1536$ LS fit evaluations.

2.3.2. *The Metropolis-Hastings-like macro step*

The Metropolis-Hastings-like macro step has the final aim of improving the estimation of the parameters (in particular of their distribution) provided by the GA. For this purpose, it uses the best-fit values estimated through the LS method, previously stored in a database, and ordered by the sum of the residuals. By assumption, the GA is supposed to provide a preliminary optimization of the parameters, so that the MCMC algorithm can consider, as input, only a restricted part of the values stored and ordered in the database. Depending on the number of iterations carried out during the genetic algorithm, n_{TGA}, the MCMC algorithm takes into consideration only the first $n_{MCMC} = n_{TGA}/n_{LSF}$ rows of the database that, in our case, is set to 8.

A single iteration of the MCMC algorithm provides for the following steps. It is selected a random row of the genetic algorithm output database between $[1, n_{MCMC}]$, and the best parameter values are subsequently absorbed. As already described in Section 2.3.1, even in this case the absorbed parameter values are supposed to be the mean of a Gaussian distribution, μ, while in this case the associated variance is defined as a fixed value $\sigma^2 = (0.2\,\mu)^2$. A random value is generated from the Gaussian distribution associated with each parameter by using the *random.normal()* function, then, the ODE system is subsequently solved. At the same time, the step provides for absorbing the array of experimental data, allowing the comparison between the ODE model output and the experimental data, based on the log-probability value.

The comparison between the ODE model output and the experimental data is carried out by considering that every single point of both the ODE solution and the experimental dataset is assumed to have a Poisson a priori distribution. Considering this hypothesis, the algorithm calculates the probability mass distribution value associated with every single experimental point using the function *scipy.stats.poisson.pmf()*

within the Python package *scipy*, version 1.5.4. The probability mass distribution is calculated considering, for each time t_i where the experimental data is available, the corresponding value provided by the ODE and by the experimental datasets. The log-probability value is obtained considering the natural logarithm of the probability mass distribution. All the log-probability values calculated singularly for each time t_i are subsequently added to each other, obtaining the final log-probability value considered as an estimator of the goodness of fit. A dedicated database stores, for each iteration, the set of randomly generated parameters and the final log-probability value.

The single iteration contributes to building the traces as follows: let us define with LP_{i-1} the log-probability value calculated for the iteration $i-1$ and i, respectively. The first iteration of the trace calculates the first LP value and stores the data in the dedicated database. From the second iteration, there is an intermediate step between the calculation of the LP_i value and the beginning of the following iteration, to evaluate which set of values should be stored in the database to ensure the convergence of the trace. A conditional (if-else) statements evaluate the following scenarios:

- If $LP_{i-1} < LP_i$ the algorithm stores in the database the LP_{i-1} value and the set of associated parameters values,
- If $LP_{i-1} > LP_i$ the algorithm stores in the database the LP_i value and the set of associated parameters values.

In this study, a total number of 20,000 iterations for each trace is considered, but in the end, only the values from the 101st to 20,000th iterations are saved. The choice of “burning” the first 100 sets of parameter values estimated in each chain is related to the higher fluctuation of the values during the first iterations. It is however supposed that the LS macro step via GA already provided a first optimization, so that after the first 100 iterations the MCMC-like algorithm is supposed to have reached a more stable convergence, providing final values with a lower uncertainty associated.

Each trace estimated according to the aforementioned process is associated with an independent chain. In this study we have considered simultaneously 384 independent chains run in a parallel algorithm using the Python 3.6.8 library *Ray*, version 1.9.2. The set of values of the traces of each chain is finally stored in a dedicated database and subsequently printed in a .csv file, for further analysis.

Remark 1. In Section 2.3.1 and 2.3.2 we have discussed the two macro-steps that constitute the core of the hybrid genetic algorithm. Besides using two different methods, the two steps compare iteratively the model output and the experimental data. In line of principle the comparison between the ODE model output and the experimental data can be carried out considering all the life stages of the insect. In other words, we may compare the output of each single ODE with the experimental data series corresponding to the specific life stages. This scenario is usually too optimistic, given that field monitoring is carried out only referring to a particular life stage. For this reason, we have oriented the algorithm to compare only the result of a single ODE with the corresponding stage monitored during the surveys that, in the case of D. suzukii, is the adult males $\mathrm{x_{Am}(t)}$. This choice does not affect the generality of the method since the algorithm can be easily changed to iteratively compare multiple stages if field data is available.

Remark 2. In this study we do not consider immigration and emigration fluxes of individuals, an additional aspect that may be relevant in field populations. This aspect concerns the theoretical framework of the model (5) and does not affect the development of the hybrid MCMC algorithm. The absence in the model of immigration and emigration rates implicitly means that: i) the population is locally in equilibrium, that is the number of outcoming and incoming individuals is perfectly balanced at each time step, and ii) the number of adult males trapped during the monitoring (see Section 2.6) is negligible with respect to the total population.

2.3.3. *Visual inspection of the traces and final adjustments*

The hybrid MCMC algorithm provides a series of independent chains containing a trace of values for each parameter to estimate. The traces corresponding to each chain are stored in specific text files and can be further analysed to obtain the final parameter values, their distribution, and their associated uncertainty. The key point of this phase of the process is to evaluate if each chain provides a set of “best” values that are “suitable” to represent, once inserted into the model, the field dataset. In case different chains lead to a set of best-fitting parameters that faithfully represent the field data, the corresponding traces can be merged to obtain the final value as the mean of their values and the uncertainty as the standard deviation.

The selection of the “best candidate” traces to merge is entrusted to a visual inspection of every single chain. A dedicated script selects separately each chain contained in the final text file and absorbs the traces associated with each parameter to estimate. The best fit value of each parameter is obtained from the chain as a mean of the values contained. After the set of best fit values obtained by the single chain under inspection is calculated, they are inserted into the model (5) and the solution is graphically represented together with the experimental dataset. If the overlap between simulations and field data is correct after a visual inspection of the simulated and the real values, the chain is stored in a dedicated file, otherwise, the chain is deleted. After selecting the best representative traces from visual inspections, the statistical distributions of the selected traces were compared to ensure that they are similar (i.e., the simulations converged to the same set of values).

The final set of parameters, accordingly, is calculated by considering this second dataset. Thus, the a posteriori distribution of the parameters is generated by merging all the traces of each of the correct simulations in a single database. The statistical distribution of the parameters is reported by its mean and the standard deviation of the traces of the database of MCMC simulations.

2.4. *Computing tools*

All the calculations were carried out using the DAFNE HPC scientific computing centre of the Università degli Studi della Tuscia. The system provides for two Hewlett Packard Enterprise (HPE) ProLiant DL560 Gen10 nodes, each one equipped with: four processors Intel Xeon Gold 5118 2.30GHz, 12 cores, 24 threads; and 512 GB of RAM. The two nodes worked in a parallel configuration managed by the Python 3.6.8 package *Ray*, version 1.9.2. All the scripts and dataset to fully reproduce the results of this work are publicly available at https://github.com/lucaros1190/LS-MCMC-hybridGenAlgo.

2.5. *Preliminary analysis with perturbed numerical solutions and given parameters*

Before testing the algorithm with real field data, we carried out a preliminary test considering a perturbed solution, obtained by assigning known parameters to the model (5). We considered the parameters listed in Table 1 to solve the model (5) and the numerical solutions were stored in a dedicated file. To reproduce a more realistic situation, we selected only the numerical solution corresponding to the adult male stage. Each value has been considered as the mean μ of a Gaussian distribution with a variance $\sigma^2 = (0.2 \cdot \mu)^2$. The series of perturbed points were randomly generated through the *random.normal()* function and stored in a separate file. Given that the usual sampling time of field surveys is one week circa, once obtained the pseudo-experimental dataset we removed some points, so that only one point for every in 7 was left in the array.

The purpose of this preliminary analysis was to understand if the hybrid MCMC algorithm was capable of correctly estimating the known values listed in Table 1. As stated in Section 2.3.1, the algorithm needs as input the set of parameters of the functions not involved in the process of fitting, the daily average temperature series, the previously obtained perturbed solution (only the adult males), and a set of initial values for the parameters to estimate with the process. The initial values were randomly assigned with no particular rationale, as it is common in a real case where the values of the parameters are usually unknown. Table 2 reports all the input values provided to the hybrid MCMC algorithm to carry out this part of the study.

The parameters a, T_L, T_M, and m of the Briére function (1) were provided by Rossini et al. (2021a, 2020a) together with their associated errors. For this phase of the study, we can consider the Briére parameters listed in Table 1 as reference values and compare them with the values provided by the hybrid MCMC algorithm.

2.6. Application to real field data

The second part of the hybrid MCMC algorithm test and application provided for the use of real field data. For this purpose, we have considered part of the dataset published by Rossini et al. (2021a, 2020a). The survey was carried out in a cherry orchard located in two municipalities of Central Italy and covered the growing seasons 2017–2019. For the sake of this study, we will focus only on a part of the aforementioned dataset, selected for the higher number of points and for the better suitability of the hybrid MCMC algorithm application. The period covered by the selected portion of the dataset, moreover, was overlapped with the presence of fruits in the orchard.

The season of interest is 2018, in particular the dataset referred to the experimental orchard located in the municipality of Montelibretti (Lazio, Central Italy). The orchard covered a surface of 2000 square metres and *D. suzukii* populations were monitored through three Droso-Trap (Biobest,Waterloo Belgium) lured with Droskidrink (Azienda Agricola Prantil, Priò, Trento, Italy). Traps remained in the field from 19th April to 12th December and were inspected weekly. During each sampling the content on the traps was analysed counting only the adult males, given their easier recognizability because of the black spots on the wings. The number of individuals of the experimental population was obtained by considering the mean value of the number of males assessed in each sampling date.

The daily average temperatures were measured by a meteorological station close to the field and managed by the ARSIAL agency (Regional Agency for the Development of Innovation and Agriculture in Lazio) (ARSIAL, 2019). The station acquired 24 temperature values in 24 h, so that the daily temperature array inserted as input in the hybrid MCMC algorithm was obtained by averaging the acquisition of each single day. Besides daily average temperature, the other inputs provided to the algorithm were the Briere's values listed in Table 2 and the following initial population values: $x_e(0) = 50$, $x_{L1}(0) = x_{L2}(0) = x_{L3}(0) = x_P(0) = x_{Am}(0) = x_{Anmf}(0) = 0$, and $x_{Amf}(0) = 97$.

Table 2
set of values provided as input to the hybrid MCMC algorithm in the preliminary analysis using perturbed numerical solutions as "field data". a and m are adimensional parameters, temperatures T_L and T_M are measured in °C.

Model function and life stage	Input parameter values
Development rate function (1) Egg to pupa	$a^{eL} = 3.30 \cdot 10^{-4}$ $T_L{}^{eL} = 1.0$ $T_M{}^{eL} = 29.0$ $m^{eL} = 2.9$
Development rate function (1) Pupa to adult	$a^P = 1.20 \cdot 10^{-4}$ $T_L{}^P = 5.0$ $T_M{}^P = 31.0$ $m^P = 4.0$
Development rate function (1) Adult survival	$a^A = 2.8 \cdot 10^{-5}$ $T_L{}^A = -1.0$ $T_M{}^A = 30.0$ $m^A = 3.7$
Mortality rate function (3)	Same as Table 1
Fertility rate function (4)	Same as Table 1

3. Results and discussion

3.1. Results of the preliminary analysis with perturbed numerical solutions and given parameters

A graphical representation of the results is provided in Fig. 3, while numerical results are listed in Table 1. Table 1 reports the best fit parameters estimated by the LS macro step, namely the set listed in the first row of the dataset, and the resulting distribution of values after the visual inspection of the traces. A total of 67 traces were merged to obtain the result of this part of the study. The uncertainties associated with the results of the LS macro step, however, were too small (at least two orders of magnitude less the expected value) and were not listed. The estimation of the distributions of the parameters, instead, has been entrusted to the MCMC macro step, and this is the reason why we reported the parameter values with their uncertainties only in the MCMC column of Table 1.

The preliminary analysis provided promising results, showing how the LS macro step correctly explored the space of the parameters, while the MCMC macro step better estimated their distribution. Overall, the simulation carried out considering the best fitting values estimated through the hybrid MCMC algorithm better represented the perturbed numerical solution, above all on the left side of the population peak (Fig. 3). As time increases, the best fitting solution tends to overestimate the population dataset of reference. Despite this overestimation for larger times, the hybrid MCMC algorithm was capable of providing results in accordance with the theoretical values listed in Table 1. Differences between theoretical and estimated parameters were assessed only on a few parameters, that is a^{eL}, T^{eL}, a^P, a^A. This difference may be responsible for the overestimation of the model observed for large times and underlines a fundamental aspect worthy of discussion.

In this study, in fact, we have provided random initial conditions in input to the hybrid MCMC algorithm with no limitation for the parameters. It is however known (Johnson and Frasier, 1985) that above all for LS fits it is possible to bind the value of each parameter to a specific range. Even though it is often difficult to have an estimation of the range of values for each single parameter, this information in some cases can be obtained in alternate ways. Let us take as an example the Briére development rate function (1) considered in this study. Among the four parameters, two of them (a and m) are empirical with no biological meaning, while the temperatures T_L and T_M represent the lower and upper temperature bounds above and below which the development of the species is not theoretically possible (Briere et al., 1999). Information about the thermal limits is usually obtained through repeated constant temperatures experiments in growth chambers (Garcia-Robledo et al., 2020), but they can also be roughly estimated considering the average weather conditions of the areas of interest measured during the monitoring survey.

Accordingly, collecting field data together with daily average temperatures (and with the other environmental parameters in case of more refined models) may provide a rough estimation of the lower and upper thermal thresholds for the development of the species, so that this information can be used to fix the bounds to the parameters in the LS process. A similar limitation can be applied to the MCMC macro step, in particular discarding the random values outside the given range. An approach of this type may surely be beneficial to increase the precision of the parameter estimation, and we deserve, in future works, to better explore this aspect in the light of ad hoc experimental trials.

An additional advantage of the perturbed numerical solution, with respect to a real case, is the number of data available from monitoring. Measurements in pest population dynamics (and for most measurements in biology), are often sparse and with large dead band zones, making the estimation of parameters for modelling purposes difficult (Petrovskaya et al., 2012). Accordingly, considering an intermediate "theoretical" step would be beneficial to have an estimation of how the algorithm behaves and fits with the more theoretical purpose of this study.

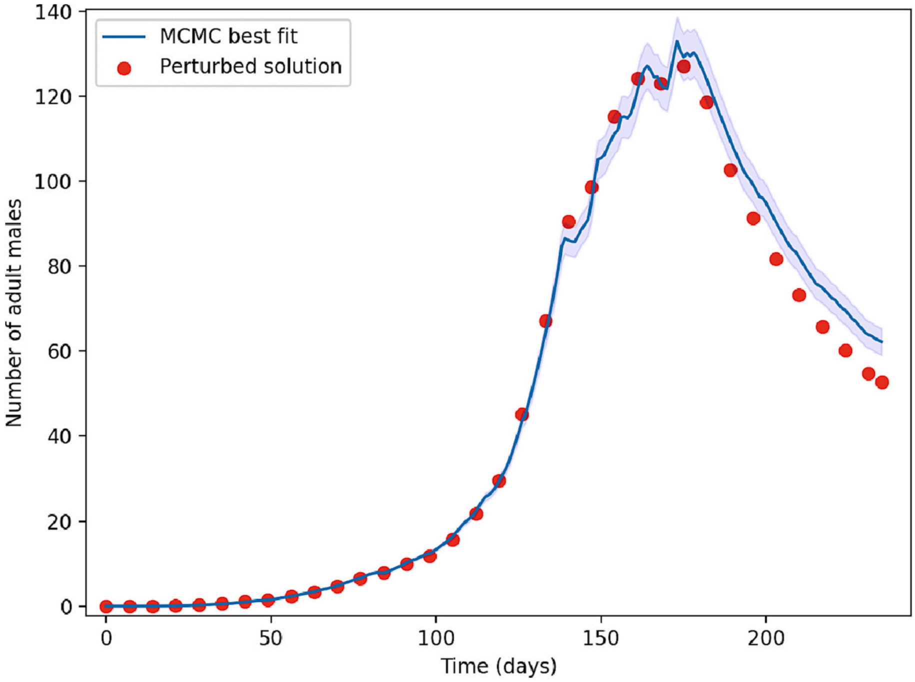


Fig. 3. Comparison between the model outputs and the perturbed numerical solution used as reference data in the preliminary evaluation of the method. The blue line represents the best fit solution resulting from the hybrid MCMC algorithm (see Table 1 for numerical values), while the blue shaded band represents the 95% confidence range.

In addition, the main assumption behind the methodology we introduced is that the model faithfully represents the life cycle of the species under study. As stated in the introduction, we have considered the specific case of *D. suzukii* described by the model of Rossini et al. (2021a), but this choice was only for illustrative purposes. This scheme is generally valid, and the code provided as supplementary material can be modified to estimate a wider set of parameters or with a more refined physiologically based model.

From this part of the results, hence, we can identify the strength of merging two apparently independent algorithms (LS via GA and MCMC), that is an overall better estimation of the expected values and of the distribution (and of the uncertainty associated with parameters, accordingly). On the other hand, there is a weak point that can surely be a great starting point for future works, namely, to understand until

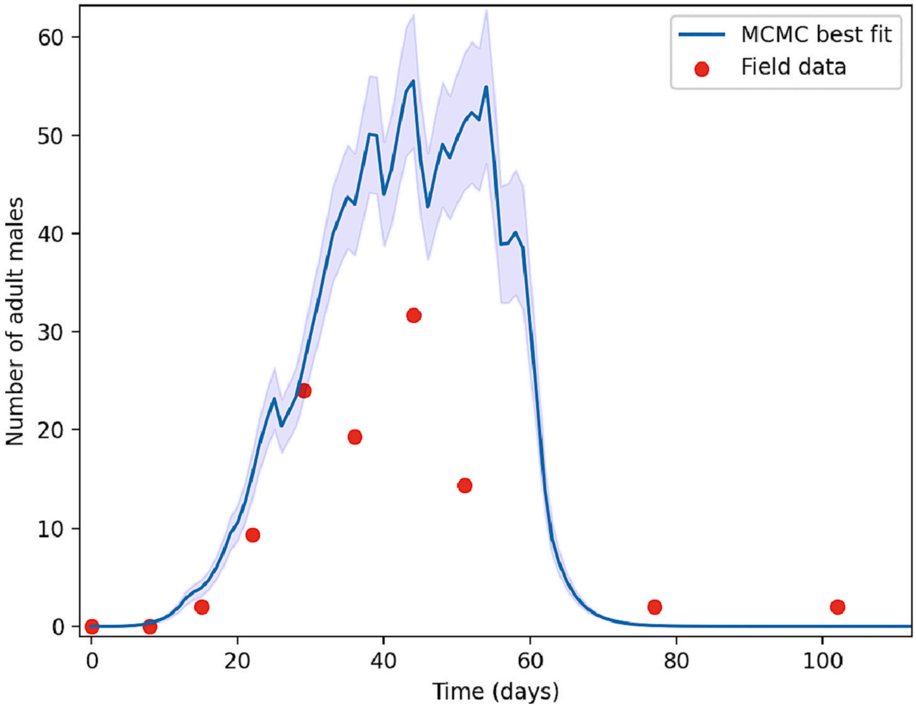


Fig. 4. Comparison between the model outputs and the field data about *D. suzukii*. The blue line represents the best fit solution resulting from the hybrid MCMC algorithm (see Table 1 for numerical values), while the blue shaded band represents the 95% confidence range.

which point the biological phenomenon is well represented by the model.

3.2. Results of the application to real field data

The test of the algorithm with field data confirmed what was already reported in Section 3.1. Even in this case, the LS macro step correctly explored the space of the parameters, while the MCMC macro step provided a better estimate of their distribution. This part of the results is graphically reported in Fig. 4, while the best fit parameters from both the LS macro step and the MCMC macro step are listed in Table 1.

The results, in this case, were obtained by merging a total of 117 chains. As already stated in Section 3.1, the errors estimated in the LS macro step were too small to be considered "reliable", for this reason were not reported in this case as well. Differently to the case of the perturbed numerical solutions, however, field data increased the variability of the parameters, highlighted by a higher standard deviation associated with the final values (Table 1).

Overall, there is only one parameter, a^P, that is not confident with the theoretical values of reference listed in Table 1. The overall overlap between the experimental adult male population and the best fit solution is good in the left side of the plot in Fig. 4, but as the simulated population approaches the peak there is an increase of variability and an overestimation of the abundance of adult males. Moreover, the best fit solution decreases with a positive time shift of about 20 days with respect to the experimental population.

As already stated in Section 3.1 for the case of the perturbed numerical solution, the use of real field data collected as usual in entomological monitoring surveys may affect the estimation of the parameters. In particular, we identify two main issues. The first issue is related to the number of data points that are often low to reach a high precision of the algorithm. To overcome this issue, ad hoc surveys should be organised considering a more frequent sampling. This is an aspect that can be easily solved, given the high number of automated monitoring tools, such as "smart traps" (Chulu et al., 2019; Lippi et al., 2022, 2021; Potamitis et al., 2017; Preti et al., 2021) that have been recently introduced on the market and on which development the scientific community is quickly advancing. The use of automated monitoring tools still highlights the convenience of working with data from field surveys instead of setting constant temperature experiments in growth chambers. Alternatively, it can be convenient to increase the number of stages monitored. Traps are usually referred to as the adult stage, but there are other sampling techniques, such as the visual inspections of random fruits, that may give other information about the dynamics of the preimaginal stages. Having the temporal dynamics of two or more stages may increase the reliability of the results, and it is surely an additional point to explore in future studies. In this study we have assumed that the portion of the field monitored is either isolated (i.e., without immigration/emigration processes) or that the process is balanced (i.e., same emigration and immigration rates), resulting in an overall zero migration process from and to neighbouring areas. If this condition is not verified, the trap counts might also be affected by the immigration/emigration rates of individuals from nearby areas. Trap counts can be influenced by the presence of fruits in the field as well, potentially increasing the efficiency of the trap (when there are no fruits in the field) or establishing a competition (where fruits are in the field). These aspects are currently not considered by the model, and their further inclusion can improve the efficiency of the parameter estimation of the hybrid MCMC algorithm. An additional approach using a sequential MCMC method, as in Bruzzone et al. (2023), or a Kalman filter (Bono Rossello et al., 2022), might help to identify when discrepancies between the observed and the modelled data occur. However, this approach needs much more data than is usually available in insect population studies.

The second issue that is worth discussing, in light of the results, is related to the high variability of the field conditions. As already stated in other papers (Bonsignore et al., 2019; Caffarra et al., 2012; Castex et al., 2018; Colinet et al., 2015), insect populations developing in natural environments are subject to a plethora of conditions that are often difficult to consider and control as well. Accordingly, the estimation of the parameters of the eq. (1) using data that may be affected by other uncontrolled parameters may reduce the reliability of the results obtained. To overcome this issue, maintaining an economic advantage from the experimental point of view, surveys conducted in semi-field environments (such as greenhouses, for instance), may be a solution. Future works should be oriented in this direction to assess, at different levels, the reliability of the hybrid MCMC algorithm in an ad hoc experimental condition.

4. Conclusions

This work presented a novel estimation algorithm that may partially substitute, through a combination of in silico and field methods, the hard work of constant temperature rearing for life tables building purposes. Additionally, this method can strongly support the parameter estimation of physiologically based models directly from field data, where the economic sustainability of the experiments is higher. Overall, the results are promising, as already shown by Chen and Gao (2010), and the methodology can be adapted to any other model with the features described in this paper. Accordingly, this study may be of great support for many other research groups working on model development and application, in light of the lack of experimental data for most of the species of agriculture and forest interest.

The use of genetic algorithms in ecology and agriculture sciences is quickly growing, and the results we showed are an additional confirmation of their utility. A similar study where the authors evaluate the combination between GA and LS was carried out by Song et al. (2012) to analyse the concentration of chlorophyll in water sources from satellite spectra. Even if the context is different from insect population dynamics, the goal of estimating parameters from experimental datasets was the same. In that sense, our study enriches the literature on this aspect, laying the foundations for further improvements.

Genetic algorithms were also used in ecology to compare and combine the outputs of different species distribution models, as in Safaei et al. (2021), or to estimate the parameters of a matrix population model of palm plants combining GA and bisection methods (Cropper et al., 2012). More entomological applications of GA were carried out by Florentino et al. (2014), that applied GA to minimize the parameters of a model describing Dengue epidemics and the biology of its vector, and by Yu et al. (2018), that combined GA with artificial neural networks to predict pest infestations. Despite these applications falling in the same topic of our study, to the best of our knowledge this is the first time that a combination of GA, LS, and MCMC is applied to estimate the parameters of physiologically based models describing populations of terrestrial arthropods.

Using the model itself to estimate parameters is not new, and many other studies have been carried out by different authors (Gilioli and Pasquali, 2007; Gillespie and Golightly, 2010; Heydari et al., 2014; Lanzarone et al., 2017; Nance et al., 2018; Pasquali et al., 2022). The difference we introduced with this study, if compared with literature works and with more theoretical papers, is a method that is a compromise to face a problem that is well-known among the entomological and ecological community. On one hand, in fact, the hybrid-MCMC algorithm uses two methods of parameter estimation that have different theoretical backgrounds (LS and MCMC) to exploit their complementary strengths. The LS method showed its validity in initially exploring the space of the parameters, but as many "try and try" methods it is conditioned by the choice of the initial parameter values to start the algorithm. Additionally, the errors are often underestimated, and this is the reason why we proposed to use MCMC to estimate the final distribution.

The greater effort that may be done in further studies is the set-up of proper experimental protocols to increase the quality of the acquired field data to allow a more reliable estimation of the parameters. Although monitoring insect populations is an expensive practice (Petrovskaya et al., 2012), it is more convenient than climatic chambers rearing for different reasons. The first reason is related to the lab equipment often required to set up classical life table experiments, the maintenance that these instrumentations require, to the experience of the researchers, and the space available in the research centres (Padmanabha and Streif, 2019). The second reason concerns the genetic variability and possible adaptation phenomena that populations continuously reared under controlled conditions may suffer (Sørensen et al., 2012). This aspect is relevant, given that it may be the cause of gaps between the response of the individuals to the external environment with a subsequent loss of reliability of the models in representing populations developing under natural conditions. The third reason, instead, is related to all the species that are difficult to rear under laboratory conditions because they have a strong dependence on the host plant, for instance (Boller and Chambers, 1977; Leppla, 2009).

This study aimed to lay the foundations to partially face these problems that affect the development of physiologically based models, by providing a method that can be of great support for their further development and optimization.

Declaration of Competing Interest

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

Data availability

Data are available on a dedicated GitHub page

Acknowledgements

The authors are grateful to the anonymous reviewers for their comments and suggestions, which have been greatly helpful for the improvement of this manuscript. L.R. is funded by Italian MUR (Ministry of University and Research) in the framework of the European Social Funding REACT-EU – National Program for the Research and Innovation 2014–2020 and by the European Commission under the Grant n. 101102281, Project “PestFinder”, call HORIZON-MSCA-2022-PF-01.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoinf.2023.102232.

References

Ainseba, B., Picart, D., Thiéry, D., 2011. An innovative multistage, physiologically structured, population model to understand the European grapevine moth dynamics. J. Math. Anal. Appl. 382, 34–46. https://doi.org/10.1016/j.jmaa.2011.04.021.

Alain, G., Bengio, Y., 2014. What regularized auto-encoders learn from the data-generating distribution. J. Mach. Learn. Res. 15, 3743–3773.

ARSIAL, 2019. SIARL - Servizio Integrato Agrometeorologico della Regione Lazio [WWW Document]. URL. http://www.arsial.it/portalearsial/agrometeo/index.asp.

Baumgärtner, J., Severini, M., 1987. Microclimate and arthropod phenologies: The leaf miner *Phyllonorycter blancardella* F. (Lep.) as an example. In: Prodi, F., Rossi, F., Cristoferi, G. (Eds.), International Conference on Agrometeorology. Fondazione Cesena Agricultura Publ, Cesena, pp. 225–243.

Bellagamba, V., di Cola, G., Cavalloro, R., 1987. Stochastic models in fruit-fly population dynamics. In: Proceedings of the CEC/IOBC International Symposium “Fruit Flies of Economic Importance 87.”, pp. 91–98.

Bellocchi, G., Rivington, M., Donatelli, M., Matthews, K., 2011. Validation of biophysical models: issues and methodologies. In: Sustainable Agriculture, vol. 2. Springer, Netherlands, Dordrecht, pp. 577–603. https://doi.org/10.1007/978-94-007-0394-0_26.

Boller, E.F., Chambers, D.L., 1977. Quality Aspects of Mass-Reared Insects, in: Biological Control by Augmentation of Natural Enemies. Springer US, Boston, MA, pp. 219–235. https://doi.org/10.1007/978-1-4684-2871-1_7.

Bono Rossello, N., Rossini, L., Speranza, S., Garone, E., 2022. State estimation of pest populations subject to intermittent measurements. IFAC-PapersOnLine 55, 135–140. https://doi.org/10.1016/j.ifacol.2022.11.128.

Bonsignore, C.P., Vono, G., Bernardo, U., 2019. Environmental thermal levels affect the phenological relationships between the chestnut gall wasp and its parasitoids. Physiol. Entomol. 44, 87–98. https://doi.org/10.1111/phen.12280.

Briere, J.-F., Pracros, P., Le Roux, A.-Y., Pierre, J.-S., 1999. A novel rate model of temperature-dependent development for arthropods. Environ. Entomol. 28, 22–29. https://doi.org/10.1093/ee/28.1.22.

Bruzzone, O.A., Utgés, M.E., 2022. Analysis of the invasion of a city by *Aedes aegypti* via mathematical models and Bayesian statistics. Theor. Ecol. 15, 65–80. https://doi.org/10.1007/s12080-022-00528-y.

Bruzzone, O.A., Logarzo, G.A., Aguirre, M.B., Virla, E.G., 2018. Intra-host interspecific larval parasitoid competition solved using modelling and bayesian statistics. Ecol. Model. 385, 114–123. https://doi.org/10.1016/j.ecolmodel.2018.07.011.

Bruzzone, O.A., Perri, D.V., Easdale, M.H., 2023. Vegetation responses to variations in climate: a combined ordinary differential equation and sequential Monte Carlo estimation approach. Ecol. Inform. 73, 101913 https://doi.org/10.1016/j.ecoinf.2022.101913.

Buffoni, G., di Cola, G., Ugolini, A., 1990. Numerical methods for the solution of PDE describing the stochastic development of an age-structured population. In: Computer Science and Mathematical Methods in Plant Protection - Proceeding of the International Workshop.

Caffarra, A., Rinaldi, M., Eccel, E., Rossi, V., Pertot, I., 2012. Modelling the impact of climate change on the interaction between grapevine and its pests and pathogens: European grapevine moth and powdery mildew. Agric. Ecosyst. Environ. 148, 89–101. https://doi.org/10.1016/j.agee.2011.11.017.

Cappio Borlino, A., di Cola, G., Marras, G., 1990. I modelli compartimentali nello studio della dinamica delle popolazioni naturali. Bollettino della società sarda di scienze naturali 27, 77–114.

Cappio Borlino, A., di Cola, G., Marras, G., 1991. Mathematical modelling of natural population dynamics. Memorie dell’Istituto Italiano di Idrobiologia 49, 127–162.

Castex, V., Beniston, M., Calanca, P., Fleury, D., Moreau, J., 2018. Pest management under climate change: the importance of understanding tritrophic relations. Sci. Total Environ. 616–617, 397–407. https://doi.org/10.1016/j.scitotenv.2017.11.027.

Chau, M., Fu, M.C., Qu, H., Ryzhov, I.O., 2014. Simulation optimization: a tutorial overview and recent developments in gradient-based methods. In: Proceedings of the Winter Simulation Conference 2014. IEEE, pp. 21–35. https://doi.org/10.1109/WSC.2014.7019875.

Chen, C., Gao, W., 2010. Estimating parameter uncertainties using hybrid Monte Carlo-Least Squares support vector machine method. In: 2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR 2010). IEEE, pp. 89–92. https://doi.org/10.1109/CAR.2010.5456735.

Chuine, I., Régnière, J., 2017. Process-based models of phenology for plants and animals. Annu. Rev. Ecol. Evol. Syst. 48, 159–182. https://doi.org/10.1146/annurev-ecolsys-110316-022706.

Chulu, F., Phiri, J., Nyirenda, M., Kabemba, M.M., Nkunika, P., Chiwamba, S., 2019. Developing an automatic identification and early warning and monitoring web based system of fall army worm based on machine learning in developing countries 1. Zambia Inform. Commun. Technol. J. 3, 13–20.

Colinet, H., Sinclair, B.J., Vernon, P., Renault, D., 2015. Insects in fluctuating thermal environments. Annu. Rev. Entomol. 60, 123–140. https://doi.org/10.1146/annurev-ento-010814-021017.

Conway, G.R., 1977. Mathematical models in applied ecology. Nature 269, 291–297.

Cropper, W.P., Holm, J.A., Miller, C.J., 2012. An inverse analysis of a matrix population model using a genetic algorithm. Ecol. Inform. 7, 41–45. https://doi.org/10.1016/j.ecoinf.2011.06.002.

Damos, P.T., Stoeckli, S.C., Rigas, A., 2018. Editorial: current trends of insect physiology and population dynamics: modeling insect phenology, demography, and circadian rhythms in variable environments. Front. Physiol. 9 https://doi.org/10.3389/fphys.2018.00336.

de Roos, A.M., Diekmann, O., Metz, J.A.J., 1992. Studying the dynamics of structured population models: a versatile technique and its application to *Daphnia*. Am. Nat. 139, 123–147. https://doi.org/10.1086/285316.

Deevey, E.S., 1947. Life tables for natural populations of animals. Q. Rev. Biol. 22, 283–314. https://doi.org/10.1086/395888.

di Cola, G., Gilioli, G., 1996. Mathematical models for age-structured population dynamics: an overview. In: 20th International Congress of Entomology, Florence, pp. 45–61.

Diekmann, O., Gyllenberg, M., Metz, J.A.J., 2020. Finite dimensional state representation of physiologically structured populations. J. Math. Biol. 80, 205–273. https://doi.org/10.1007/s00285-019-01454-0.

Dorazio, R.M., 2016. Bayesian data analysis in population ecology: motivations, methods, and benefits. Popul. Ecol. 58, 31–44. https://doi.org/10.1007/s10144-015-0503-4.

Durgabai, R.P.L., Bhargavi, P., Jyothi, S., 2018. Pest management using machine learning algorithms: a review. Int. J. Comput. Sci. Eng. Inform. Technol. Res. 8, 13–22.

Emiljanowicz, L.M., Ryan, G.D., Langille, A., Newman, J., 2014. Development, reproductive output and population growth of the fruit fly pest *Drosophila suzukii* (Diptera: Drosophilidae) on artificial diet. J. Econ. Entomol. 107, 1392–1398. https://doi.org/10.1603/ec13504.

Florentino, H.O., Cantane, D.R., Santos, F.L.P., Bannwart, B.F., 2014. Multiobjective genetic algorithm applied to dengue control. Math. Biosci. 258, 77–84. https://doi.org/10.1016/j.mbs.2014.08.013.

Forouzanfar, F., Reynolds, A.C., 2014. Joint optimization of number of wells, well locations and controls using a gradient-based algorithm. Chem. Eng. Res. Des. 92, 1315–1328. https://doi.org/10.1016/j.cherd.2013.11.006.

Garcia-Robledo, C., Kuprewicz, E.K., Dierick, D., Hurley, S., Langevin, A., 2020. The affordable laboratory of climate change: devices to estimate ectotherm vital rates under projected global warming. Ecosphere 11. https://doi.org/10.1002/ecs2.3083.

Gilioli, G., Pasquali, S., 2007. Use of individual-based models for population parameters estimation. Ecol. Model. 200, 109–118. https://doi.org/10.1016/j.ecolmodel.2006.07.017.

Gillespie, C.S., Golightly, A., 2010. Bayesian inference for generalized stochastic population growth models with application to aphids. J. R. Stat. Soc. Ser. C Appl. Stat. 59, 341–357. https://doi.org/10.1111/j.1467-9876.2009.00696.x.

Govindan, Hutchison, 2020. Influence of temperature on age-stage, two-sex life tables for a Minnesota-acclimated population of the brown marmorated stink bug (*Halyomorpha halys*). Insects 11, 108. https://doi.org/10.3390/insects11020108.

Gutierrez, A.P., Baumgärtner, J., Summers, C.G., 1984. Multitrophic models of predator-prey energetics: i. age-specific energetics models – pea aphid *Acyrthosiphon pisum* (Homoptera: Aphidae) as an example. Can. Entomol. 116, 923–932. https://doi.org/10.4039/Ent116923-7.

Gutierrez, A.P., Mills, N.J., Schreiber, S., Ellis, C.K., 1994. A physiologically based tritrophic perspective on bottom-up-top-down regulation of populations. Ecology 75, 2227–2242. https://doi.org/10.2307/1940879.

Gutierrez, A.P., Ponti, L., Gilioli, G., Baumgärtner, J., 2017. Climate warming effects on grape and grapevine moth (*Lobesia botrana*) in the Palearctic region. Agric. For. Entomol. https://doi.org/10.1111/afe.12256.

Harcourt, D.G., 1969. Development and use of life tables in study of natural insect populations. Annu. Rev. Entomol. 14, 175. https://doi.org/10.1146/annurev.en.14.010169.001135.

Heydari, J., Lawless, C., Lydall, D.A., Wilkinson, D.J., 2014. Fast Bayesian parameter estimation for stochastic logistic growth models. Biosystems 122, 55–72. https://doi.org/10.1016/j.biosystems.2014.05.002.

Holst, N., Ruggle, P., 1997. A physiologically based model of pest–natural enemy interactions. Exp. Appl. Acarol. 21, 325–341. https://doi.org/10.1023/A:1018415509349.

Ikemoto, T., Kiritani, K., 2019. Novel method of specifying low and high threshold temperatures using thermodynamic SSI model of insect development. Environ. Entomol. 48, 479–488. https://doi.org/10.1093/ee/nvz031.

Johnson, M.L., Frasier, S.G., 1985. [16] Nonlinear Least-Squares Analysis, pp. 301–342. https://doi.org/10.1016/S0076-6879(85)17018-7.

Kegl, T., Kovač Kralj, A., 2020. Multi-objective optimization of anaerobic digestion process using a gradient-based algorithm. Energy Convers. Manag. 226, 113560 https://doi.org/10.1016/j.enconman.2020.113560.

Kim, D.S., Lee, J.H., 2003. Oviposition model of *Carposina sasakii* (Lepidoptera: Carposinidae). Ecol. Model. 162, 145–153. https://doi.org/10.1016/S0304-3800(02)00402-7.

Lanzarone, E., Pasquali, S., Gilioli, G., Marchesini, E., 2017. A Bayesian estimation approach for the mortality in a stage-structured demographic model. J. Math. Biol. 75, 759–779. https://doi.org/10.1007/s00285-017-1099-4.

Leppla, N.C., 2009. Rearing of insects. In: Encyclopedia of Insects. Elsevier Inc, pp. 866–869. https://doi.org/10.1016/B978-0-12-374144-8.00227-7.

Lessio, F., Alma, A., 2021. Models applied to grapevine pests: a review. Insects 12, 169. https://doi.org/10.3390/insects12020169.

Lippi, M., Bonucci, N., Carpio, R.F., Contarini, M., Speranza, S., Gasparri, A., 2021. A YOLO-based pest detection system for precision agriculture. In: 2021 29th Mediterranean Conference on Control and Automation (MED). IEEE, pp. 342–347. https://doi.org/10.1109/MED51440.2021.9480344.

Lippi, M., Carpio, R.F., Contarini, M., Speranza, S., Gasparri, A., 2022. A data-driven monitoring system for the early pest detection in the precision agriculture of hazelnut orchards. In: 7th IFAC Conference on Sensing, Control and Automation Technologies for Agriculture. Munich.

Nance, J., Fryxell, R.T., Lenhart, S., 2018. Modeling a single season of *Aedes albopictus* populations based on host-seeking data in response to temperature and precipitation in eastern Tennessee. J. Vector Ecol. 43, 138–147. https://doi.org/10.1111/jvec.12293.

Naranjo, S.E., Ellsworth, P.C., 2017. Methodology for developing life tables for sessile insects in the field using the whitefly, *Bemisia tabaci*, in cotton as a model system. J. Vis. Exp. 2017, 1–12. https://doi.org/10.3791/56150.

Orlandini, S., Magarey, R.D., Park, E.W., Sporleder, M., Kroschel, J., 2018. Methods of Agroclimatology: Modeling Approaches for Pests and Diseases, pp. 453–488. https://doi.org/10.2134/agronmonogr60.2016.0027.

Otero, M., Solari, H.G., Schweigmann, N., 2006. A stochastic population dynamics model for *Aedes aegypti*: formulation and application to a city with temperate climate. Bull. Math. Biol. 68, 1945–1974. https://doi.org/10.1007/s11538-006-9067-y.

Padmanabha, M., Streif, S., 2019. Design and validation of a low cost programmable controlled environment for study and production of plants, mushroom, and insect larvae. Appl. Sci. 9, 5166. https://doi.org/10.3390/app9235166.

Pasquali, S., Soresina, C., Marchesini, E., 2022. Mortality estimate driven by population abundance field data in a stage-structured demographic model. The case of *Lobesia botrana*. Ecol. Model. 464, 109842 https://doi.org/10.1016/j.ecolmodel.2021.109842.

Petrovskaya, N., Petrovskii, S., Murchie, A.K., 2012. Challenges of ecological monitoring: estimating population abundance from sparse trap counts. J. R. Soc. Interface 9, 420–435. https://doi.org/10.1098/rsif.2011.0386.

Petrovskii, S., Bearup, D., Ahmed, D.A., Blackshaw, R., 2012. Estimating insect population density from trap counts. Ecol. Complex. 10, 69–82. https://doi.org/10.1016/j.ecocom.2011.10.002.

Ponti, L., Gutierrez, A.P., de Campos, M.R., Desneux, N., Biondi, A., Neteler, M., 2021. Biological invasion risk assessment of *Tuta absoluta*: mechanistic versus correlative methods. Biol. Invasions 5. https://doi.org/10.1007/s10530-021-02613-5.

Potamitis, I., Eliopoulos, P., Rigakis, I., 2017. Automated remote insect surveillance at a global scale and the Internet of Things. Robotics 6, 19. https://doi.org/10.3390/robotics6030019.

Preti, M., Verheggen, F., Angeli, S., 2021. Insect pest monitoring with camera-equipped traps: strengths and limitations. J. Pest. Sci. 2004 (94), 203–217. https://doi.org/10.1007/s10340-020-01309-4.

Quinn, B.K., 2017. A critical review of the use and performance of different function types for modeling temperature-dependent development of arthropod larvae. J. Therm. Biol. 63, 65–77. https://doi.org/10.1016/j.jtherbio.2016.11.013.

Rathod, S., Singh, K., Arya, P., Ray, M., Mukherjee, A., Sinha, K., Kumar, P., Shekhawat, R.S., 2017. Forecasting maize yield using ARIMA-Genetic Algorithm approach. Outlook Agric. 46, 265–271. https://doi.org/10.1177/0030727017744933.

Ratkowsky, D.A., 1993. Principles of nonlinear regression modeling. J. Ind. Microbiol. 12, 195–199. https://doi.org/10.1007/BF01584190.

Ratkowsky, D.A., Reddy, G.V.P., 2017. Empirical model with excellent statistical properties for describing temperature-dependent developmental rates of insects and mites. Ann. Entomol. Soc. Am. 110, 302–309. https://doi.org/10.1093/aesa/saw098.

Rebaudo, F., Rabhi, V.-B., 2018. Modeling temperature-dependent development rate and phenology in insects: review of major developments, challenges, and future directions. Entomol. Exp. Appl. 166, 607–617. https://doi.org/10.1111/eea.12693.

Rossi, V., Giosuè, S., Caffi, T., 2010. Precision Crop Protection - The Challenge and Use of Heterogeneity. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-90-481-9277-9.

Rossini, L., Severini, M., Contarini, M., Speranza, S., 2019a. A novel modelling approach to describe an insect life cycle vis-à-vis plant protection: description and application in the case study of *Tuta absoluta*. Ecol. Model. 409, 108778 https://doi.org/10.1016/j.ecolmodel.2019.108778.

Rossini, L., Severini, M., Contarini, M., Speranza, S., 2019b. Use of ROOT to build a software optimized for parameter estimation and simulations with Distributed Delay Model. Ecol. Inform. 50, 184–190. https://doi.org/10.1016/j.ecoinf.2019.02.002.

Rossini, L., Contarini, M., Giarruzzo, F., Assennato, M., Speranza, S., 2020a. Modelling *Drosophila suzukii* adult male populations: a physiologically based approach with validation. Insects 11, 751. https://doi.org/10.3390/insects11110751.

Rossini, L., Contarini, M., Severini, M., Talano, D., Speranza, S., 2020b. A modelling approach to describe the *Anthonomus eugenii* (Coleoptera: Curculionidae) life cycle in plant protection: a priori and a posteriori analysis. Fla. Entomol. 103, 259–263. https://doi.org/10.1653/024.103.0217.

Rossini, L., Severini, M., Contarini, M., Speranza, S., 2020c. *EntoSim*, a ROOT-based simulator to forecast insects' life cycle: description and application in the case of *Lobesia botrana*. Crop Prot. 129, 105024 https://doi.org/10.1016/j.cropro.2019.105024.

Rossini, L., Speranza, S., Contarini, M., 2020d. Distributed Delay Model and Von Foerster's equation: different points of view to describe insects' life cycles with chronological age and physiological time. Ecol. Inform. 59, 101117 https://doi.org/10.1016/j.ecoinf.2020.101117.

Rossini, L., Bono Rosselló, N., Speranza, S., Garone, E., 2021a. A general ODE-based model to describe the physiological age structure of ectotherms: description and application to *Drosophila suzukii*. Ecol. Model. 456, 109673 https://doi.org/10.1016/j.ecolmodel.2021.109673.

Rossini, L., Speranza, S., Mazzaglia, A., Turco, S., 2021b. *EntoSim*, an insects life cycle simulator enclosing multiple models in a Docker container. Environ. Eng. Manag. J. 20, 1703–1710.

Rossini, L., Bono Rosselló, N., Contarini, M., Speranza, S., Garone, E., 2022a. Modelling ectotherms' populations considering physiological age structure and spatial motion: a novel approach. Ecol. Inform. 70, 101703 https://doi.org/10.1016/j.ecoinf.2022.101703.

Rossini, L., Bruzzone, O.A., Contarini, M., Bufacchi, L., Speranza, S., 2022b. A physiologically based ODE model for an old pest: modeling life cycle and population dynamics of *Bactrocera oleae* (Rossi). Agronomy 12, 2298. https://doi.org/10.3390/agronomy12102298.

Ryan, G.D., Emiljanowicz, L., Wilkinson, F., Kornya, M., Newman, J.A., 2016. Thermal tolerances of the spotted-wing drosophila *Drosophila suzukii* (Diptera: Drosophilidae). J. Econ. Entomol. 109, 746–752. https://doi.org/10.1093/jee/tow006.

Sadovski, A., 2020. Precision agriculture through agroecological approach and mathematical modeling. Ecol. Eng. Environ. Protect. 63–69 https://doi.org/10.32006/eeep.2020.2.6369.

Safaei, M., Rezayan, H., Zeaiean Firouzabadi, P., Sadidi, J., 2021. Optimization of species distribution models using a genetic algorithm for simulating climate change effects on Zagros forests in Iran. Ecol. Inform. 63, 101288 https://doi.org/10.1016/j.ecoinf.2021.101288.

Severini, M., Gilioli, G., 2002. Storia e filosofia dei modelli di simulazione nella difesa delle colture agrarie. Notiziario sulla protezione delle piante 15, 9–29.

Severini, M., Baumgärtner, J., Ricci, M., 1990. Theory and practice of parameter estimation of distributed delay models for insect and plant phenologies. Meteorol. Environ. Sci. 674–719.

Shang, Y., Zhu, Y., 2018. Research on intelligent pest prediction of based on improved artificial neural network. In: 2018 Chinese Automation Congress (CAC). IEEE, pp. 3633–3638. https://doi.org/10.1109/CAC.2018.8623592.

Sharov, A., 1996. Modelling forest insect dynamics. In: Mukkela, H., Salonen, T. (Eds.), Caring for the Forest: Research in a Changing World. Gummerus Printing, Tampere, pp. 6–12.

Shi, P.-J., Reddy, G.V.P., Chen, L., Ge, F., 2017. Comparison of thermal performance equations in describing temperature-dependent developmental rates of insects: (I) empirical models. Ann. Entomol. Soc. Am. 110, 113–120. https://doi.org/10.1093/aesa/saw067.

Sinclair, T.R., Seligman, N.G., 1996. Crop modeling: from infancy to maturity. Agron. J. 88, 698. https://doi.org/10.2134/agronj1996.00021962008800050004x.

Son, Y., Lewis, E.E., 2005. Modelling temperature-dependent development and survival of *Otiorhynchus sulcatus* (Coleoptera: Curculionidae). Agric. For. Entomol. 7, 201–209. https://doi.org/10.1111/j.1461-9555.2005.00260.x.

Song, K., Lu, D., Li, L., Li, S., Wang, Z., Du, J., 2012. Remote sensing of chlorophyll-a concentration for drinking water source using genetic algorithms (GA)-partial least square (PLS) modeling. Ecol. Inform. 10, 25–36. https://doi.org/10.1016/j.ecoinf.2011.08.006.

Sørensen, J.G., Addison, M.F., Terblanche, J.S., 2012. Mass-rearing of insects for pest management: challenges, synergies and advances from evolutionary physiology. Crop Prot. 38, 87–94. https://doi.org/10.1016/j.cropro.2012.03.023.

Tochen, S., Dalton, D.T., Wiman, N., Hamm, C., Shearer, P.W., Walton, V.M., 2014. Temperature-related development and population parameters for *Drosophila suzukii* (Diptera: Drosophilidae) on cherry and blueberry. Environ. Entomol. 43, 501–510. https://doi.org/10.1603/EN13200.

Vansickle, J., 1977. Attrition in distributed delay models. IEEE Trans. Syst. Man. Cybern. 7, 635–638. https://doi.org/10.1109/TSMC.1977.4309800.

Voulgaris, S., Stefanidakis, M., Floros, A., Avlonitis, M., 2013. Stochastic modeling and simulation of olive fruit fly outbreaks. Proc. Technol. 8, 580–586. https://doi.org/10.1016/j.protcy.2013.11.083.

Wang, X.J., Ma, C. sen, 2022. Can laboratory-reared aphid populations reflect the thermal performance of field populations in studies on pest science and climate change biology? J. Pest. Sci. 2004 https://doi.org/10.1007/s10340-022-01565-6.

Wang, X.-G., Serrato, M.A., Son, Y., Walton, V.M., Hogg, B.N., Daane, K.M., 2018. Thermal performance of two indigenous pupal parasitoids attacking the invasive *Drosophila suzukii* (Diptera: Drosophilidae). Environ. Entomol. 47, 764–772. https://doi.org/10.1093/ee/nvy053.

Winkler, A., Jung, J., Kleinhenz, B., Racca, P., 2021. Estimating temperature effects on *Drosophila suzukii* life cycle parameters. Agric. For. Entomol. https://doi.org/10.1111/afe.12438 afe.12438.

Zhang, L., Carpenter, B., Gelman, A., Vehtari, A., 2022. Pathfinder: parallel quasi-Newton variational inference. J. Mach. Learn. Res. 23, 1–49. https://doi.org/10.48550/arXiv.2108.03782.