

# Towards pest outbreak predictions: Are models supported by field monitoring the new hope?

Nicolás Bono Rosselló<sup>a,1</sup>, Luca Rossini<sup>b,c,\*</sup>, Stefano Speranza<sup>c</sup>, Emanuele Garone<sup>b</sup>

<sup>a</sup> Namur Digital Institute, University of Namur, 5000 Namur, Belgium

<sup>b</sup> Service d'Automatique et d'Analyse des Systèmes, Université Libre de Bruxelles (ULB), Av. F.D. Roosevelt 50, CP 165/55, Brussels 1050, Belgium

<sup>c</sup> Dipartimento di Scienze Agrarie e Forestali, Università degli Studi della Tuscia, Via San Camillo De Lellis snc 01100, Viterbo, Italy

## ARTICLE INFO

### Keywords:

Decision support systems  
Precision agriculture  
Pest monitoring  
Physiologically-based models  
Estimation update  
Predictive models  
Integrated pest management

## ABSTRACT

Physiologically-based models are the core of Decision Support Systems (DSS) for insect pest and disease control in cultivated fields. However, the large-scale use of DSS remains scarce and limited, despite the continuous update and formulation of new models by the literature. The main reason behind this lack of real-world use relates to the purely descriptive approach of these models, which are usually validated a posteriori. The major limiting factors that preclude the use of these tools for prediction purposes are their dependence on time zero and initial abundance to start the simulations. In this study, we present a theoretical framework that includes field monitoring data as an active part of a pest population density model simulation, which helps to overcome these obstacles. More specifically, we propose the application of an estimator scheme in the form of an Extended Kalman Filter (EKF) to a revised physiologically-based model from the literature. In the paper, we carry out a preliminary test of the theoretical framework applied to the case of *Drosophila suzukii*. This case study shows that the dependence of the simulations on the initial conditions and time zero is strongly reduced by using the EKF. Overall, the outcome of this research indicates that an estimator scheme is a necessary step to move from description to prediction in the pest population modelling field.

## 1. Introduction

The mathematical interpretation of biological phenomena is gaining popularity in many fields of research (Mesarovic et al., 2004). The potential of reliable mathematical models, and the practical applications that these models may provide (Plant and Mangel, 1987; Sinclair and Seligman, 1996), is arousing interest among the scientific community working on the fields of biology and ecology.

Decision Support Systems (DSS) are among the fastest growing applications of mathematical biology (Murray, 2012). Namely, in the field of ecology and agricultural sciences, the predictive potential of pest and disease models fits perfectly with the Precision Agriculture (PA) paradigm (Rupnik et al., 2019). PA aims to optimize the human inputs in the management of cultivated fields (e.g., agrochemicals, water, fertilisers), so that they are provided only where and when they are needed (Srbínovska et al., 2015; Stafford, 2015). This reduction of resources employed provides obvious economic advantages and allows to

safeguard the environment and the biodiversity of agroecosystems (Barrett and Rose, 2020), which is in line with the goals of many public institutions and government agencies worldwide.

Insects pests are one of the main reasons for the extensive use of agrochemicals and pesticides, which are two important examples of the aforementioned field inputs in intensive cultivation (da Silva et al., 2019). Additionally, the large spectrum of active ingredients applied seriously endangers the population of helpful and beneficial organisms, such as pollinators or predators and parasitoids of insect pests (da Silva et al., 2019). According to Integrated Pest Management (IPM) framework, one way to reduce this side effect is by rotating the use of active ingredients, knowing which are the most infested portions of the field, and when the pests are the most susceptible. IPM guidelines suggest the use of less environmentally damaging control strategies. However, most of these strategies are only efficient during specific phases of the life cycle or up to a certain population density threshold. All these requirements make the monitoring of the pests life cycle fundamental

\* Corresponding author at: Service d'Automatique et d'Analyse des Systèmes, Université Libre de Bruxelles (ULB), Av. F.D. Roosevelt 50, CP 165/55, Brussels 1050, Belgium.

E-mail address: [luca.rossini@unitus.it](mailto:luca.rossini@unitus.it) (L. Rossini).

<sup>1</sup> These authors share the first name.

(Bange et al., 2004; Rossi et al., 2019), and motivate the development of accurate monitoring and predictive tools.

Regarding monitoring and predictive tools, physiologically-based models are gaining in importance in the description of insect pest population dynamics. The reason for their success lies on their mathematical description of insect populations developing over time and through life stages (Ponti et al., 2021), while considering the characteristic ectotherm behaviour of terrestrial arthropods (Gutierrez et al., 2017). The stage-maturation time of ectotherms strongly depends on environmental parameters (e.g., temperature, relative humidity) and physiologically-based models are accurately formulated to consider this aspect (Lessio and Alma, 2021).

Even though some physiologically-based models in the existing literature have been successfully validated, we identify two main problems that still hinder their implementation in the DSS framework: *i*) it is difficult to identify a “time zero” to start the simulations (i.e., when the development of the population begins) and *ii*) it is difficult to estimate the initial population abundance (i.e., how many individuals are into each stage). In other words, the estimation of the initial conditions is one of the crucial points of the model application, which is amplified by the high variability under which populations develop in cultivated fields.

Given this precondition, the predictive potential and accuracy of physiologically-based models is heavily reduced if simulations are carried out in “open loop” (i.e., model simulations with no field measurement corrections) (Bono Rosselló et al., 2022). This problem does not come from the accuracy and precision of the models themselves, but from their dependence on a proper estimation of the initial conditions (Pasquali et al., 2019). This dependence, and the absence of field measurement corrections, means that a wrong estimation of the initial conditions is never corrected in the prediction, leading to large inaccuracies during the peak of individuals.

Monitoring carried out through field measurements, instead, provides empirical assessments and predictions regarding pest infestation trends (Onufrieva and Onufriev, 2021; Rossini et al., 2022a). However, field monitoring is a highly time-consuming activity, and often relies on the experience of farmers and technicians. Accordingly, it is hard to have large series of data with a high sampling frequency (Rossini et al., 2022a). Field measurement data provides a picture of the infestation from the past to date, and it is common to use them in open-loop simulations to validate models (e.g., (Bellocchi et al., 2011; Gilioli et al., 2022; Pasquali et al., 2022; Rossini et al., 2020a; Rossini et al., 2020b; Rossini et al., 2020c; Rossini et al., 2020d; Rossini et al., 2021a)).

Models and field measurements have complementary strengths and weaknesses. Accordingly, the next step is to include field measurements as an active part of the modelling process, contributing to improve the model predictions. This approach is common in many fields of engineering, and, in this study, we suggest and demonstrate how the combination of these two different sources of information improves the overall estimation of the pest population trend.

This paper is inspired by analogous works in the agricultural field that already included measurements into mathematical models applied to agriculture. Early examples of application were presented by Liu et al. (Xiaoyun et al., 2023) and by Sorensen (Sorensen, 2002) in agricultural economics. These works were pioneers in the use of estimation methods outside the field of pure engineering. Other applications, closer to the engineering field, concern the use of GPS data to improve the positioning information of the machines, such as in the case of automated driving systems of tractors (Gomez-Gil et al., 2013). To the best of our knowledge, there are very few works using estimation methods applied to pest population dynamics, which is the main focus of this study.

A recent theoretical work of Bono Rosselló et al. (Bono Rosselló et al., 2022) showed how control system theory can be applied to pest population models as well. This study aims to extend this theoretical framework by improving the physiologically-based model recently introduced by Rossini et al. (Rossini et al., 2021b), and to present how different

measurement techniques commonly applied to pest monitoring can be incorporated in the presented framework. The theoretical framework is then completed with a practical application example. This application combines trap measurements and model predictions, showing how the proposed method clearly improves the model accuracy by compensating, through field measurements, the initial conditions uncertainty. For this purpose, we will use the same dataset published by Rossini et al. (Rossini et al., 2020e; Rossini et al., 2021b), providing a comparison between the open and closed loop simulations subject to initial conditions uncertainty. The improvements presented by the proposed approach are manifold: *i*) it accounts for the problem of estimating the insect population's initial conditions, *ii*) it helps to include the possible migrations into the field during the season, and *iii*) it allows to combine sparse and different counting systems into the pest abundance estimation framework.

## 2. Materials and methods

The theoretical framework we introduce in this study is organized in different logical steps. The starting point is the pest population density model describing the biology of most species of agricultural interest. This model should be capable of simulating the number of individuals that develop over time through the different life stages. These quantities are often measurable using conventional techniques that we will detail in the following sections. The idea is to combine the model with an iterative estimation scheme that improves the estimations based on the field measurements. However, to achieve this combination it is necessary to modify the original model to consider the uncertainties of its parameters, often measured through laboratory experiments. Subsequently, field measurements are also considered as non-deterministic values affected by noise and measurement uncertainties.

Based on that, the estimation algorithm will combine both sources of information given the uncertainties associated with both model and measurements, providing the optimal estimate. In what follows, we will detail all the steps we briefly described hereby, while considering the different kinds of information that various techniques of insect pest monitoring provide.

### 2.1. The pest population density model

As a starting point, let us consider the general physiologically-based model presented in (Rossini et al., 2021b), which has recently been extended to the spatial context (Rossini et al., 2022b). While referring the most interested readers to the cited literature for more detailed information, in this section we report only the essential features of the model.

A population of insects or, more in general, of terrestrial arthropods that develops over time, and through discrete and well identified life stages, can be described through a compartmental scheme (Bellagamba et al., 1987; Borlino et al., 1990; Borlino et al., 1991; Cola and Gilioli, 1996; Manetsch, 1976; Severini et al., 1990; Vansickle, 1977). Accordingly, the life cycle is mathematically represented by a set of chained stages each of which corresponds to an “entomologically identifiable stage”. In other words, we consider single compartments for the egg stage  $e$ , for each preimaginal stage (larval instars of nymphal stages)  $L_1, L_2, \dots, L_n$ , for adult males  $A_m$ , and for adult females. The latter is in turn divided in two substages  $A_{f_1}$  and  $A_{f_2}$  to take into account, for instance, the different gonotrophic cycles, the subdivision between non-mated and mated females, or to consider cyclicity in reproduction.

Each life stage  $i$  has an associated state  $x_i(t)$  that describes the number of individuals that at time  $t$  populates the stage  $i$ , where  $i = e, L_1, \dots, L_n, A_m, A_{f_1}, A_{f_2}$ . Moreover, each life stage is described by an Ordinary Differential Equation (ODE) that considers the number of incoming individuals developing from the previous stage and the outgoing individuals due to maturation and death. Mortality is slightly different in

adult stages, given that after adults there are no further stages. In these stages, mortality functions are generally composed of a survival rate term indicating the ageing of adults, and of a second term involving any other type of mortality (if sub-models are available) (Rossini et al., 2021b; Rossini et al., 2022a; Rossini et al., 2022b). Besides development and mortality, reproduction is defined as the number of offspring (eggs) that adult females produce.

The incoming and outgoing flows of individuals among the stages are associated with specific transition rates, hereafter denoted as development, mortality, and fertility rate functions. These transition rates are assumed to depend, in the most general case, on both population density and environmental parameters (Rossini et al., 2021b; Rossini et al., 2022b).

The previous assumptions lead to the following mathematical set of equations:

$$\left\{ \begin{array}{l} \frac{d}{dt}x_e(t) = G_{A_{f_1}}(t)\beta_1(t)x_{A_{f_1}}(t) + G_{A_{f_2}}(t)\beta_2(t)x_{A_{f_2}}(t) - G_e(t)x_e(t) - M_e(t)x_e(t) \\ \frac{d}{dt}x_{L_1}(t) = G_e(t)x_e(t) - G_{L_1}(t)x_{L_1}(t) - M_{L_1}(t)x_{L_1}(t) \\ \vdots \\ \frac{d}{dt}x_{L_n}(t) = G_{L_{n-1}}(t)x_{L_{n-1}}(t) - G_{L_n}(t)x_{L_n}(t) - M_{L_n}(t)x_{L_n}(t) \\ \frac{d}{dt}x_{A_m}(t) = (1 - S_R(t)) \cdot G_{L_n}(t)x_{L_n}(t) - M_{A_m}(t)x_{A_m}(t) \\ \frac{d}{dt}x_{A_{f_1}}(t) = S_R(t)G_{L_n}(t)x_{L_n}(t) - G_{1 \rightarrow 2}(t)x_{A_{f_1}}(t) - M_{A_{f_1}}(t)x_{A_{f_1}}(t) + \\ \quad + G_{1 \rightarrow 2}(t)x_{A_{f_2}}(t) \\ \frac{d}{dt}x_{A_{f_2}}(t) = G_{1 \rightarrow 2}(t)x_{A_{f_1}}(t) - M_{A_{f_2}}(t)x_{A_{f_2}}(t) - G_{1 \rightarrow 2}(t)x_{A_{f_2}}(t). \end{array} \right. \quad (1)$$

A detailed description of the functions and variables involved in the model (1) is provided in Table 1. Note that the number of ODEs depends on the number of preimaginal stages of the species under study, which

**Table 1**  
List of the variables and functions involved in the model (1) and (2).

| Function                 | Description                                                                                    |
|--------------------------|------------------------------------------------------------------------------------------------|
| $x_e(t)$                 | Number of individuals in the egg stage at time $t$ .                                           |
| $x_{L_i}(t)$             | Number of individuals in the $i^{\text{th}}$ larval stage at time $t$ .                        |
| $x_{A_m}(t)$             | Number of individuals in the adult male stage at time $t$ .                                    |
| $x_{A_{f_1}}(t)$         | Number of individuals in the female adult substage 1 at time $t$ .                             |
| $x_{A_{f_2}}(t)$         | Number of individuals in the female adult substage 2 at time $t$ .                             |
| $M_e(t)$                 | Mortality rate of the eggs at time $t$ .                                                       |
| $M_{L_i}(t)$             | Mortality rate of the $i^{\text{th}}$ larval instar at time $t$ .                              |
| $M_{A_m}(t)$             | Mortality rate of adult males at time $t$ .                                                    |
| $M_{A_{f_1}}(t)$         | Mortality rate of adult females substage 1 at time $t$ .                                       |
| $M_{A_{f_2}}(t)$         | Mortality rate of adult females substage 2 at time $t$ .                                       |
| $\beta_1(t)$             | Fertility rate of the adult females substage 1 at time $t$ .                                   |
| $\beta_2(t)$             | Fertility rate of the adult females substage 2 at time $t$ .                                   |
| $G_e(t)$                 | Development rate function of the egg stage.                                                    |
| $G_{L_i}(t)$             | Development rate function of the $i^{\text{th}}$ larval stage at time $t$ .                    |
| $G_{A_{f_1}}(t)$         | Development rate function of the adult females substage 1 at time $t$ .                        |
| $G_{A_{f_2}}(t)$         | Development rate function of the adult females substage 2 at time $t$ .                        |
| $S_R(t)$                 | Sex ratio of the species: $S_R(t)$ for females, $1 - S_R(t)$ for males.                        |
| $G_{1 \rightarrow 2}(t)$ | Transition rate of adult females to substage 2 from the substage 1 at time $t$ .               |
| $G_{1 \rightarrow 2}(t)$ | Transition rate of adult females to return in the substage 1 from the substage 2 at time $t$ . |

can be defined once we know the biological traits of the species (Rossini et al., 2021b; Rossini et al., 2022b).

A fundamental feature of this model concerning the theoretical framework presented in this paper refers to the fact that each state variable represents a single stage. This allows us to obtain a direct correspondence between the quantity we simulate (i.e., the number of individuals over time in each life stage) and what we can potentially measure in the field (i.e., the number of individuals in one or more stages over time).

The model (1) is hereby presented as an apparently linear model. However, it is worth pointing out that in general we may modify the rate functions involved to consider the density-dependence as well. For instance, let us briefly compare the case of a population of insects developing in a cultivated field and in a more natural environment. In the first case, the non-linearity of the model, such as the scarcity of food or the dependence of fertility on the population density of males, can be neglected. The main reason is that cultivated fields have, by definition, conditions that are suitable for the development of a given species, mainly if specialized for the cultivated host plant. On the other hand, in more natural environments such as forests there are several additional factors that affect the growth of the population, such as the presence of natural enemies or scarcity of food. Even though there might be non-linearities, the theoretical framework introduced is still valid. We refer the most interested reader to Rossini et al. (Rossini et al., 2021b) for further details about potential non-linearities of the general model (1).

State observers and estimators are algorithms that provide an estimate of the internal states of a given system (Anderson and Moore, 1979), based on some measurements of the real phenomenon and the knowledge of the system dynamics. In this paper, the dynamics of the system are given by the mathematical model (1), and the measurements are associated with different field monitoring techniques, as we have already stated at the beginning of the current section.

It must be noted that the sensors or techniques that measure the states of a real system (e.g., the number of individuals in a given stage over time) are, by definition, not perfect, but affected by noise and inaccuracies. This fact should be considered when estimating the current state of the system. Usually, the uncertainty associated with the measurements is not deterministically known, and it needs to be statistically estimated. This estimated noise is thus introduced into the model of the system as a stochastic disturbance.

Similarly, a dynamic mathematical model is often a simplified representation of the natural phenomenon, as in (1). Accordingly, it has a certain degree of inaccuracy in its predictions. These uncertainties and other perturbations can also be incorporated into the system as stochastic process noise.

The following section introduces the modifications required to adapt the model (1) to the estimator's scheme. Two steps are foreseen: *i*) modelling the uncertainties associated with the model dynamics, and *ii*) modelling the uncertainties associated with the field data.

## 2.2. Stochastic model

As previously mentioned, rate functions regulate the number of individuals entering or leaving the different life stages, and generally depend on environmental parameters (e.g., temperature or relative humidity). While the mathematical form of the rate functions is general, their parameters characterize the species, and their estimation requires ad hoc laboratory experiments (Damos and Savopoulou-Soultani, 2012; Quinn, 2017; Ratkowsky and Reddy, 2017; Shi et al., 2017). Accordingly, the biological information included in the model is affected by the uncertainties associated with the estimation of the parameters (Bellocchi et al., 2011). These uncertainties are modelled as stochastic noise and included into the equations of the model (1).

Besides the uncertainty introduced by the rate functions' parameters, we should consider that simulations are based on field measurements of the environmental variables (e.g., temperature) subject to noise

(Didham et al., 2023; Jactel et al., 2023). The uncertainty related to the computation of the rates is introduced in each ODE of (1) as an additive term  $w_{ij}(t)$ , associated with the stage  $i = e, L_1, \dots, L_n, A_m, A_{f1}, A_{f2}$  and the rate function  $j$ , that multiplies the state variable  $x_i(t)$ . Mathematically:

$$w_{ij}(t) = w_{ij}^1(t) + w_{ij}^2(t),$$

with  $w_{ij}^1(t) \sim N(0, Q_{ij}^1)$  the uncertainty related to parameters estimation and  $w_{ij}^2(t) \sim N(0, Q_{ij}^2)$  the noise that the measurements of the environmental parameters introduces. If we assume that both terms are Gaussian with zero means, it holds the following property for the expected values

$$E[w_{ij}(t)] = E[w_{ij}^1(t)] + E[w_{ij}^2(t)] = 0,$$

and for the covariance

$$\text{Cov}(w_{ij}(t)) = Q_{ij}^1 + Q_{ij}^2 = Q_{ij},$$

with  $w_{ij}(t) \sim N(0, Q_{ij})$ .

By introducing these uncertainties into the model (1), we obtain the following mathematical description of insect populations developing under field conditions:

$$\left\{ \begin{aligned} \frac{d}{dt}x_e(t) &= \left( G_{A_{f1}}(t)\beta_1(t) + w_{e,1}(t) \right) x_{A_{f1}}(t) + \left( G_{A_{f2}}(t)\beta_2(t) + \right. \\ &\quad \left. w_{e,2}(t) \right) x_{A_{f2}}(t) - \left( G_e(t) + w_{e,3}(t) \right) x_e(t) - \\ &\quad \left( M_e(t) + w_{e,4}(t) \right) x_e(t) \\ \frac{d}{dt}x_{L_1}(t) &= \left( G_e(t) + w_{e,1}(t) \right) x_e(t) - \left( G_{L_1}(t) + w_{L_1,2}(t) \right) x_{L_1}(t) \\ &\quad - \left( M_{L_1}(t) + w_{L_1,3}(t) \right) x_{L_1}(t) \\ \frac{d}{dt}x_{L_n}(t) &= \left( G_{L_{n-1}}(t) + w_{L_{n-1},1}(t) \right) x_{L_{n-1}}(t) - \left( G_{L_n}(t) + w_{L_n,2}(t) \right) x_{L_n}(t) \\ &\quad - \left( M_{L_n}(t) + w_{L_n,3}(t) \right) x_{L_n}(t) \\ \frac{d}{dt}x_{A_m}(t) &= (1 - S_R(t)) \cdot \left( G_{L_{n-1}}(t) + w_{A_m,1}(t) \right) \cdot x_{L_{n-1}}(t) \\ &\quad - \left( M_{A_m}(t) + w_{A_m,2}(t) \right) x_{A_m}(t) \\ \frac{d}{dt}x_{A_{f1}}(t) &= S_R(t) \left( G_{L_{n-1}}(t) + w_{A_{f1},1}(t) \right) x_{L_{n-1}}(t) - \left( G_{1 \rightarrow 2}(t) \right. \\ &\quad \left. + w_{A_{f1},2}(t) \right) x_{A_{f1}}(t) - \left( M_{A_{f1}}(t) + w_{A_{f1},2}(t) \right) x_{A_{f1}}(t) \\ \frac{d}{dt}x_{A_{f2}}(t) &= \left( G_{1 \rightarrow 2}(t) + w_{A_{f2},1}(t) \right) x_{A_{f1}}(t) - \left( G_{2 \rightarrow 1}(t) \right. \\ &\quad \left. + w_{A_{f1},2}(t) \right) x_{A_{f2}}(t) - \left( M_{A_{f2}}(t) + w_{A_{f2},2}(t) \right) x_{A_{f2}}(t). \end{aligned} \right. \quad (2)$$

In the general presentation of the model (1), Section 2.1, we assumed that in the case of cultivated fields the model can be considered as a linear system. However, the introduction of the uncertainty  $w_{ij}(t)$  related to the estimation of the parameters introduces a non-linearity. In what follows, we will therefore consider the model as non-linear, with the advantage of having a more general theoretical framework that can also be applied to more complex versions of the model (1).

The system (2) is an autonomous dynamical system where the rate functions are time dependent and there is no controlled input variable.

Moreover, (2) is subject to stochastic process noise, and it can be represented in compact form as

$$\dot{x}(t) = f(x(t), w(t)), \quad (3)$$

with the state vector

$$x(t) = \begin{bmatrix} x_e(t) \\ x_{L_1}(t) \\ \vdots \\ x_{A_m}(t) \\ x_{A_{f1}}(t) \\ x_{A_{f2}}(t) \end{bmatrix},$$

and the noise vector

$$w(t) = \begin{bmatrix} w_{e,1}(t) \\ \vdots \\ w_{A_{f2},2}(t) \end{bmatrix},$$

where  $w(t) \sim N(0, Q)$  with  $Q \in \mathbb{R}^{m \times m}$ . As previously stated in Section 2.1, the number  $m$  varies according to the preimaginal stages considered in the model when applied to a particular case study.

Field measurements are usually discrete in time, since field pest monitoring is carried out at regular (or irregular, in the case of random inspections) time ranges. By defining a constant sampling time  $T_s$ , we obtain the following discrete time system:

$$x_{k+1} = g(x_k, w_k), \quad (4)$$

with  $x_k = x(kT_s)$ . The sampling time  $T_s$  can take any value in the positive real numbers space,  $\mathbb{R}_+$ , but for the sake of this study we set  $T_s = 1 \text{ day}$ , as it is the most recurrent time unit in the ecological literature.

### 2.2.1. Sensing model

The second step concerns the definition of a model that describes which information from the model (1) we can directly measure. In that model, we have also to represent the uncertainties associated with these measurements, as it was done for the parameters of the physiologically-based model.

The state vector  $x_k \in \mathbb{R}^m$  represents the number of individuals at each life stage at the time instant  $k$ . The measurement matrix  $C \in \mathbb{R}^{l \times m}$  defines which of these quantities are measured and accessible at that time instant. This matrix maps the information from the states of the system (i.e., the life stages) that is provided by the measurements  $y_k \in \mathbb{R}^l$ . This mapping is described by the following equation

$$y_k = Cx_k. \quad (5)$$

We may assume that the state of the population can be measured at given time instants through different techniques. The insect monitoring technique is hence important since the shape of the matrix  $C$  will directly depend on the measurement technique. Generally speaking, the mapping can be defined as

$$y_k = C_k x_k, \quad (6)$$

where at each time instant the time-variant measurement matrix  $C_k$  might change based on the technique used.

The mathematical aspects we will develop in this section are general and can be adapted to any kind of monitoring carried out in entomology. To give some practical examples, we may consider classical techniques such as monitoring with traps, visual inspections, branch shaking, or more modern techniques such as automated traps or similar automated counting systems (Ebrahimi et al., 2017; Lippi et al., 2021; Zha et al., 2021). We can generally say that most of the monitoring techniques track only certain life stages. The mathematical aspects that follow might be applied to the cases where we have information from all life stages, but for the sake of exposition we will consider the case of tracking only adults (Preti et al., 2021).



Depending on the measurement system used, the kind of information that can be obtained about the system is different. In the case of real-time vision counting systems or techniques like frapping (branch shaking), the number of insects that is obtained is proportional to the population at the measurement time instant. On the other hand, when using traps, the information obtained corresponds to the cumulative number of insects captured during the time period that the trap remained in the field. Thus, the sensing action is modelled in two different ways: instantaneous measurements, or cumulative captures. A second aspect to consider when measuring the number of insects concerns the identifiability within the adult life stages, which might vary depending on the species and on the technique considered (Rossini et al., 2020e). In other words, the fact of being able to count: *i*) how many females and males are present or, *ii*) only the total number of adults. Let us consider a very general example, just to fix the ideas. Traps lured with sexual pheromones have the feature to attract only a specific stage (only adult males or only adult females, for instance), while chromotropic traps attract both males and females. In the first case, we have the information on a specific stage, while in the second case it depends on how difficult the distinction is between males and females. If the discrimination between males and females is difficult, for example because there is no stereomicroscope to observe some specific morphological traits of the specimen trapped, the only possible action is to count all the individuals. The mathematical details of instantaneous and cumulative measurements, and the different identifiability aspects follow in the next paragraphs.

**Instantaneous measurements.** This case considers measuring techniques that provide an instantaneous value of the current population at the moment of the measurement action (e.g., branch shaking, visual inspections or automated counting systems). From a general point of view, it can be initially assumed that the 3 adult stages of the general model can be measured independently. This provides the following observation matrix

$$C = \begin{bmatrix} 0_{1 \times p} & \kappa_1 & 0 & 0 \\ 0_{1 \times p} & 0 & \kappa_2 & 0 \\ 0_{1 \times p} & 0 & 0 & \kappa_3 \end{bmatrix},$$

where  $p$  is the number of non-adult stages associated with the species under study and  $\kappa_i \in \mathbb{R}_+ \forall i = 1, 2, 3$  are the estimated counting efficiencies, providing the percentage of insects out of the whole population that can be counted at time  $k$ .

It is worth remarking that the definition of measurable life stages will depend on the species, i.e., biological traits that allow their distinction, and the measurement devices available. To give a practical example, let us consider the case of *frapping*, where the individuals falling from the branch are subsequently caught. When the caught individuals are classified, they can be divided into males and females (mated plus not mated) or, in some cases, in males, non-mated females, and mated females.

In other cases, we can measure a percentage of the whole adult population with no distinction between the adult substages. This case can be mathematically described as

$$C = [0_{1 \times p} \quad \kappa_1 \quad \kappa_2 \quad \kappa_3].$$

Note that in other cases it may happen that  $\kappa_1 = \kappa_2 = 0$ , where only one adult stage can be measured.

Sensing techniques are also subject to measurement noise and inaccuracies associated with the estimated counting efficiency. Given the nature of the measurements (insect counting) and of the system itself (a positive system, as the number of individuals will always be greater or equal to zero) the measurements are always positive. Additionally, in the case of a larger amount of insects, the variance in the efficiency of the counting will vary.

A simple way to formalize these aspects while keeping the assumption of uncorrelated noise,  $\mathbb{E}[x_k \nu_k] = 0$ , is by considering the measure-

ment noise primarily associated with the estimation of the capturing/counting efficiency  $\kappa_i$  as

$$y_k = (C + \nu_k)x_k,$$

where  $\nu_k \in \mathbb{R}^l$  represents the stochastic sensor noise which is assumed  $\nu \sim N(0, R)$  with  $R \in \mathbb{R}^{l \times l}$  the noise covariance matrix. Note that in the case of  $\kappa > R_{ii} \forall i = 1, \dots, l$ , where  $R$  is a diagonal matrix and  $R_{ii}$  are its diagonal elements, it holds that with high probability the outcome  $y_k$  will be non-negative.

**Cumulative measurements.** This case refers mainly to the use of traps as measuring devices. Traps are deployed and remain in the field for a given amount of time, during which they are inspected and emptied at more or less regular intervals. Accordingly, their measurements consider the cumulative number of insects that were trapped during the period of time between two inspections. This can be formalized by adding extra states to the model that are associated with the life stages caught by the traps, obtaining the extended vector state

$$x(t) = \begin{bmatrix} x_e(t) \\ x_{L_1}(t) \\ \vdots \\ x_{A_{J_2}}(t) \\ x_{trap, A_m}(t) \\ x_{trap, A_{J_1}}(t) \\ x_{trap, A_{J_2}}(t) \end{bmatrix},$$

where  $x_{trap, A_m}(t)$ ,  $x_{trap, A_{J_1}}(t)$  and  $x_{trap, A_{J_2}}(t)$  describe the quantity of adult males, adult females in stage 1, and adults females in stage 2 that are caught by the trap at time  $t$ . The evolution of these 3 extra states can be mathematically represented as

$$\begin{cases} \dot{x}_{trap, A_m}(t) = M_{trap}(t)x_{A_m}(t) - r(t) \cdot x_{trap, A_m}(t) \\ \dot{x}_{trap, A_{J_1}}(t) = M_{trap}(t)x_{A_{J_1}}(t) - r(t) \cdot x_{trap, A_{J_1}}(t) \\ \dot{x}_{trap, A_{J_2}}(t) = M_{trap}(t)x_{A_{J_2}}(t) - r(t) \cdot x_{trap, A_{J_2}}(t), \end{cases} \quad (7)$$

where  $M_{trap}(t)$  is the efficiency of the trap and the input action  $r(t)$  represents the reset action of emptying or changing the trap. This action is modelled as a Dirac delta impulse such that

$$r(t) = \delta(t - t_i) = \begin{cases} \text{undefined} & \text{if } t = t_i \\ 0 & \text{if } t \neq t_i, \end{cases}$$

which is constrained to satisfy the equality

$$\int_{-\infty}^{+\infty} \delta(t - t_i) dt = 1$$

and where  $t_i$  represents the time instant where the trap is reset.

Additionally, there is also an uncertainty associated with the efficiency of the trap, that leads to the following set of stochastic equations

$$\begin{cases} \dot{x}_{trap, A_m}(t) = (M_{trap}(t) + w_{trap, 1})x_{A_m}(t) - r(t) \cdot x_{trap, A_m}(t) \\ \dot{x}_{trap, A_{J_1}}(t) = (M_{trap}(t) + w_{trap, 2})x_{A_{J_1}}(t) - r(t) \cdot x_{trap, A_{J_1}}(t) \\ \dot{x}_{trap, A_{J_2}}(t) = (M_{trap}(t) + w_{trap, 3})x_{A_{J_2}}(t) - r(t) \cdot x_{trap, A_{J_2}}(t), \end{cases} \quad (8)$$

where  $w_{trap}(t) \sim N(0, Q)$  with  $Q \in \mathbb{R}^{3 \times 3}$ .

Then, assuming that the measurement provides the current value of the additional trap states, we obtain the following observation matrix

$$C = \begin{bmatrix} 0_{1 \times p+3} & 1 & 0 & 0 \\ 0_{1 \times p+3} & 0 & 1 & 0 \\ 0_{1 \times p+3} & 0 & 0 & 1 \end{bmatrix},$$

where  $p$  is the number of preimaginal stages.

Similarly to the case of instantaneous measurements, there might be

cases where the adult stages are trapped but the specimens are not distinguishable in the measurement process. This provides the following matrix

$$C = [0_{1 \times p+3} \quad 1 \quad 1 \quad 1].$$

The shape of the observation matrix is similar to the previous case

$$y_k = (C + \nu_k)x_k,$$

where  $\nu_k \in \mathbb{R}^l$  represents the stochastic sensor noise which is assumed  $\nu \sim N(0, R)$  with  $R \in \mathbb{R}^{l \times l}$  the noise covariance matrix.

An additional peculiarity of entomological measurements worthy of consideration is their intermittence. This is due to the amount of resources, in terms of manpower and measurement tools, required in each counting action (Rossini et al., 2022b). This is a fundamental aspect to formalize, and for this purpose we introduce the binary variable  $\gamma_k \in (0, 1)$  which takes the value 1 if the measurement is carried out at time  $k$ , and 0 otherwise.

Thus, the general measurement equation is finally defined as

$$y_k = \gamma_k(Cx_k + \nu_k x_k),$$

so that the overall pest population dynamics can be compactly written as

$$\begin{cases} x_{k+1} = g(x_k, w_k) \\ y_k = h(x_k, \nu_k, \gamma_k). \end{cases} \quad (9)$$

### 2.3. Estimator: the Extended Kalman filter

The estimation of the state of the population is carried out using an Extended Kalman Filter (EKF) (Ljung, 1979) with intermittent observations. The Kalman Filter (Kalman, 1960) provides a recursive scheme to compute the estimates of the internal states of a system based on the knowledge of its dynamics and noisy measurements. The Kalman Filter (KF) is an algorithm that uses the dynamic model of a system, e.g. (1), and a series of measurements observed over time, including statistical noise and other inaccuracies. This algorithm computes the population density more accurately than single measurements or single model predictions, as it combines this information at each instant.

The KF is a two-steps algorithm, composed of a prediction and a correction step. In the prediction step, the KF produces estimates of the

current state variables based on the dynamic model of the system. Once we obtain the outcome of the next measurement, the estimates are updated based on the uncertainty associated with the model and the measurements, respectively. The EKF is similar to its linear version, but in this case we linearize the system around the estimate of the state at each instant, see Fig. 1.

To apply the estimation scheme, we assume to know the initial probability distribution of the system with  $p(x_0) \sim N(\bar{x}_0, P_0)$ . The initial estimate and the covariance of the EKF are denoted as

$$\hat{x}_{0|0} = \bar{x}_0 \quad \text{and} \quad P_{0|0} = P_0,$$

respectively.

The estimation scheme of the EKF can be mathematically formalized as follows. The prediction step is

$$\hat{x}_{k|k-1} = g(x_{k-1}, 0) \quad (10)$$

$$P_{k|k-1} = F_{k-1}P_{k-1|k-1}F_{k-1}^T + L_{k-1}QL_{k-1}^T, \quad (11)$$

which relies on the known dynamics of the system, and where

$$L_{k-1} = \left. \frac{\partial g}{\partial w} \right|_{\hat{x}_{k-1|k-1}}, \quad F_{k-1} = \left. \frac{\partial g}{\partial x} \right|_{\hat{x}_{k-1|k-1}}, \quad (12)$$

are the linearized dynamics around the previous estimate.

The correction step, where the prediction is corrected based on the measurements, is defined as

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \gamma_k K_k (y_k - C\hat{x}_{k|k-1}) \quad (13)$$

$$K_k = P_{k|k-1}C^T(CP_{k|k-1}C^T + M_kRM_k^T)^{-1} \quad (14)$$

$$P_{k|k} = P_{k|k-1} - K_k\gamma_k CP_{k|k-1}, \quad (15)$$

with

$$M_k = \left. \frac{\partial h}{\partial v} \right|_{\hat{x}_{k|k-1}}, \quad (16)$$

where  $\hat{x}_{k|k}$  is the estimated value of the state at time  $k$  given the information available at time  $k$ , and  $P_{k|k}$  is the covariance matrix of the error

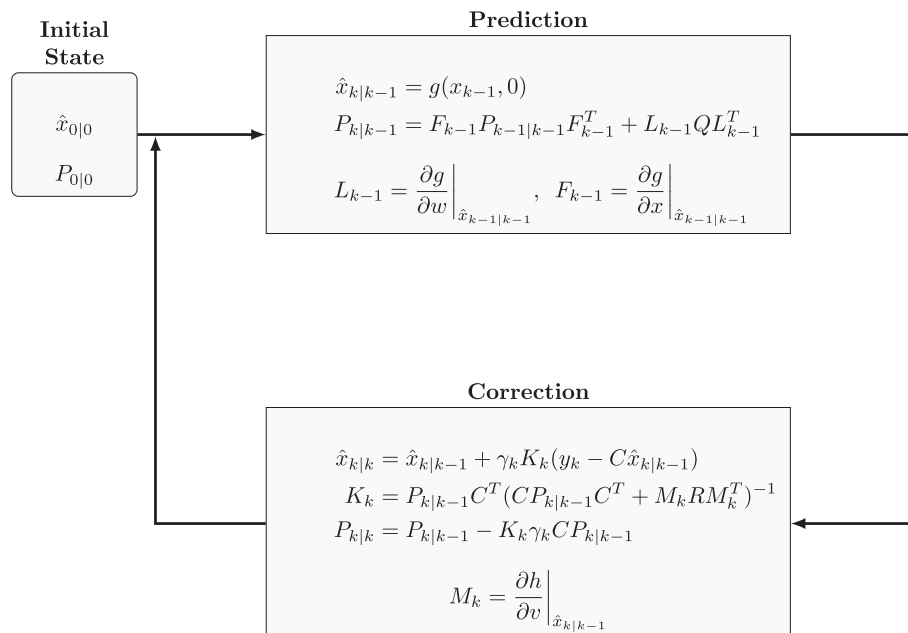


Fig. 1. Example of the Extended Kalman Filter working scheme.

associated with the estimation at time  $k$ .

It is worth remarking that when  $\gamma_k = 0$ , i.e., when there are no new measurements, the correction step does not modify the prediction step, relying uniquely on the estimate given by the model dynamics. Specifically,

$$\begin{aligned}\hat{x}_{k+1|k+1} &= \hat{x}_{k+1|k} \\ P_{k+1|k+1} &= P_{k+1|k}.\end{aligned}$$

This is a relevant property that supports the choice of the EKF as estimation scheme for pest monitoring, since it can work even in the case of measurements carried out irregularly. This aspect is specially important given the high costs of entomological monitoring, which makes irregular sampling ranges very common in the case of monitoring carried out by farmers for pest management purposes.

## 2.4. The case study of *Drosophila suzukii*

### 2.4.1. General model and parameters

The general model (1) can be applied to a case study only after the definition of a species and of its biological traits (Rossini et al., 2021b; Rossini et al., 2022b). For this purpose, we choose the spotted wing drosophila *Drosophila suzukii* (Matsumura) to assess the improvements that the methodology introduced in Section 2.3 brings to the theory. This pest is responsible for several yield losses in soft fruit cultivation worldwide (Asplen et al., 2015), which explains the large amount of information in the current literature that supports its model development.

An open loop validation of the model (1) has already been carried out in (Rossini et al., 2021b). Accordingly, considering the same dataset makes it easier to highlight the improvements that the present study introduces. Hereafter, only the essential information about the biology of the species, of the model parameterization, and of the field experimentation will be reported.

The life cycle of *D. suzukii* is composed of an egg stage,  $x_e(t)$ , three larval instars,  $x_{L1}$ ,  $x_{L2}$ , and  $x_{L3}$ , a pupa stage,  $x_p(t)$ , an adult male,  $x_{Am}$ , and an adult female stages (Winkler et al., 2021). Adult females are divided in two substages that we identify as “non-mated females”,  $x_{Amf}$ , and “mated females”,  $x_{Amf}$ . This subdivision leads to set  $\beta_2(t) = 0$  (Bono Rosselló et al., 2022; Rossini et al., 2021b; Rossini et al., 2022b). According to the existing literature, adult females mate only once in their life cycle (Emiljanowicz et al., 2014) that is, mathematically,  $G_{1 \rightarrow 2}(t) = 0$ . The transition rate between non mated and mated females, instead, can be roughly estimated considering that: *i*) in a cultivated field there are the optimal conditions in terms of food availability and habitat, and accordingly *ii*) there is a high probability that males and females will mate. In other words, all the non-mated females will become mated except for the portion that die, namely  $G_{1 \rightarrow 2}(t) = 1 - M[T(t)]$ . The sex ratio, on the other hand, is a constant value of  $S_R = 0.5$  (1:1, males:females) (Emiljanowicz et al., 2014). In this study we consider that the development, mortality, and fertility rates depend exclusively on temperature, already considered as the main driving variable in (Rossini et al., 2020e; Rossini et al., 2021b). Let us detail the functions involved.

Regarding the development rate function, in (Rossini et al., 2020e; Rossini et al., 2021b) the authors proposed the use of the Briere equation (Briere et al., 1999):

$$G[T(t)] = aT(t)(T(t) - T_L)(T_M - T(t))^{\frac{1}{m}}, \quad (17)$$

where  $a$  and  $m$  are empirical parameters, and  $T_L$  and  $T_M$  are the lower and upper temperature thresholds above and below which the development is theoretically not possible. The parameter values were already estimated in (Rossini et al., 2020e; Rossini et al., 2021b) based on the life tables data provided by Tochen et al. (Tochen et al., 2014). It is worth reminding that the function (17) covers the egg-adult stage in the case of *D. suzukii*. In fact, given the literature information available (Ryan et al., 2016; Tochen et al., 2014; Winkler et al., 2021), it is not possible to

estimate the parameters of the Eq. (17) singularly for each life stage.

Temperature-dependent mortality rates are defined by the so-called “bathtub function” of Wang et al. (Wang et al., 2002). The daily loss of individuals due to mortality is mathematically described by the following fourth-order polynomial function:

$$M[T(t)] = a_1T(t)^4 + b_1T(t)^3 + c_1T(t)^2 + d_1T(t) + e_1, \quad (18)$$

with  $a_1$ ,  $b_1$ ,  $c_1$ ,  $d_1$  and  $e_1$  empirical parameters with no biological meaning. As already stated for the development rate, the mortality function also covers the egg-adult stage, with no distinction among the life stages.

At this point, it is worth clarifying how mortality is introduced in the adult stages. The concept of “development” is not completely suitable for adult stages, as there are no further stages to go in. This is the reason why in the last three equations of the model (1) we only consider the mortality rate function  $M(t)$ . At the same time,  $M(t)$  can be considered as the combination of two terms, one describing the adult temperature-dependent longevity (identified by the Briere function (17)) and the temperature-dependent mortality (18). Accordingly, while mortality for the egg-pupa stages is only described by the Eq. (18), in the case of adults we have

$$\begin{cases} M^{Am}[T(t)] = G[T(t)] + M[T(t)] \\ M^{Amf}[T(t)] = G[T(t)] + M[T(t)] \\ M^{Amf}[T(t)] = G[T(t)] + M[T(t)]. \end{cases} \quad (19)$$

Temperature-dependent fertility indicating the number of eggs per day produced by the adult females is described by the following Gaussian-like function (Ryan et al., 2016):

$$\beta_1[T(t)] = \begin{cases} \alpha \left[ \frac{\gamma + 1}{\pi \lambda^{2\gamma+2}} (\lambda^2 - ([T(t) - \tau]^2 + \delta^2))^{\gamma} \right] & \text{if } T_{min} < T(t) < T_{max} \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

where the parameters  $\alpha$ ,  $\gamma$ ,  $\lambda$ ,  $\delta$  and  $\tau$  in Eq. (20) are empirical, while  $T_{min}$  and  $T_{max}$  are the lower and upper boundary below and above which egg production is not theoretically foreseen, respectively.

All the parameters from Eqs. (17)–(20) specific for *D. suzukii* are listed in Table 2, together with their respective literature of reference.

According to the previous assumptions, the model (1) applied to *D. suzukii* is the following:

**Table 2**

List of the parameters of the development, fertility and mortality rates specific for the case of *Drosophila suzukii*. The standard error (SE) is reported only if available from the cited literature.

| Rate function     | Parameter $\pm$ SE                  | Reference                                                           |
|-------------------|-------------------------------------|---------------------------------------------------------------------|
| Brière<br>(17)    | $a = (1.20 \pm 0.15) \cdot 10^{-4}$ | (Rossini et al., 2020e; Rossini et al., 2021b; Tochen et al., 2014) |
|                   | $T_L = 3 \pm 2$                     |                                                                     |
|                   | $T_M = 30 \pm 1$                    |                                                                     |
|                   | $m = 6 \pm 3$                       |                                                                     |
| Mortality<br>(18) | $a_1 = (-5 \pm 1) \cdot 10^{-5}$    | (Ryan et al., 2016)                                                 |
|                   | $b_1 = (5 \pm 8) \cdot 10^{-4}$     |                                                                     |
|                   | $c_1 = 0.1 \pm 0.2$                 |                                                                     |
|                   | $d_1 = (2.2 \pm 0.3) \cdot 10^{-5}$ |                                                                     |
|                   | $e_1 = 1.3 \pm 0.9$                 |                                                                     |
|                   | $\alpha = 659.06$                   |                                                                     |
| Fertility<br>(20) | $\gamma = 88.53$                    | (Ryan et al., 2016)                                                 |
|                   | $\lambda = 52.32$                   |                                                                     |
|                   | $\delta = 6.06$                     |                                                                     |
|                   | $\tau = 22.87$                      |                                                                     |
|                   | $T_{min} = 5$                       |                                                                     |
|                   | $T_{max} = 30$                      |                                                                     |

$$\begin{cases}
\frac{d}{dt}x_e(t) = G(t)\beta_1(t)x_{A_2}(t) - G(t)x_e(t) - M[T(t)]x_e(t) \\
\frac{d}{dt}x_{L_1}(t) = G(t)x_e(t) - G(t)x_{L_1}(t) - M[T(t)]x_{L_1}(t) \\
\frac{d}{dt}x_{L_2}(t) = G(t)x_{L_1}(t) - G(t)x_{L_2}(t) - M[T(t)]x_{L_2}(t) \\
\frac{d}{dt}x_{L_3}(t) = G(t)x_{L_2}(t) - G(t)x_{L_3}(t) - M[T(t)]x_{L_3}(t) \\
\frac{d}{dt}x_P(t) = G(t)x_{L_3}(t) - G(t)x_P(t) - M[T(t)]x_P(t) \\
\frac{d}{dt}x_{A_m}(t) = (1 - S_R(t)) \cdot G(t) \cdot x_P(t) - [G(t) + M[T(t)]]x_{A_m}(t) \\
\frac{d}{dt}x_{A_{f_1}}(t) = S_R(t)G(t)x_P(t) - x_{A_{f_1}}(t) \\
\frac{d}{dt}x_{A_{f_2}}(t) = [1 - G(t) - M[T(t)]]x_{A_{f_1}}(t) - [G(t) + M[T(t)]]x_{A_{f_2}}(t).
\end{cases} \quad (21)$$

## 2.5. Field data, sensing model and parameters

The model (1) was validated in open loop in (Rossini et al., 2021b) using data collected during a three-year survey (2017, 2018, and 2019) conducted at two experimental orchards located in the Sabina area (Lazio, Central Italy) (Rossini et al., 2020e; Rossini et al., 2023). For the sake of this study, we will consider only the year 2018 in our simulations, as it is the year with the higher number of measurements.

The 2018 field survey was carried out from 19<sup>th</sup> April to 15<sup>th</sup> December in two experimental organic orchards located in the municipality of Monterotondo and Montelibretti (Lazio, Central Italy). The orchards received standard agronomical practices and no insecticide treatment was carried out during the duration of the survey. Referring the most interested readers to (Rossini et al., 2020e; Rossini et al., 2021b; Rossini et al., 2023), each orchard was monitored through three Drosophila-Traps (Biobest, Waterloo Belgium) lured with Drosophila drink (Azienda Agricola Prantil, Priò, Trento, Italy) (Grassi et al., 2015; Stacconi et al., 2019; Tait et al., 2018), and inspected weekly. Traps provided only the adult male population measurements, as they were easier to recognize thanks to the black spots on their wings (Hauser, 2011; Ibouh et al., 2019).

Daily average temperatures, calculated considering 48 measurements per day, were provided by two meteorological stations respectively close to the experimental fields, and managed by the ARSIAL agency (Regional Agency for the Development of Innovation and Agriculture in Lazio) (ARSIAL, 2022).

Note that the model (21) needs to be adapted to the theoretical framework presented in this study. For this purpose, given that only the adult males are monitored (Rossini et al., 2020e; Rossini et al., 2023), a ninth state is added to the set of Eq. (21)

$$\dot{x}_{trap,A_m}(t) = (\varepsilon(t) + w_{trap,i}(t))x_{A_m}^i(t) - r_i(t) \cdot x_{trap,A_m}^i(t),$$

where  $\varepsilon(t)$  is the trap efficiency, and the additional mortality in the case of males is formalized as

$$\dot{x}_{A_m}(t) = (1 - S_R(t)) \cdot (G_{L_{n-1}}(t) + w_{A_m,1}(t)) \cdot x_{L_{n-1}}(t) - (\overline{M}_{A_m}(t) + w_{A_m,2}(t))x_{A_m}(t),$$

where  $\overline{M}_{A_m}(t) = M_{A_m}(t) + M_{trap}(t)$ .

Mortality and trap efficiency are not necessarily the same phenomenon, as some insects might be attracted but not caught by traps (Onufrieva and Onufriev, 2021). Therefore, it can be assumed that there is an additional mortality associated with the traps such that  $\varepsilon(t) > M_{trap}(t)$ , where  $\varepsilon(t)$  is the trap efficiency and  $M_{trap}(t)$  the trap mortality. The term  $M_{trap}(t)$  is commonly neglected if traps are sparse and used only for monitoring purposes. In other words, we are assuming that the portion of caught individuals is substantially lower than the total population on the whole field. This also might imply that the

orchard under study is not a closed system, and that there is an exchange of individuals with the neighbouring orchards.

There are two main situations that may occur when using traps. The first situation occurs when incoming and outgoing flows of individuals are in equilibrium. In this case, even though some individuals are trapped, this very low value is assumed to be balanced by individuals migrating from neighbouring orchards. The opposite case occurs when traps are used for mass trapping control actions. In that case, the mortality associated with traps is no longer negligible and should be considered by the model. A lower value of  $M_{trap}(t)$ , if compared to  $\varepsilon(t)$ , might imply that traps attract additional insects from outside the population of interest, missing the condition of equilibrium previously explained.

Despite being fundamental pieces of information, the efficiency and the mortality rates associated with a single trap are usually unknown due to the lack of techniques to estimate these values (Onufrieva and Onufriev, 2021). For the sake of this study, we have roughly estimated the trap efficiency and mortality of the traps used in the field as  $\varepsilon(t) = 0.2$  and  $M_{trap}(t) = \varepsilon(t) \cdot 0.8$ , respectively. These values were tuned by using the field data available. Simulations carried out with different parameter values were compared with field data to assess their accuracy and influence. Traps were inspected weekly, so the measuring and reset periods for the traps is set to 7 days. The parameters involved in the sensing part of the model are summarized in Table 3.

The use of traps and the monitoring of only adults males, described by the additional state  $x_{trap,A_m}(t)$ , provides the following observation matrix

$$C = [0_{1 \times 8} \quad 1],$$

where all the individuals caught by the trap are measured, and the measurement equation

$$y_k = \gamma_k(Cx_k + \nu_k x_k), \quad (22)$$

with  $y_k \in \mathbb{R}$  and  $x_k \in \mathbb{R}^9$ .

## 2.6. Numerical solutions

The model (2) applied to the case of *D. suzukii* has been solved according to the procedure already detailed in (Rossini et al., 2021b; Rossini et al., 2022b). The procedure has been encoded in a Matlab (vers. R2018b) script publicly available at the GitHub page [https://github.com/Niboros91/Ectotherms\\_estimation\\_field](https://github.com/Niboros91/Ectotherms_estimation_field).

## 2.7. Model validation: open versus closed loop and related issues

The classical way of validating entomological models is the comparison between simulations and field data (Pasquali et al., 2019; Rossini et al., 2020b; Rossini et al., 2022a). Different indices are usually considered to quantify the distance between the experimental and the simulated populations (Bellocchi et al., 2011; Ikemoto and Kiritani, 2019; Ratkowsky and Reddy, 2017). The consideration of the noise associated with the field measurements and their active use in the estimator scheme does not allow us to validate the outcome of the simulations in such a way. Field data in ecology are often difficult to obtain and no accurate measurements exist that can be assumed as

**Table 3**  
List of the model parameters used in the simulation.

| Parameter          | Value                   | Description                             |
|--------------------|-------------------------|-----------------------------------------|
| $\varepsilon$      | 20%                     | Trap efficiency.                        |
| $M_{trap}$         | $\varepsilon \cdot 0.8$ | Trap mortality.                         |
| $t_{trap}$         | 7 days                  | Time between two trap inspections.      |
| $x_{ini}$          | $10^6$                  | Initial conditions.                     |
| $\pm \sigma_{ini}$ | 20%                     | Perturbation of the initial conditions. |



ground truth in this scenario. Therefore, the only data available is the sparse and noisy data from trap measurements. For the sake of this study, the validation of the estimation scheme focuses mainly on the improvement of the estimation of the adult male population based on the observed trends in the trap measurements, and the convergence of the estimation.

The second aspect we evaluated in the simulations is related to the initial conditions. In ecology, open loop models tend to provide reasonable results under the strong assumption that the initial conditions are well known. In real applications, however, this kind of information is missing and has to be estimated. Unfortunately, very few works in the literature focus on the choice of the initial conditions of ecological models (Pasquali et al., 2019), which makes us suggest that this part of the modelling process needs to be further explored. The improvements introduced by the EKF help to overcome this common issue related to the initial conditions. Thus, to emphasize this aspect of the proposed scheme, we have carried out simulations by randomly perturbing the initial conditions estimate in 100 different simulations.

The expected outcome from these simulations is a higher variability in the model response when simulations are carried out in an open loop with respect to the EKF scheme. This comparison should highlight how the EKF scheme presented in this study does not only improve the model estimation based on field measurements but also overcomes the problem of assigning accurate initial conditions to the system. Moreover, this part of the study will show how the implementation of the EKF scheme on the model (1) is necessary to effectively apply a “real time” monitoring and estimation scheme, instead of the classical simulations based on a posteriori comparison between field data and model solutions.

To summarize, we compare:

- Estimation based on the open loop model and noisy temperature measurements (traditional approach).
- Estimation based on the EKF using noisy counting measurements and noisy temperature measurements (presented approach).

Note that given the absence of a ground truth regarding the adult

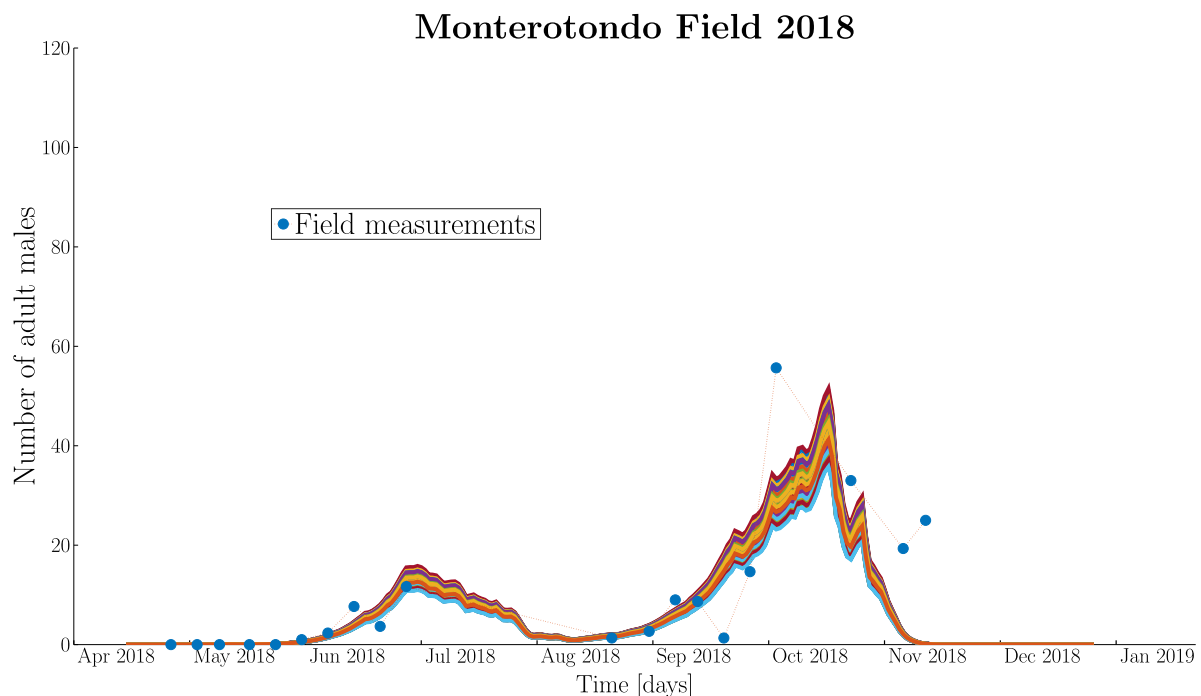
male population, the evaluation remains at the qualitative level, but informative enough to understand how the theoretical framework performs. One of the main aspects evaluated is the sensibility of the simulations to the modification of the estimated initial conditions and how they converge to similar values besides these initial errors.

### 3. Results and discussion

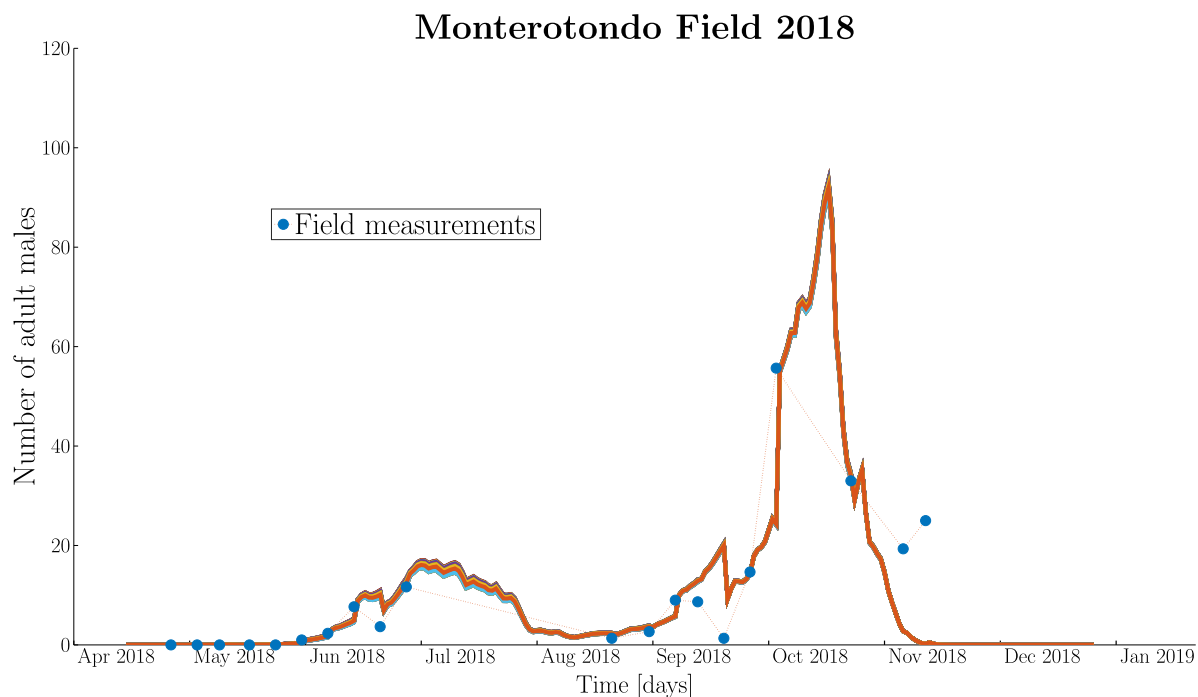
As stated in the previous sections, we have carried out simulations using *D. suzukii* as case study. We will present the results related to the two orchards of interest during the 2018 growing season. In these simulations both scenarios, open loop and EKF, are subject to the same perturbations of the initial conditions. We will focus on two aspects of the male population trends: i) the dependence of the open loop and EKF scheme on the initial conditions in terms of variability of the simulation outcome, and ii) the capability of the EKF to better represent the field data compared to the open loop scheme.

The Monterotondo field simulations using an open loop approach and using the EKF are depicted in Fig. 2 and Fig. 3, respectively. These simulations show a great disparity between the open loop and the EKF case. In the open loop case there are large differences between the simulations, highlighting the strong variability associated with the initial estimated values provided as input to the model (Fig. 2). It is worth pointing out that the differences between the simulations are higher as the population increases. As the population approaches its maxima, the differences between the simulations corresponding to various sets of initial conditions grow. This is mainly due to the error on the estimate, i.e., the difference between the model estimate and the actual state, which depends on the evolution of the system (21). When the system is stable, the differences are reduced, while when the system is unstable (around the peaks) the differences are larger.

This high dependence on the initial conditions is not present in the EKF case (Fig. 3). This fact demonstrates that the estimation of the population is effectively corrected, step by step, by the measurements. Accordingly, the corrections carried out by the EKF scheme end up providing a similar evolution in the estimated values for the 100



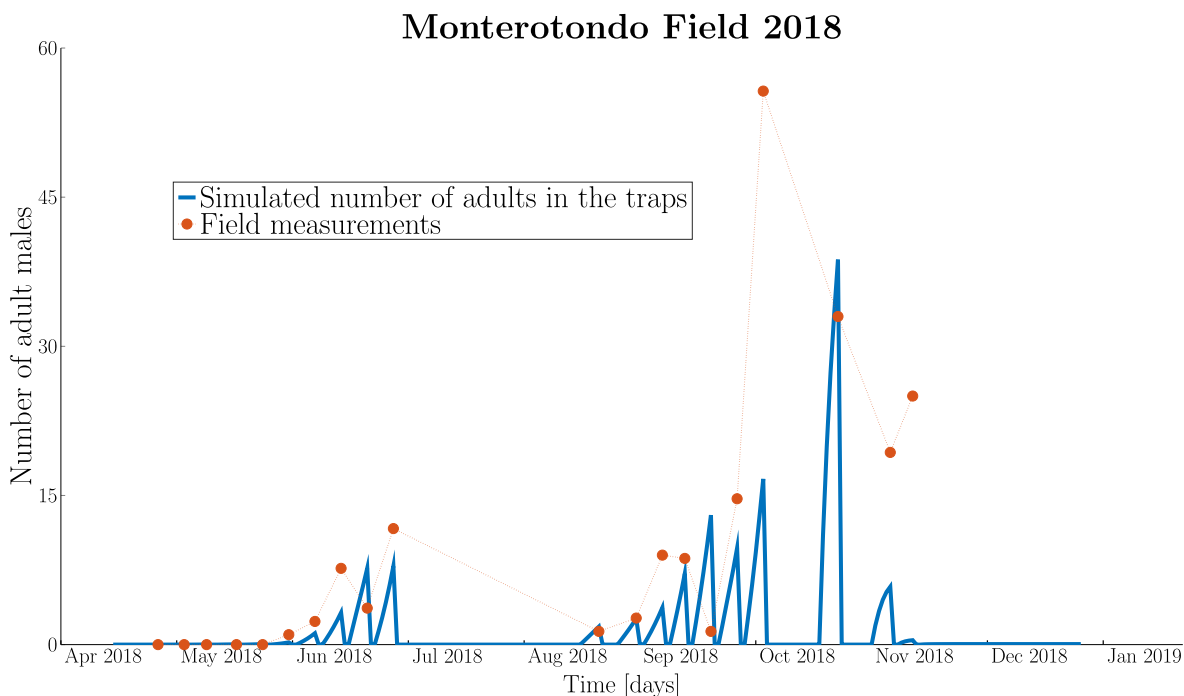
**Fig. 2.** Evolution of the number of adult males at the Monterotondo field in 2018 for the open loop simulation. Each line represents a simulation with different estimated initial conditions. Blue dots represent field measurements recorded during 2018. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 3.** Evolution of the number of adult males at the Monterotondo field in 2018 for the EKF simulation. Each line represents a simulation with different estimated initial conditions. Blue dots represent field measurements recorded during 2018. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

simulations, despite being initialized with different estimated initial conditions. In this sense, it is interesting to note that, in the case of the EKF simulations, there is more variability in the first peak of the population than in the second (and more abundant) peak. This situation can be explained by the quantity of measurements (field information) that are incorporated to the EKF scheme, which is higher as more field data is

collected. Additionally, this behaviour allows us to qualitatively demonstrate the efficiency of the estimator scheme where, contrary to the open loop simulations, a higher population does not necessarily provide a higher variability. As expected, this result provides the very interesting conclusion that if the estimator is properly tuned, the problem of estimating the initial conditions can be heavily reduced.



**Fig. 4.** Evolution of the number of adult males trapped at the Monterotondo field in 2018. The blue line represents the evolution of the additional state variable of the model describing trapped adult males while the red dots represent the filed measurements. For the sake of exposition only one of the simulations is depicted. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

By comparing Fig. 2 and Fig. 3 we can state that the introduction of the EKF scheme provides a better estimation of the field population of males than the open loop case. According to the results of the first orchard, we can claim that the EKF improves the description of the population trend and the robustness of the estimation, meant as the reduction of the dependence on the initial condition values.

It is worth pointing out that a quantitative evaluation of the improvements introduced by the EKF scheme is still not possible at this stage, and that we need further investigations to formulate a more refined protocol of evaluation. Despite this fact, the theoretical framework already showed its reliability in the case of Monterotondo orchard.

An additional piece of information that supports the improvements introduced by the EKF scheme is provided in Fig. 4. This plot shows the only feasible comparison between the trap measurements and the estimated trap values from the additional state (representing the catches). Fig. 4 shows that the EKF estimation follows quite accurately the trend in the measurements with the only difference at the high peak occurred in September. This difference may be due to uncontrolled effects related to the traps, or to the high variability of the conditions that usually affect open field experiments.

The data from the Montelibretti orchard provided similar results, confirming the outcomes already observed in the Monterotondo dataset. The outcomes of these simulations are provided in Fig. 5 and Fig. 6.

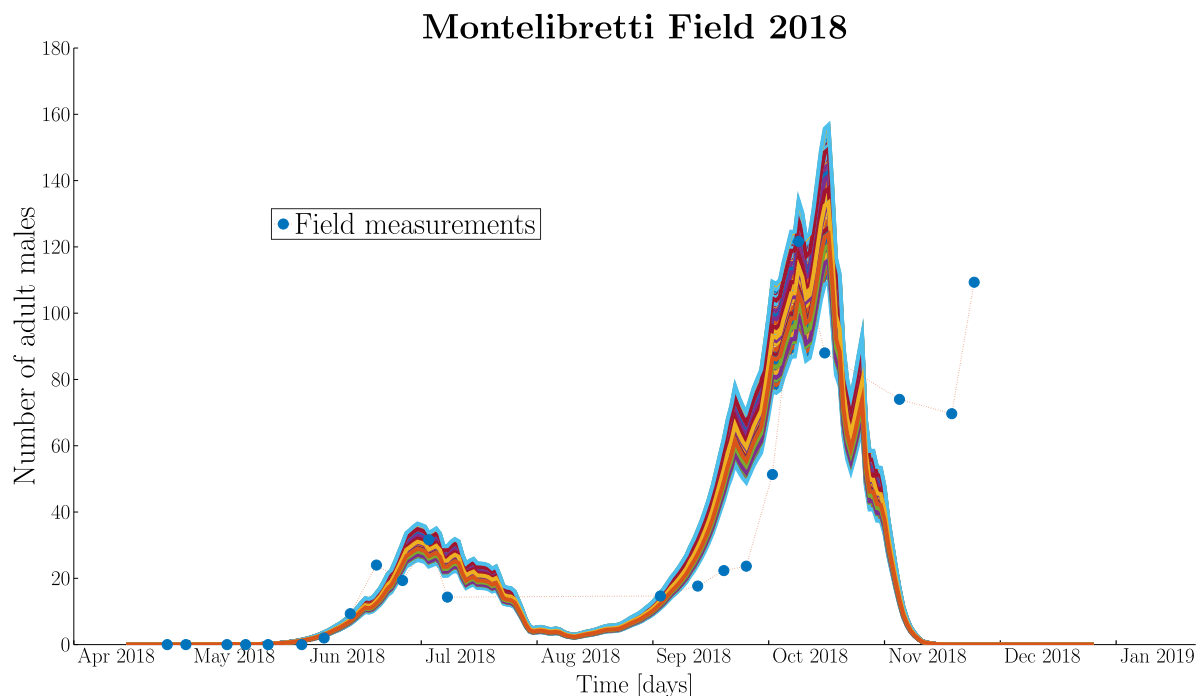
Also in this case, the first aspect to point out concerns the dependence of the simulations on the initial conditions, which is even stronger than in the Monterotondo case previously shown in Fig. 2. Coherently, the variability of the simulations due to different set of initial conditions is higher as the population increases approaching to its maxima. On the other hand, simulations in Fig. 6 confirm that the EKF scheme correctly updates the estimations using field data, with a subsequent reduction of the variability under different sets of initial conditions. Also in this case, the highest difference between simulations depicted in Fig. 6 occurs at the first peak of population densities, and not at their maxima. These results confirm the first results obtained from the simulations in the Monterotondo field.

The simulated trend of trapped individuals at the Montelibretti

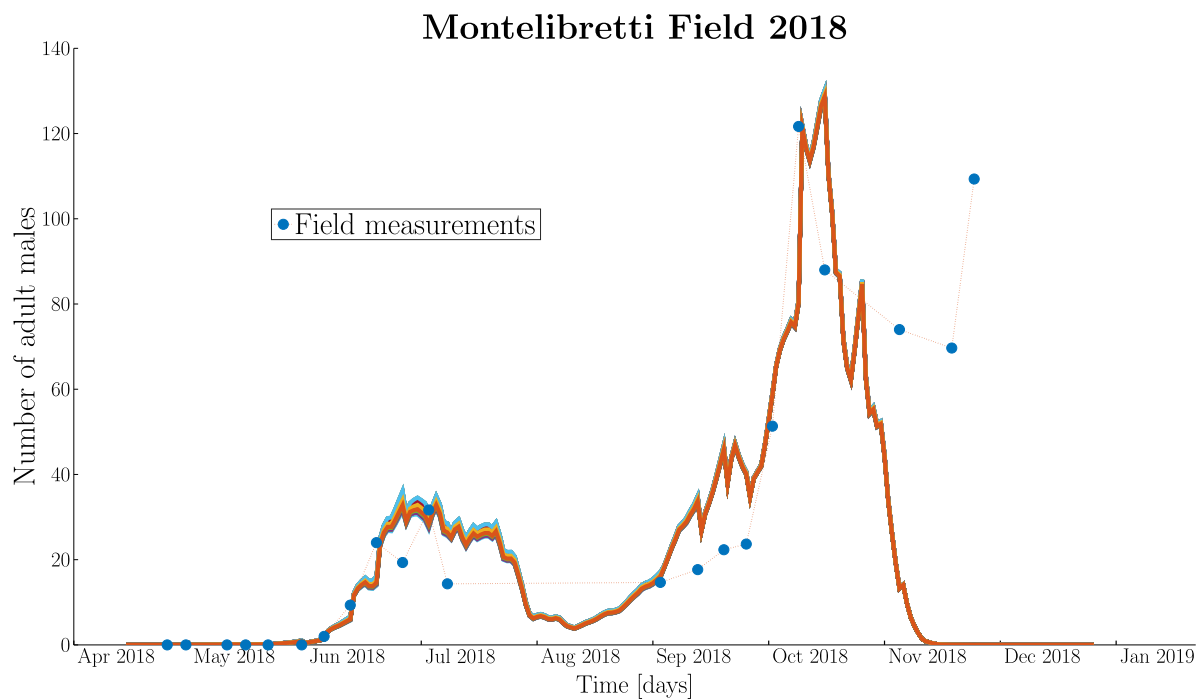
orchard is reported in Fig. 7. The measurements are well represented by the estimated value, except for the last three measured data points of the season. There are two possible explanations for this discrepancy. The first source of discrepancy can be the model parameterization. It is in fact possible that the rate functions considered to describe the biology of *D. suzukii* tend to overestimate both development and mortality in certain conditions. In low temperatures conditions, as in late autumn, the model possibly overestimates the population at the point that an observer scheme cannot correct anymore the model prediction. This discrepancy suggests the need for the improvement, in further studies, of the parameterization of the model and to better investigate its behaviour at low temperature conditions.

A second possible explanation could be associated with the increase in efficiency of the traps during late autumn. *D. suzukii* was monitored through traps activated by a food-based lure (Grassi et al., 2015). Two main problems are associated with this kind of lure, even though not enough demonstrated by the scientific community. Firstly, the efficiency of the lure attracting individuals may depend on the environmental conditions, with a subsequent increase or decrease as the season goes on. Secondly, the efficiency of the traps may be conditioned by the presence of fruits in the orchard. When there are fruits on the trees or on the soil, the individuals are more attracted by the fruits than by the lure. This condition creates a *competition* between fruits and trap lures, with a possible decrease of trap efficiency. On the other hand, at the end of the growing season when there are no more fruits from cultivated nor wild plants, the trap lures are the only attractant. Based on this precondition, it might be possible that the model parameterization is still valid, but measured data are biased by the aforementioned uncontrolled effects.

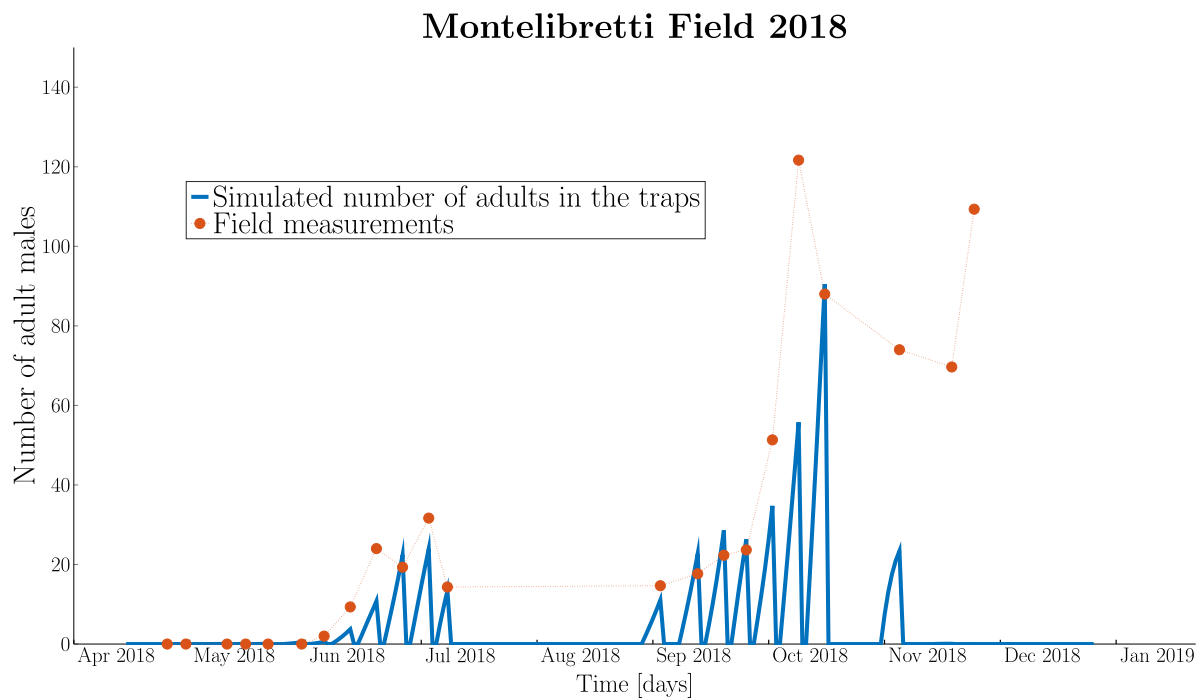
Overall, the results obtained show that, with a state prediction and correction scheme, we can provide a more accurate estimation of the population over time, given that the open loop approach strongly depends on the initial conditions of the simulation. Accordingly, in open loop a wrong estimation of the initial conditions provides a not reliable overall estimation, leading the users of a DSS to wrong decisions, while an EKF approach can overcome this issue. Note that from an experimental point of view two main problems make the estimation of the



**Fig. 5.** Evolution of the number of adult males at the Montelibretti field in 2018 for the open loop simulation. Each line represents a simulation with different estimated initial conditions. Blue dots represent field measurements recorded during 2018. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 6.** Evolution of the number of adult males at the Montelibretti field in 2018 for the EKF simulation. Each line represents a simulation with different estimated initial conditions. Blue dots represent field measurements recorded during 2018. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 7.** Evolution of the number of adult males trapped at the Montelibretti field in 2018. The blue line represents the evolution of the additional variable describing traps adult males while the red dots represent the filed measurements. For the sake of exposition only one of the simulations is depicted. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

initial conditions difficult: *i*) the knowledge of the initial distribution of the individuals into the different age classes, information that is practically impossible to accurately assess in natural environments, and *ii*) the time “zero” at which the simulation should start.

While the former point has been already discussed, the latter deserves more attention. Most physiologically-based models present in the

literature describe the development of terrestrial arthropods under “normal” conditions, namely when the environmental conditions are favourable for their development. However, it is known that when adverse environmental conditions occur (e.g., winter) the individuals activate diapause or dormancy mechanisms (Saunders, 2014). These mechanisms are not described by the model and small amount of



information is available in the literature from the modelling point of view (Baumgärtner et al., 2012). This leads model scientists to empirically define a “day zero” to start simulations (Ponti et al., 2021), which is often based on rough field monitoring or on empirical daily average temperature thresholds. The conventional “day zero” is a concept that worked well for degree day accumulation models and it is commonly named “BioFix”, but it is not applicable to physiologically-based models. To the best of our knowledge, few authors (e.g., (Pasquali et al., 2019)) conducted detailed studies on the effect of the initial conditions in the outcome of the simulations. It is worth pointing out that we are not referring to stability analyses, a common mathematical practice carried out in theoretical papers (e.g. (Jensen et al., 2019; Pappalardo et al., 2021; Singh and Emerick, 2021; Tang and Zhou, 2005; Zhang and Xiao, 2016)). We are instead referring to the comparison of different simulation outputs with the same model parameters when the initial values for the simulations vary. The EKF approach introduced in this study overcomes this issue, since it reduces the dependence on the set of initial values, and it corrects the estimation as new measurements are available.

The EKF approach overcomes an additional aspect commonly neglected by other mathematical models: the influence of immigration or emigration of individuals from and towards neighbouring fields. Most models, including (1), implicitly assume that the population is closed, i. e., there are no exchanges of individuals. However, this situation resembles more to an equilibrium in the exchange of individuals, where the amount of insects that moves to the neighbouring fields is balanced by the same incoming flux. Even though this condition may be valid for homogeneous fields, it may cause disturbances in the normal process of open-loop validation and application. The EKF approach, instead, corrects the model estimation based on field observations, so that even when the equilibrium between the incoming and outgoing flow of individuals is perturbed, the model updates accordingly.

As we stated in the introduction of this study, the use of estimators is not totally new in agriculture. However, their application has been more related to the engineering and the economical aspects of agriculture than to the biological systems themselves. We believe that this study represents a significant step forward in agricultural management, thanks to the application of this theoretical scheme to a biological problem that can be further extended to more natural ecosystems. This study follows the direction of previously published studies that initially suggested to use KF for biological modelling purposes.

An example of these applications can be found in the work of Ennola et al. (Ennola et al., 1998), where the KF was applied to improve the performance of a zooplankton population model. The authors showed that the KF was effective in reducing the measurement noise and in estimating the parameters of the model. Even though their application concerns a different biological system, their results are equivalent to the ones we observed with pest populations. Other examples of applications of KF to biological systems were presented by Gauthier et al. (Gauthier et al., 2007) in the case of populations of greater snow goose, and by Sullivan (Sullivan, 1992) in the case of fish populations. However, there are very few applications of KF to pest populations besides Bono Rosselló et al. (Bono Rosselló et al., 2022). A pioneering study was conducted by Zavaleta and Dixon (Zavaleta and Dixon, 1982) on *Ceratomyia trifurcata* (Forster), a Coleoptera Chrysomelidae that infests bean leaves. The authors pointed out the potential that the application of KF could have in the monitoring of pest population, which was hereby confirmed by our improved theoretical framework. Thus, based on the cited literature and the encouraging results of the present study, we argue that the application of estimators should be further improved and extended to other species of both agriculture and forest interest.

#### 4. Conclusion

This work introduced a theoretical framework to update the estimation of physiologically-based models using an Extended Kalman

Filter approach. The theoretical development started from a population density model describing insect populations, which was adapted to the estimator scheme so to improve its estimation performance.

We believe that the outcomes of this study can be of great support for the further development of DSS. To date, field data have been only used for validation purposes or to estimate model parameters, such as in the works of Pasquali et al. (Pasquali et al., 2022), Gilioli et al. (Gilioli and Pasquali, 2007) or Nance et al. (Nance et al., 2018). However, we believe that field data can also take an active part in the prediction activities. To demonstrate their application in a case of agricultural interest, we have focused on two main scenarios of data acquisition. First, the use of traps where the cumulative values obtained are commonly a source of bias, which showed promising results in the case of *D. suzukii*. Second, we theoretically showed how the presented framework can also be used in the case of instantaneous measurements, thus covering the largest part of the monitoring techniques in ecology.

Our study confirms that open loop simulations are purely valid as descriptive field populations studies, not as real-time monitoring and predictive tools. Also, we show that an empirical estimation of the initial conditions does not ensure the best starting point for simulations, as it can greatly affect the outcome of the simulation. The EKF method, instead, provides an indirect solution to this problem, as it presents a more robust approach that fits better DSS systems paradigm.

It must be noted that open loop validation is still a necessary step for model development. The open loop model validation has a well-defined protocol of application that allows to quantify the goodness of fit of a given model by comparing recorded field data and simulation. Several authors proposed indices to quantify this goodness of fit, as for instance in (Bellocchi et al., 2011; Ikemoto and Kiritani, 2019; Ratkowsky and Reddy, 2017; Rossini et al., 2020c). In this sense, a mathematical model that describes well a given phenomenon is a good starting point for building a predictive system to include in DSS (Grimm et al., 2020). Once a model is validated, and its capability of describing the natural phenomena under study is assessed, it can be included in the EKF scheme for decision support activities.

Moreover, it is worth mentioning that the quality of the estimation depends on the quality of the data, which points to an additional aspect that should be considered in future works. The current state of the art presents several techniques to monitor the insect population abundance over time. However, more rigorous information on their efficiency and on the degree of uncertainty of the values measured (measurement noise) should be provided. Specific studies on this aspect would be beneficial to define standards in field data acquisition and to increase their overall quality, with a subsequent increase of the quality of the model validation and estimation.

To conclude, we may say that this work makes a great step forward towards a proper inclusion of physiologically-based models in the DSS framework, while further research will be focused on the EKF approach strengths and weaknesses in the entomological field. The case study of *D. suzukii* is certainly a good starting point, but future works will focus on two main aspects. First, the theoretical scheme should be applied to more monitoring campaigns of *D. suzukii* and other species of agricultural interest. Second, the theory should be applied to practical cases of the same species but considering additional monitoring techniques so to observe any difference in performance.

#### Declaration of Competing Interest

The authors declare that they have no conflict of interest.

#### 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. The project was carried out in the framework of the project “Smart testing”, funded by the Fonds de la Recherche Scientifique (FNRS) under the grant number 40003443. L.R. is funded by the European Commission, MSCA-PF-2022 project “Pest-Finder” n. 101102281 and 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.

## References

- Anderson, B., Moore, J.B., 1979. Optimal filtering. In: Prentice-Hall Information and System Sciences Series.
- ARSIAL, 2022. Sialr - servizio integrato agrometeorologico della regione lazio [online].
- Asplen, M.K., Anfora, G., Biondi, A., Choi, D.-S., Chu, D., Daane, K.M., Gibert, P., Gutierrez, A.P., Hoelmer, K.A., Hutchison, W.D., Isaacs, R., Jiang, Z.-L., Kárpáti, Z., Kimura, M.T., Pascual, M., Phillips, C.R., Plantamp, C., Ponti, L., Véték, G., Vogt, H., Walton, V.M., Yu, Y., Zappalà, L., Desneux, N., 2015. Invasion biology of spotted wing drosophila (*Drosophila suzukii*): a global perspective and future priorities. *J. Pest. Sci.* 88, 469–494. <https://doi.org/10.1007/s10340-015-0681-z>.
- Bange, M., Deutscher, S., Larsen, D., Linsley, D., Whiteside, S., 2004. A handheld decision support system to facilitate improved insect pest management in Australian cotton systems. *Comput. Electron. Agric.* 43, 131–147. <https://doi.org/10.1016/j.compag.2003.12.003>.
- Barrett, H., Rose, D.C., 2020. Perceptions of the fourth agricultural revolution: what's in, what's out, and what consequences are anticipated? *Sociol. Rural. O* <https://doi.org/10.1111/soru.12324>.
- Baumgärtner, J., Gutierrez, A.P., Pesolillo, S., Severini, M., 2012. A model for the overwintering process of European grapevine moth *Lobesia botrana* (denis & schiffmuller) (Lepidoptera, Tortricidae) populations. *J. Entomol. Acarol. Res.* 44, 2. <https://doi.org/10.4081/jeur.2012.e2>.
- Bellagamba, V., Cola, G.D., Cavalloro, R., 1987. Stochastic models in fruit-fly population dynamics, pp. 91–98.
- Bellocci, G., Rivington, M., Donatelli, M., Matthews, K., 2011. Validation of biophysical models: Issues and methodologies, vol. 2. Springer Netherlands. [https://doi.org/10.1007/978-94-007-0394-0\\_26](https://doi.org/10.1007/978-94-007-0394-0_26).
- Bono Rosselló, N., Rossini, L., Speranza, S., Garone, E., 2022. State estimation of pest populations subject to intermittent measurements. In: Sensing, Control and Automation Technologies for Agriculture - 7th AGRICONTROL 2022.
- Borlino, A.C., Cola, G.D., Marras, G., 1990. I modelli compartimentali nello studio della dinamica delle popolazioni naturali. *Bollettino della Società Sarda di Scienze Naturali* 27, 77–114.
- Borlino, A.C., Cola, G.D., Marras, G., 1991. Mathematical modelling of natural population dynamics. *Memorie dell'Istituto Italiano di Idrobiologia* 49, 127–162.
- Briere, J.-F., Pracros, P., Roux, A.-Y.L., 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>.
- Cola, G.D., Gilioli, G., 1996. Mathematical models for age-structured population dynamics: An overview. In: 20th International congress of Entomology, Florence, pp. 45–61.
- da Silva, G.J., Dahlet, G., Takagi, M., DelGrossi, M., de Lima, P., Ceolin, S., 2019. From Fome Zero to Zero Hunger. FAO. <https://doi.org/10.4060/CA5524EN>.
- Damos, P., Savopoulou-Soultani, M., 2012. Temperature-driven models for insect development and vital thermal requirements. *Psyche* 2012. <https://doi.org/10.1155/2012/123405>.
- Didham, R.K., Basset, Y., Collins, C.M., Leather, S.R., Littlewood, N.A., Menz, M.H.M., Müller, J., Packer, L., Saunders, M.E., Schönrogge, K., Stewart, A.J.A., Yanoviak, S. P., Hassall, C., 2023. Interpreting insect declines: seven challenges and a way forward. *Insect Conserv. Divers.* 13, 103–114. <https://doi.org/10.1111/icad.12408>.
- Ebrahimi, M., Khoshaghaz, M., Minaei, S., Jamshidi, B., 2017. Vision-based pest detection based on SVM classification method. *Comput. Electron. Agric.* 137, 52–58. <https://doi.org/10.1016/j.compag.2017.03.016>.
- 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>.
- Ennola, K., Sarvala, J., Dévai, G., 1998. Modelling zooplankton population dynamics with the extended kalman filtering technique. *Ecol. Model.* 110, 135–149. [https://doi.org/10.1016/S0304-3800\(98\)00057-X](https://doi.org/10.1016/S0304-3800(98)00057-X).
- Gauthier, G., Besbeas, P., Lebreton, J.-D., Morgan, B.J.T., 2007. Population growth in a snow geese: a modelling approach integrating demographic and survey information. *Ecology* 88, 1420–1429. <https://doi.org/10.1890/06-0953>.
- 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>.
- Gilioli, G., Sperandio, G., Simonetto, A., Ciampitti, M., Gervasio, P., 2022. Assessing the risk of establishment and transient populations of *Spodoptera frugiperda* in Europe. *J. Pest. Sci.* <https://doi.org/10.1007/s10340-022-01517-0>, 6.
- Gomez-Gil, J., Ruiz-Gonzalez, R., Alonso-Garcia, S., Gomez-Gil, F., 2013. A kalman filter implementation for precision improvement in low-cost gps positioning of tractors. *Sensors* 13, 15307–15323. <https://doi.org/10.3390/s131115307>.
- Grassi, A., Anfora, G., Maistri, S., Gottardello, A., Maddalena, G., Cristofaro, A.D., Savini, G., Ioriatti, C., 2015. Development and efficacy of droskidrink, a food bait for trapping *Drosophila suzukii*. *IOBC/WPRS Bull.* 109, 197–204.
- Grimm, V., Johnston, A.S.A., Thulke, H.-H., Forbes, V.E., Thorbek, P., 2020. Three questions to ask before using model outputs for decision support. *Nat. Commun.* 11, 4959. <https://doi.org/10.1038/s41467-020-17785-2>.
- 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>, 8.
- Hauser, M., 2011. A historic account of the invasion of *Drosophila suzukii* (matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. *Pest Manag. Sci.* 67, 1352–1357. <https://doi.org/10.1002/ps.2265>.
- Iboub, K., Oreste, M., Bubici, G., Tarasco, E., Stacconi, M.V.R., Ioriatti, C., Verrastro, V., Anfora, G., Baser, N., 2019. Biological control of *Drosophila suzukii*: efficacy of parasitoids, entomopathogenic fungi, nematodes and deterrents of oviposition in laboratory assays. *Crop Prot.* 125, 104897. <https://doi.org/10.1016/j.cropro.2019.104897>.
- 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>.
- Jactel, H., Bonifacio, L., van Halder, I., Vétillard, F., Robinet, C., David, G., 2023. A novel, easy method for estimating pheromone trap attraction range: application to the pine sawyer beetle *Monochamus galloprovincialis*. *Agric. For. Entomol.* 21, 8–14. <https://doi.org/10.1111/afe.12298>.
- Jensen, G.G., Uekermann, F., Sneppen, K., 2019. Multi stability and global bifurcations in epidemic model with distributed delay sirns-model. *Eur. Phys. J. B* 92, 28. <https://doi.org/10.1140/epjb/e2018-90562-1>.
- Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. *J. Basic Eng.* 82 (1), 35–45. <https://doi.org/10.1115/1.3662552>.
- 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), pp. 342–347. <https://doi.org/10.1109/MED51440.2021.9480344>.
- Ljung, L., 1979. Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems. *IEEE Trans. Autom. Control* 24 (1), 36–50. <https://doi.org/10.1109/TAC.1979.1101943>.
- Manetsch, T.J., 1976. Time-varying distributed delays and their use in aggregative models of large systems. *IEEE Trans. Syst. Man Cybern.* SMC-6, 547–553. <https://doi.org/10.1109/TSMC.1976.4309549>.
- Mesarovic, M.D., Sreenath, S.N., Keene, J.D., 2004. Search for organising principles: understanding in systems biology. *Syst. Biol.* 1, 19–27. <https://doi.org/10.1049/sb:20045010>.
- Murray, J.D., 2012. Vignettes from the field of mathematical biology: the application of mathematics to biology and medicine. *Interface Focus* 2, 397–406. <https://doi.org/10.1098/rsfs.2011.0102>.
- 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>.
- Onufrieva, K.S., Onufriev, A.V., 2021. How to count bugs: a method to estimate the most probable absolute population density and its statistical bounds from a single trap catch. *Insects* 12, 932. <https://doi.org/10.3390/insects12100932>.
- Pappalardo, S., Villa, M., Santos, S.A., Benhadi-Marín, J., Pereira, J.A., Venturino, E., 2021. A tritrophic interaction model for an olive tree pest, the olive moth — *Prays oleae* (Bernard). *Ecol. Model.* 462, 109776. <https://doi.org/10.1016/j.ecolmodel.2021.109776>.
- Pasquali, S., Soresina, C., Gilioli, G., 2019. The effects of fecundity, mortality and distribution of the initial condition in phenological models. *Ecol. Model.* 402, 45–58. <https://doi.org/10.1016/j.ecolmodel.2019.03.019>.
- 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>.
- Plant, R.E., Mangel, M., 1987. Modeling and simulation in agricultural pest management. *SIAM Rev.* 29, 235–261.
- 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>, 8.
- Preti, M., Verheggen, F., Angeli, S., 2021. Insect pest monitoring with camera-equipped traps: strengths and limitations. *J. Pest. Sci.* 94 (2), 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>.
- 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>.
- Rossi, V., Sperandio, G., Caffi, T., Simonetto, A., Gilioli, G., 2019. Critical success factors for the adoption of decision tools in ipm. *Agronomy* 9, 710. <https://doi.org/10.3390/agronomy9110710>.

- Rossini, L., Speranza, S., Contarini, M., 2020a. Distributed delay model and von foerster's equation: different points of view to describe insects' life cycles with chronological age and physiological time. *Eco. Inform.* 59, 101117. <https://doi.org/10.1016/j.ecoinf.2020.101117>.
- Rossini, L., Contarini, M., Severini, M., Speranza, S., 2020b. Reformulation of the distributed delay model to describe insect pest populations using count variables. *Ecol. Model.* 436, 109286. <https://doi.org/10.1016/j.ecolmodel.2020.109286>.
- Rossini, L., Contarini, M., Severini, M., Talano, D., Speranza, S., 2020c. 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., 2020d. 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., Contarini, M., Giarruzzo, F., Assennato, M., Speranza, S., 2020e. Modelling *Drosophila suzukii* adult male populations: a physiologically based approach with validation. *Insects* 11, 751. <https://doi.org/10.3390/insects11110751>.
- Rossini, L., Virla, E.G., Albarracín, E.L., Nieuwenhove, G.A.V., Speranza, S., 2021a. Evaluation of a physiologically based model to predict *Dalbulus maidis* occurrence in maize crops: validation in two different subtropical areas of south america. *Entomol. Exp. Appl.* 169, 597–609. <https://doi.org/10.1111/eea.13066>.
- Rossini, L., Bono Rosselló, N., Speranza, S., Garone, E., 2021b. 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., Bruzzone, O.A., Contarini, M., Bufacchi, L., Speranza, S., 2022a. 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>.
- Rossini, L., Bono Rosselló, N., Contarini, M., Speranza, S., Garone, E., 2022b. Modelling ectotherms' populations considering physiological age structure and spatial motion: a novel approach. *Eco. Inform.* 70, 101703. <https://doi.org/10.1016/j.ecoinf.2022.101703>.
- Rossini, L., Bruzzone, O.A., Speranza, S., Delfino, I., 2023. Estimation and analysis of insect population dynamics parameters via physiologically based models and hybrid genetic algorithm mcmc methods. *Eco. Inform.* 77, 102232. <https://doi.org/10.1016/j.ecoinf.2023.102232>.
- Rupnik, R., Kukar, M., Vračar, P., Košir, D., Pevac, D., Bosnić, Z., 2019. Agrodss: a decision support system for agriculture and farming. *Comput. Electron. Agric.* 161, 260–271. <https://doi.org/10.1016/j.compag.2018.04.001>.
- 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>.
- Saunders, D.S., 2014. Insect photoperiodism: effects of temperature on the induction of insect diapause and diverse roles for the circadian system in the photoperiodic response. *Entomol. Sci.* 17, 25–40 cartaceo: cartella 4. <https://doi.org/10.1111/ens.12059>.
- 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.
- 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>.
- Singh, A., Emerick, B., 2021. Generalized stability conditions for host–parasitoid population dynamics: implications for biological control. *Ecol. Model.* 456, 109656. <https://doi.org/10.1016/j.ecolmodel.2021.109656>.
- Sorensen, C., 2002. Modeling seasonality in agricultural commodity futures. *J. Futur. Mark.* 22, 393–426. <https://doi.org/10.1002/fut.10017>.
- Srbínovska, M., Gavrovski, C., Dimcev, V., Krkoleva, A., Borozan, V., 2015. Environmental parameters monitoring in precision agriculture using wireless sensor networks. *J. Clean. Prod.* 88, 297–307. <https://doi.org/10.1016/j.jclepro.2014.04.036>.
- Stacconi, M.V.R., Grassi, A., Ioriatti, C., Anfora, G., 2019. Augmentative releases of *Trichopria drosophilae* for the suppression of early season *Drosophila suzukii* populations. *BioControl* 64, 9–19. <https://doi.org/10.1007/s10526-018-09914-0>.
- Stafford, J., 2015. Essential technology for precision agriculture. <https://doi.org/10.2134/1996.precisionagproc3.c74>.
- Sullivan, P.J., 1992. A kalman filter approach to catch-at-length analysis. *Biometrics* 48, 237. <https://doi.org/10.2307/2532752>.
- Tait, G., Grassi, A., Pfab, F., Crava, C.M., Dalton, D.T., Magarey, R., Ometto, L., Vezzulli, S., Rossi-Stacconi, M.V., Gottardello, A., Pugliese, A., Firrao, G., Walton, V. M., Anfora, G., 2018. Large-scale spatial dynamics of *Drosophila suzukii* in trentino, Italy. *J. Pest. Sci.* 91, 1213–1224. <https://doi.org/10.1007/s10340-018-0985-x>.
- Tang, Y., Zhou, L., 2005. Hopf bifurcation and stability of a competition diffusion system with distributed delay. *Publ. Res. Inst. Math. Sci.* 41, 579–597. <https://doi.org/10.2977/prims/1145475224>.
- 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>.
- Wang, K.S., Hsu, F.S., Liu, P.P., 2002. Modeling the bathtub shape hazard rate function in terms of reliability. *Reliab. Eng. Syst. Saf.* 75 (3), 397–406. [https://doi.org/10.1016/S0951-8320\(01\)00124-7](https://doi.org/10.1016/S0951-8320(01)00124-7).
- 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.
- Xiaoyun, L., Xiaoling, Z., Xian, X., 2023. Application of kalman filter in agricultural economic forecasting, vol. 3. *IEEE*, pp. 2185–2189. <https://doi.org/10.1109/ICSMC.1996.565487>.
- Zavaleta, L.R., Dixon, B.L., 1982. Economic benefits of kalman filtering for insect pest management. *J. Econ. Entomol.* 75, 982–988. <https://doi.org/10.1093/jee/75.6.982>.
- Zha, M., Qian, W., Yi, W., Hua, J., 2021. A lightweight YOLOv4-based forestry pest detection method using coordinate attention and feature fusion. *Entropy* 23 (12), 1587. <https://doi.org/10.3390/e23121587>.
- Zhang, G., Xiao, A., 2016. Exact and numerical stability analysis of reaction-diffusion equations with distributed delays. *Front. Math. China* 11, 189–205. <https://doi.org/10.1007/s11464-015-0506-7>.