



UNIVERSITÀ
DEGLI STUDI DELLA
TUSCIA

Università degli Studi della Tuscia

DAFNE

Dipartimento di Scienze Agrarie e Forestali

Corso di Dottorato di Ricerca in
Scienze delle Produzioni Vegetali e Animali - XXXIII Ciclo

Development of stochastic models for plant protection

December 2020

Academic year 2019-2020

PhD Candidate:

Luca Rossini

Supervisor:

Prof. Stefano Speranza

PhD course Coordinator:

Prof. Roberta Bernini

S.S.D.:

AGR/11

Mignolo: "Cosa facciamo questa sera, Prof.?"
Prof.: "Quello che facciamo tutte le sere, Mignolo.
Tentare di conquistare il mondo!"

- Mignolo e Prof.

"Io ho bisogno di credere
che qualcosa di straordinario
sia possibile."

- A beautiful mind.

Contents

1	Introduction	9
2	A novel mathematical formulation for the Distributed Delay Model	13
2.1	<i>Abstract</i>	14
2.2	Introduction	14
2.3	Materials and Methods	16
2.3.1	Distributed Delay Model	16
2.3.2	Analytical solutions	19
2.3.3	Development rate functions	20
2.3.4	Numerical solutions of both the DDM versions	22
2.3.5	Comparison between the DDM versions	22
2.4	Case studies: <i>Anthonomus eugenii</i> Cano and <i>Lobesia botrana</i> (Denis & Schif- fermüller)	23
2.4.1	Greenhouse scenario: <i>A. eugenii</i>	23
2.4.2	Open field scenario: <i>L. botrana</i>	24
2.5	Results and discussion	25
2.5.1	<i>Anthonomus eugenii</i>	25
2.5.2	<i>Lobesia botrana</i>	26
2.6	Conclusions	29
3	The milestones of a novel software for model applications in entomology	33
3.1	<i>Abstract</i>	34
3.2	Introduction	34
3.3	Materials and Methods	36
3.4	Results	39
3.5	Discussion and Conclusions	44
4	Towards a new model: limitations of the Distributed Delay Model	47
4.1	<i>Abstract</i>	48
4.2	Introduction	48
4.3	Materials and Methods	49
4.3.1	Connections between Distributed Delay Model and Von Foerster equation	49
4.3.2	Verification of expression (4.8)	53

4.3.3	<i>Lobesia botrana</i> : biology and field survey	54
4.4	Results	55
4.5	Discussion and Conclusions	57
5	A novel model: the “Generalised Von Foerster equation”	59
5.1	<i>Abstract</i>	60
5.2	Introduction	60
5.3	Materials and Methods	61
5.3.1	Development rate functions	62
5.3.2	The population density model	65
5.3.3	The simulator	69
5.3.4	Experimental design for validation	71
5.4	Results and Discussion	71
5.5	Conclusions	75
6	A protocol to apply and validate the model: <i>a priori</i> and <i>a posteriori</i> analysis	77
6.1	<i>Abstract</i>	78
6.2	Introduction	78
6.3	Materials and Methods	79
6.3.1	Population modelling	79
6.3.2	Data analysis	80
6.3.3	Experimental design for validation	81
6.4	Results	82
6.5	Discussion	85
6.6	Conclusions	86
7	A revised version of <i>EntoSim</i>: new models and new functions	87
7.1	<i>Abstract</i>	88
7.2	Introduction	88
7.3	Materials and Methods	89
7.3.1	Parameters estimation	89
7.3.2	Population modelling	91
7.4	Data collection and comparison between simulations and monitoring	93
7.5	Results	93
7.5.1	Parameter estimation from life tables	93
7.5.2	Simulations and validations	95
7.6	Discussion and conclusions	97
8	A second novel model: the integro-differential generalised Von Foerster equation	99
8.1	<i>Abstract</i>	100
8.2	Introduction	100

8.3	Materials and Methods	101
8.3.1	The population density model	101
8.3.2	Interaction between species and environment	105
8.3.3	Numerical solutions and model validation	106
8.3.4	The European grapevine moth <i>Lobesia botrana</i>	107
8.3.5	Field data acquisition and analysis	108
8.4	Results and discussion	108
8.5	Conclusions	110
9	Conclusions	113

Chapter 1

Introduction

“L’ho già detto che se rinasco faccio l’idraulico?”

Mathematics is not only a powerful tool in several scientific fields but she is, also, the language of Mother Nature. So, if we wish to query the Nature, we need to speak as her. Physics is historically the first discipline who have understood the importance to schematize and translate natural phenomena in mathematical language.

Nowadays we are living in a century where Biology is behaving like a queen. Scientists, in fact, are fascinated by several aspects belonging to this discipline, and they are taking up challenges which need a synergy among different scientific fields, to reach the goals. Thereby, Physicists, Mathematicians and Computer Scientists are working with Biologists and Ecologists to understand how to schematize the mechanisms of living organisms, and which are the basic laws driving their processes. According to Ageno [1], the main difference between the Physics and Biology is the lack of a theory which clearly describes, with mathematics, a series of basic laws in Biology. Accordingly, anyone who wants to contribute to this field is faced with a choice: *i)* to accept the compromise of simplify as much as possible the complex systems studied, or *ii)* to explore each aspect of a particular phenomenon seeking basic laws.

Among the modelling approaches in Biology, one in particular is arousing the interest of the insiders also for its potential applications: the description of ectotherms’ populations. The class of ectotherms, more specifically, involves several organisms such as plants, bacteria, fungi and the main subject of this thesis: insects. A deep knowledge of insects’ physiology, and its clever schematization can be helpful to formulate reliable mathematical models. Even if models can be helpful for direct applications, widely discussed in this thesis, the master aim for a model scientist have to be *“the most faithful representation of the biological processes”*. Aware of this precondition, surely applications are the impact measuring instrument for this kind of research. The power of reliable mathematical models, in fact, is their application as Decision Support Systems (DSS), a helpful tool to evaluate different types of risks affecting ecosystems and agro-ecosystems.

Agriculture, probably, is the biggest beneficiary of this field of research, since every year

several economic losses are due to pests and diseases. Besides the cultivated fields, there is the need, also, to preserve as much as possible the biodiversity and human and environmental health. Accordingly, mathematical models are helpful not only for agricultural field, but also for Forest, and Urban ecosystems, to cite some examples. Hence, a mathematical representation of a biological phenomenon can be used as a forecasting tool: in fact, providing a scenario of the future risks, entomologists, farmers, technicians, ecologists and forest scientists can formulate the most beneficial control strategy. Thereby, in simple words, a model which correctly catches the basic laws for ectotherms' development, is, then, a reliable forecasting tool to be used as decision support system. In fact, the monitoring activities are only important to understand the field situation "from the past to the present" but "without saying anything about the future".

This thesis summarize a three-year work on the development of mathematical models to describe insect pest populations considering their development through the life stages and over time. The work can be considered as a continuum of the Author Master's thesis [2]. In fact, the first issue was to explore the existing models applied in entomology to describe insect's populations, understanding their strengths and weaknesses.

In this regard, the first model analysed was the "Distributed Delay Model", in its original formulation. However, a deeper study suggested that there was a need to reformulate this model, in order to make it more suitable to represent field data. The new formulation of the Distributed Delay Model showed a better response in comparison with its previous version, but it still presents difficulties due to field applications and validations. In this regard a second idea, flanked to the development of a mathematical model to represent ectotherms, was born: the development of a dedicated software to work out the models and to analyse results. Applications and validations, in fact, requires long calculations and the use of numerical solutions, since the environmental parameters are time-varying and measured directly with specific tools.

The switch towards the new mathematical theory developed as main result of the present work, happened considering the analogies between the Distributed Delay Model and another famous and important equation: the "Von Foerster's equation". On one hand the Distributed Delay Model considers some aspects of fundamental importance in describing ectotherms, despite its application is complex, but on the other hand the Von Foerster's equation is more compact and easier to apply, despite lacks of some fundamental information.

This precondition, indicates the aim of the study: the development of a model (or more models) reliable in describing ectotherm populations. In particular an ideal model *i)* must consider the development of the species towards its life stages, *ii)* must consider the relationship between species and external environment, *iii)* must be spatialized, *iv)* must consider the interaction between insects and host plants and between insects and natural enemies, *v)* must be easily implemented in a calculation software, *vi)* must be simple to apply and validate.

To pursue this aim, some valid assumptions were retrieved from the Distributed Delay Model, and other ones from the Von Foerster's equation. This has led to two new physiologically-based models which response to the main requests for applications to ectotherms.

But it is not enough. In fact, a model application and validation provides for a dedicated software on one hand, and of a statistical analysis on the other hand. All the models were

enclosed in a specialised software developed combining C++ programming language and the ROOT's software libraries. More specifically, the aim of this program is to work out the equations with numerical schemes. Since numerical schemes are the only alternative to obtain solutions, it makes difficult the comparison with field data. Moreover, the equations were developed to host a series of sub-models which describe the interaction between the species and the environment. Since there are several proposals in literature for these sub-models, it is often difficult to choose the best representing for a specific case study. Hence, part of the mathematical theory concerned the development of a specific protocol of application for the models, providing, also, a series of indicators to evaluate their reliability.

The main logic line of the work developed during these three years was organised in compartments for an easier development. In other words, each compartment, corresponding to a specific aspect of the project, was treated as independent from the others, and connected only in a second moment. Hence, the aim of this work was pursued with a series of steps which connected "something from the past with something from the present, going towards the future". Each part of the theory was, therefore, developed, tested with field data and published on IF peer reviewed journals. The strong contribution to the publications are highlighted by the authorship and correspondence of the Author in each article. Therefore, since the work was organised compartmentally, the choice was to present it as a collection of publications from the Author "as they are", and organized according to a logic line ranging from the theoretical formulation of the models to the validation in several case studies of agricultural interest.

Chapter 2

A novel mathematical formulation for the Distributed Delay Model

*“Nun potemo annà su Marte,
se prima nunn’amo capito
come funziona la Luna!”*

- Our office neighbour

This first article [3] represents the connection among “*the past, the present and the future*”. The Distributed Delay Model was widely applied to describe ectotherms populations: more specifically it was one of the first models applied to insect pests. The mathematical theory showed in this chapter was developed into the Author’s Masters thesis [2], and highlights the difference between *flows* and *counts* of individuals in field applications. In fact, besides the mathematical details, this first article analyse the different response of the model with the support of field data.

Article details:	
Title:	Reformulation of the Distributed Delay Model to describe insect pest populations using count variables.
Authors:	Luca Rossini, Mario Contarini, Maurizio Severini and Stefano Speranza
Corresponding author:	Luca Rossini
Journal:	Ecological Modelling

Reformulation of the Distributed Delay Model to describe insect pest populations using count variables.

2.1 Abstract

Among the models used to describe insect pest populations, the Distributed Delay Model has been applied in several case studies in recent years. Its success is due mainly to its simplicity, and its versatility to be easily included in software to calculate numerical solutions. In its original formulation, the Distributed Delay Model provides, as a solution, the distribution of the insects' maturation flow; then, this is compared with monitoring in field applications. A different form of the model can be obtained, with the same assumptions, to describe the distribution of the number of individuals which are in a specific life stage at time t . The first aim of this work was to show the mathematical details in order to obtain the second form of the Distributed Delay Model, and to calculate its analytical solutions. The second aim was to analyse the model's behaviour in describing insect pest's population in varying environmental conditions, specifically in terms of temperature. To pursue this second aim, two case studies of noteworthy relevance in agriculture were considered: the pepper weevil, *Anthonomus eugenii* and the European grapevine moth, *Lobesia botrana*. For each case study, field populations were simulated with both the Distributed Delay Model versions, and the results were compared to determine the most appropriate model for application in the case of insect pest populations. Both the case studies highlighted that the novel formulation presented in this work significantly improves simulation, providing a more reliable representation of field data.

2.2 Introduction

Modelling insect pest populations has been an interest of agricultural researchers for a long time [4], [5]. The primary reason to delve into and improve this argument is for the more effective use of these models as Decision Support Systems (DSS) in crop protection, besides their support in the comprehension of the biological mechanisms [6]. Predictions have a key role to play in Integrated Pest Management (IPM), because they allow one to forecast pest outbreaks [7], [8]. Accordingly, a projection of the insect pest population dynamics provides *i*) a scenario of the risk of economic losses [9] and *ii*) time to formulate the most efficient control strategy [10].

In agriculture, several population dynamics models have been introduced in pursuit of this aim [11]–[13]. To date, the most reliable models to describe insect populations are physiologically-based models [4]. These models have the capacity to describe ectotherm populations considering the relationship between species growth and external environment [14].

A pioneering physiologically-based model is the Distributed Delay Model (DDM), introduced for the first time by Manetsch [15] (1976) and revised the subsequent year by VanSickle (1977) [16]. The DDM, in fact, provided the possibility of linking external sub-models (usually defined as phenological models) describing the relationship between species and environment [17] for the first time. Hence, with the DDM, phenological and population dynamics models were first

considered as a unique block [18]. Although this advantage, the DDM does not have the capacity to describe the spatial patterns of the outbreaks as well as, for instance, the more modern models based on network analysis citeStrona2018, Strona2020, but in some simple cases it still represents a valuable candidate to describe ectotherms life cycle. Typical applications, in fact, concern fields with homogeneous microclimatic conditions or controlled environments such as greenhouses. Under the aforementioned hypotheses, the whole environment may be represented with a single point where ectotherm populations develop, and this is the implicit assumption of all the applications of DDM available in literature. In addition, the recent implementations in GIS systems combined with remote sensing is overcoming the issue of the spatialization in areas where the aforementioned assumptions are not respected, as showed by Gutierrez et al. [19], [20].

The innovative idea behind the DDM was the possibility of representing an insect life cycle (or part of it, if there is a particular stage of interest) as a series of non-observable chained stages [21], usually considered as the “degree of delay” or “degree of stochasticity” of the model [22]. As a solution, the DDM provides a distribution of the insects’ maturation flows [21], [23] representing, in other words, the flow of individuals that complete their life cycle. However, the hypotheses behind its formulation [15], [16] allow to obtain a second mathematical form which, at difference of the classical DDM, provides as solution the count of individuals belonging to the last stage of their life cycle. This second mathematical form was never tested, but it is more similar, from a conceptual point of view, to the modern physiologically-based models considering systems of ordinary differential equations [24], [25] or partial differential equations [4], [14], [26]–[31]. These models, in fact, describe the temporal variation of population’s individuals, and not their maturation flows.

A wide range of works highlight the reliability of the DDM in describing ectotherms such as plants and insects. For example, cotton (*Gossypium hirsutum* L.) cultivations were modelled by Gutierrez and Pizzamiglio [23]. Then Gutierrez et al. [32] successfully extended the study to bean, tomatoes and cassava cultivations. Some years later, Severini et al. [21] used the DDM to describe the development of apricot flowers. A slightly different application of the DDM in plant modelling was provided by Wermelinger et al. [33]; in this case the DDM was used to describe the allocation of carbon and nitrogen in grapevines, dividing the single plant into sub-populations of leaves, shoots, roots and fruits.

Alongside plant modelling, other successful applications of the DDM concern insect pest populations. The case study of the European grapevine moth *Lobesia botrana* (Denis & Schiffmüller) was considered by Severini et al. [34] and by Gutierrez et al. [19] with successful results. In particular, the first work describes *L. botrana* populations, validating the model with field data, while in the second work the DDM was used to predict European grapevine moth’s populations in a climate change scenario. Other relevant applications are discussed by Schaub et al. [35] with regard to the European corn borer *Ostrinia nubilalis* (Hübner), by Limonta et al. [22] with regard to the rice moth *Corcyra cephalonica* Stainton, and by Baumgärtner and Severini [36] with regard to the case of *Phyllonorycter blancardella* (Fabricius). Moreover, Allen and Byrd [37] extended the applicability of the DDM to forensic entomology.

The applications listed above compare experimental and simulated maturation flows of insect pest populations. Besides applications in agricultural and environmental sciences, the DDM was employed also for flow forecasting by Spolia et al. [38] and included in a SIRnS-model to represent the spread of infectious diseases by Jensen et al. [39].

However, from an experimental point of view, it is difficult to directly measure the maturation flows or, in other words, the exact moment of life stage transitions. During experimentations, researchers usually count the number of individuals in each life stage by direct samplings or using monitoring traps. At this point, one may wonder if a novel mathematical form of DDM describing the count of individuals into the life stages, and demonstrated starting from the same assumptions of the classical one, is more suitable to represent ectotherm populations. This question is reasonable, since to date the DDM is still applied in simulating insect pest populations as some recent works show [19], [20], [40].

The aim of this work, hence, is to show how obtain the novel form of the DDM which describes the temporal variation of the number of individuals into each stage, instead of the temporal variation of the maturation flows. After the mathematical description, a comparison between the two versions of the DDM was conducted, investigating which of the two is the better and more effective form for entomological applications. In this regard, two applications were compared: the case study of the European grapevine moth *Lobesia botrana* and the case study of the pepper weevil *Anthonomus eugenii* Cano. Data about these insects' population were retrieved from two already published works [41], [42]; this data was selected to investigate the behaviour of the two DDM versions in cases of different species and different development environments. *A. eugenii* populations were monitored in an experimental greenhouse, while *L. botrana* was monitored in open field conditions.

2.3 Materials and Methods

2.3.1 Distributed Delay Model

Considering the scheme showed in Fig. 2.1, an ectotherm's life cycle can be represented with a series of h chained substages.

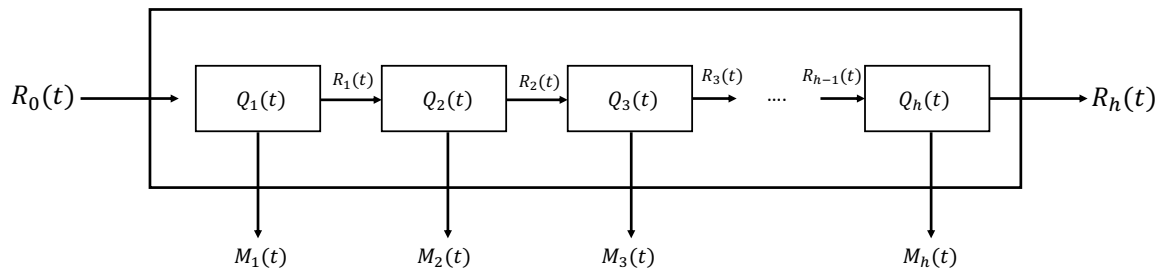


Fig. 2.1: Schematic representation of an ectotherm's life cycle according to the Distributed Delay Model.

Each i^{th} substage is characterized by an entering flow of individuals denoted with $R_{i-1}(t)$,

an exit flow of matured individuals $R_i(t)$ and by a flux of dead $M_i(t)$. An additional function, $Q_i(t)$, describes the temporal variation of the individuals into the stage.

As reported by Manetsch [15] and VanSickle [16], a mathematical description of each substage consists in three equations. The first one (2.1) is a balance equation which links the variation of the individuals into each substage with the entering and exit flows, while the second (2.2) and the third (2.3) ones define the flows $R_i(t)$ and $M_i(t)$. In formulas:

$$\frac{d}{dt}Q_i(t) = R_{i-1}(t) - R_i(t) - M_i(t) \quad (2.1)$$

$$R_i(t) = \frac{h}{D(t)}Q_i(t) \quad (2.2)$$

$$M_i(t) = AR(t) \cdot Q_i(t) \quad (2.3)$$

$D(t)$ and h in the equation (2.2) represent the expected emergence time of the whole population and the number of substages, respectively. $AR(t)$ in expression (2.3) is the mortality function according to the formulation of VanSickle [16].

Classical formulation $R_i(t)$

Manetsch [15] and VanSickle [16] obtained the generic DDM equation, written with $R_i(t)$ variables, substituting the expression (2.2) and (2.3) into the equation (2.1):

$$\frac{d}{dt}R_i(t) = \frac{h}{D(t)} \left[R_{i-1}(t) - R_i(t) \left(1 + AR(t) \frac{D(t)}{h} + \frac{d}{dt} \frac{D(t)}{h} \right) \right] \quad (2.4)$$

The scheme presented in Fig.2.1 assumes that it is possible to represent a single observable stage or the whole life cycle just varying the number of the equations, h , of the system. This is one of the major criticalities in applying the DDM, since h itself is one of the unknown of the problem. However, even if there is not a direct correspondence between the observable life stages and the chained boxes described in Fig. 2.1, there are two possible ways to apply the DDM. The first possibility is to consider a system of chained DDMs, where each system of equations represents a single observable stage [43]. The second possibility is to consider the whole life cycle as a single stage described by a single system of DDM equations. At first sight, the first proposal seems to be more realistic in representing the real situation, but at the same time provides for the estimation of h for each system involved. Moreover, it is more expensive from a computational point of view, besides a more complicated validation. The second possibility, instead, provokes loss of biological information, but at the same time it is easier to manage and apply; it is a good compromise when there is an interest in modelling only a particular life stage. This second representation has been chosen for the case studies considered in the present work. A common condition to both the aforementioned ways to apply the DDM is that although may not be a direct correspondence between observable and non-observable stages alongside the life cycle represented, surely the first and the last substage of the process in Fig. 2.1 can be directly associated with observable stages. For the purpose of this work, hence, the first substage

represents the cohort of eggs entering within the process, while the last substage are the adults. The hypotheses discussed above were maintained for both the classical DDM version (2.4) and the novel one that we are going to introduce in the following section.

Novel formulation $Q_i(t)$

The $Q_i(t)$ version of equation (2.4) can be obtained following an analogous logic scheme. With a direct substitution of the expressions (2.2) and (2.3), expression (2.1) becomes:

$$\frac{d}{dt}Q_i(t) = R_{i-1}(t) - \frac{h}{D(t)}Q_i(t) - AR(t) \cdot Q_i(t) \quad (2.5)$$

Now, let us consider a transformation of the term $R_{i-1}(t)$ in the equation (2.5). With reference to Fig. 2.1, it is possible to say that each substage with $i \geq 1$ can be described with their specific equations (2.1)-(2.3). This means that even for the substage $i - 1$ a triad of equations similar to the equations (2.1)-(2.3) will exist [2]. More specifically, the $R_{i-1}(t)$ term can be obtained considering the equivalent of the expression (2.2), referred to the substage $i - 1$:

$$R_{i-1}(t) = \frac{h}{D(t)}Q_{i-1}(t) \quad (2.6)$$

Substituting the expression (2.6) into the equation (2.5), the $Q_i(t)$ version of DDM is obtained:

$$\frac{d}{dt}Q_i(t) = \frac{h}{D(t)} \left[Q_{i-1}(t) - Q_i(t) \left(1 + \frac{D(t)}{h} AR(t) \right) \right] \quad (2.7)$$

The complete system, finally, is:

$$\begin{cases} \frac{d}{dt}Q_1(t) = \frac{h}{D(t)} \left[Q_0(t) - Q_1(t) \left(1 + \frac{D(t)}{h} AR(t) \right) \right] \\ \frac{d}{dt}Q_2(t) = \frac{h}{D(t)} \left[Q_1(t) - Q_2(t) \left(1 + \frac{D(t)}{h} AR(t) \right) \right] \\ \vdots \\ \frac{d}{dt}Q_h(t) = \frac{h}{D(t)} \left[Q_{h-1}(t) - Q_h(t) \left(1 + \frac{D(t)}{h} AR(t) \right) \right] \end{cases} \quad (2.8)$$

System (2.8) represents the main mathematical results of this work, written in its most general form. In the next sections, aspects related to analytical and numerical solutions will be treated, before a direct application to the cited case studies. In fact, the primary question of this paper now revolves around which DDM version is more appropriate for field applications, and whether there are particular cases where no differences between the two models can be identified.

2.3.2 Analytical solutions

Classical formulation $R_i(t)$

Concerning the $R_i(t)$ version, several works have faced this problem [21]. A particular case without mortality and with constant $D(t)$ term was considered by Manetsch [15], while VanSickle [16] extended the calculation considering a more general case with constant mortality. Both cases lead to an analytical solution found through the use of Laplace transform, assuming that $R_0(t)$, the input function in the system, is represented by a Dirac delta function, namely $R_0(t) = \delta(t)$. The analytical solution, assuming the hypotheses described, is the Erlang Probability Density Function (Erlang PDF):

$$R_h(t) = \left(\frac{h}{D}\right)^h \cdot \frac{t^{h-1}}{(h-1)!} \cdot \exp\left[-\left(AR + \frac{h}{D}\right)t\right] \quad (2.9)$$

This special case means that in the general equation of the $R_i(t)$ version, the equation (2.4), the term $\frac{d D(t)}{dt h}$ is equal to zero, since the function $D(t)$ is assumed to be a constant.

Novel formulation $Q_i(t)$

It is the task of this section to show that, with the same assumptions, it is possible to find an analytical solution for the system (2.8), namely the $Q_i(t)$ version of the DDM. Again, through the use of Laplace transform [2] it is possible to convert the system (2.8) into an algebraic system:

$$\begin{cases} \mathcal{L}[Q_1(t)] = \frac{h}{D} \cdot \frac{1}{s + \frac{h}{D} + AR} \\ \mathcal{L}[Q_2(t)] = \frac{h}{D} \cdot \frac{1}{s + \frac{h}{D} + AR} \cdot \mathcal{L}[Q_1(t)] \\ \vdots \\ \mathcal{L}[Q_h(t)] = \frac{h}{D} \cdot \frac{1}{s + \frac{h}{D} + AR} \cdot \mathcal{L}[Q_{h-1}(t)] \end{cases} \quad (2.10)$$

System (2.10) is solvable with a series of chained substitutions which lead to the following expression:

$$\mathcal{L}[Q_h(t)] = \left(\frac{h}{D}\right)^h \left(\frac{1}{s + \frac{h}{D} + AR}\right)^h \quad (2.11)$$

Now, applying the inverse of the Laplace transform to equation (2.11) it can be demonstrated that the solution of the $Q_i(t)$ version of the DDM is again the Erlang PDF, if and only if the input function $Q_0(t) = \delta(t)$ and $D(t)$ and $AR(t)$ are assumed to be constant:

$$Q_h(t) = \left(\frac{h}{D}\right)^h \cdot \frac{t^{h-1}}{(h-1)!} \cdot \exp\left[-\left(AR + \frac{h}{D}\right)t\right] \quad (2.12)$$

Moreover, the expression (2.11) can also be obtained considering that in the frequency domain the transfer function of the system is:

$$\frac{Q_h(s)}{Q_0(s)} = \left(\frac{h}{D}\right)^h \left(\frac{1}{s + \frac{h}{D} + AR}\right)^h \quad (2.13)$$

where $\mathcal{L}[Q_h(t)] = Q_h(s)$ and $\mathcal{L}[\delta(t)] = Q_0(s)$. The main result to be taken from this section concerns the similarity between the two versions of the DDM. In particular conditions, represented by the assumptions formulated at the beginning of the calculation, they provide the same solution. Hence, it is possible to assess that when the DDM is used in conditions of constant temperature, namely with constant $D(t)$ [21] and with a constant mortality term, there is no clear preference between the two versions of the model. However, this situation is not verified in the most general case, where a numerical solution is needed to solve both the DDM versions. In this case, a comparison of the model response is required, directly applying both DDM versions to the case studies.

2.3.3 Development rate functions

Biological parameters which clearly identify the species represented can be introduced in the DDM models through development rate functions. In fact, the mean development time $D(t)$ included in the system (2.8) and in its related $R_i(t)$ version is connected to environmental parameters, temperature in particular, through the following mathematical relation [44]

$$G[T_i] = \frac{1}{D[T_i]} \quad (2.14)$$

The expression (2.14) connects the mean development time $D[T_i]$, related to a particular temperature, to its corresponding development rate $G[T_i]$. Moreover, it is implicitly intended that there is time dependence in $D[T_i]$ and $G[T_i]$ in varying temperature conditions. Accordingly, writing the mean development time as $D(t)$ it usually means $D[T(t)]$. Before discussing the development rate expressions involved in the present work, it is helpful to point out the meaning of constant temperature from the point of view of the DDM. As detailed in the previous section, one of the assumptions needed to calculate analytical solutions for both the DDM versions requires $D(t)$ as time-invariant. Expression (2.14) directly connects the mean development time with the daily average temperature, a physical dimension subject to changes in the natural environment. Accordingly, it is possible to consider $D(t)$ as time-invariant only in conditions of controlled temperature, as, for instance, in a climatic chamber. Hence, the application of the DDM in varying temperature environments implies a usage of the models in their general form, requiring numerical solutions. The literature concerning development rate functions has been enriched over the years, and several authors have proposed expressions to connect insects'

stage duration with the temperature; reviews on this argument were presented by Damos and Savopoulou-Soultani [44] and Ikemoto and Kiritani [45]. These functions are formulated both on empirical and physical bases, but a common feature between the various proposals is the need to estimate parameters with experimental data. To pursue the aim of this work, the Logan and Brière expressions will be considered. Mathematically speaking, the Logan expression [46] is represented by the following formula:

$$G[T(t)] = \psi \cdot \left[\exp(\rho T(t)) - \exp\left(\rho T_M - \frac{T_M - T(t)}{\Delta T}\right) \right] \quad (2.15)$$

where T_M is the lower temperature threshold for insects' growth, ΔT is the temperature range between the maximum of the function and T_M , ψ and ρ are empirical parameters. The Brière development rate function [47], instead, is mathematically described as follows:

$$G[T(t)] = a \cdot T(t) \cdot (T(t) - T_L) (T_M - T(t))^{\frac{1}{m}} \quad (2.16)$$

with T_L and T_M being the lower and upper temperature thresholds, and a and m empirical parameters. The introduction of the development rate function completes the meaning of the DDM. In fact, in the light of the methodology presented in this section, the general version of the system (2.8) (and its corresponding $R_i(t)$ version) can be rewritten as follows:

$$\begin{cases} \frac{d}{dt}Q_1(t) = h \cdot G[T(t)] \cdot [Q_0(t) - Q_1(t) (1 + h \cdot G[T(t)] \cdot AR(t))] \\ \frac{d}{dt}Q_2(t) = h \cdot G[T(t)] \cdot [Q_1(t) - Q_2(t) (1 + h \cdot G[T(t)] \cdot AR(t))] \\ \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ \frac{d}{dt}Q_h(t) = h \cdot G[T(t)] \cdot [Q_{h-1}(t) - Q_h(t) (1 + h \cdot G[T(t)] \cdot AR(t))] \end{cases} \quad (2.17)$$

$G[T(t)]$ in system (2.17) can be substituted with the expression (2.15) or (2.16) on the basis of the specific case study.

Moreover, the experimental data used to estimate the development rate function parameters, defined as “life tables” in the entomological jargon [48], contributes to choose the way of application of the DDM discussed in Section 2.3.1. In fact, if data about the average development time at different constant temperatures are provided for all the observable stages, then it is possible to estimate the parameter of development rate function for each observable stage. In this case, a series of chained DDMs is applicable. On the other hand, if data allow the estimation of a single development rate function for the whole life cycle (egg-adult), the DDM application should be oriented towards a single DDM. The species considered in this study lie in this latter case.

Analysing the system (2.17) in greater depth, it can be noted that there is no possibility to solve it analytically in the most general case. In fact, despite expressions (2.15) and (2.16) can be manageable from a mathematical point of view, they contain the unknown function $T(t)$, which, in line of principle, describes the variation of the daily average temperature. This physical dimension can be measured directly with meteorological stations or data loggers and

predicted for a few days after; however, there is not an analytical function which describes its trend throughout the year [42], [49]. For this purpose, a numerical solution is required for the system (2.17) and for its analogous classical version $R_i(t)$.

2.3.4 Numerical solutions of both the DDM versions

The issues described in the previous section lead to the use of numerical algorithms to find solutions of both DDM versions, before their application in case studies. The simplest way to implement the systems in a calculation software is the application of the Euler's finite difference scheme to each equation of (2.17). Concerning the classical version $R_i(t)$, a numerical scheme of this type was provided by Manetsch, VanSickle and Alilla et al. [15], [16], [50]. Therefore, in order to not distract the attention of the reader, only a discrete version of system (2.17) will be discussed. Let us focus on the generic equation of the system (2.17):

$$\frac{d}{dt}Q_i(t) = h \cdot G[T(t)] \cdot [Q_{i-1}(t) - Q_i(t) (1 + h \cdot G[T(t)] \cdot AR(t))] \quad (2.18)$$

The time variable t can be approximated with a sequence of values t_n with the condition $t_n < t_{n+1}$ [31]. Hence, defining the discrete time step $j = t_{n+1} - t_n$ it is possible to apply the finite difference scheme to the equation (2.18). To distinguish the continuous variables from their discrete versions, a change of notation will be adopted: Q_i^j , $G[T^j]$, AR^j . In the light of the assumptions discussed, each equation of the system (2.17) can be transformed as follows:

$$Q_i^{j+1} = j h G[T^j] \cdot Q_{i-1}^j + Q_i^j \cdot (1 + j h G[T^j] - j AR^j) \quad (2.19)$$

This numerical scheme and its respective R_i^j version were implemented in a specialised software, detailed in the works of Rossini et al. [42], [49]. This was used to carry out all the operations required to pursue the aims of this paper. The last issue before the application of the numerical scheme (2.19), and of its corresponding classical version, concerns the number of equations of the system, namely h . As stated in Section 2.3.1, h represents one of the unknown of the problem, needing field data for its estimation. The DDM applications listed in the introduction of the present study, usually assign a defined value to h , but the reasons behind are often difficult to understand. Hence, to overcome the problem, a methodology was proposed by Rossini et al. [49], [51]. Their proposal was to solve a high number of equations (2.19) considering the h value such that the convergence of the numerical scheme is ensured. Then, since the first and the last stages of the scheme in Fig. 2.1 are overlapped to the first and the last observable life stage, it is possible to compare each numerical solution with field data, in order to find which is the last equation of the system. This comparison can be evaluated with the same χ^2 -function (2.20).

2.3.5 Comparison between the DDM versions

After the model description and discretization, a methodology is needed to compare the results with field data. In particular, for each case study considered in the present work, it is necessary

to determine which DDM version is more appropriate for accurately representing the field data. A methodology to compare simulation outputs and field data was discussed by Rossini et al. [30], [41] on the basis of considerations introduced by Ikemoto and Kiritani [45] and Mirhosseini et al. [52]. The reliability of a simulation can be evaluated considering:

- The χ^2 function, with the meaning of indicator of the distance between n field data, O_i , and simulations E_i :

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (2.20)$$

- The coefficient of determination R^2 , which provides an estimation of the reliability of the simulation in representing field data:

$$R^2 = 1 - \frac{\sum_{i=1}^n (O_i - E_i)^2}{\sum_{i=1}^n (O_i - E_i)^2} \quad (2.21)$$

Coherently with the expression (2.20), O_i in this case indicates the observed data and E_i simulations.

According to this methodology, defined as “*a posteriori analysis*” [41], the DDM version which provides the lower χ^2 value and an R^2 nearest to 1 is the most reliable to describe field data.

2.4 Case studies: *Anthonomus eugenii* Cano and *Lobesia botrana* (Denis & Schiffermüller)

2.4.1 Greenhouse scenario: *A. eugenii*

The first comparison between the two DDM versions concerns the semi-field case of the pepper weevil, *A. eugenii*. In fact, a semi-controlled environment is fundamental to understand the response of the models because of the ability to manage factors affecting population development, such as natural enemies or other source of stresses.

During the 2014 growing season, a pepper weevil population was sampled in a greenhouse located in Fondi, (Lazio region, Italy). These data have been published in the work of Rossini et al. [41] to validate a different physiologically based model. Despite the data acquisition and the experimental scheme being well detailed in the cited literature, it is opportune to mark, in the present article, some additionally fundamental information. The *A. eugenii* population was assessed using chromothropic sticky traps, deployed along the diagonal of the greenhouse; these were checked on a weekly basis. An uncertainty of ± 3.5 days, half of the time between two inspections, was assumed for the experimental population. Moreover, during the monitoring period, the daily average temperature was recorded at a meteorological station. These

temperature data were then included in the calculation to estimate, through the development rate function, the corresponding value to insert in the discretized version of the system (2.17) and into its corresponding $R_i(t)$ version. Data about the biology and the different response to the external temperature were, instead, provided by Toapanta et al. [53] and were fundamental to estimate the development rate function parameters. Since the work of Rossini et al. [41] shows a better response of the Logan function (2.15) in representing life tables, this choice will be maintained for the application of the DDM.

In the first part of the experimentation the greenhouse was left open to promote the natural infestation by pepper weevil adults. Once ascertained the presence of the insect, the greenhouse was isolated from the external environment with an anti-aphid net. The day of the greenhouse isolation was considered as day zero for simulations (20th of October), assuming that the initial cohort of eggs was laid in that moment.

2.4.2 Open field scenario: *L. botrana*

The second case study selected for this work concerns another insect pest of significant importance within agriculture, specifically in grape production: the European grapevine moth *L. botrana*. Also in this case the choice was made to use data published in the work of Rossini et al. [42] related to a three-year survey (2014, 2015 and 2016 seasons) conducted in an experimental vineyard located in Castelli Romani area (Lazio region, Italy). The investigation provided the three yearly generations which are usually reported in the area, and for the three growing seasons surveyed. Adults were caught with pheromone delta traps (BIOGARD S.r.l., 20834 Nova Milanese (MB), Italy) equipped with Selure-Lb dispenser (Shin-Etsu Chemical Co., Ltd., Tokyo) and checked weekly. An uncertainty of ± 3.5 days, half of the time between two inspections, was assumed for the experimental population. During the survey, a meteorological station (AWS310 model, Vaisala company, Helsinki, Finland,) recorded 48 temperature values per day, which were used to estimate the corresponding daily rate to insert into the discretized version of the model (2.8) and into its corresponding $R_i(t)$ version. As regards the development rate functions, the protocol followed by Rossini et al. [42] was maintained in the present study. More specifically, the Brière expression (2.16) was used in temperature ranges $(6.7 - 13)^\circ C$ and $(25 - 30)^\circ C$ while Logan (2.15) in the ranges $(13 - 25)^\circ C$ and $(30 - 34)^\circ C$. Since the damages usually reported are not relevant for this investigation, the first generation of the year was not simulated [54]. However, the first generation is fundamental in driving the simulation of the second and third generations and, accordingly, can be used as an initial condition for both the DDM versions. More specifically, the day in which the peak of the first generation was ascertained by monitoring was considered as day zero for simulations, hypothesising that a cohort of eggs was laid in that day. In numbers, 22nd May, 24th April and 19th June were selected for 2014, 2015 and 2016 growing seasons, respectively.

2.5 Results and discussion

2.5.1 *Anthonomus eugenii*

The Logan function parameters were retrieved directly from the work of Rossini et al. [41] and are listed in Table 2.1. These parameters will be common to both the DDM versions, as well as the daily average temperature recorded in the experimental greenhouse. The other common value inserted as input in both the model versions is $h = 10$, following the procedure described in Section 2.3.4.

Table 2.1: *Anthonomus eugenii* case study. Logan development rate function parameters retrieved from the work of Rossini et al. [41]. Parameters were estimated through life tables from Toapanta et al. [53]. Additional information includes the χ^2 -value, R^2 -value and the number of degrees of freedom (NDF).

Rate function	Parameters ($\pm SE$)	χ^2 -value	R^2 -value	NDF (n)
Logan	$\psi = 0.003 \pm 0.001$	0.0017	0.99	3
	$\rho = 0.11 \pm 0.01$			
	$T_M = 35.0 \pm 0.6$			
	$\Delta T = 3.5 \pm 0.7$			

For this specific case study, simulations showed different responses describing the field population. Overlapping the numerical solution from both the models with field data, it is possible to assess the different trends of the classical, $R_i(t)$, and the novel, $Q_i(t)$, DDM versions in varying temperature conditions.

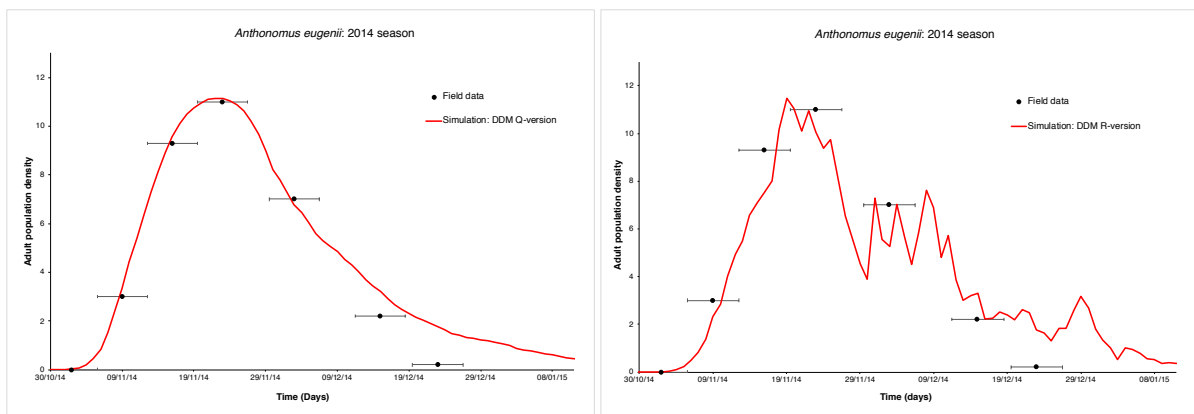


Fig. 2.2: *Anthonomus eugenii* case study. Field population and simulation using the $Q_i(t)$ version of DDM (left plot) and using the $R_i(t)$ version of DDM (right plot).

The population peaks were well described in both the models, even if the $Q_i(t)$ version had more reliable results. In fact, monitoring data indicated an experimental population peak

occurring on 23rd November, while the $Q_i(t)$ and $R_i(t)$ DDM versions forecasted the same on 22nd and 19th November, respectively. A graphical comparison of the results is provided in Fig. 2.2. A point to highlight is the higher variability of the classical DDM version $R_i(t)$ (right plot on Fig. 2.2) which underlines the meaning at the basis of the model formulation. The varying temperature conditions, in fact, provoke abrupt changes in the forecasted flow of emerging adults, at first sight undetected by the field data available. A more reliable description, instead, was provided by the novel DDM version $Q_i(t)$ (left plot in Fig. 2.2), which, in general, seems to better represent the experimental population.

A numerical estimation of the reliability of both the models was provided analysing the χ^2 and R^2 values, listed in Table 2.2. With this methodology, in fact, it is possible to determine which DDM version is the most effective for the experimental scheme adopted to retrieve the field population.

Table 2.2: *Anthonomus eugenii* case study. χ^2 and R^2 values resulting from the comparison between simulations with both DDM versions and field data.

DDM version	χ^2 -value	R^2 -value
$Q_i(t)$ -version	1.76	0.97
$R_i(t)$ -version	3.03	0.92

Focusing on the values listed in Table 2.2, it is possible to assess that the $Q_i(t)$ -version confirms its reliability with a lower χ^2 value and with a R^2 nearest to 1 respect to the $R_i(t)$ -version. The open field condition will be discussed in the following section.

2.5.2 *Lobesia botrana*

As shown in the *A. eugenii* case study, the first step is the Logan and Brière development rate functions parameters estimation. The values (Table 2.3) were estimated with life tables from Moshtaghi Maleki et al. [55] and used for both the DDM versions.

Concerning the simulation, for each year of survey, the same development rate function parameters, the same daily average temperature and the same values h were inserted. More specifically, h was estimated, also for the present case study, with the procedure described in Section 2.3.4. Values used for the simulations are listed in Table 2.4.

Let us focus on the 2014 season simulations. The second generation reported an experimental peak on 17th June, well represented by both the DDM versions. Advancing by order, the novel version $Q_i(t)$ (left plot on Fig. 2.3) forecasted a peak on 14th June, while the classical version $R_i(t)$ (right plot on Fig. 2.3) indicated the same on 15th June. Regarding the third generation of 2014, monitoring recorded a population peak on 19th August. Simulations provided a maximum peak on 18th August in the case of $Q_i(t)$ version, while the $R_i(t)$ provided a forecast for 19th August. On the whole, the values resulting from both the DDM versions are close to the field data, showing their good response in representing a population developing in an open field scenario.

Table 2.3: *Lobesia botrana* case study. Logan and Brière development rate function parameters retrieved from the work of Rossini et al. [42]. Parameters were estimated through life tables from Moshtaghi Maleki et al. [55]. Additional information includes the χ^2 -value, R^2 -value and the number of degrees of freedom (NDF).

Rate function	Parameters ($\pm SE$)	χ^2 -value	R^2 -value	NDF (n)
Logan	$\psi = (7.0 \pm 0.5) \cdot 10^{-4}$	0.0019	0.92	2
	$\rho = 0.18 \pm 0.04$			
	$T_M = 34.1 \pm 0.8$			
	$\Delta T = 4 \pm 2$			
Brière	$a = (3.4 \pm 0.2) \cdot 10^{-5}$	0.0036	0.87	2
	$T_L = 6 \pm 1$			
	$T_M = 32.6 \pm 0.5$			
	$m = 5 \pm 1$			

Table 2.4: *Lobesia botrana* case study. h -values used for each generation and each year of survey. Values are common for both the DDM versions.

Year of survey	Generation	h -value
2014	2	13
	3	13
2015	2	20
	3	8
2016	2	15
	3	21

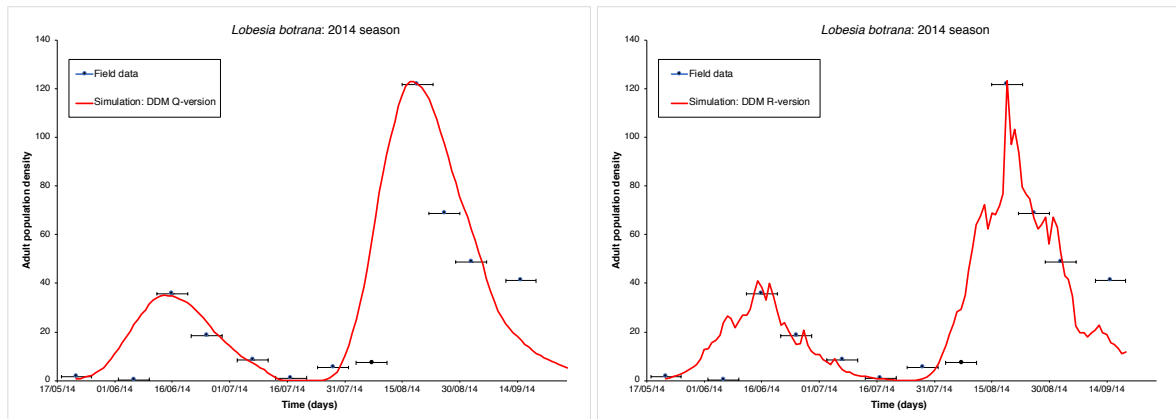


Fig. 2.3: 2014 season, *Lobesia botrana* case study. Field population and simulation using the $Q_i(t)$ version of DDM (left plot) and using the $R_i(t)$ version of DDM (right plot).

The results highlighted in the first season of survey were confirmed in the years after. In 2015, the second generation occurred on 16th June, as observed in the field survey. Simulations from both the models reported values close to the experimental population peak; more specifically the novel version $Q_i(t)$ (left plot on Fig. 2.4) indicated a population peak on 15th June, while the classical version $R_i(t)$ (right plot on Fig. 2.4) on 18th June. The third generation's population peak of 2015 season occurred on 19th August, while both the DDM versions provided a forecasted value on 11th August, as shown in Fig. 2.4. Thus, in this case both the models provided a reliable response in describing the field population.

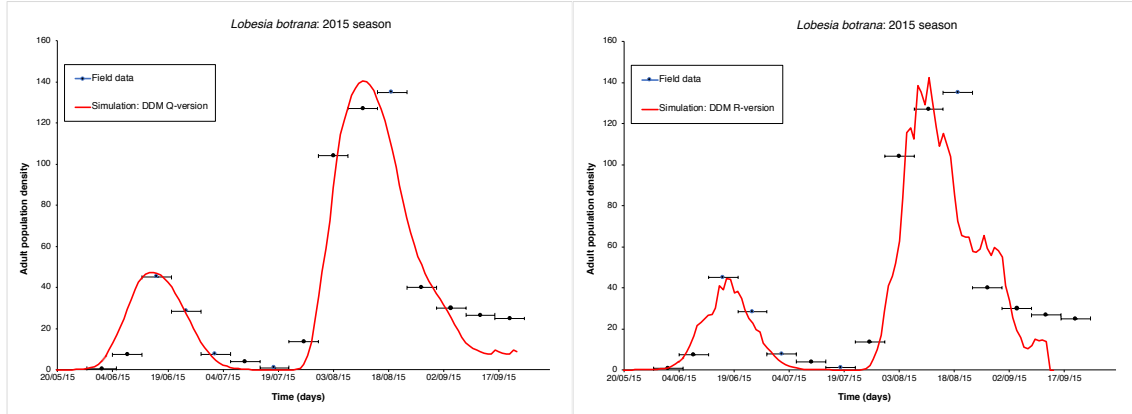


Fig. 2.4: 2015 season, *Lobesia botrana* case study. Field population and simulation using the $Q_i(t)$ version of DDM (left plot) and using the $R_i(t)$ version of DDM (right plot).

The 2016 season was the last year of the survey for the *L. botrana* case study. The second generation's peak occurred on 11th August, a value simulated on 10th August by the novel version $Q_i(t)$ (left plot on Fig. 2.5) and on 9th August by the classical version $R_i(t)$. Concerning the third generation, there was no possibility to determine exactly when the experimental population peak occurred, since the same highest value was reported on two days, namely, 14th and 21st September. Simulations with $Q_i(t)$ model indicated a population peak on 17th September (left plot on Fig. 2.5) while $R_i(t)$ version provided the same on 16th September.

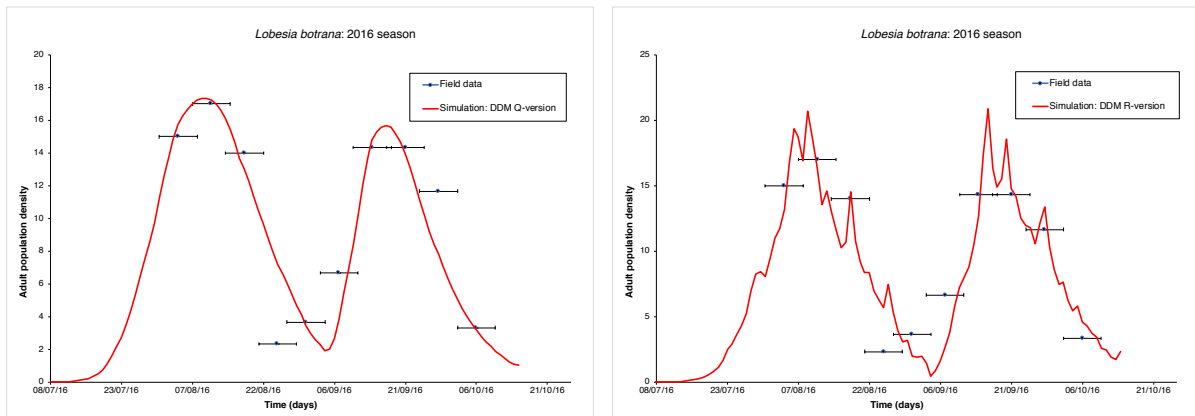


Fig. 2.5: 2016 season, *Lobesia botrana* case study. Field population and simulation using the $Q_i(t)$ version of DDM (left plot) and using the $R_i(t)$ version of DDM (right plot).

The values indicated by both DDM versions, on the whole, represented the experimental population well, but it is difficult to determine the most reliable model version without using numerical indicators. As already discussed in the case of *A. eugenii*, a cross check between χ^2 and R^2 values also answers to the main question of this article in the case study of *L. botrana*. Referring to Table 2.5, it is possible to notice that the lower χ^2 -values and R^2 -values nearest to 1 were reported by the $Q_i(t)$ DDM version. The only case where a higher R^2 -value was reported in the case of $R_i(t)$ instead of $Q_i(t)$ version occurred in the 2016 season (despite the χ^2 value being higher than the $Q_i(t)$ version). This unusual result, if compared to the others, may be justified by the lack of a clearly identified experimental population peak in the third generation (Fig. 2.5), probably located between the two higher values recorded by the monitoring. On the other hand, there is no significant difference between the two R^2 values, and this makes the χ^2 values associated to each DDM version reasonable.

Table 2.5: *Lobesia botrana* case study. χ^2 and R^2 values resulting from the comparison between simulations with both DDM versions and field data.

Year of survey	DDM version	χ^2 -value	R^2 -value
2014	$Q_i(t)$ -version	198.85	0.87
	$R_i(t)$ -version	211.43	0.77
2015	$Q_i(t)$ -version	241.99	0.91
	$R_i(t)$ -version	618.59	0.76
2016	$Q_i(t)$ -version	7.33	0.88
	$R_i(t)$ -version	10.44	0.89

2.6 Conclusions

Results showed that the novel Distributed Delay Model version $Q_i(t)$ introduced in the present work is more reliable in representing insect pest populations. The new implications of this paper contribute towards an improvement in the historical reliability of the DDM, testing a second formulation never used before.

Another important question contained in this work may concern the meaning of data acquired with traps. The two DDM versions provide a conceptual distinction between flux and count of individuals, although with the simple relation (2.2), however, to explore thoroughly the differences, ad hoc experiments at controlled conditions, such as in climatic chambers, should be conducted. In this way, it will be possible to control the life cycle of each individual, obtaining more detailed information about the newly emerged individuals to simulate with both the DDM versions.

One of the concerns highlighted by the results of the present work, moreover, is the time range between two field samplings: even if results are oriented towards a higher reliability of the novel DDM version $Q_i(t)$, this cannot be stated with certainty. In fact, the high fluctuations of the classical DDM are not always appreciated by field data, and one might wonder if fluctuations are wide as described by the simulations of the present work or not.

This question is reasonable, since Colinet et al. [56] assessed that the daily contribution received by ectotherms developing in fluctuating temperature conditions leads to an average temperature value, T_{DT} , which can be calculated at the end of the life cycle. Hence, it may be possible to assume that changes in development rates are not immediate, and depend on the past history of insects' living environment. The results of the present work do not allow to investigate this aspect, but given its importance, future studies on DDM should be addressed towards this direction. An investigation on ectotherms' "thermal inertia" may provide not only clarifications about the questions raised in the present study, but may also contribute to understand which is the optimal time scale to simulate ectotherms with physiologically-based models in general.

Although some questions are still open, the present study contributed also to fill a theoretical gap of knowledge, that is the connection between the classical DDM and the Von Foerster equation [57]:

$$\begin{cases} \frac{\partial}{\partial t} N(t, a) + \frac{\partial}{\partial a} N(t, a) = -M(t, a) \cdot N(t, a) \\ N(t, 0) = \int_0^\infty \beta(t, a') \cdot N(t, a) da' \\ N(0, a) = n^0(a) \end{cases} \quad (2.22)$$

Equation (2.22) describes the density of a population $N(t, a)$ developing over time, t , divided in continuous age classes, a , and with a mortality rate $M(t, a)$. $N(0, a)$ is the initial condition indicating the distribution of the individuals in the age classes at $t = 0$, while the boundary condition $N(t, 0)$ provides the new-born's profile on the basis of the population density $N(t, a)$ and the fertility rate function $\beta(t, a)$. Plant and Wilson [58] showed that replacing the age variable a with a sequence of values a_n , which respect the condition $a_n < a_{n+1}$, it is possible to obtain a similar form of the DDM. This discretization, in fact, leads to an approximation of the partial derivative in the a variable [14], [27], [51], providing a system of ordinary differential equations similar to the DDM:

$$\begin{cases} \frac{d}{dt} N_1(t) = \frac{1}{A} N_0(t) - N_1(t) \left[\frac{1}{A} + M_1(t) \right] \\ \frac{d}{dt} N_2(t) = \frac{1}{A} N_1(t) - N_2(t) \left[\frac{1}{A} + M_2(t) \right] \\ \vdots \quad \quad \quad \vdots \\ \frac{d}{dt} N_i(t) = \frac{1}{A} N_{i-1}(t) - N_i(t) \left[\frac{1}{A} + M_i(t) \right] \end{cases} \quad (2.23)$$

In the system (2.23) $A = a_n - a_{n-1}$ is the size of the discrete age classes which are, by definition, chained.

Now, let us assume that the discrete age classes have the same size, and that they divide the average duration of the life cycle $D(t)$ (which can be assumed time-dependent, if the population

develops under varying-temperature conditions) in h parts. It follows that $\frac{1}{A} = \frac{h}{D(t)}$, and considering, also, $M_1(t) = \dots = M_i(t) = AR(t)$ the system (2.23) is similar to the novel formulation of the DDM (2.8) introduced in the present study. It is possible to show, also, that the system (2.23) provides an Erlang PDF as analytical solution, if the assumptions stated in Section 2.3.1 are made [2].

Hence, the discretization of the Von Foerster equation does not lead to the general form of the classical DDM (2.4) in the $R_i(t)$ variables, but to the novel formulation introduced in the present study. Nevertheless, the mathematical details discussed in Section 2.3 clarify how to pass from $Q_i(t)$ variables to $R_i(t)$ and vice versa.

The situation is different if the assumptions discussed in Section 2.3.1 are made. In fact, in this case has been shown that the two DDM versions have the same behaviour and analytical solutions. This may be confusing if connected with the Von Foerster's equation, and needs a clarification. The discretization of the equation (2.22), which leads to the system (2.23), does not have assumptions restrictive enough as in Section 2.3.1, since the discretised mortality rate function $M_i(t)$ is still time dependent, as well as the mean development time $D(t)$. Hence, the discretization of the equation (2.22) with the method of the lines leads to the general form of the DDM, if no assumptions of constant mean development time and mortality rate are made. Accordingly, the mathematical aspects discussed in this work complete, also, the connection between the Von Foerster equation [57] and the Manetsch [15] - VanSickle [16] DDM formulation, providing, at the same time, the missing step in the demonstration.

Chapter 3

The milestones of a novel software for model applications in entomology

*“When you are solving a problem, don’t worry.
Now, after you have solved the problem,
then that’s the time to worry”*

- Richard P. Feynman

The development of new models and, above all, their application is not possible without support of computational tools. The development of a dedicated software, hence, was a parallel activity to the aim of this thesis: the mathematical formulation. In this second chapter is reported the article [49] where the foundations of an *ad hoc* software for entomological applications of mathematical models were discussed. In particular, this article discusses the milestones of *EntoSim*, the software developed as support of the mathematical development. Even if in this first presentation of *EntoSim* the model included is the Distributed Delay Model, during the years all the four models developed were included, together with statistical tools for models evaluation, visualisation and validation.

Article details:	
Title:	Use of ROOT to build a software optimized for parameter estimation and simulations with Distributed Delay Model
Authors:	Luca Rossini, Maurizio Severini, Mario Contarini and Stefano Speranza
Corresponding author:	Luca Rossini
Journal:	Ecological Informatics

Use of ROOT to build a software optimized for parameter estimation and simulations with Distributed Delay Model

3.1 Abstract

ROOT is a software package developed by a CERN project started in 1994 by René Brun for statistical analysis in high-energy physics. This software package can also be used in the field of protecting plants against pest insects. Indeed, for a long time there has been a significant use of the Distributed-Delay Model, but there is no specific software available to date that is useful for following the research from the first step to simulations and field validations. This work, through ROOT's libraries, builds a series of macros that consent to do non-linear fits with functions such as Erlang PDF, linear-rate, Logan, Brière, Sharpe and De Michele, thereby giving support to the parameters-estimate step in laboratory sessions and then numerically solving the Distributed-Delay Model equations. This study supplies results both graphically and numerically.

3.2 Introduction

In Agricultural, Forest and Environmental Sciences, and especially in the context of risk management, modelling has become an important tool. In particular, entomologists started using several models to forecast pest insects' life cycles for two primary reasons: safeguarding human and environmental health, and because of the restrictive laws regarding pesticide use. The difference between a monitoring system and a forecasting system results from the possibility of having an idea of the risk situation of an agricultural field or an urban and peri-urban area [59]. There is a need to develop a control strategy different from the conventional one. Despite monitoring and forecast systems seeming similar, in reality there is a conspicuous difference between them. In the first case, data collected with traps placed in fields provides empirical information on insect presence. Capinera defines the term “*monitoring*” as “*careful observation of pest abundance and damage*” [60]. The inconvenience is that monitoring provides no information about the future trends of the monitored population. On the other hand, a simulation could not report the exact field situation, but it is helpful to know the future development of the population. In other words, forecasting is “*the process of making predictions based on past and present data*”. These predictions can support a series of research projects that involve using natural enemies such as predators, parasitoids, entomopathogenic fungi and bacteria for the control of insect pests. Indeed, natural enemy complex plays a preeminent role in controlling invasive and autochthonous species in different environmental contexts. For example, the entomopathogenic fungus *Entomophaga maimaiga* in areas where it has been introduced is effective against the gipsy moth *Lymantria dispar* (Lepidoptera: Erebidæ) populations [61], encouraging further experimentation aimed at diffusing this pathogen in new areas across the world [62]. In the agronomic context, the biological control applications of the entomopathogenic bacteria *Bacillus thuringiensis* [63], [64] or the predator *Zelus obscuridorsis* (Hemiptera: Heteroptera:

Reduviidae) [65] in the tomato leaf miner, *Tuta absoluta* (Lepidoptera: Gelechiidae) might be mentioned. The inconvenience of this type of control is that it requires highly qualified technicians, quick response and accurate knowledge of the most susceptible life stage of the insect pest if one is to maximize the efficacy of the use of biological control. On the other hand, if it is impossible to apply a biological control strategy, to know the pests' population trends helps one perform chemical control to reduce the quantity of the pesticide in question. One of the widely used models is the Distributed Delay Model (DDM).

Beginning with its introduction by Manetsch [15], the DDM has garnished interest from entomologists and environmental scientists to describe the life cycle of poikilothermic organisms. One of the strengths concerns the possibility of linking environmental and species' parameters to the description of a population dynamics. Several applications have shown the DDM's adaptability to different pest insect species [18], [22], [59], [66]–[71]. A typical application of the DDM for a species with a poorly known biological cycle starts with a series of constant-temperature laboratory rearing. This is the first phase or “*laboratory session*”: entomologists build cohorts of insects (all laid on the same day) and put them in climatic cells. Researchers follow each cohort's individual day by day during its life cycle, collecting data of the duration of each stage. There were similar observations for other insect pests belonging to different orders, including the beetles (Coleoptera) *Rhynchophorus ferrugineus* (Dryophthoridae) and *Anthonomus eugenii* (Curculionidae) [53], [72], the butterflies (Lepidoptera) *Lobesia botrana* (Tortricidae), *Corcyra cephalonica* (Pyralidae), and *Tuta absoluta* (Gelechiidae) [55], [73], [74], the wasp (Hymenoptera) *Diadegma anurum* (Ichneumonidae) [75], and the true bug (Hemiptera) *Halyomorpha halys* (Pentatomidae) [76].

Life tables are important because of the description of the relationship between the temperature (the main environmental driving variable for poikilothermic organisms) and the average developmental time, $D[T]$, for each stage composing the life cycle. It is common in practical cases to convert the mean developmental time $D[T]$ in mean development rate $R[T]$, (specific for each temperature), by the expression [77], [78]:

$$R[T] = \frac{1}{D[T]} \quad (3.1)$$

The reason behind this choice is the trend reported in a plot rate-temperature. In fact, with this conversion, one can obtain an increasing-decreasing profile which highlights the thermal optimum for the species as well as its lower and upper thermal thresholds. The second phase of a simulation process concerns the simulation: all the estimated parameters are inserted in a model together with the measured average daily temperature. This obtains the forecasted population dynamics. Subsequently, a validation is made comparing the results of a simulation, with data collected via in-field monitoring. Despite wide application, one of the main problems reported by the insiders is the lack of a specific software that allows researchers to collect data during each step of the project. This work introduces a new application of the ROOT software as a base to build a helpful tool for all simulation processes with DDM. More specifically, ROOT is “*a modular scientific software framework. It provides all the functionalities needed to deal with*

big data processing, statistical analysis, visualization and storage. It is mainly written in C++ but integrated with other languages such as Python and R.” (<http://root.cern.ch>) [79]. The following sections describe the specifics of each operation and the corresponding ROOT’s macro developed to manage all the required calculations.

3.3 Materials and Methods

Manetsch first introduced DDM, though Vansickle [16] performed it for the first time one year afterward, at the same time introducing the mortality coefficient as attrition. More specifically, Vansickle’s improvements concern comparing the mortality to a friction force: in his work he considered that for each life stage there is a specific mortality coefficient (such as the attrition coefficient in physics) which is directly proportional to population density. Accordingly, the rate of dead is higher as the population density increases. In the case of insects, it is possible to indicate the entire life cycle with a series of chained boxes, each of which is described by an ordinary differential equation. If we call $Q_i(t)$ the number of individuals in the i -th stage, $D[T(t)]$ the average developmental time (depending on the temperature), and h the number of steps for each prediction. The mathematical expression of Manetsch’s model is the following:

$$\begin{cases} \frac{d}{dt}Q_1(t) = \frac{h}{D[T(t)]} [Q_0(t) - Q_1(t)] \\ \frac{d}{dt}Q_2(t) = \frac{h}{D[T(t)]} [Q_1(t) - Q_2(t)] \\ \vdots \\ \frac{d}{dt}Q_h(t) = \frac{h}{D[T(t)]} [Q_{h-1}(t) - Q_h(t)] \end{cases} \quad (3.2)$$

The system (3.2) can be solved analytically for constant temperatures rearing in the case of a Dirac delta function $\delta(t)$ as input function $Q_0(t)$. In the most general case, indicated in the system (3.2), temperature is variable over time, so to assert that it is constant leads one to consider $D[T]$ instead of $D[T(t)]$. This is a mathematical means of representing a cohort of individuals that join the process at the same time, as well as what is done to build life tables with constant temperature rearing. The analytical solution is represented by the Erlang Probability Density Function (Erlang PDF) [15], [16], [21]:

$$f(t) = N \cdot \left(\frac{h}{D[T]} \right)^h \cdot \frac{t^{h-1}}{(h-1)!} \cdot e^{-\frac{h}{D[T]}t} \quad (3.3)$$

Several works [21], [35], [50], [58] have shown that the development over time for a cohort is well described from the expression (3.3) in which N is a normalization constant considered as the number of initial individuals in a cohort. In fact, the temperature causes variations in development time of a poikilothermic organism, making it directly linked to the average development time $D[T]$. Expression (3.3) is helpful in the laboratory phase: since the emerging

time follows an Erlang PDF trend, it is possible to do a non-linear fit operation to estimate the parameters h , $D[T]$ and N . In particular, $D[T]$ is the most important as it represents the expected value of the emerging time; it can be converted into development rate through the expression (3.1). A series of pairs of points $(T_i, R[T_i])$ is obtainable by repeating rearing with different constant temperatures. This mathematical description inspired the development of the first two macros using the ROOT's libraries: *graph2Dpunti*. This was done to plot experimental points as well as *fiterlang* to work out the non-linear fit operations with the Erlang PDF.

The series of experimental pairs of points $(T_i, R[T_i])$ is useful for building the development rate function [44], [80]. This function represents a driving function for the development through stages, at the same time providing the corresponding rate for each temperature. For a simulation process, the role of the development rate function is crucial in that it links the population dynamics to the environmental parameters. In other words, it represents a driving function for the development through the stages and provides the corresponding rate (or time, if converted with the expression (3.1)) for each temperature measured in the open field. In fact, the open field is definable as the environment in which the mean daily temperature is variable, $T(t)$. The typical trend reported in life tables starts from a lower temperature threshold, T_L , under which the development rate is equal to zero. There is then an increasing range until it reaches a maximum peak that represents the optimal temperature for development. On the right of the maximum a decreasing trend is until a maximum threshold value T_M ; after that, the development rate is equal to zero. Figure 3.2 presents the case of *A. eugenii* as an example [53].

One of the main problems of the development rate function is the lack of points: in fact, the build of the life tables requires significant time and a conspicuous number of individuals to compose the cohort. Furthermore, one of the hardest parts is the rearing of insects at temperatures close to the upper and lower thresholds. The number of individuals for the starting cohort in this case have to be higher than the other cases (e.g. rearing near the optimum temperature) because of high mortality. Significant errors thereby affect data from the “laboratory phase” owing to the genetic variability of the population, as well as other systematic errors due to the instrumentation, to the insiders’ expertise, and to diet (artificial or natural).

To date there is no preferred mathematical expression to describe the relationship between rate and temperature, but over the course of time scientists have reported several functions that insiders have used widely on the basis of the available life tables. The most common are:

- the linear rate function [44], [78]:

$$R[T] = aT + b \quad (3.4)$$

with a and b the empirical parameters;

- the Brière rate function [47]:

$$R[T] = a \cdot T \cdot (T - T_L) (T_M - T)^{\frac{1}{m}} \quad (3.5)$$

with a and m empirical parameters, T_0 the lower development threshold, T_L the upper development threshold.

- the Logan rate function [46]:

$$R[T] = \psi \cdot \left[\exp(\rho T) - \exp\left(\rho T_M - \frac{T_M - T}{\Delta T}\right) \right] \quad (3.6)$$

with ψ and ρ empirical parameters, T_M the maximum threshold and ΔT the distance between the maximum peak of the function and T_M .

- the Sharpe and De Michele rate function [81]

$$R[T] = \frac{T \cdot \exp\left(A - \frac{B}{T}\right)}{1 + \exp\left(C - \frac{D}{T}\right) + \exp\left(E - \frac{F}{T}\right)} \quad (3.7)$$

in which A, B, C, D, E, F are parameters related to the enzyme's kinetics.

Other than the linear rate function (3.4), the other functions all have more than two parameters to estimate through the non-linear fit model method. The proposed macro named *fittasso*, makes this operation simpler. The idea is that the experimenter inserts his/her life tables and executes the non-linear fitting operations using the preferred development rate function between the listed above. To determine which is the best expression related to the experimental data, it is necessary to cross check between the R^2 and χ^2 -values, and their calculation is therefore an obliged request for the related macro. The second phase concerns the simulation, and using Manetsch's model estimates all parameters and other inputs discussed afterwards that forecast the insects' life cycle (3.2). The first problem encountered is related to the lack of an analytical function that exactly describes the daily average temperature trend. Simulations can usually be used while inserting theoretical or measured temperature data, but an analytical function would not describe both cases. Theoretical values generate on the basis of historical data or meteorological models as a series of pairs of points $[t, T(t)]$. In the context of applications in the field of entomology, measuring this data through the use of a meteorological station or a datalogger placed is required in a specific point of interest for the simulation if one is to predict the populations of the insect pests to be controlled. Furthermore, to perform the simulations, it is possible to insert forecasted temperatures for the following in according with a weather forecasting system. All this implies it is impossible to solve the system (3.2) in an analytical form (because of the varying daily temperature) - a numerical solution is needed. The developed simulator uses the finite difference method to obtain a discrete form of the DDM [34]. One of the inputs needed for the simulation, is the mean development time value for the specific daily average temperature measured. Performing this operation requires a macro with the task to input all the estimated parameters related to the development rate function, the daily temperature, as well as to provide the corresponding development rate as output. Moreover, to complete the operation, it is necessary to convert this result in development time, $D[T]$, using the expression

(3.1). A dedicated macro to work out this process has been developed: *functioncalculator*. The need to have a fast screening to determine the average development time on the basis of recent temperatures drives the decision to exclude this result from the simulator. In this way the simulation software has only one input that encloses all the results of the laboratory and acquisition data phase. At this point, the simulation process involves only the numerical solutions for the system (3.2). Typical research concerning the use of a modelling approach to forecast insects' life cycle ends with a validation process. It is necessary to compare the series of simulated curves provided by the simulator with field data, monitoring the results. Determining the individuation of which is the most faithful simulated curve to the experimental data is difficult, and manual realization takes a long time. To optimize time and resources, this study developed a conclusive macro. In the results section, we show the details of each developed macro.

3.4 Results

More macros that are organized by the experimental guidelines articulate the “*DSimulator*” (Delay Simulator). The logical succession represents every step followed by an entomologist while estimating parameters, simulations and validations with field data. One of the advantages of using ROOT's libraries is to develop every macro independently with respect to one another. Each can be launched and used independently through ROOT's command line. In addition, this study developed a main interface that collects all of these tools to permit multiple actions when all data are available.

Even if ROOT is not difficult to use, a little knowledge is required with regards to ROOT's basic commands, such as “*.x DSimulator.c*” to load and run macros. A peculiarity of this software is its speed in reaching results: every output and input file becomes available to all tools that need such information. Users can report their data in specific *.txt* files, then at the end of every process, they can find results in a clear way printed in *.txt* output files, in ROOT's shell, and in a “*canvas*” with charts. To better understand the capabilities of this structure, we show every single macro in detail:

- *graph2Dpunti*: this is the first and the simplest macro built, and it is very close to ROOT's basic functions. This is a solution to a typical need during experimental trials: creating charts with acquired points and corresponding error bars. The user can report data in a file named “*graph.txt*”, dividing them into four ordered columns which represent x-axis, y-axis, x-values errors and y-values errors. Launching this tool also creates a “*canvas*” containing the corresponding chart. Some ROOT's operations are available through the toolbar, such as linear, gauss, polynomial and exponential fitting. One can change the plotted points and adjust them by selecting different styles. There is finally a possibility to export the graph as a *.pdf* file.
- *fiterlang*: after a cohort of insects ended its lifecycle in climatic cells, it is possible to estimate the corresponding developmental time. The easiest thing to do is to calculate the average value and its corresponding standard deviation. A procedure like this is often

justified because big data is relatively unavailable, and the point's shape seems to be a normal distribution. Nevertheless, many publications [21], [27], [33], [82] have shown that the Erlang PDF-function (3.3) is more appropriate for describing the exit times of an insect population whose individuals all started their lifecycle in the same day. The approach considered introduces greater precision in experimental results but complicates the calculation operations. With ROOT it is possible to build a macro in which scientists can input a .txt file called “erl.txt” containing four columns ordered as described above for the “graph2Dpunti” tool: time-values, number of individuals, and their corresponding errors. It is complicated to give an estimation of the error for the number of individuals, so it is common to consider only time as a variable, which varies with respect to the sampling rate. After macro's launch, there is a possibility to choose between a no mortality Erlang function [15] and a mortality Erlang function [16], [83]. The main difference is that the first is like (3.3), with three parameters to estimate, whereas the second has a similar shape but contains another parameter, AR , which represents the mortality coefficient meant as an attrition:

$$f(t) = N \cdot \left(\frac{h}{D[T]} \right)^h \cdot \frac{t^{h-1}}{(h-1)!} \cdot e^{-\left(AR + \frac{h}{D[T]} \right) t} \quad (3.8)$$

Given “erl.txt” in input, “fiterlang” tool returns a .txt file named “Erl_fitris.txt” containing the initial values chosen to start the fitting process, the best fit parameters value with their corresponding errors, the χ^2 -value, the adjusted R^2 -value, and the number of degrees of freedom. The correlation matrix is shown only on ROOT's console. Another output is a canvas containing a double chart, with the possibility to export it: only the experimental points are plotted on the left, and the right represents the fit results (points, best fit function and a resume of the values).

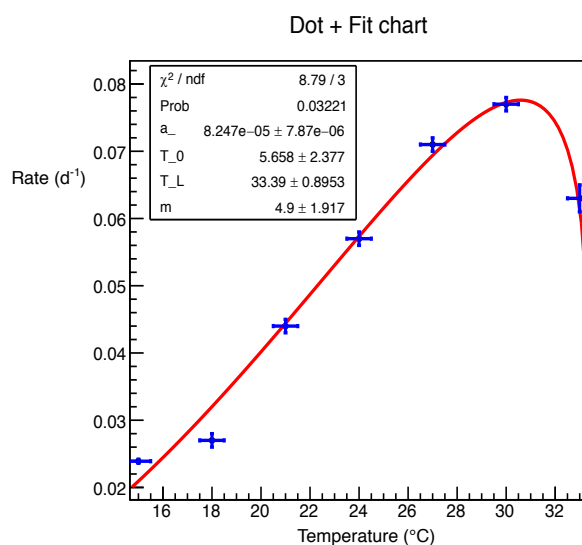


Fig. 3.1: Example of non-linear fit using “fittasso” macro. This is a screenshot from the .txt file in output, corresponding to the graph in Fig. 3.2. Data for the calculation are from Toapanta et al. [53].

- *fittasso*: it operates in the same way as “*fiterlang*” described above. There are many input files based on the type of function with which users want to operate: “*LSDM.txt*” is specific for Logan and Sharpe and De Michele development rate-function, while “*briere.txt*” and “*tassolin.txt*” are dedicated respectively to Brière and linear development rate-functions. When macro is loaded, it is possible to choose which kind of fit is necessary, thereby obtaining results in .txt files (“*LSDM_fitris.txt*”, “*fitlin_ris.txt*”, “*Logan_fitris.txt*”). This indicates the initial values for the fitting process, the best fit parameters with their errors and, lastly, χ^2 -value, adjusted R^2 -value, and the number of degrees of freedom (Fig. 3.1). Additional helpful information includes the coordinate (T_{max} , $R[T]_{max}$) related to the maximum of each function discussed. Information such as this is important because it indicates the optimal temperature for a species’ development. In addition, the associated errors are reported. Two charts containing a canvas provides a graphical view (Fig. 3.2), available in the same way as through the “*fiterlang*” tool.

```

      Briere FIT FUNCTION: fit a Briere function of the type R[T] = [a]*T*(T-
[T_0])*pow((T_L-T), (1/[m]))

Initial value setted for a = 1.0e-03
Initial value setted for T_0 = 1.2e+01
Initial value setted for T_L = 3.4e+01
Initial value setted for m = 3.0e+00

FIT RESULTS:

*****

a =          8.2469e-05  +/-  7.8687e-06
T_0 =        5.6576e+00  +/-  2.3766e+00
T_L =        3.3390e+01  +/-  8.9511e-01
m =          4.9002e+00  +/-  1.9167e+00

ChiSquared = 8.790
Coefficient of determination (R^2) = 0.985
Adjusted coefficient of determination (R^2_adj) = 0.955
Degrees of freedom = 3

Optimal temperature:

T_opt =      31.659366  +/-  1.436632

Optimal rate value R[T_opt]:

R[T_opt] =    0.997120  +/-  0.014256

*****

Created by Luca Rossini
Universita' degli Studi della Tuscia, Viterbo

```

Fig. 3.2: Example of non-linear fit using “*fittasso*” macro, which is the most important of the two charts. This is a Brière fit function with data for the calculation from Toapanta et al. [53]. The results are provided in graphical mode.

- *fitlineare*: linear regression is one of the most common operations when working with data. Even if ROOT has a specific tool for linear fitting, and even we also added a function of this type in the last macro, we deem it appropriate to include a tool on the logic line of the others. In a .txt file named “*fitlin.txt*”, it is possible to insert the x-values, y-values, and their respective errors into four columns. The outputs are a canvas with two charts and a .txt file, “*lin_fitris.txt*”, with the same list mentioned in “*fittasso*”.

- *functioncalculator*: a development rate-function is needed for a simulation, because when recording the daily temperature, it is possible to know the daily development-rate (or the daily mean development-time). The task of this tool as follows: it takes for its input the daily average-temperature and evaluates the rate function of these points. The input file is called “dailytemp.txt”, and it is composed of two columns. The first accounts for days, and the second reports the corresponding average temperature. After launch, the user can choose which kind of development rate-function to use before inserting the corresponding best-fit parameters (evaluated with “fittasso” macro). A final question is whether the user wants only to calculate rates and average times, or to also visualize the results in a chart (Fig. 3.3). Results can be found in two .txt files, “DailyRate.txt” and “MeanDeveloptime.txt”. A specific control in the code eliminates temperatures greater or less than the development thresholds for the analysed species so as to reduce distortions in simulation due to abnormal values.

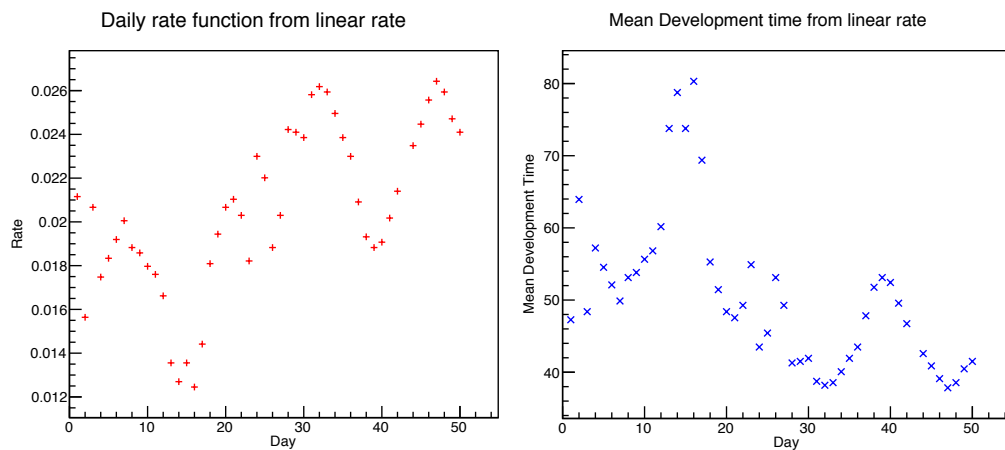


Fig. 3.3: Daily rate function (left) and daily mean development time (right) from linear rate function, worked out with “*functioncalculator*” macro. A linear rate function has been calculated with data from Toapanta et al. [53]. The rate is relative to the daily temperature registered during a period of 50 days, the information for which a meteorological station provides.

- *Simulatore 1.0*: one of the needs introduced is to solve the system (3.2) numerically, and now, with previous tools, we have all the input files we need. The time-scale is daily, and it corresponds to each step in Euler’s method [84]. All the equations are located in columns, then incremented row by row. A configuration of this type means that every row corresponds to the daily value for each h -th equation. There are different input files that merit discussion here: the first is the “cohort.txt” file, which reports the initial population if the information is available. If such information is not available, there is a “standard” file that concentrates a cohort on the day zero, meaning that the other days have 0 (zero) values. A configuration of this type causes results to be read as probability densities values. The peaks will therefore represent the most probable developmental period. The second input is “MeanDeveloptime.txt”, the output from “*functioncalculator*”, which provides the corresponding daily-value of the mean development time, $D[T(t)]$, shown in the system (3.2).

Because of how “*functioncalculator*” is built, there are many “MeanDevelopTime.txt” files that are ordered in different folders. When the user loads “*simulatore*”, the program asks them which kind of developmental rate-function is preferable, thereby ensuring the right one is used. Outputs in this case are different. A folder stores a series of .txt files called “h=number.txt” in which “number” is indicated by the corresponding equation’s number. Each file is composed of two columns, one relative to the number of days, and the other to the corresponding Erlang’s value, calculated numerically. Before the end of the process, it is possible to choose which files to represent in a canvas composed of four charts (hence must be chosen four *h*-values), and then which files to represent in a single chart (Fig. 3.4). The guideline to have numerical and graphical results is maintained here. In addition, between the outputs there is a last file named “*simulerlang.txt*” in which all the single *h*-files are gathered into columns. Before the calibration, it is difficult to understand which is the most correct file, but afterward the user has the “best *h*-value” available.

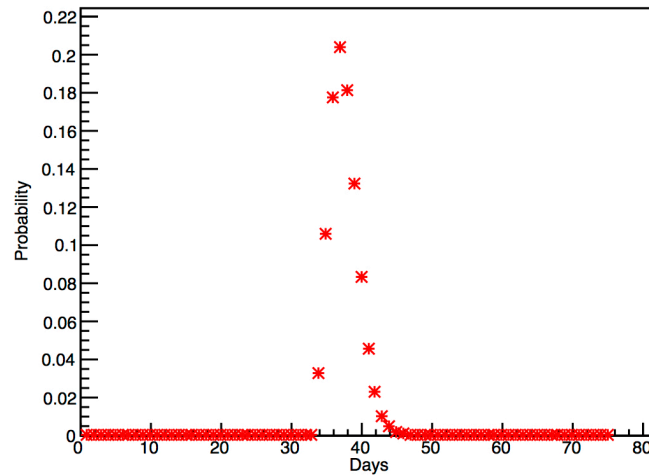


Fig. 3.4: A simulation output from the “*simulatore*” macro. Data relate to the lifecycle of *Anthonomus eugenii* at constant temperature (24°C), with a Brière rate function in input. Data for the Brière’s rate function parameter estimation are from Toapanta et al. [53] (Fig. 3.1).

- *calibratore*: in the previous section we indicated there is a need to know which *h*-value corresponds to reality. The Euler’s method application to Manetsch’s equations has the inconvenience of needing to solve all equations because it lacks a formula that relates the experimental *h*-value with the *h*-th system equation. To resolve this problem, there is “*calibratore*”, which takes as its input “*obserlang.txt*”, a file containing field validation data organized in two columns: days and the number of individuals. The sum of the standard deviation between each point and the nearest *h*-value is printed on the ROOT’s console serve to compare every *h*-file resulting from “*simulatore*”. That value is the one to read in every future simulation.
- *DS terminal interface*: a primary interface collects and presents all macros described above.

The user can choose what they intend to use by typing the corresponding letter. In that way, it is possible to save more time between operations.

3.5 Discussion and Conclusions

In the context of entomological modelling, several studies have discussed building software using one of the existing programming languages, as well as which among the languages is most adaptable. The operation presented in this article, if entrusted to a software company, can absorb many economic resources (i.e. to pay developers or software licenses [85]). Furthermore, a common problem is that there are communication difficulties between users and software developers - typically belonging to different scientific fields. Misunderstandings are often responsible for time delays due to software adjustments [86], [87]. The economic aspects, together with the possible but not immediate availability of the professional roles required, penalizes small scientific groups, above all in developing countries [88]. With this work, we present an application in the field of entomology of an existing and powerful open-source software that does not require a deep knowledge of programming language. This current development is only an early version, and future developments are likely to include several additional functions, such as mortality and fecundity rates, as well as an estimation of the errors beginning from the firsts steps until the simulation. It is important to indicate that this study has not considered two aspects: the mortality and fecundity rate. However, these represent a simple modification to add in future works without making major changes in what we have illustrated. An additional modification would be integration with other common models used in population dynamics, such as the Von Foerster equation [57] and other models based on Markov chains [27], [89], [90]. The user in this way could have a broader choice so as to align with their needs. An improvement of this kind would be an extension of software usefulness in ecology. Last but not least is the will to transform the “*DS terminal interface*” into a graphic one that is simple to use and that can allow management of more operations. ROOT represents a good basis for our aims. It poses wide documentation and a large community of users, and it covers all the most common mathematical operations through its various functions [91]. The importance of modelling approaches in Integrated Pest Management has been increasing. A good forecast is important as a Decision Support System to show to Agronomists and Technicians which is the better moment for chemical treatments, and for the launch of natural enemies in the case of biological control. Furthermore, Gutierrez et al., in their recent works [19], [20], [92] illustrated that modelling in agriculture is important also for projections of the risks in the case of an increase of temperatures due to climate change [93]. This aspect has great importance in the twenty-first century; in plant protection it enforces the role of simulations, but it is also important in relation to the problem of invasive species. Such alien species establishing and spreading are the consequences of increased human mobility and global trade [94], and this problem’s prevalence is evidence of the far-reaching effects of global warming [95], [96] that has exponentially sped up [97], [98]. The introduction of invasive species represents a threat for composition and processes of ecosystems, causing a decrease of native species abundance and richness [99]. This also leads to an increase of costs associated

with their management and damage [100]. Simulations are of crucial interest to provide a quick response to scientists, farmers and technicians to plan an eradication or control strategy of the non-indigenous species. This study uses applications of the described ROOT's libraries in the case of two alien species to Italy (*A. eugenii* and *T. absoluta*), and they will be the object of further publications.

Chapter 4

Towards a new model: limitations of the Distributed Delay Model

*“I think theories are just a way to hide
the difficulties under the carpet”*

- Richard P. Feynman

The Distributed Delay Model, despite its wide use, presents issues related to its calibration and validation, above all if applied in varying-temperature conditions. In fact, the main issue is the lack of a direct correspondence between the number of insects’ life stages (or preimaginal stages) and the number of model’s equations. This third article [51] analyses these criticalities and proposes an alternative methodology to estimate the h parameter. The results and considerations emerged from this study led to the fomulation of two new models, the main goals reached during the PhD work.

Article details:	
Title:	Distributed Delay Model and Von Foerster’s equation: Different points of view to describe insects’ life cycles with chronological age and physiological time
Authors:	Luca Rossini, Stefano Speranza and Mario Contarini
Corresponding author:	Luca Rossini
Journal:	Ecological Informatics

Distributed Delay Model and Von Foerster's equation: different points of view to describe insects' life cycles with chronological age and physiological time

4.1 Abstract

Mathematical models are a powerful tool when used to describe ectotherms' life cycles, above all for their suitability in being utilised for decision support systems. In particular, two models continue to arouse the interest of the scientific community and inspire new developments: the Manetsch-VanSickle Distributed Delay Model and the Von Foerster equation. Even though these models have been widely studied, discussed and applied, some aspects relating to their different points of view in representations of the same life cycle are yet to be explored. One of the main issues open for ongoing investigation is the different modes of division in preimaginal stages, which leads to different interpretations of the concept of age between the two models. The Distributed Delay Model considers a subdivision in h chained preimaginal stages with the same size, based on the concept of physiological time, in which the development of the species is related to the daily average temperature. On the other hand, the Von Foerster equation considers chronological age, defined commonly as a time with a different scale. This work highlights the analogies between the two models and shows, using the case study of *L. botrana*, how to obtain the number of the h stages considered by the Distributed Delay Model, from the number of observed preimaginal stages of the Von Foerster equation. To make the models comparable, the upwind scheme has been applied to the Von Foerster equation, leading to a system of ordinary differential equations that is similar to the Distributed Delay Model.

4.2 Introduction

Nowadays, the availability of reliable mathematical models describing the life cycle of ectotherms is broad, as are their applications in decision support systems. Among this variety of choice, two models in particular continue to arouse the interest of the scientific community, though they were introduced several years ago: the Von Foerster equation (VFE) [57] and the Distributed Delay Model (DDM) [15], [16]. Their distinguishing feature is their ability to represent a cohort of individuals developing over time and through their life stages, making them suitable to describe populations of ectotherms. After their first introduction, in fact, these models were widely applied in several case studies, ranging from plants and insects to fish populations, and they were an important milestone for the development of more recent physiologically-based models [31], [101], [102]. The VFE considers age as a time with a different scale, while the DDM connects the mean development time of the population with a series of sub-models, which describe the relationship between the species and environmental parameters. Therefore, the DDM has the added value of considering age not as a time with a different scale, but as a "physiological time" [52], [103], essentially defined as the units of degree days [58], [104]. The units of degree days are represented in different ways, ranging from the simplest definition of "thermal summation"

[25], [78], to more accurate development rate functions [45]. With this precondition, VFE and DDM seem at first sight to be two different models and thus not correlated; however, the work of Plant and Wilson [58] showed that it was possible to obtain the DDM from the VFE. Their work suggests that these two models only differ from each other in having different point of view in representing the same life cycle, suggesting a connection between chronological age and physiological time. A mathematical congruence between DDM and VFE was subsequently argued in a more straightforward way by Gilbert et al. [14]. The VFE and the DDM represent an insect's life cycle as a black box made up of chained boxes, known as substages [105], [106]. These substages are driven by the environmental parameters in the case of the DDM but this is not the case in the VFE. Accordingly, from an experimental point of view, the modelling approach provided by the DDM is more suitable to describe ectotherms, while the VFE is more appropriate for homeotherms. On the other hand, the DDM requires the additional step of calibration to estimate the parameters before the validation [42] and consequently involves greater effort during the experimentations. The issue implicit in the DDM is the h parameter, defined as the number of stages in the delay process [107]. The biological meaning is not completely clear, since there is not a direct connection with the number of observable stages in insect life cycles. In the first description provided by Manetsch [15], h also represents the number of the system's equations. Hence, h is one of the unknown values of the model, and further experimental data are needed for its estimation [35]. Chronological age in the VFE, however, has the advantage of being fixed depending on the case study. Initially it could be considered reasonable to determine the age according to the number of observable stages in an insect's life cycle, but there is no information about the duration (in time units) of each stage. Expanding from the work of Plant and Wilson [58], this study proposes a way to overcome the problem of the estimation of h in the DDM, using its connection with the VFE. Moreover, to test the reliability of this new methodology, an example of its application in the case study of the European grapevine moth, *Lobesia botrana* (Denis & Schiffermüller) (Lepidoptera: Tortricidae), is given.

4.3 Materials and Methods

4.3.1 Connections between Distributed Delay Model and Von Foerster equation

From a theoretical point of view, insects' life cycles can be described using a single big block, where input and output flows illustrate the distribution of eggs and adults respectively. However, this approach lacks information about the preimaginal stages, for which reason several authors [21], [25], [31], [43], [105] set out a more realistic scheme, detailed in Fig. 4.1. The preimaginal stages can be schematised by a series of chained boxes contained in the single big block, making up the whole life cycle. Manetsch and VanSickle [15], [16] used this assumption to introduce the DDM, specifying that all the substages have the same duration. The total duration of the life cycle is divided into h parts where the individuals of the initial cohort are distributed [21]. Mathematically speaking, the DDM is represented by a system of chained Ordinary Differential

Equations (ODEs), one for each substage drawn in Fig. 4.1 [2], [49]:

$$\begin{cases} \frac{d}{dt}Q_1(t) = \frac{h}{D(t)}Q_0(t) - Q_1(t) \left[\frac{h}{D(t)} + M_1(t) \right] \\ \frac{d}{dt}Q_2(t) = \frac{h}{D(t)}Q_1(t) - Q_2(t) \left[\frac{h}{D(t)} + M_2(t) \right] \\ \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ \frac{d}{dt}Q_h(t) = \frac{h}{D(t)}Q_{h-1}(t) - Q_h(t) \left[\frac{h}{D(t)} + M_h(t) \right] \end{cases} \quad (4.1)$$

The variables contained in the system (4.1) are: $Q_n(t)$, the number of individuals into the n -th stage; h , the number of substages; $D(t)$, the expected life cycle's duration; $M_n(t)$, the specific mortality rate function for the n -th substage. $D(t)$ and $M_n(t)$ are usually estimated with laboratory experiments, while a different situation concerns the parameter h .

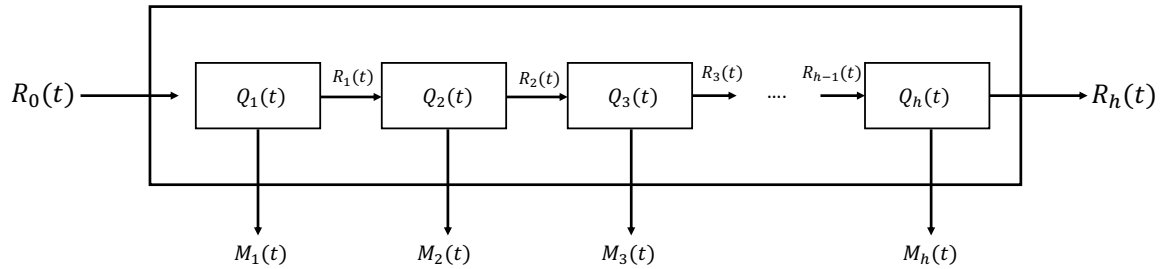


Fig. 4.1: Schematic representation of an ectotherm organism's life cycle. The whole life cycle can be considered as a single big block composed of a series of chained substages. Considering the n -th stage, $R_{n-1}(t)$ and $R_n(t)$ are the input and output flows respectively, $M_n(t)$ is the mortality flow and $Q_n(t)$ is the number of individuals into the next stage.

Manetsch [15], VanSickle [16] and Severini et al. [21] indicate h both as the number of substages involved in the process and as the number of equations in the model (4.1). However, experience suggests that the preimaginal stages composing an insect's life cycle do not necessarily have the same duration: for example *L. botrana* passes most of its life cycle as fifth instar larva and pupa [55]. Accordingly, there is no correspondence between h and the observable life stages and this means that h is one of the unknowns of the DDM, whose determination requires a calibration experiment. A method to estimate h from field data was proposed by Rossini et al. [49]. In particular, they proposed: *i*) to numerically solve a high number of equations of the system (4.1), and *ii*) to consider the h value appearing in the fraction $\frac{h}{D(t)}$ such that the convergence of the numerical scheme is ensured. Subsequently, each numerical solution is compared with the adults' population density provided by a field experiment. The reliability of this method is based on the assumption that the only observable life stages having a direct correspondence with the DDM equations are the first (the distribution of eggs) and the last (the distribution of adults). In fact, referring to Fig. 4.1, it is possible to see that the left and right borders of the big black box (the whole life cycle) are the first and the last substages,

respectively. The unknown data is how many substages there are between the first and the last substages, and for this reason they can be defined as “non observable substages” [108]. The overlap between field data and numerical solutions [30], [42] is evaluated by the χ^2 -function, defined as:

$$\chi^2 = \sum_{n=1}^{n=N} \frac{(D_n - S_n)^2}{S_n} \quad (4.2)$$

and where S_n and D_n represent the simulated and field data, respectively. Equation (4.2) compares the N daily values of the field population with each corresponding daily value of numerical solutions calculated for DDM equations. In force of the aforementioned assumptions, the last equation (the h -th) of the system (4.1) is the equation with the lower χ^2 -value, and it will provide any further simulation for the species. However, this method is highly expensive from both an IT and experimental point of view, and that is why an alternative method is required. Let us introduce the McKendrick-Von Foerster equation [57], [109],

$$\frac{\partial}{\partial t} N(t, a) + \frac{\partial}{\partial a} N(t, a) = -M(t, a) \cdot N(t, a) \quad (4.3)$$

which describes the density of a population $N(t, a)$ evolving over time t , chronological age a and with mortality rate $M(t, a)$. As evidenced by Plant and Wilson [58], equation (4.3) can be discretized in the age variable a using the method of lines, also known as “upwind scheme” [110]. Next, there is a sequence of values a_n , with the condition $a_n < a_{n+1}$, with the aim to replace the continuous variable a [31]. Moreover, considering the interval $A_n = a_n - a_{n-1}$ the partial derivative on the age variable can be substituted with its discrete approximation:

$$\frac{\partial}{\partial a} N(t, a) = \frac{N(t, a_n) - N(t, a_{n-1})}{A_n} \quad (4.4)$$

The discretization can be also extended to mortality $M(t, a)$ which is now $M(t, a_n)$. To avoid confusion, let us change the notation from $N(t, a_n)$ to $N_n(t)$ for population density, and from $M(t, a_n)$ to $M_n(t)$ for mortality. Hence, the VFE (4.3) is now transformed into a series of chained ODEs, similar to Manetsch’s DDM (4.1):

$$\begin{cases} \frac{d}{dt} N_1(t) = \frac{1}{A_1} N_0(t) - N_1(t) \cdot \left[\frac{1}{A_1} + M_1(t) \right] \\ \frac{d}{dt} N_2(t) = \frac{1}{A_2} N_1(t) - N_2(t) \cdot \left[\frac{1}{A_2} + M_2(t) \right] \\ \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ \frac{d}{dt} N_i(t) = \frac{1}{A_i} N_{i-1}(t) - N_i(t) \cdot \left[\frac{1}{A_i} + M_i(t) \right] \end{cases} \quad (4.5)$$

Since chronological age a is defined as a time with a different scale, this assumption is maintained also for its discrete form. Before continuing, two additional assumptions are needed: *i*) each equation in the system (4.5) identifies only an observable life stage and, *ii*) A_n is the

duration, in time units, of each observable stage. The analogies between the system (4.1) and (4.5) suggest a comparison of the h -th and i -th equations, respectively, namely the last ones of both the systems which describe the temporal distribution of adults. A graphical support is provided in Fig. 4.2, where the two different subdivisions of the same life cycle adopted by the systems (4.1) and (4.5) are compared.

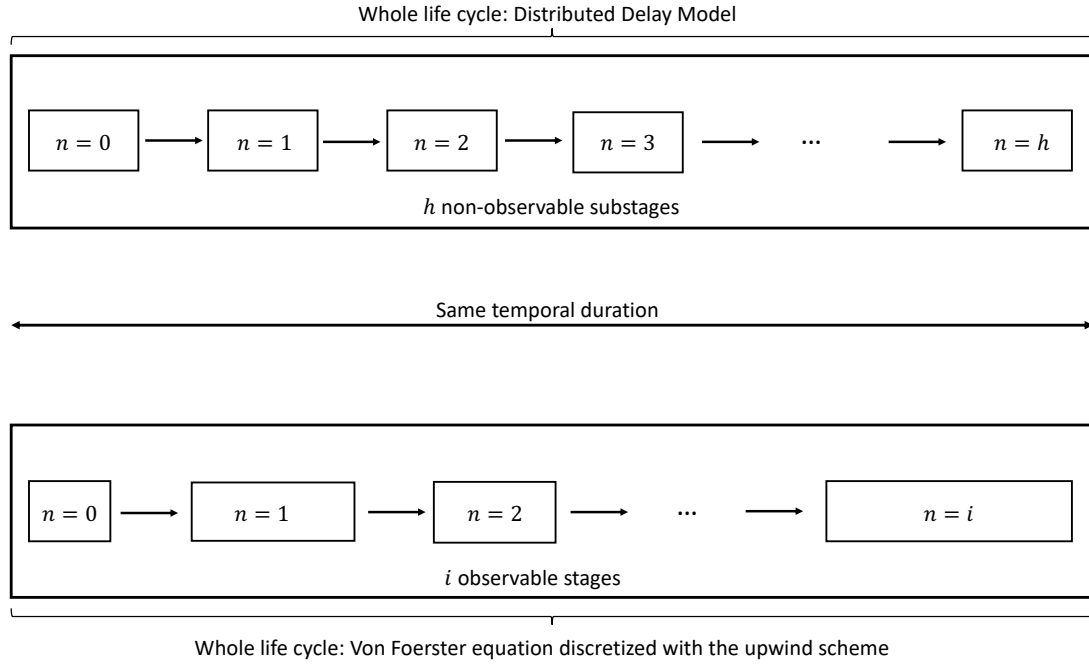


Fig. 4.2: Comparison between the representation of an insect's life cycle using the Distributed Delay Model and the Von Foerster equation discretized with the upwind scheme. The duration of the whole life cycle is the same, but the DDM considers a life cycle composed by a series of non-observable substages of the same size, while the Von Foerster equation considers a defined number of observable preimaginal stages, but with different temporal durations to each other.

Hence, either h or i provide the distribution of adults. In the light of the assumptions, comparing the last equation of the system (4.1) and (4.5), the following equivalence is obtained:

$$\frac{1}{A_i} = \frac{h}{D(t)} \quad (4.6)$$

that, if inverted, provides the value of h :

$$h = \frac{D(t)}{A_i} \quad (4.7)$$

However, a final consideration is required regarding the expression (4.7). As already discussed, the A_n values have different durations to each other. Let us define the “stage unit” as the expected time of each observable life stage in the system (4.5). In other words, it is possible to indicate the A_n values as a sequence of natural numbers $A_1 = 1$, $A_2 = 2$, ..., $A_i = i$ which clearly identifies the observable life stage described at the bottom of Fig. 4.2. This definition

leads to the final version of the expression (4.7):

$$h = \frac{D(t)}{i} \quad (4.8)$$

which provides a method to estimate the h value in the DDM, once the number of observable stages for an insect's species is known.

4.3.2 Verification of expression (4.8)

Working with the observable stages i allows their number to be fixed from biological observations. Hence, the experimental part of the present work aims to confirm the number of observable stages of *L. botrana* from a simulation with the DDM (4.1). If the number of observable stages i , resulting from the experimentation detailed below, is consistent with the theoretical value obtained from biological observations, then expression (4.8) is a reliable method to estimate h .

The application of the DDM requires, as input, the daily average temperatures $T(t)$ recorded in the experimental field. These values will be processed by the development rate functions, a family of expressions which puts in relation the mean development time $D(t)$ and the temperature [49]. Even if there are various development rate functions available in the literature [44], [45], the Logan formula [46], [52] was used for this study. The application of the development rate functions in the case study of *L. botrana* was explored by Rossini et al. [30], [42], as well as an estimation of parameters. Mathematically speaking, the Logan formula is presented as follows:

$$\frac{1}{D(t)} = \psi \cdot \left[\exp(\rho T(t)) - \exp\left(\rho T_M - \frac{T_M - T(t)}{\Delta T}\right) \right] \quad (4.9)$$

where $T(t)$ is the daily average temperature recorded in the field; ψ and ρ are empirical parameters; T_M is the upper temperature threshold above which the species cannot develop and; ΔT indicates the range of abrupt collapse due to high temperatures.

Expression (4.9) is subsequently inserted in a form of the system (4.1) discretized with the Euler's difference method, as reported by Manetsch and VanSickle [15], [16], and computed with the tools presented by Rossini et al. [49]. To test the expression (4.8), only the third and last *L. botrana* generations recorded in the field were simulated (see Section 2.3), while the flights of the second generation were used as initial conditions for the numerical scheme, as already reported by Rossini et al. [42]. The next step was the determination, through the expression (4.2), of the numerical solution which fit better with the field data.

The parameters h and $D(t)$ were estimated from the numerical solution chosen in the previous step, in the light of the following reasoning. As highlighted by expression (4.9), the expected time $D(t)$ depends on the daily average temperature recorded in the field. Since there is not an analytical expression to represent the temporal variation of the daily average temperature along the year [41], [42], a numerical scheme for the system (4.1) is the only reasonable solution. On the other hand, considering the definition of physiological time [52], [82], the daily increase in the insect's development depends on the daily average temperature. Each of these daily

contributions designs, at the end of the life cycle, an average temperature value T_{DT} , called “total development temperature” [56], associated with the mean generation time $D(T_{DT})$. This consideration allows the introduction of the Erlang Probability Density Function (Erlang PDF) which is, as showed in literature [21], [34], [38], the analytical solution of the time-invariant version of the system (4.1):

$$f(t) = N \cdot \left(\frac{h}{D(T_{DT})} \right)^h \cdot \frac{t^{h-1}}{(h-1)!} \cdot e^{-\frac{h}{D(T_{DT})}t} \quad (4.10)$$

Hence, fitting the numerical solution selected by the expression (4.2) with expression (4.10) it is possible to estimate N , the total of the individuals within the population, h , the number of non-observable substages in the DDM, and $D(T_{DT})$, the mean generation time. The non-linear fit was calculated with the tool indicated by [49], and then evaluated with the adjusted R^2 -value, and through a χ^2 test. The final step of the analysis provided for the insertion of h and $D(T_{DT})$ into the expression (4.8), solved for i :

$$i = \frac{D}{h} \quad (4.11)$$

to check if the number of observable stages, i , is consistent with its theoretical value, i_{th} , derived from biological observations. Since the non-linear fit (4.10) provides the best fit parameters, N , h , D and their associated standard errors σ_N , σ_h , σ_D , using the propagation of uncertainty formula:

$$\begin{aligned} \sigma_i &= \sqrt{\left[\frac{\partial}{\partial h} \left(\frac{D}{h} \right) \cdot \sigma_h \right]^2 + \left[\frac{\partial}{\partial D} \left(\frac{D}{h} \right) \cdot \sigma_D \right]^2} = \\ &= \frac{1}{h} \sqrt{\frac{D^2}{h^2} \cdot \sigma_h^2 + \sigma_D^2} \end{aligned} \quad (4.12)$$

it has been possible to calculate the error associated to i in the expression (4.11).

4.3.3 *Lobesia botrana*: biology and field survey

The validation of the method described in Section 4.3.2 was conducted in the case study of the European grapevine moth, *Lobesia botrana*. This choice of species was primarily determined by the high levels of interest surrounding this insect, due to the high rate of damage that it causes in grape fields. Its harmfulness lead scientists to delve into its biology: the relationship between temperature and mean development time was discussed by Brière and Pracros [111], and Moshtaghi Maleki et al. [55]. In particular, they provided the life tables, a common tool in entomology to analyze the response of poikilothermic organisms to thermal variations [48]. Meanwhile, Rossini et al. [30], [42] used data from Moshtaghi Maleki et al. [55] to test various development rate functions and in addition provided the parameters for the Logan expression (4.9). The other biological information needed to validate the expression (4.8) concerns the number of observable stages in the life cycle of the European grapevine moth. From the specific

literature [55], [111], [112] it emerges that *L. botrana* has a total of $i_{th} = 7$ observable stages, divided into egg, 5 larval instars, pupa and adult, considering the egg stage as $i = 0$. Alongside this, data about the field population were provided by Rossini et al. [42]. In particular, they conducted a three-year survey in the Castelli Romani area (Lazio, Italy), monitoring the adult flights of the three generations usually reported in that area. The validation of the expression (4.8) involved only the second and the third generation of the experiment year 2015 in applying the procedure detailed in Section 4.3.2. As reported by Rossini et al. [42], adult males were captured with three pheromone sticky traps (Biogard BDT Delta trap, Biogard CBC Europe S.r.l.), which were deployed in the vineyard and checked weekly. More specifically, traps were hung in the plant's canopy and left in the field from April to September. The daily average temperature was recorded using an automated meteorological station (AWS310, Vaisala company, Helsinki, Finland) placed in the vineyard (which recorded 48 measurements per day).

4.4 Results

The first adult of the second generation was caught on 1st June 2015 and, accordingly, this day was fixed as the starting point for simulations. As explained in Section 4.3.2, these data were inserted as initial condition into the numerical scheme, together with the daily average temperature values recorded by the meteorological station in the field. Third generation's adults, however, appeared on 15th July, provided a maximum peak on 16th August and ended on 23rd of September, determining the end of the simulation. Moreover, the daily average temperature was inserted into the discrete version of the system (4.1) using the Logan expression (4.9), whose parameters are listed in Table 4.1. To ensure the convergence of the numerical scheme, a value $h = 9$ has been assigned, and a number of 250 equations for the system (4.1) were solved. Each solution was then compared, using the expression (4.2), with the field data of the third generation, in order to detect which was the nearest. With a $\chi^2 = 37.994$, the 21st solution was selected: a graphical comparison between the elected simulation and the third generation's flights is showed in Fig. 4.3.

Table 4.1: Parameters of the Logan development rate function used for the simulation with their corresponding standard errors. Additional information is provided by the χ^2 -value, adjusted R^2 -value and the number of degrees of freedom (NDF). These data were retrieved from Rossini et al. [42].

Logan rate function parameters ($\pm SE$)	χ^2 -value	R^2_{adj} -value	NDF (n)
$\psi = (7.0 \pm 0.5) \cdot 10^{-4}$			
$\rho = 0.18 \pm 0.04$			
$T_M = 34.1 \pm 0.8$	0.0019	0.925	2
$\Delta T = 4 \pm 2$			

From the plot in Fig. 4.3, it is possible to notice the accordance between the simulation and observed data: the field population's peak occurred on 16th August, while the simulation

indicated the same value on the 53rd day, namely 15th August. Following the procedure described in Section 4.3.2, fitting the numerical solution with Erlang PDF (4.10) with the computational tools presented by Rossini et al. [49], the parameters N , h , and $D(T_{DT})$ were obtained. The best fit values are listed in Table 4.2, while a graphical comparison is plotted in Fig. 4.4.

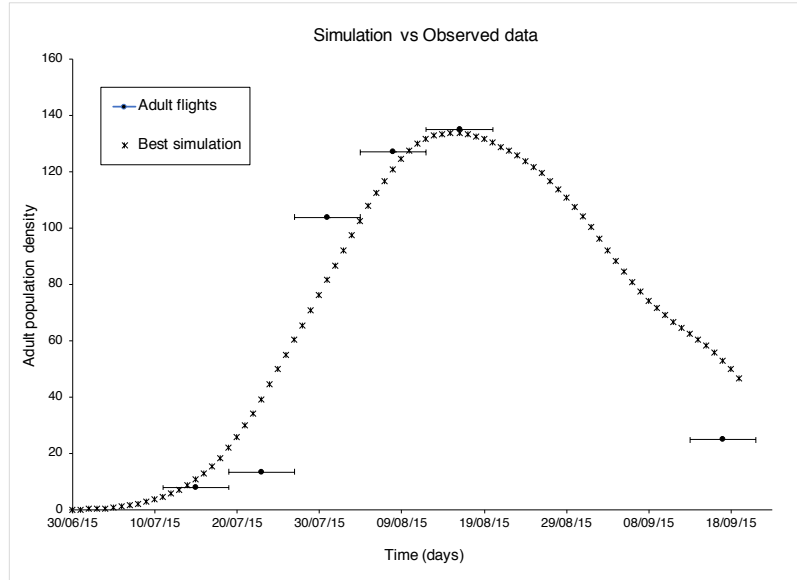


Fig. 4.3: Comparison between the third-generation flights occurred in the experimental field and the simulation. Error bar on field data is fixed to ± 3.5 days, half of the sampling range [42].

Table 4.2: Erlang PDF (4.10) best fit parameters and their standard errors. Additional information are provided by the χ^2 -value, by the adjusted coefficient of determination R_{adj}^2 and by the number of degrees of freedom NDF.

Erlang PDF parameters ($\pm SE$)	χ^2 -value	R_{adj}^2 -value	NDF (n)
$N = 6500 \pm 10$			
$h = 9.0 \pm 0.6$	39.47	0.969	85
$D = 61.8 \pm 0.7$			

The R_{adj}^2 -value (listed in Table 4.2) close to 1 and a χ^2 -test, which reported $P < 0.001$, indicated a good accordance between the Erlang PDF (4.10) and the numerical solution of the system (4.1). The last step of the present study provided the calculation of i and its corresponding error σ_i using the expressions (4.11) and (4.12), respectively:

$$i = 6.8 \pm 0.3 \quad (4.13)$$

This value is consistent with the theoretical value derived by biological observations which is, as reported in Section 4.3.3, $i_{th} = 7$.

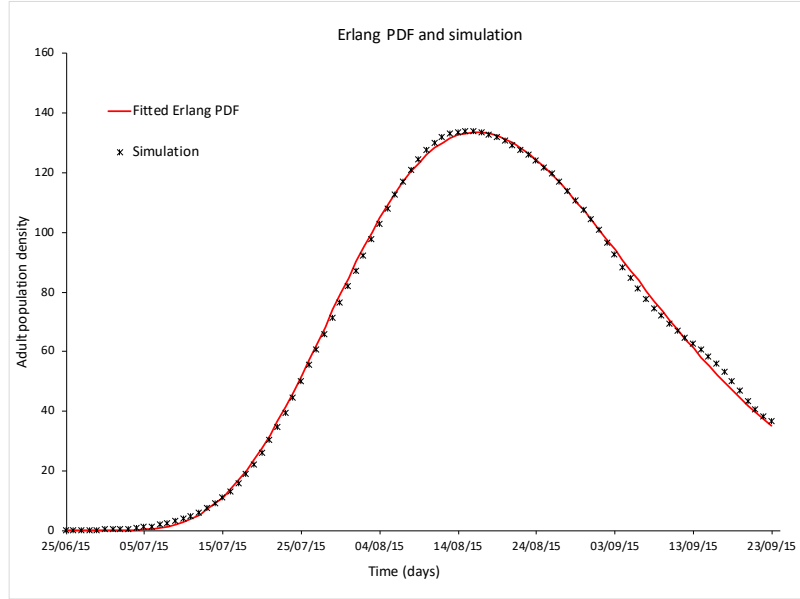


Fig. 4.4: Best fit function with expression (4.10) using “*fiterlang*” macro from Rossini et al. [49].

4.5 Discussion and Conclusions

The results above provide various types of information about the relationship between insects’ life cycles, the Distributed Delay Model and the Von Foerster Equation. The main result is surely contained in the expression (4.13), where the consistency between the results obtained applying the expression (4.11) and its theoretical value i_{th} has been demonstrated. The advantage of this result lies in the applicability of the expression (4.8) to estimate the h value, which is helpful during the application of the DDM. In fact, the value i_{th} derives from biological observations and it is usually determined with laboratory rearing [48]. Moreover, it is possible to estimate a mean generation time $D(T_{DT})$ even if the daily average temperature fluctuates. A possible strategy, in fact, is to analyse the daily average temperatures during the days of the experimentation or simulation, and to find an average value T_{DT} . Hence, through the development rate function it is possible to obtain $D(T_{DT})$, to substitute in the expression (4.8). At this point, the value h for the DDM application is provided.

The second relevant result of this work concerns the possibility of approximating an insect’s development in varying temperature conditions, with an insect’s development in constant temperature conditions. In particular, this evidence has been shown through fitting the numerical solution of the system (4.1) with the Erlang PDF. The Erlang PDF is the analytical solution of the system (4.1), and a constant temperature development implies that $D(t)$ is time-invariant. On the other hand, the numerical solution is required in case of time-varying temperature, because in this case the mean development time $D(t)$ changes day by day. However, at the end of the development, it seems that each daily contribution leads to an average temperature value, T_{DT} , confirming, also, the hypotheses of Colinet et al. [56]. Hence, numerical solutions are required because the daily average temperature changes on time, and there is need of direct measurements. The Erlang PDF describes the numerical solution of the system (4.1) accurately;

this implies that the organisms accumulate heat units day by day, leading to a final averaged value. In other words, this result shows that high fluctuations in the daily average temperatures do not provoke abrupt increases or decreases in insects' development.

This work has investigated the connections between the DDM and the VFE, showing how effectively physiological time is a clear indication (in time units) of insects' life stages. This connection extends the applicability of the Von Foerster equation to the world of ectotherms. The considerations discussed in this work, in fact, led Rossini et al. [31] to formulate a model which includes some principia of the VFE (such as a PDE-based age-structured model), and some principia of the DDM (the dependence of stage duration on environmental parameters). In fact, on the one hand, a method emerges from this study to estimate h in the DDM; on the other hand, the difficulty in applying both the DDM and VFE to ectotherms remains. Hence, merging the issues highlighted in this study, it is possible to conclude that: *i*) reliable models to describe insects' life cycles need to consider physiological age instead of physiological time or chronological age, and *ii*) there is a need to extend the concept of the development rate function including other environmental parameters.

Chapter 5

A novel model: the “Generalised Von Foerster equation”

*“Do. Or do not.
There is no try.”*

- Master Yoda

The times are ripe. After the deep exploration of the Distributed Delay Model and of the Von Foerster equation, a novel model was formulated in this article [31]. The result is a modified version of the Von Foerster equation where a function called “generalised development rate function” is introduced. More specifically, this function act as a trojan horse introducing the relationship between insects’ development and environmental parameters. Moreover, the main advantages of this new equation are *i)* its suitability to be included in *ad hoc* calculation software, and *ii)* an easier validation, since there is a direct correspondence between insect’s life stages and physiological age. Hence, this article represents the first main goal of the PhD work.

Article details:	
Title:	A novel modelling approach to describe an insect life cycle vis-à-vis plant protection: Description and application in the case study of <i>Tuta absoluta</i>
Authors:	Luca Rossini, Maurizio Severini, Mario Contarini and Stefano Speranza
Corresponding author:	Luca Rossini
Journal:	Ecological Modelling

A novel modelling approach to describe an insect life cycle vis-à-vis plant protection: Description and application in the case study of *Tuta absoluta*

5.1 Abstract

Tuta absoluta (Meyrick) (Lepidoptera: Gelechiidae), commonly known as the tomato leaf miner, is arousing interest throughout the scientific community because of its aggressiveness against *Lycopersicon* spp., *Capsicum* spp. and *Solanum* spp. Significant economic losses surrounding the tomato plant have been reported in different parts of the world and these are, above all, the result of the immense capacity of the tomato leaf miner to reproduce within its life cycle, which spans an average of 30 days. The aim of this paper is the demonstration of a novel mathematical model capable to describe ectotherms life cycle, and to test it in the case of *T. absoluta*. A mathematical description can help farmers and stakeholders to forecast when the stages at which the insect is most susceptible to control strategies, will be reached. This will allow timely planning for the performance of the most efficient and environmentally friendly control strategy. The model is based on a first order partial differential equation, which can describe the density of population in function of the time t , and physiological age x , and which is driven by the environmental parameters and fundamental conditions for poikilothermic organisms. The result is a simulation of the life cycle, which depicts that the development of the insect is largely due to daily average temperature. This study is strengthened by double validation with semi-field data collected in an experimental greenhouse.

5.2 Introduction

Tuta absoluta (Meyrick) (Lepidoptera: Gelechiidae) is one of the most prominent pests in tomato farms. Its origins are found in South America [113], where it has been responsible for economic loss for many decades. In recent years, however, the presence of the tomato leaf miner has been reported in other countries. In Europe, the first reporting was in Spain in 2007, followed by Italy in 2008 [114]–[116]. The most affected plant genera from the Solanaceae family belong mainly to the *Lycopersicon* genus, but *T. absoluta* becomes increasingly threatening when aggressively infests other genera, such as *Capsicum* and *Solanum* [117], [118]. In Italy, susceptible target zones are areas cultivated with tomatoes, potatoes and peppers, such as the Fondi area in the Lazio region. This area provides an example of how the combined action of two aggressive alien species, namely *Anthonomus eugenii* Cano [65] and *T. absoluta* [116], can cause a transformation of the agricultural landscape and, in turn, economic loss [115], [119].

Currently, the cost of traditional control strategies against phytophagous agents has a significant impact on the total production costs and this is not compatible with sustainable-agriculture ideologies. For this reason, entomologists and Integrated Pest Management scientists have found a series of alternative control strategies that allow for organic agricultural practices that are in line with national and regional directives.

One of the main challenges to control strategies is the biological structure of the tomato leaf miner. Its life cycle begins with an egg: as reported by Biondi et al. [115] around 260 eggs are laid typically by the female on the underside of the leaves, with a hatching time of 4-5 days. After the eggs hatch, larvae penetrate the leaves' mesophyll and start their feeding activity. The larval stage represents the longest stage experienced in the total life span: its duration is approximately 20 days in optimal conditions (30°C) [115], or 60 – 70% of the total life cycle. This stage comprises of 4 instars, at the end of which the pupa stage takes over. Pupation has an average duration of 4-7 days, after which the adults fly [116], [120]. The tendency for a life cycle to end within 30 days, together with high fertility rates, means that *T. absoluta* can produce up to 9 or 10 generations in a single year [115]–[117], [120]. Damage is concentrated mainly on the leaves and can cause the total loss of the cultivated surfaces if not controlled. In more advanced cases, feeding activity can affect fruits and flowers as well [115].

As regards control strategies, chemical pesticides have been the only solution for decades. Nevertheless, several recent studies have focused on the use of alternative products. Scientific literature has recorded examples of the use of predators [121] and parasitoids, or entomopathogenic fungi and bacteria [63] [115] [64] [122]. The role of modelling is crucial in this context: a forecast of the pest population dynamic based on environmental parameters helps to plan a strategy that can maximise efficiency [21], [34], [35], [66], [123], [124]. The aim of this study is to develop a new mathematical description of the poikilothermic organisms' life cycle by using a modified version of the Von Foerster's equation [57], [83], [125], and then to test the resulting model in the tomato leaf miner case of study. The main reason for reformulating an age-structured model is to mathematically represent all the life-cycle stages for poikilothermic organisms, but also to establish a link between their developmental stages (more in general, the physiological age) and environmental factors. This condition is fundamental for a correct description in the case of poikilothermic organisms. The mathematical description laid out in the next sections, involved the division of the study into four phases: the formulation of the model; the tomato leaf miner parameters estimation; the simulation; and validation with data collected in an experimental greenhouse. In order to ensure the reliability of the work, the validation process was performed twice, and the results are reported in the conclusive part of this article. To perform all the calculations required, the new model has been inserted in a specialised software using the C++ programming language [49], and libraries from the open source software ROOT were incorporated [79], [126].

5.3 Materials and Methods

The mathematical description of a poikilothermic organism life cycle can be divided conceptually in two parts: the interaction between the environment and the species; and the population density in function of time, t , and physiological age, x . This last represents the main focus of this work, but before, an introduction about the most common models used to mathematically describe the effect of the environmental parameters (temperature mainly) on ectotherm organisms is summarized.

5.3.1 Development rate functions

Temperature, humidity and photoperiod play a crucial role in the insect’s development. Humidity and photoperiod, in particular, become more influential when environmental conditions become critical, such as happens with the change of the seasons (from hot season to cold season and vice versa). However, in absence of extreme thermal conditions changes, it can be postulated that temperature is the main driving variable. The dependence of the mean development time from the temperature can be examined with in laboratory experiments in which the so-called “life tables” are built [48]. The procedure to estimate the life tables involves a series of constant temperature rearing in a climatic chamber. A cohort of eggs laid in the same day is positioned in a climatic cell. Each egg is kept in a Petri dish to maintain the individuality of the cohort’s members. After this positioning, each insect is followed for the duration of the life cycle. At the end of the life cycle, the mean value of the development time at a specific temperature is recorded [119]. The mathematical translation of the data [44], [45], [78], is based on a conversion of the mean development times $D(T_i)$ in mean development rates $R[T_i]$ by the use of the expression (5.1):

$$R[T_i] = \frac{1}{D(T_i)} \quad (5.1)$$

The reason for this conversion is the advantage of obtaining a typical profile from the experimental data, reported in Fig. 5.1.

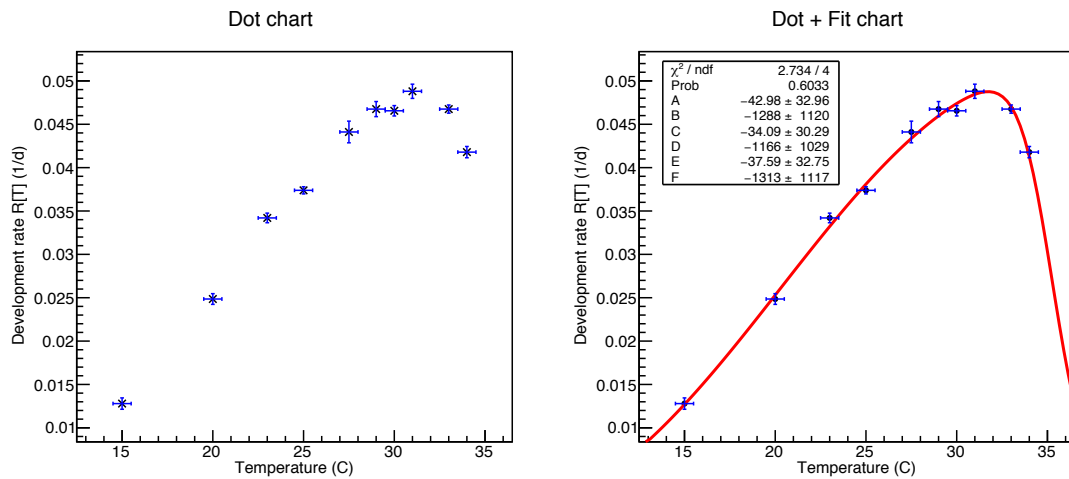


Fig. 5.1: Experimental life tables in the case of the tomato leaf miner *Tuta absoluta*. Data are obtained from the experimental life tables reported in the work of Ozgokçe et al. [74]. On the left-hand side are the mean development rates $R[T_i]$ corresponding to the constant temperature rearing at T_i . Data have been converted into rates by the use of expression (5.1), while for the associated error, the propagation of uncertainty formula has been used, reported in expression (5.2). In the chart on the right-hand side, data are fitted with the Sharpe and De Michele development rate function. Calculations have been worked out with *EntoSim*.

Temperatures start to influence development above the lower threshold point, T_L , below which the life of the insect is theoretically impossible. Beyond this critical point, there is a phase of quasi-linear increasing that is often associated with the optimum-development-temperature range for the species, which continues up to a maximum peak. This peak corresponds to the optimal temperature, T^* , at which the species moves rapidly towards the end of its life cycle. A decreasing range in the development rate function, due to high temperatures, follows, until a maximum threshold, T_M , is reached. After this point the life is once again theoretically not possible. The experimental mean development time $D(T_i)$ presents an associated experimental error, usually obtained with a standard deviation $\sigma_{D(T_i)}$. By using this uncertainty formula, it is possible to calculate the associated error for the mean development rates $\sigma_{R[T_i]}$. The result is the equation (5.2).

$$\sigma_{R[T_i]} = \frac{\sigma_{D(T_i)}}{[D(T_i)]^2} \quad (5.2)$$

The most common expressions used in mathematical descriptions of life tables are formulated mainly by empirical assumptions except in one case, which will be discussed further on. To finalise the application in the case of the tomato leaf miner, the following development rate functions have been considered:

- The linear rate function [44], [78]:

$$R[T] = \frac{1}{S}(T + T_L) \quad (5.3)$$

where S is the thermal summation, T is the daily average temperature and T_L is the lower temperature threshold. Although this expression may be unable to describe the typical decreasing trend in the life tables, it has been tested in this study for the purpose of comparison with the following one.

- The Brière rate function [47]:

$$R[T] = aT(T - T_L)(T_M - T)^{\frac{1}{m}} \quad (5.4)$$

where a and m are empirical parameters, T is the daily average temperature and T_L and T_M are the lower and upper temperature thresholds, respectively. The Brière development rate function has the feature to provide information about the thermal life spectrum of the species, despite its formulation on empirical assumptions.

- The Logan rate function [127]:

$$R[T] = \psi \left[\exp(\rho T) - \exp\left(\rho T_M - \frac{T_M - T}{\Delta T}\right) \right] \quad (5.5)$$

where ψ and ρ are empirical parameters, T_M is the upper temperature threshold beyond which development is not possible, and ΔT is the interval width between the maximum

peak of the function and T_M . With respect to the Brière expression, this development rate function does not provide information about the lower threshold, but has the advantage of providing other important information, located in the parameter ΔT . ΔT describes the width of the decreasing trend starting from the optimum temperature, and this helps to understand the capability of the insect to react in the case of increasing temperatures.

- Sharpe and De Michele rate function [45], [77], [81], [128]:

$$R[T] = \frac{T \cdot \exp\left(\frac{\phi - \frac{\Delta H_A}{T}}{R}\right)}{1 + \exp\left(\frac{\Delta S_L - \frac{\Delta H_L}{T}}{R}\right) + \exp\left(\frac{\Delta S_H - \frac{\Delta H_H}{T}}{R}\right)} \quad (5.6)$$

Where: T is the temperature in Kelvin (K), R is the universal gas constant ($1.987 \text{ cal} \cdot K^{-1} \cdot \text{mol}^{-1}$), ϕ is a conversion factor with no thermodynamic meaning, with the aim of rendering the exponential dimensionless, ΔH_A is the enthalpy of activation of the reaction that is catalysed by the enzyme ($\text{cal} \cdot \text{mol}^{-1}$), ΔS_L is the change in entropy associated with low temperature inactivation of the enzyme ($\text{cal} \cdot K^{-1} \cdot \text{mol}^{-1}$), ΔH_L is the change in enthalpy associated with low temperature inactivation of the enzyme ($\text{cal} \cdot \text{mol}^{-1}$), ΔS_H is the change in entropy associated with high temperature inactivation of the enzyme ($\text{cal} \cdot K^{-1} \cdot \text{mol}^{-1}$), ΔH_H is the change in enthalpy associated with high temperature inactivation of the enzyme ($\text{cal} \cdot \text{mol}^{-1}$). The advantage of this expression is its formulaic foundation on physical, chemical and biological hypotheses. On the other hand, the calculation effort required for the parameter estimation is higher than that required in the Brière and Logan expression. The main difference is due to the presence of six parameters to be estimated, instead of the four contained in the other expressions. Despite this difficulty, when the life tables contain a sufficient number of points, they form the most representative expression to link the function of temperature to the biology of the species. To perform the calculation, the expression can be reformulated (5.6) with the introduction of new variables. If we reformulate in the following way:

$$\begin{aligned} \frac{\phi}{R} &= A; & \frac{\Delta H_A}{R} &= B; & \frac{\Delta S_L}{R} &= C; \\ \frac{\Delta H_L}{R} &= D; & \frac{\Delta S_H}{R} &= E; & \frac{\Delta H_H}{R} &= F \end{aligned} \quad (5.7)$$

Expression (5.6) becomes:

$$R[T] = \frac{T \cdot \exp\left(A - \frac{B}{T}\right)}{1 + \exp\left(C - \frac{D}{T}\right) + \exp\left(E - \frac{F}{T}\right)} \quad (5.8)$$

Expression (5.8) will be used in this form and applied in the case of the tomato leaf miner.

We can't choose *a priori* which is the best development rate function to use, and the strategy followed in this study is the estimation of all the parameters related to the expressions (5.3), (5.4), (5.5), (5.8), along with a cross-check of the R^2 and χ^2 -values to isolate the optimal performing function.

5.3.2 The population density model

Let us now introduce the population density model, the main focus of this article. There are several models widely used to describe a population as it evolves through life stages and across time. The Distributed Delay Model introduced by Manetsch and performed with the addition of a mortality coefficient by Vansickle [15], [16] has been applied in numerous case studies in entomology (e.g. Allen and Byrd, [37], Galeano-Vasco et al., [68], Limonta et al., [22], Wang et al., [71]) and not only (e.g. Sipahi et al., [129], Spolia [38]). This model, is linked to the well-known Von Foerster equation [57], [109], [130] as shown by Plant and Wilson [58], but both present some problem areas in the conceptual description of age. In the case of poikilothermic organisms, development is driven by environmental parameters, which conceptually means that age is a quantity that accumulates over time. The assumption of the cited models is that the growth of the cohort's individuals increases linearly with the time [131]. In other words, age is considered similar to time. The formulation of the model is founded on considerations in the work of Vansickle [83], and further discussed and developed by other authors [11], [24], [27], [110]. A life cycle could be represented in a schematic way, as shown in Fig. 5.2, with a series of chained stages.

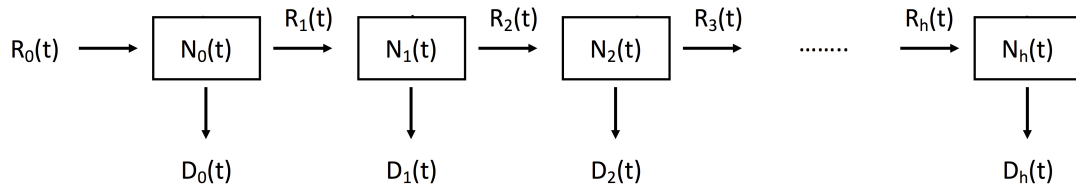


Fig. 5.2: Schematic representation of a life cycle divided in stages. $R_i(t)$ represents the fluxes of individuals that enter in a particular stage, and after the maturation exit and advancement to the next stages. $N_i(t)$ represents the number of individuals in the stage, and $D_i(t)$ is the exit flux from each stage due to mortality.

Each stage has an entering flux of individuals and an exit flux of individuals. After maturation, the individuals move onto the next stage. This flux is ruled by a development rate function which considers, in the most general case, both the interaction between the species and the environment, and the reduction of the rate due to aging. In a practical context, the development rate functions inserted in the model will be one between the expressions (5.3), (5.4), (5.5), (5.8), but, theoretically, it can be extended to a more general case. The reason is that once a population-dynamics model has been formulated, which takes into account age as being driven by a general development rate function, it is possible to focus on the development

of this kind of expression. This improves the reliability of the model and the integrity of the phenomena described. Each stage also contains an exit flux indicating mortality, which is ruled by a mortality rate function with similar aspects to the described development rate function. A state function is introduced to count the number of individuals presents in the life stages. Considering an infinitesimal stage, as represented in Fig. 5.3, it is possible to write a balance equation which considers the temporal variation of the number of individuals $N(t)$ related to: the entering flux $R_{x-\Delta x}(t)$; the exit flux $R_x(t)$ due to maturation; and to the exit flux $D_x(t)$ due to mortality:

$$\frac{d}{dt}N(t) = R_{x-\Delta x}(t) - R_x(t) - D_x(t) \quad (5.9)$$

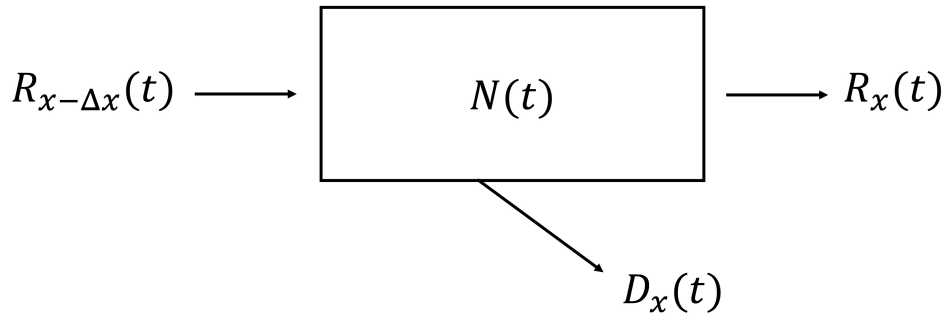


Fig. 5.3: Schematic representation of an infinitesimal stage. $R_{x-\Delta x}(t)$ is the entering flux of individuals, $R_x(t)$ is the exit flux of the matured individuals, $D_x(t)$ is the exit flux due to the mortality and $N(t)$ is the stage function which counts the individuals present in the stage.

Let us introduce the function of two variables which describes the density of a population that evolves across time and through different stages $N(t, x)$. The terms of the equation (5.9) can be reformulated, considering that:

- The number of individuals in the current stage, $N(t)$, corresponds to the following integration:

$$N(t) = \int_{x-\Delta x}^x N(t, x') dx' \quad (5.10)$$

- The entering flux is ruled by a *generalised development rate function*, $G(t, x)$, which acts as a filter for the population density:

$$R_{x-\Delta x}(t) = G(t, x - \Delta x) \cdot N(t, x - \Delta x) \quad (5.11)$$

- The exit flux due to the maturation of individuals is also ruled by the development rate function $G(t, x)$:

$$R_x(t) = G(t, x) \cdot N(t, x) \quad (5.12)$$

- The mortality flux is ruled by a mortality rate function $M(t, x)$:

$$D_x(t) = M(t, x) \cdot N(t, x) \quad (5.13)$$

Substituting the new form (5.10), (5.11), (5.12) and (5.13) in the balance equation (5.9), and dividing both the members for the stage duration Δx :

$$\frac{d}{dt} \frac{1}{\Delta x} \int_{x-\Delta x}^x N(t, x') dx' = \frac{G(t, x - \Delta x) \cdot N(t, x - \Delta x) - G(t, x) \cdot N(t, x)}{\Delta x} - \frac{M(t, x) \cdot N(t, x)}{\Delta x} \quad (5.14)$$

Expression (5.14) is now divided in three terms, and each of them can be rewritten in a different way. The hypothesis expressed on the size of the stage described in Fig. 5.3 allows the mean value theorem for integrals to be applied to the left term:

$$\frac{d}{dt} \frac{1}{\Delta x} \int_{x-\Delta x}^x N(t, x') dx' \approx \frac{d}{dt} N(t, x) \quad (5.15)$$

Concerning the first term on the right of the (5.14), it is possible to approximate the fraction considering the Taylor series expansion of the function $G(t, x) \cdot N(t, x)$, truncating on the first order:

$$\frac{G(t, x - \Delta x) \cdot N(t, x - \Delta x) - G(t, x) \cdot N(t, x)}{\Delta x} \approx -\frac{\partial}{\partial x} [G(t, x) \cdot N(t, x)] \quad (5.16)$$

The last term on the right of the (5.14) represents mortality in the stage. Assuming that into every single stage, mortality is equiprobable, the fraction $\frac{M(t, x)}{\Delta x}$ represents an estimation of the average mortality, $M^i(t, x)$, in line with these considerations:

$$\frac{M(t, x) \cdot N(t, x)}{\Delta x} \approx M^i(t, x) \cdot N(t, x) \quad (5.17)$$

Let us substitute the expressions (5.15), (5.16), (5.17) into the equation (5.14). Considering the derivative of the time variable a partial derivative, has been obtained a “generalized” Von Foerster equation in which the development across all stages is ruled by the development rate function:

$$\frac{\partial}{\partial t} N(t, x) + \frac{\partial}{\partial x} [G(t, x) \cdot N(t, x)] = -M^i(t, x) \cdot N(t, x) \quad (5.18)$$

Von Foerster’s considerations on the initial and boundary conditions [57], [102] can be maintained also for the equation (5.18). Using the notation $\beta(t, x)$ for the fertility rate function, the boundary condition provides the new-borns’ profile on the basis of the density of population:

$$N(t, 0) = \int_0^\infty \beta(t, x') \cdot N(t, x') dx' \quad (5.19)$$

The initial condition, instead, represents the starting population’s profile, in such a way that it indicates the initial division of the individuals through the stages:

$$N(0, x) = N^0(x) \quad (5.20)$$

Expressions (5.18), (5.19) and (5.20), together, represent a well posed problem. One of the problems that obstructs practical application is the daily average temperature profile. In a real context, the daily average temperature changes over time, and there is not a specific mathematical rule to describe the profile over the course of any given year. This means that this is a quantity that should be measured every day in the place of interest, by a meteorological station or a data logger. In order to use the model, it is necessary to have a numerical solution and to use a calculation software programme. A discretization technique usually applied to partial differential equations as (5.18), provides to combine the upwind scheme and the finite difference method [24], [27], [82], [110]. More specifically, once the discrete variables are introduced, the partial derivative on age, $\frac{\partial}{\partial x} [G(t, x) \cdot N(t, x)]$, is approximated backward while $\frac{\partial}{\partial t} N(t, x)$ forward. Continuous variables t and x in equation (5.18) can be represented with a sequence of values t_n, x_n with the condition $t_n < t_{n+1}$ and $x_n < x_{n+1}$ respectively. Accordingly, the discrete variables $i = t_{n+1} - t_n$ for time, and $h = x_n - x_{n-1}$ for age are defined. Furthermore, the method provides also the discretization of development, mortality and fertility rate functions: they are indicated as G_h^i , M_h^i and β_h^i , respectively. As result, the corresponding discrete form of the equation (5.18) is represented by:

$$\begin{cases} N_h^{i+1} = \frac{i}{h} G_h^i N_{h-1}^i + N_h^i \left[1 - 2\frac{i}{h} G_h^i + \frac{i}{h} G_{h-1}^i - i M_h^i \right] \\ N_0^i = \sum_{i=0}^{\infty} \beta_h^i N_h^i h \\ N_h^0 = N_0 \end{cases} \quad (5.21)$$

Before moving on to the next section, a clarification of how temperature is inserted into the model is needed. One of the novelties introduced into the model, is the “generalised” development rate function $G(t, x)$. In a wider scenario, this latter expresses the relationship between species and environment, and the effect of aging on the growth: it can be considered as a “bridge” to insert other physiological models into an age structured model. Because of the paucity of literature on this side, an obliged choice is to use one between the expression (5.3), (5.4), (5.5), (5.8). In other words, the simplification is that the function $G(t, x)$ is dependent only on t , and more specifically it is assumed that:

$$G(t) = R[T(t)] \quad (5.22)$$

Expression (5.22), in its discrete version become $R[T_i]$, where T_i is the daily average temperature measured in poikilothermic organisms’ development environment. Following these as-

sumptions, the expression (5.21) is now reduced to:

$$\begin{cases} N_h^{i+1} = \frac{i}{h} R[T_i] N_{h-1}^i + N_h^i \left[1 - \frac{i}{h} R[T_i] - i M_h^i \right] \\ N_0^i = \sum_{i=0}^{\infty} \beta_h^i N_h^i h \\ N_h^0 = N_0 \end{cases} \quad (5.23)$$

Expression (5.23) is now ready to be inserted into a specialised software, which receives in input the daily average temperature, and the specific parameters for the tomato leaf miner. These last correspond to the development rate function chosen. The output is a representation of the life cycle for all the stages, given in function of the time.

5.3.3 The simulator

In this section, we will introduce a schematic representation of the simulator built to manage all the calculations required by the process of application of the model. The basic idea can be found in a previous work by the authors [49]. To maintain the same logic scheme described in the previous section, the simulator, called “*EntoSim*” is divided in two parts. The first part has been named “*fittasso*” and its task is to follow the users in the phase of parameters estimation. More specifically, after the laboratory session in which the life tables have been estimated, scientists need to estimate the parameters of the development rate functions (5.3), (5.4), (5.5), (5.8) that will be a feature for the species examined. This part of the software includes in a C++ program the ROOT libraries related to the non-linear fits. At the beginning, the user has to report the experimental life tables and their errors in a .txt file. Then the program can be launched. Once the fitting function has been chosen, a starting point for each parameter is required for estimation and potential necessity to restrict them within a particular range. This operation has been added to increase the integrity of the results, together with a calculation of the R^2 and χ^2 values. Another tool inserted is the calculation of the maximum of the rate functions, which corresponds to the thermal optimum, T^* , for the development of the species. The results are provided in graphical (Fig. 5.1) and textual formats, which can be useful for further calculations. In particular, the parameter results with their associated errors, R^2 and χ^2 values, initial values inserted to start the fitting operations and the coordinates $(T^*, R[T^*])$ of the maximum, are recorded. To simply explain the decision process surrounding the identification of the best representative expression to be used for the simulation, a plot with experimental data and fitted function is provided (Fig. 5.4).

The second part of “*EntoSim*” is called “*simulatore*”, and it is the tool responsible for providing the final results of the simulations. In a C++ code and with the support of the ROOT’s libraries, the discrete form of the model (5.23) was inserted, in which the development rate function is dependent only on the external daily average temperature (and hence on time). The

dependence of the rate on aging has not been considered. Before the launch, the user has to insert in specific `.txt` files: a vector containing the daily average temperature measured by a data logger or a meteorological station, the fertility and mortality rates for all the life stages, and finally, the initial profile of the population - once it is possible to be estimated. If this final piece of information is missing, a predefined setting considers a cohort of eggs, that means, all the stages are empty except the first. All the parameters estimated in “*fittasso*” are automatically absorbed by “*simulatore*”.

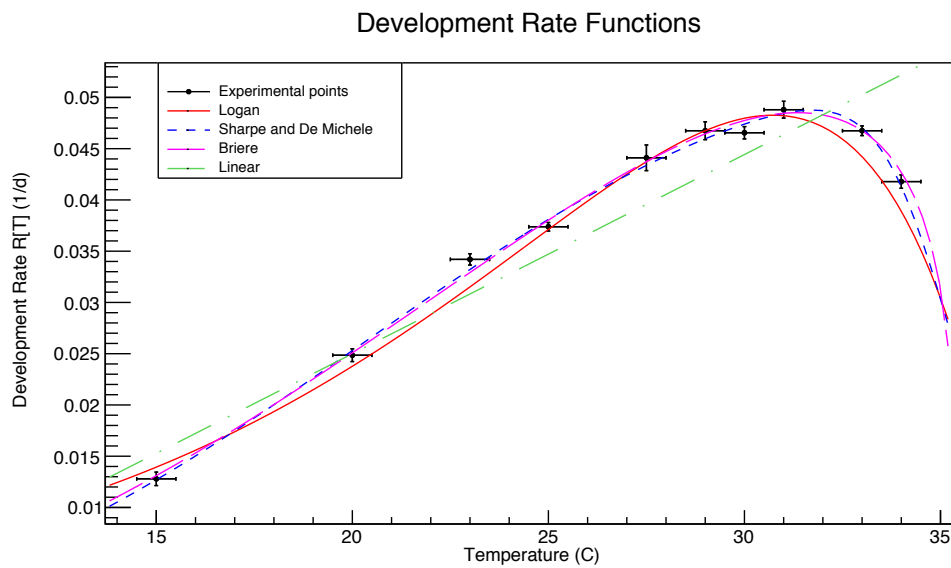


Fig. 5.4: Comparison plot for all of the development rate functions calculated in the case of the tomato leaf miner, using the “*fittasso*” tool contained in the “*EntoSim*” software. Data for calculation are provided from the experimental life tables reported by Ozgokçe et al. [74]. A first visual screening shows how the results from the linear rate function are unreliable for a simulation. Concerning the Briere and Logan rate function, the results have an average reliability, while the most representative development rate function for the data is by Sharpe and De Michele. The results of this first phase, also enforced by a cross check between the R^2 and χ^2 values, suggests the aptness of the Sharpe and De Michele development rate function for simulations in the tomato leaf miner case study.

After the launch, the user is requested to select which kind of development rate function he or she wishes to use: it is possible to use more than one expression between (5.3), (5.4), (5.5), (5.8), but in different ranges that need to be specified before the simulation. If the tool records two or more functions selected for the same range, it cancels the operation with a warning message. The results of this tool are provided in graphical format (see supplementary material) and stored in `.txt` files. These files can be archived or managed with other software programmes and their content considers the temporal evolution of the density of population in each stage of the life cycle.

5.3.4 Experimental design for validation

The validation process involves a double semi-field experiment in a greenhouse; isolated from the external environment to avoid the dispersal of *T. absoluta* adults from the moment flight is possible. For the first experiment, 35 *Lycopersicon esculentum* Mill. cv. “Roma” tomato plants were placed in a greenhouse box that was labelled “box 1”. In the second experiment, 21 tomato plants were placed in a second box, labelled “box 2”. The plants were sown in pots with a diameter of 22 cm, and placed on tables at 90 cm from the ground. Each greenhouse box had a $3 \times 3 \times 3$ m dimension. 70 adults were obtained from Università della Tuscia’s Plant Protection Laboratory, 24 hours old. Then 35 unsexed adults were released into each box containing the tomato plants, which were at an advanced phase of blossoming. Adults oviposited on the leaves, starting new generations. The experiment was conducted between the 5th of September 2018 (the day of the adults’ release), and the 18th of October 2018 (the day of last capture). From the point of initial release, the temperature was measured using a data logger with a sensitivity of $\pm 0.1^\circ\text{C}$. The daily average temperature was calculated with 24 readings for each day (one per hour). The activity of hatched larvae in each box was monitored daily, until the first report of the new adults: from this moment, the daily number of new adults was counted. After the count, each adult was removed manually from the box to prevent future days’ counts from being affected. This operation allowed the creation of the experimental population density to be used in the validation of the model.

5.4 Results and Discussion

The experimental life tables for the tomato leaf miner can be found in the literature of Marchioro et al. [119], and Ozgokçe et al. [74]. Both articles report the experimental data, as well as an estimation of all the parameters related to various development rate functions. In this study, we estimated the parameters of the expressions, mentioned in Materials and Methods section, taking into account data provided by Ozgokçe. All the calculations were worked out using the “fittasso” tool contained in “EntoSim”, and the study by Marchioro and colleagues was used for a first comparison. All the results are reported in Table 5.1. Additionally, the chart reported in Fig. 5.5 shows a graphical representation comparing the starting points of the experiment with the results.

A cross check between the R^2 and χ^2 values reported in Table 5.1, suggests that the best development function to be used in the simulation is the Sharpe and De Michele. As indicated in the work of Ikemoto and Kiritani [45], to evaluate the goodness of this kind of non-linear regression the choice lies between the adjusted coefficient of determination R^2 , the residual sum of squares RSS, or the root mean square error RMSE. Instead of the last two, for this work we have also considered the χ^2 test. Examining only the R^2 value, the preference is the Sharpe and De Michele, but χ^2 test reverses the situation in favor of the Brière. In fact, despite the lower χ^2 value belongs to the Sharpe and De Michele, in this last the number of degrees of freedom is lower. Accordingly, the results are $P = 0.73$ for the Brière against $P = 0.60$ for the Sharpe and De Michele. Notwithstanding, the latter expression lies its formulation on

physical, chemical and biological considerations, as discussed in the Materials and Methods section. This aspect increases the reliability of the simulations’ results, despite the non-linear fit operation for parameters estimation that required more attention. This is due to the strong correlation between the parameters [128]. To reduce as much as possible inconveniences during the parameters estimation process, a series of restriction and controls were inserted into the “fittasso” tool, such as the inserction of starting values and interval limits for each parameter. The use of a non-empirical rate function constitutes one of the main results of this study, since it is not a common practice in simulations of pest populations. However, despite the reliable value for R^2 and χ^2 , the difficulties implied by the use of the Sharpe and De Michele function left their trace in the estimation of the errors associated with the parameters, as shown in Table 1.

Table 5.1: Results from the linear and non-linear parameters estimation process through the use of “fittasso” macro. The parameters’ values for linear, Brière, Logan and Sharpe and De Michele rate functions are reported with their corresponding errors. Additional information is provided by the χ^2 -value, R^2 -value and the number of degrees of freedom (NDF). Data are obtained from the experimental life tables reported in the work of Ozgokçe et al. [74].

Rate function	Parameters ($\pm SE$)	χ^2 -value	R^2 -value	NDF (n)
Linear	$\frac{1}{S} = (1.94 \pm 0.06) \cdot 10^{-3}$ $\frac{T_L}{S} = -0.013 \pm 0.001$	135.96	0.849	8
Brière	$a = (4.0 \pm 0.5) \cdot 10^{-5}$ $T_L = 6 \pm 2$ $T_M = 35.3 \pm 0.9$ $m = 3 \pm 1$	3.59	0.998	6
Logan	$\psi = (3.3 \pm 0.7) \cdot 10^{-3}$ $\rho = 0.14 \pm 0.05$ $T_M = 37.0 \pm 0.1$ $\Delta T = 5 \pm 2$	11.86	0.988	6
Sharpe and De Michele	$A = -42 \pm 32$ $B = -1288 \pm 1120$ $C = -34 \pm 30$ $D = -1166 \pm 1029$ $E = -37 \pm 32$ $F = -1313 \pm 1117$	2.73	0.999	4

An alternative choice could be the use of the Brière function, which also displayed reliable results from the cross check between R^2 and χ^2 , but with lower values for the errors associated with the parameters. More details to motivate the choice of the Sharpe and De Michele are provided as supplementary material, together with an analysis of the simulation outputs varying the expression in input. Notwithstanding, this result is significant because it is not common

to find experimental life tables with a sufficient number of points to apply the Sharpe and De Michele function. A first comparison with Marchioro and colleagues' work confirmed the reliability of the results. Hence it allowed us to proceed to the simulation step and the results are shown in Fig. 5.6 for “box 1”, and Fig. 5.7 for “box 2”.

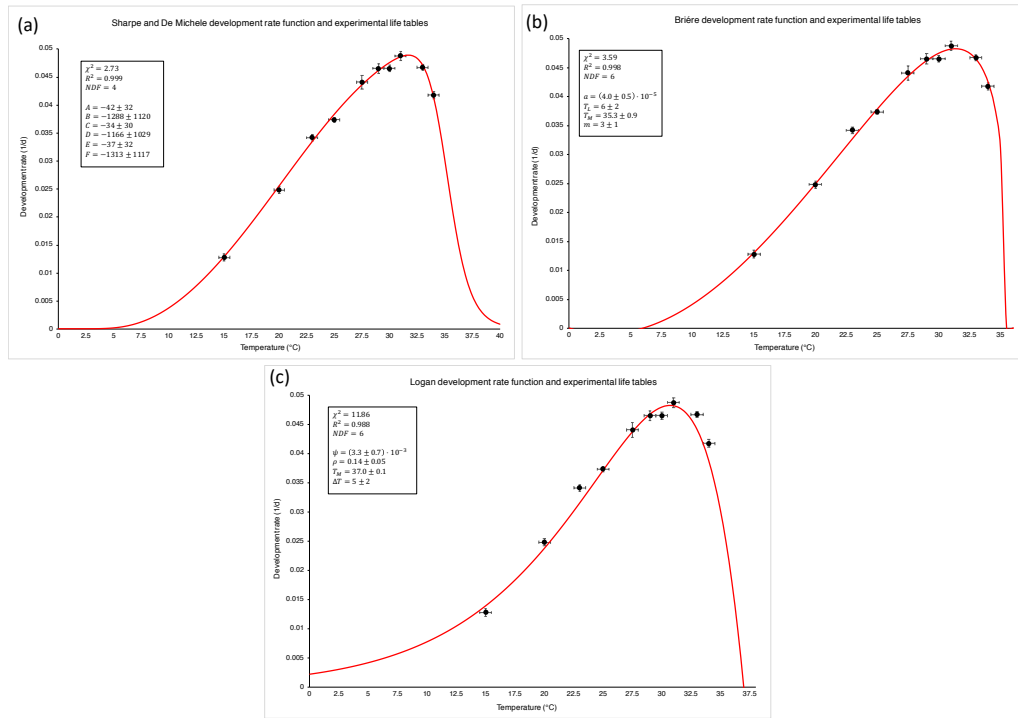


Fig. 5.5: Graphical representation of the best fit development rate functions listed in Table 5.1. Plot (a) represents the Sharpe and De Michele, plot (b) the Briere and plot (c) the Logan.

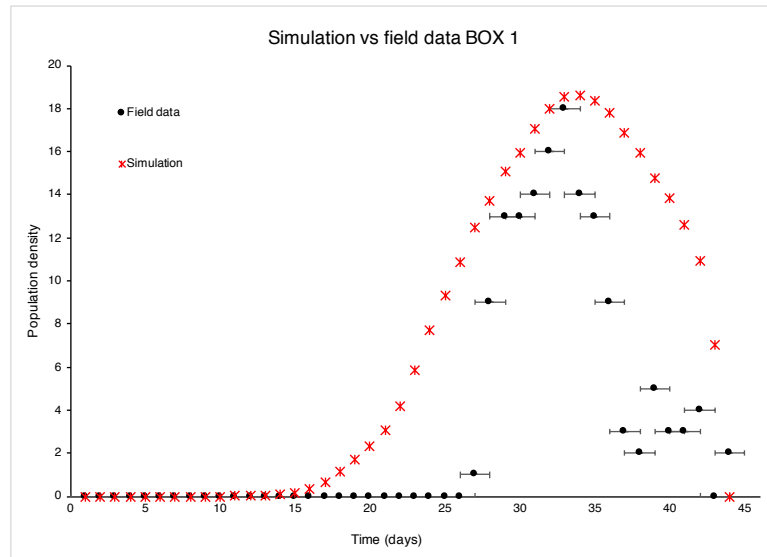


Fig. 5.6: Comparison between simulations and semi-field data collected in the experimental greenhouse. Data are relative to the first validation, “box 1”. The error bars are related to the sampling interval, 1 day.

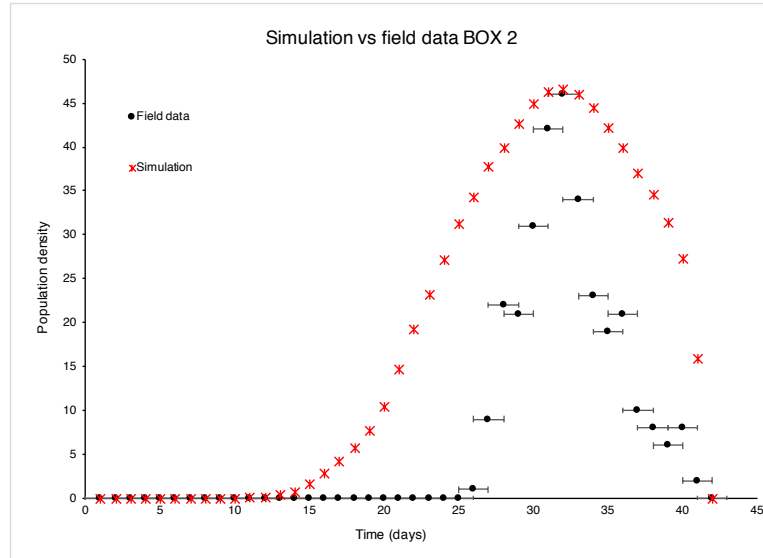


Fig. 5.7: Comparison between simulations and semi-field data collected in the experimental greenhouse. Data are relative to the first validation, “box 2”. The error bars are related to the sampling interval, 1 day.

Along with the simulations’ results, the respective validation data was also reported. The comparison between simulations and semi-field data showed a good accordance in the description of the maximum population dynamic’s peak. The main differences are the width of the two shapes and the total number of individuals in the populations (the integral of the two numerical solutions). Proceeding by order, Fig. 5.6 and Fig. 5.7 both show how the prediction has a tendency to anticipate the increase of the population density, while the right-hand side of the plots the results are in agreement with the experimental population. The disagreement between the predicted and experimental population at the beginning of the cohort’s development could be due to the combined effect of photoperiod and temperature, which both decreased during the period of the validation (September-October). This particular environmental condition could be responsible for a temporal shift in the preimaginal stage development, causing a delay in the adult emerging [132]. The differences between the total number of individuals of the population in “box 1” and “box 2” can be justified by the number of adults released at the beginning of the experimentation (35 for each box). Despite in literature a sex ratio female biased is reported [133], this condition could not be maintained for low numbers of individuals. This means that in “box 1” a higher proportion of males could be released and, accordingly, the progeny was less than “box 2”. Otherwise, the release of a higher number of adults at the beginning of the experimentation might produce a highest population density, introducing stress factors such as the competition for the choice of oviposition sites or for nutrition sources. Furthermore, other environmental parameters (time-dependent) could be inserted into the development rate function. It should also be noted that the age factor has not yet been explored sufficiently by the scientific community, and more efforts should be made to understand how aging slows down growth. The general theory on development rate function hypothesises that time and physiological age are the determiners of the development through the life stages. For this reason,

it is common practice (e.g. in the works of Ainseba et al. [134], Gilioli et al. [4], Limonta et al. [22]) to use the described expressions (5.3), (5.4), (5.5) and (5.8), with no dependence on $G(t, x)$ by the variable x . Other possible reasons for the disagreement at the beginning and at the end of the development could be due to the growth condition of a semi-field environment. Despite the temperature constituting a variable both in open field and semi-field scenarios, in the latter there is an absence of stress factors, such as natural enemies or competitors. This aspect could increase the width of the experimental population, in addition to a corrected balance in the number of plants and the population size. In fact, the number of adults released in each box to start the monitored generation was individuated considering the potential fertility of each female (around 260 eggs as reported by Biondi et al. [115]), and the probability of each egg hatching (and the total life cycle survivor rate). The result is a low density population in contact with a significant number of plants to be infested and this ultimately means low stress levels during the growth period of the new insect generations.

5.5 Conclusions

In conclusion, it is possible to notice that the population dynamics, represented by the first order partial differential equation, effectively represents the development of a cohort of individuals as they evolve through the stages of their life cycle, and in function of the time. Nevertheless, one of the aspects to improve upon in future works with a view to making the model presented more reliable, is the development rate function. In fact, this represents the kernel of the model, because it is responsible for the mathematical description of the interaction between the species and the environment. The development of this argument is not only helpful for making predictions in agricultural and environmental sciences, but also can help us to understand biological mechanisms and describe these mathematically. Mathematical models in biology tend to be simplified and based on empirical assumptions or on experiments, and the lack of basic laws [1] forces model scientists to calibrate and tune models after the validation. This operation has undertaken to introduce other systematic errors derived from calibration experiments, which will allow for future works to analyse biological mechanisms in greater depth. Another interesting aspect could be explored by explicating the derivative on the age variable in the equation (5.18):

$$\frac{\partial}{\partial x} [G(t, x) \cdot N(t, x)] = G(t, x) \cdot \frac{\partial}{\partial x} N(t, x) + N(t, x) \cdot \frac{\partial}{\partial x} G(t, x) \quad (5.24)$$

The explicit expression (5.22), details how the knowledge of dependence between the development rate function and the age provides a term, $N(t, x) \frac{\partial}{\partial x} G(t, x)$, similar to a mortality rate. This aspect could simplify the understanding of mortality rate, which is usually based on coefficients estimated during the construction of life tables. In other words, one of the advantages when shifting from a chronological age to a physiological age - in the most complete form of the development rate function - is that it allows us to divide the contribution of the mortality rate in two parts: features of the species and behaviour as a response to progress through life stages; and external factors.

The advantages of using predictive models in agricultural and forest management merge with the aim to apply more sustainable control methods for pest populations and plant diseases. An application of this type is helpful in several contexts, as discussed in the introduction. Organic farms, agricultural suburbs and protected areas such as national parks can benefit from plans for alternative treatments against natural enemies such as predators, parasitoids and entomopathogenic fungi and bacteria. Furthermore, an increasing problem in pest and disease management is the introduction of harmful alien species, such as *T. absoluta* described in this paper. Prompt circumscription action of the introduced species is crucial for eradication programmes, allowing agronomists, farmers and technicians to plan the best strategy. In other cases, a modelling approach is helpful for territorial programmes in preparation for climate changes: studies about this theme are widespread across current literature [135]–[138], and most of them have been enriched with the incorporation of simulations. Therefore, the comprehension of biological mechanisms and their translation into mathematical language is fundamental if we are to improve the reliability of predictive results.

Chapter 6

A protocol to apply and validate the model: *a priori* and *a posteriori* analysis

*“Professional: one who believes that if
something is worth doing,
its worth doing right”*

- USS Midway, San Diego (CA)

The previous article [31] presented a new physiologically-based model and its application to the case study of *Tuta absoluta*. However, a series of issues related to the generalised Von Foerster equation, and to its application are still not discussed. Hence, in this article [41] two aspects in particular were considered. The first one concerns the formulation of a reliable methodology to select the best representing development rate function to use for simulations. The second aspect concerns an evaluation of the reliability of the simulations in representing field data. These analyses were presented in a new application of the model to the case study of the pepper weevil *Anthonomus eugenii*.

Article details:	
Title:	A modelling approach to describe the <i>Anthonomus eugenii</i> (Coleoptera: Curculionidae) life cycle in plant protection: <i>a priori</i> and <i>a posteriori</i> analysis
Authors:	Luca Rossini, Mario Contarini, Maurizio Severini, Daniele Talano and Stefano Speranza
Corresponding author:	Luca Rossini, Stefano Speranza
Journal:	Florida Entomologist

A modelling approach to describe the *Anthonomus eugenii* (Coleoptera: Curculionidae) life cycle in plant protection: *a priori* and *a posteriori* analysis

6.1 Abstract

Anthonomus eugenii Cano (Coleoptera: Curculionidae) is one of the most severe pests for sweet and hot varieties of pepper (*Capsicum* spp.; *Solanaceae*). The species is distributed widely, principally in Central America, but in 2013 it was detected for the first time in the Lazio region of Italy. Modelling plays a key role in reducing chemical treatments used on *Capsicum* spp., but reliable predictions of pest populations require adjusted tools, as well as intense knowledge of the insect's biology and its typical environment. The main goal of this work is to describe the life cycle of *A. eugenii* with a physiologically based model, which links the population dynamics with the environmental parameters. More specifically, this analysis focuses on the different response of the age-structured model in relation to the development rate function in input. Two methodologies to determine the best representative development rate function suitable for simulations are proposed; the first is “*a priori analysis*”, whereas the second is the “*a posteriori analysis*”. Simulations were compared with semi-field data, collected in a controlled experimental greenhouse where *A. eugenii* developed in varying temperature conditions. Results showed that the model used is adequate to describe *A. eugenii* population dynamics and highlighted how the *a posteriori* analysis can be essential to *i)* analyze the simulation outputs, and *ii)* determine the best representing development rate function, if the *a priori* analysis does not provide this information sufficiently clearly.

6.2 Introduction

The pepper weevil *Anthonomus eugenii* Cano (Coleoptera: Curculionidae) represents an increasing problem for sweet and hot varieties of pepper cultivations (*Capsicum* spp.), not only in Central America [114], its land of origin, but also in Europe where this insect has been discovered recently [121]. The first Italian outbreak (2013) [121] of the pepper weevil was recorded in Fondi and Monte San Biagio, two municipalities in the province of Latina (Lazio region). The short development time of *A. eugenii* in favorable conditions and its destructive feeding caused considerable economic losses among pepper producers, arousing the interest of Integrated Pest Management (IPM) scientists.

Because *A. eugenii* may develop on both wild and cultivated *Solanum* and *Capsicum* (both *Solanaceae*) species, insect control strategies often are difficult to implement successfully, because nightshade-residing populations are able to re-infest cultivated pepper fields the following season [139]. Damages are inflicted primarily by the trophic activities of the larvae and adults. In addition, the oviposition behavior of adults cause abortion and fall of flowers [140] and infested fruits, causing extensive losses. A complete generation usually requires 20 to 30 days [141], though this is strictly dependent on the optimum climatic conditions. In a laboratory, *A.*

eugenii may reach 8 complete life cycles; however, 3 to 5 generations develop in most pepper weevil native areas [142]. Indeed, it is common to find adults in the fields from March until June in the native areas of the pepper weevil, but a certain amount of adults also can be recovered throughout the year, except for the months of December and January [141]. Toapanta et al. [53] conducted a detailed study describing the response of the pepper weevil to external environmental temperatures, which provided life tables [48]. An external temperature of 30°C seems to be optimal for egg to adult development time, reproduction, fecundity, and fertility [53].

A mathematical description of the life cycle of pepper weevil may be of significant importance in planning control strategies to protect pepper crops. In fact, forecasts in agriculture have the potential to provide a general picture of what will happen in the following period. This feature is typical of the Decision Support System (DSS); once simulation results show a trend of an insect pest's development, experts can fix a threshold beyond which a series of consequences have to be activated in order to implement the appropriate control strategy.

The principal goal of this paper is to describe mathematically the life cycle of pepper weevil, applying the innovative approach described in the work of Rossini et al. [31]. In particular, this study highlighted the reliability of the model applied in semi-field contexts such as greenhouses, where pepper is usually cultivated in Italy. In addition, an application protocol is proposed in order to improve future applications of physiologically based models. For this purpose, simulation results are compared with unpublished data from an experimental greenhouse located in Fondi (Latina, Italy), in the epicentre of the area that is typically infested.

6.3 Materials and Methods

6.3.1 Population modelling

The mathematical details and the computational tools applied in this study are described in the works of Rossini et al. [31], [42], [49]. In particular, the model proposed by these authors describes an insect life cycle considering its development over time and through the life stages, and associated with the environmental temperature. The mathematical representation of the model is as follows:

$$\left\{ \begin{array}{l} \frac{\partial}{\partial t} N(t, x) + \frac{\partial}{\partial x} [G(t, x) \cdot N(t, x)] = -M^i(t, x) \cdot N(t, x) \\ N(t, 0) = \int_0^{x_m} \beta(t, x') \cdot N(t, x') dx' \\ N(0, x) = N^0(x) \end{array} \right. \quad (6.1)$$

where x and t indicate physiological age and time, respectively. The expressions $G(t, x)$, $\beta(t, x)$, and $M(t, x)$ represent, in order, the development, fertility, and mortality rate functions usually estimated through laboratory experiments. The physiological age, as described in the

work of Rossini et al. [31], may be associated with the preimaginal stages of the life cycle of *A. eugenii*.

The expression $G(t, x)$ introduces different phenological models, often represented by the development rate functions. In the case of insects, the established practice is to consider temperature as the most significant environmental variable, which drives the development of the individuals through their life stages [45], [52]. Accordingly, only the dependence of $G(x, t)$ on time t is considered. The parameters of the development rate function usually are estimated with the life tables, as described by [42], [49]. To distinguish the generalized form of the development rate function from its simplification, the notation changes as follows:

$$G(t, x) = R[T(t)] \quad (6.2)$$

where $T(t)$ is the external environmental temperature recorded in the field. In this work, the choice was made to test equation (6.1) for the *A. eugenii* life cycle with two particular development rate functions. The first one considered was the Brière rate function [47]:

$$R[T] = aT(T - T_L)(T_M - T)^{\frac{1}{m}} \quad (6.3)$$

where a and m are empirical parameters, $T(t)$ is the daily average temperature, and T_L and T_M are the lower and upper temperature thresholds, respectively, under and beyond which development is no longer possible for *A. eugenii*. The second development rate function to be considered is the Logan rate function [127]:

$$R[T] = \psi \left[\exp(\rho T) - \exp\left(\rho T_M - \frac{T_M - T}{\Delta T}\right) \right] \quad (6.4)$$

where ψ and ρ are empirical parameters, T_M is the upper temperature threshold beyond which development is no longer possible, and ΔT is the interval width between the maximum peak of the function and T_M .

Hence, there is a synergy between two apparently independent models: equation (6.1) manages the population dynamics, whereas equations (6.3) and (6.4) describe the insect's response to the daily average temperature. Even if equations (6.3) and (6.4) are used equally in literature, their mathematical features may provoke different responses in the simulations. In particular, the development rate functions' parameters are estimated with the same data, but the non-linear fit results may underline a preference for one expression instead of the other. In other cases, it can be difficult to highlight which is the most effective development rate function to use to represent the data. Hence, only a comparison between simulations and field data may determine these differences. This aspect will be discussed in the following section.

6.3.2 Data analysis

Once the non-linear fit results have provided the functions' parameters, the choice of the best representative development rate function between equations (6.3) and (6.4) will be determined with a χ^2 -test and by analyzing the adjusted R^2 -value. This procedure, together with com-

putational aspects, was described by Rossini et al. [31]. To estimate the parameters of the expressions (6.3) and (6.4) the ROOT's "*Minuit*" tool [91] was included in an *ad hoc* C++ program, as reported by Rossini et al. [42]. As mentioned previously, it may happen that the χ^2 and R^2 values resulting from the non-linear regression with equations (6.3) and (6.4) are similar. When this condition is verified, an additional operation is required to assess which is the best representative development rate function for the specific application. In a preliminary analysis, if the non-linear fits results have minor or otherwise negligible differences, the user could make a direct choice of which development rate function to use. This statement is supported by the empirical formulation of both Brière and Logan formulas.

On the other hand, an alternative is to run simulations considering both the Brière and Logan expressions as development rate functions as input into equation (6.1). Simulation outputs then will be compared with field data, and the most effective development rate function to use in further applications will be determined a posteriori. The comparison between simulation outputs and field data will be conducted using the methodology described in Rossini et al. [31], [42]. The authors reported that the distance between the simulated and the actual experimental populations would be evaluated, again considering the χ^2 function. As an indicator of the distance, a lower χ^2 -value means a more accurate overlap between simulations and field data. Accordingly, if the best development rate function to use as input in equation (6.1) cannot be selected using the criteria above, it can be helpful to compare simulation outputs with field data. Thereafter, the development rate function that provides a lower χ^2 -value will be applied in further simulations.

6.3.3 Experimental design for validation

The model validation was conducted during the 2014 growing season in an experimental greenhouse located in Fondi (Latina, Italy) at the company GENISTA S.r.l., which specializes in genetic improvement. The greenhouse had a length of 40 m, a width of 10 m, a minimum height of 1.95 m, and a maximum height of 3.5 m at the center of the roof. Pepper plants were transplanted into the greenhouse in August, which was opened to promote *A. eugenii* infestations.

Once the presence of pepper weevils was ascertained inside the experimental field by visual observations, the greenhouse was isolated with an anti-aphid net (Agrintech s.r.l., Eboli, Salerno, Italy) to avoid the further entrance or exit of *A. eugenii* individuals. At this point, the adults enclosed in the greenhouse mated and laid eggs on pepper plants, producing the generation that was monitored to validate and analyze equation (6.1).

The monitoring activity focused on pepper weevil adults. For this purpose, 5 yellow chromotropic sticky traps (Serbios s.r.l., Badia Polesine, Rovigo, Italy) were deployed along the diagonal axis of the field, and checked each week from October to December. The experiment was interrupted at the end of December when the pepper fruit was harvested. The environmental temperature in the greenhouse was recorded from the date the pepper plants were transplanted until the interruption of the experiment. More specifically, a meteorological station placed in the greenhouse provided the daily average temperature based on 24 measurements (1 every hour).

The error associated with the experimental population (dots in Figs. 6.1 and 6.2) is related

to sampling time. Normally, this value is fixed at half the temporal range between the two field surveys, which is ± 3.5 days.

6.4 Results

The life tables provided by Toapanta et al. [53] facilitated the estimation of the development rate function parameters. The best-fit parameters for equations (6.3) and (6.4) are listed in Table 6.1, together with χ^2 and R^2 values. Fig. 6.3 and 6.4 provide a graphical representation of the best-fit functions. Additional information (correlation matrix, variances and covariances matrix and a plot of the adjustment curves with 95% confidence intervals) are reported in the supplementary materials. The criteria for selecting the best development rate function to use for simulations, as described in the “Data Analysis” section, requires a cross check between the χ^2 -test and the R^2 -value, if the χ^2 and R^2 values are not too similar between the two functions. The values listed in Table 6.1 underscore that this criteria cannot be applied, since a χ^2 -test indicates a $P < 0.01$ for both equations (6.3) and (6.4). In addition, the respective coefficients of determination R^2 are similar. On the basis of the χ^2 -test, there is a preference for the Brière equation, but considering the R^2 , the situation is overturned. These first results highlight that both Logan and Brière equations represent the life tables from Toapanta et al. [53] effectively; nevertheless, it remains difficult to choose one of the two equations. From a biological point of view, it could be more interesting to select the Brière equation instead of the one by Logan. The reason for this is that the information about the lower temperature threshold for the development of the species, T_L , is not indicated by the Logan equation.

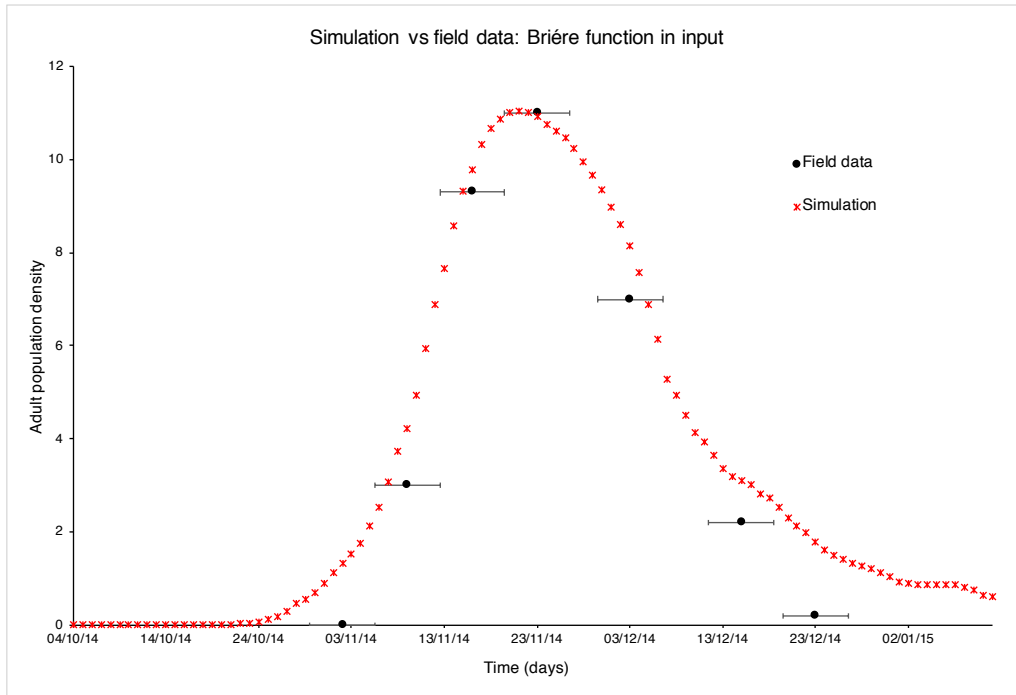


Fig. 6.1: Simulation output from the model (6.1) evaluating the daily average temperature with the Brière development rate function.

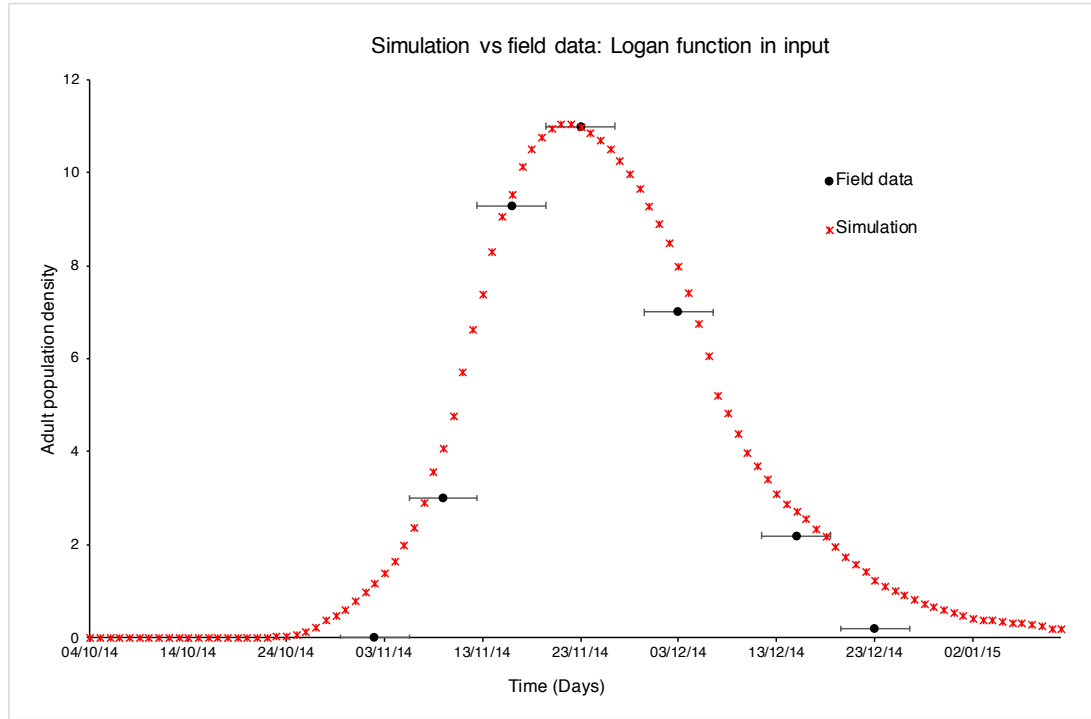


Fig. 6.2: Simulation output from the model (1) evaluating the daily average temperature with the Logan development rate function.

Table 6.1: . The parameter values for Brière and Logan rate functions reported with their standard errors, calculated with the methodology reported by Rossini et al. [31], [42], [49], using data from Toapanta et al. [53]. Additional information includes the χ^2 -value, the one-tailed probability of the χ^2 distribution, R^2 -value, and the number of degrees of freedom (NDF).

Rate function	Parameters ($\pm SE$)	χ^2 -value	χ^2 -probability	R^2 -value	NDF (n)
Brière	$a = (8.2 \pm 0.7) \cdot 10^{-5}$	0.0011	0.9999903	0.985	3
	$T_L = 5.6 \pm 0.2$				
	$T_M = 33.3 \pm 0.9$				
	$m = 4.9 \pm 0.1$				
Logan	$\psi = 0.003 \pm 0.001$	0.0017	0.9999814	0.991	3
	$\rho = 0.11 \pm 0.01$				
	$T_M = 35.0 \pm 0.6$				
	$\Delta T = 3.5 \pm 0.7$				

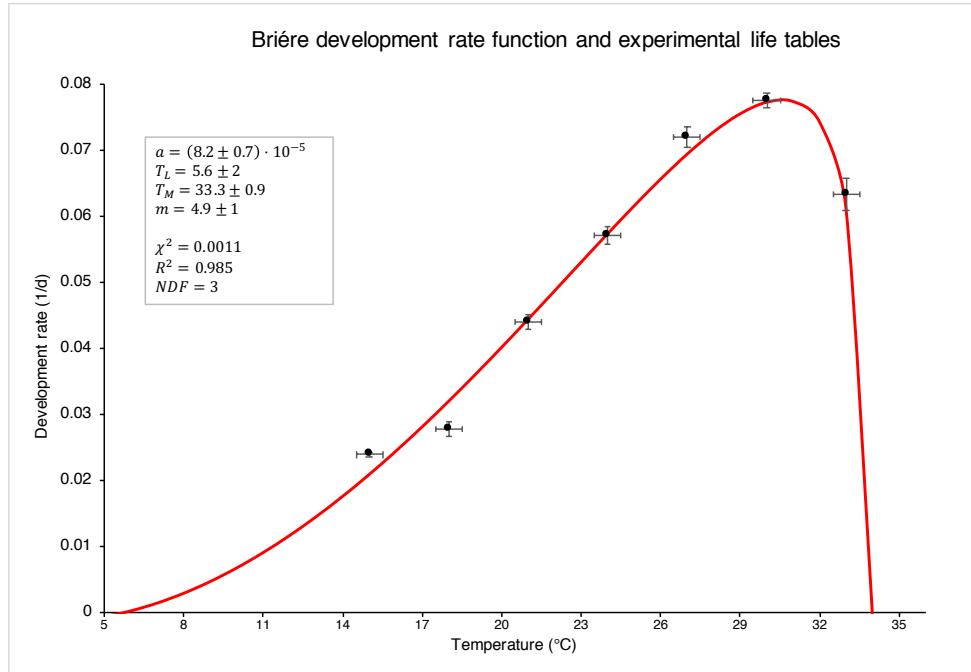


Fig. 6.3: Briere development rate function compared with life tables point from Toapanta et al. [53].

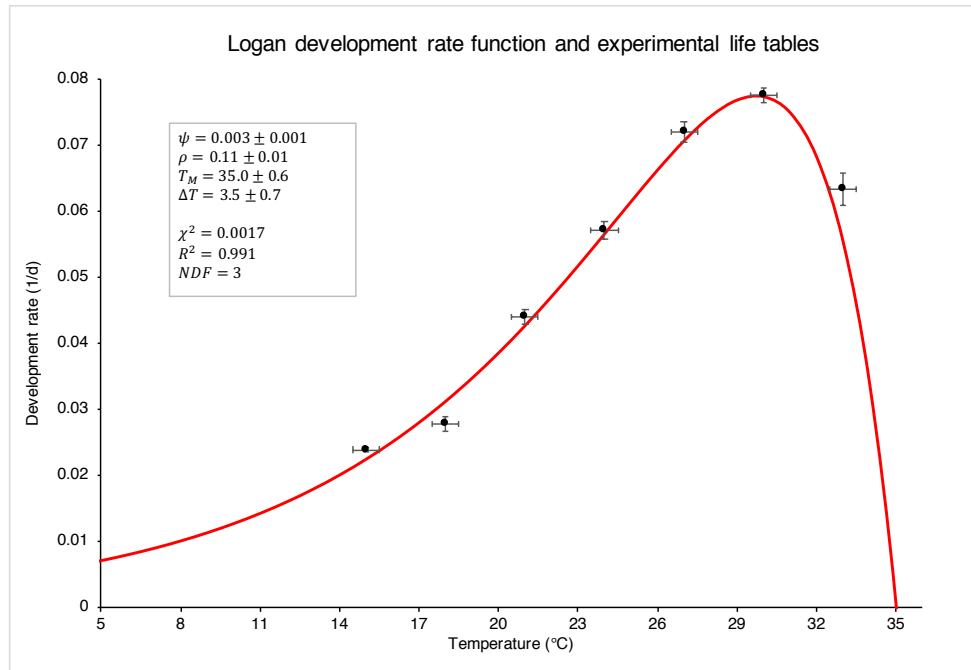


Fig. 6.4: Logan development rate function compared with life tables point from [53].

On the other hand, the aim of this work is to provide an adjusted tool to use as a decision support system; hence, the best representative development rate function will be determined a posteriori, comparing the simulations with field data. An additional and important piece of biological information provided by this first part of the analysis concerns the optimal thermal

requirements for the development of *A. eugenii*. In particular, an optimal temperature of $(31 \pm 1)^\circ\text{C}$ has been calculated for the Brière equation, whereas $(29 \pm 3)^\circ\text{C}$ has been calculated for the Logan equation. Considering the errors associated with the optimal temperatures provided by the development rate function equations, it is possible to determine that these results are, in fact, in accordance with each other. Simulation results are reported in Fig. 6.1 and Fig 6.3. In particular, Fig. 6.1 reports the simulation output compared with field data considering the Brière equation as input. In this case, the projections indicate a maximum peak of the population on 22nd November, whereas field data confirmed that the maximum occurred on 23rd November. The same situation was reported for the Logan equation, as shown in Fig. 6.4. The principal differences between the two simulations are concentrated in the right-hand side of the plot, on the decreasing tails of the graph. Table 6.2 reports the χ^2 and R^2 values: whereas the first value indicates the distance between simulations and field data, the second value is helpful for understanding the extent to which the simulations represent the actual greenhouse population. On the basis of the values in Table 6.2, it is possible to assess that simulations are more reliable if, in equation (6.1), the response of *A. eugenii* to the daily average temperature is evaluated by the Logan equation.

Table 6.2: . Numerical evaluation of the differences between the simulation outputs and the field data. In this case, χ^2 indicates the distance between simulations and field data, whereas R^2 indicates how much the simulation represents field data.

Development rate function in input into the model	χ^2 -value	R^2 -value
Briere	1.8231	0.9604
Logan	1.1135	0.9797

6.5 Discussion

The results obtained highlight that the Logan equation (inserted into equation 6.1) is more appropriate to describe *A. eugenii* populations. This is not true for all insect species, and it depends on several aspects. For example, the non-linear fit results with equations (6.3) and (6.4) are strongly dependent on the number of experimental points in the life tables. Accordingly, enriching the life tables with additional points can increase the accuracy of the first part of the analysis. Besides demonstrating and analyzing these results, this work aims to provide an operating protocol to apply physiologically based models, which requires a combination of a population dynamics model and a phenological model. Specifically for this work, population dynamics is managed by equation (6.1), while the phenology (in this case, the relationship between insects and environmental temperature) is managed by the development rate function expressions. An additional result, as demonstrated above, is the utility of the *a posteriori* analysis. It has, in fact, been shown that when it is impossible to decide prior to the simulation

the most appropriate development rate function to use as input in the population dynamics model, the *a posteriori* analysis is the only viable means of reaching a decision in this respect. As already anticipated, both Brière and Logan equations are formulated on an empirical basis; there are no reasons to prefer one instead of the other because they represent the life tables in a similar way. On the other hand, the *a posteriori* analysis showed that even if they represent the life tables in a similar way, the simulation outputs can report more discernible differences.

6.6 Conclusions

The considerations resulting from this study are helpful for model scientists who aim to provide tools for direct applications in integrated pest management. The use of mathematical models as a decision support system is growing, and its application has been found in several case studies [143], most of which are sponsored by government programs [144]. For instance, there are models which consider not only the potential abundance of pests in the field, but also include population density in relation to certain economic thresholds [145]. Despite the existence of various proposals for the mathematical representation of population dynamics [25], [31], [134], the reliability of the simulations is strongly dependent on the function which relates the species with its environmental parameters. Accordingly, it is not the best method to choose *a priori* a specific development rate function. Instead, the development rate functions should be evaluated twice: *i) a priori* to exclude wide differences in representing life tables data, and *ii) a posteriori* to evaluate potential differences between different functions in input.

Chapter 7

A revised version of *EntoSim*: new models and new functions

“*Cento misure
e n’tajo solo*”

- *Nonno Ferruccio*

Mathematical theory and software development are the two main goals of the thesis. Hence, each mathematical novelty introduced in the previous articles (i.e. the generalised Von Foester equation and the *a priori* and *a posteriori* analysis) needs a software to work out calculations. It has been demonstrated [49] that the combination of C++ programming language and ROOT’s software libraries are a suitable tool to develop *ad hoc* programs for models applications in entomology. In this article [42] the improvements of *EntoSim* were presented, as well as an application to the case study of the European grapevine moth *Lobesia botrana*. Therefore, this work represents not only a software description, but it is also an additional validation for the generalised Von Foerster equation.

Article details:	
Title:	<i>EntoSim</i> , a ROOT-based simulator to forecast insects’ life cycle: description and application in the case of <i>Lobesia botrana</i>
Authors:	Luca Rossini, Maurizio Severini, Mario Contarini and Stefano Speranza
Corresponding author:	Luca Rossini
Journal:	Crop Protection

EntoSim*, a ROOT-based simulator to forecast insects' life cycle: description and application in the case of *Lobesia botrana

7.1 Abstract

In the context of plant protection strategies against insect pests, the role of modelling is arousing interest among entomologists and environmental scientists. This work aims to provide an application of a computational tool, which includes a modified version of Von Foerster's equation to perform simulations for crop protection. More specifically, this software, named *EntoSim*, follows scientists and final users in each step of a simulation process. For this purpose, the program is divided into two blocks: the first contains a “laboratory phase”, in which scientists rear cohorts of insects at different constant temperatures to obtain life tables. The second block contains a simulation process and validation using field data. The developed tool was applied to the grapevine moth *Lobesia botrana* (Denis & Schiffermüller) alongside a 3-year validation process (2014, 2015, 2016 seasons), using data related to a monitoring activity in a vineyard located in the Castelli Romani area (Lazio, Italy). Results are a series of simulations of *L. botrana* life cycle, using daily average temperature as main driving variable for insect development. Comparison between simulations and field data showed a positive accordance, particularly in the 2016 season, confirming the reliability of both population dynamics model and computational tool.

7.2 Introduction

Grape producers worldwide are increasingly focused on high quality deliverables that are produced in an environmentally-friendly way [146]–[148], in order to satisfy consumers' requests [149]. Even if conspicuous efforts to reduce the rate of chemical inputs such as fertilizers and pesticides have been made [8], a significant number of agrochemicals are often required to control weeds, plant diseases and prevent infestations from phytophagous agents, raising environmental concerns and production costs [150]. In this context, crop pest forecasts may play a fundamental role in modern agriculture providing a helpful tool as a decision support system [143] in order to plan optimized control strategies. Moreover, their common feature is to use past and present climatic data to outline a future risk scenario.

Several mathematical models have been proposed in agriculture and ecological sciences in general, together with software tools, for decision support. Software tools represent the final “ready-to-use” tool, which helps farmers and technicians in solving various issues and, at the same time, can help researchers to understand the dynamics of the phenomena [151]. Specialised software can be used in different phases of the agricultural cycles. Han et al. [9], for instance, developed a software for decision making in crop production in an increasing temperature scenario. Similar examples are provided in other works [152]–[154] in which the importance of technological tools in water and soil resource management is discussed. In plant protection, decision support systems have been developed to evaluate risks during the growing season, as reported

by Bange et al. [155] for cotton systems in Australia. In particular, they describe a tool used to simulate populations of *Helicoverpa* spp. (Lepidoptera: Noctuidae) on the basis of the weather conditions. Decision support systems have also been developed for stored product pests [156].

The possibility of building a forecasting tool to describe insects' life cycle has been discussed by Rossini et al. [49]: they combined C++ programming language codes with libraries from the ROOT software [79], [126]. The basic idea was to provide a methodology to use an existing open source software, originally developed for high energy physics, in agriculture and ecology. According to Rossini et al. [49], C++ was suggested because of its diffusion, notoriety and compatibility with old and new operating systems, while ROOT is a modular scientific software which provides several mathematical, statistical and graphical tools to manage experimental data. This combination led to a series of macros which composed "DSimulator", divided into two main blocks: one related to laboratory data analysis and one to simulations. Laboratory data analysis involves the estimation of the parameters of development rate functions usually employed to describe life tables [44], [45]. The latter are an important tool to describe the response of ectotherms to variations in temperature [48].

Besides the application, a generalised version of the Von Foerster's equation (GVFE) [31], [57], [83], [109] is used here instead of the Distributed Delay Model (DDM) [15], [16]. The advantage of using the GVFE is that it can describe an insect population that evolves over time and different life stages. In addition, problems derived by the need to calibrate the model using field data, as happens for Manetsch's DDM [35], are avoided. On the basis of these considerations, this work aims *i)* to report some improvement on "DSimulator", which is now renamed "*EntoSim*", and *ii)* to provide an application and a three-year validation with the European grapevine moth *Lobesia botrana* (Denis & Schiffermüller) (Lepidoptera: Tortricidae), one of the most harmful insect pests for the European vineyards [28], [157]. With regard to the choice of *L. botrana* as a case study, insecticides are frequently used to reduce damage [54]. As alternative control methods to the use of insecticides, some farmers apply different control strategies based on biological methods [158], [159]. Such a process requires farmers to know exactly when the population of insects is in its most susceptible stage, in relation to the control strategy planned [160]. In this regard, the setup of a specialised software to forecast the grapevine moth's life cycle is of crucial importance.

7.3 Materials and Methods

7.3.1 Parameters estimation

The work of Rossini et al. [49] produced "*fittasso*", a computational tool that is helpful for evaluating the development rate function parameters. In this section, improvements and changes to the original version of "*fittasso*" will be described, while the expressions involved are listed in the supplementary material.

The first improvement of "*fittasso*" concerns the possibility of inserting directly in a `.txt` file the list of the experimental temperatures T_i , the corresponding mean development times, $D[T_i]$, and their related errors σ_{T_i} , $\sigma_{D[T_i]}$. These data are automatically converted in rates $R[T_i]$

applying the expression

$$R[T_i] = \frac{1}{D[T_i]} \quad (7.1)$$

Concerning $\sigma_{D[T_i]}$, its conversion is possible using the propagation of uncertainty formula. In its general form, if x_i independent variables have an associated standard deviation σ_{x_i} , and they are connected by a function $f(x_1, x_2, \dots, x_i)$, the standard deviation of the latter, $\sigma_{f(x_1, x_2, \dots, x_i)}$, is given by:

$$\sigma_{f(x_1, x_2, \dots, x_i)} = \sqrt{\sum_{n=1}^i \left(\frac{\partial}{\partial x_n} \right)^2 \cdot \sigma_{x_n}^2} \quad (7.2)$$

The application of formula (7.2) to the expression (7.1) led to the following result:

$$\begin{aligned} \sigma_{R[T_i]} &= \sqrt{\left[\frac{\partial}{\partial D[T_i]} \left(\frac{1}{D[T_i]} \right) \right]^2} = \\ &= \frac{1}{(D[T_i])^2} \cdot \sigma_{D[T_i]} \end{aligned} \quad (7.3)$$

where $\sigma_{R[T_i]}$ is the uncertainty of the experimental development rates. The new list of data $(T_i, R[T_i])$ and their respective uncertainty $(\sigma_{T_i}, \sigma_{R[T_i]})$ is absorbed from a second tool included in “*fittasso*”, with the task of estimating the parameters of the development rate functions using a non-linear fitting operation.

At this point, the user can choose to evaluate a development rate function between those listed in the supplementary material. The second improvement of “*fittasso*” aims to mitigate the low number of experimental points. For this purpose, it is possible to limit each parameter of the selected expression within a specific range, enhancing the ability to obtain reliable results. This operation is simplified because there is a physical or biological explanation for most of them. In addition, this improvement merges with the already existing possibility to choose a series of initial values from which the operation of non-linear regression starts.

C++ code manages the operation of conversion from times to rates and the .txt files reading/writing processes, while non-linear regression, parameter limitation and insertion of initial values derives from ROOT. The reason for this division is the reliability of the algorithms: ROOT’s libraries contain “*Minuit*”, a tool which provides the best fit parameters minimizing the χ^2 function [91]. Other advantages include the presentation of the fitting results; again, in ROOT there is the possibility to display the results in a plot, including in the legend the number of degrees of freedom, the χ^2 -value, the parameters’ values and their respective errors. These results are then included in a second part of the code in order to automatically calculate other helpful information to interpret results: the R^2 -value, the adjusted R^2 -value and the coordinates of the maximum peak with their corresponding errors.

The third improvement of “*fittasso*” concerns the analytical expressions (instead of the previous numerical method) to calculate the maximum peak of Brière and Logan rate functions

(see supplementary material) and their errors. In the case of Brière rate function, placing the first derivative at zero $ddTR[T] = 0$ the formula for the optimum temperature, corresponding to the T_B^* coordinate of maximum peak is [47]:

$$T_B^* = \frac{2mT_M + T_L(m+1) + \sqrt{4m^2T_M^2 + T_L^2(m+1)^2 - 4mT_MT_L}}{4m+2} \quad (7.4)$$

$R[T_B^*]$ value is obtained by evaluating the Brière development rate function in T_B^* . During the fitting process, ROOT's libraries provide the errors associated to the parameters estimated. Accordingly, applying the propagation of uncertainty formula (7.2) to expression (7.4) and to Brière development rate function, it is possible to obtain the associated errors to the maximum peak's coordinates $(T_B^*, R[T_B^*])$. With the same process, the resulting expression for the T_L^* coordinate of maximum peak for Logan rate function is:

$$T_L^* = T_M - \frac{\Delta T}{\rho \Delta T - 1} [\ln(\rho) + \ln(\Delta T)] \quad (7.5)$$

$R[T_L^*]$ coordinate is provided, again, reading Logan expression in T_L^* value. The propagation of uncertainty formula (7.2) is used also in this case to calculate the associated errors. These results are fundamental in understanding the behaviour of the species.

As in its original version, the user can read results in two specific files: the first is a text file that reports all numerical values, while the second is a double plot that reports experimental data, and the fitted curve with experimental data.

7.3.2 Population modelling

The density of population and its development over time and age can be described by a model which includes the development rate functions. Originally, *EntoSim* involved Manetsch's DDM [15], [16], but its application required the additional step of calibration [35]. As a result, the previous model has been replaced with a generalized version of the Von Foerster's equation, which includes the development rate functions. A detailed description of the model can be found in the work of Rossini et al. [31]. Let us introduce the concept of "physiological age" x , which, in the case of insects, is used to refer to a variable, which clearly identifies the life stages. In the most general case, the development rate function depends on two variables, time t and physiological age x , and considers the effects of both environmental factors and aging on growth. With this precondition, the Von Foerster's equation assumes a different form [83], and it is included in the "*simulator*" tool:

$$\begin{cases} \frac{\partial}{\partial t} N(t, x) + \frac{\partial}{\partial x} [G(t, x) \cdot N(t, x)] = -M^i(t, x) \cdot N(t, x) \\ N(t, 0) = \int_0^{x_m} \beta(t, x') \cdot N(t, x') dx' \\ N(0, x) = N^0(x) \end{cases} \quad (7.6)$$

In the model (7.6), $N(t, x)$ is the density of population, $G(t, x)$ is the “generalized” development rate function, $M(t, x)$ is the “generalized” mortality rate function, $\beta(t, x)$ is the “generalized” birth rate and $N^0(x)$ is the profile of the initial population. Furthermore, the boundary condition $N(t, 0)$ expresses the new-borns’ trend. To adapt equation (7.6) in the case study presented in this work, a necessary simplification led us to consider $G(t, x) = R[T(t)]$: in other words, the aging effect on growth is not considered. Moreover, temperature T in the development rate function depends on time and is measured with a meteorological station. Accordingly, the only way to implement equation (7.6) in a simulator with variable temperature is through an algorithm for numerical solutions.

A combined use of the upwind scheme and the finite difference method [31] allows one to obtain a discrete form of the problem (7.6) that is ready to be inserted into C++ software:

$$\begin{cases} N_h^{i+1} = N_h^i - \frac{i}{h} R[T_i] (N_h^i - N_{h-1}^i) - i N_h^i M_h^i \\ N_0^i = \sum_{i=0}^{x_m} \beta_h^i N_h^i h \\ N_h^0 = N_0 \end{cases} \quad (7.7)$$

During the discretization process, the variables t and x are substituted by their respective discrete form i and h to make a conceptual distinction between continuous and discrete variables. The model (7.7) is the kernel of “*simulatore*”, the second part of the software improved in this work.

The parameters resulting from “*fittasso*”, related to the development rate functions, are absorbed as input in this case, together with specific text files which include the daily average temperature, the specific fertility and mortality rates for each life stage [49].

In order to improve the simulations, the possibility to insert data from previous field monitoring was added. If the latter information is missing or not required, a default option assumes that the adults of the previous generation concentrate oviposition on day zero, namely the first day of the simulation.

After the launch of “*simulatore*”, the user must select the development rate function and the temperature range where the latter is evaluated. At this point, the simulation process automatically starts, employing equation (7.7). The program reads the daily average temperature from its specific text file and then each value is processed and assigned to its specific development rate function by a filter.

Subsequently, “*TFormula*” tool, included in the ROOT’s libraries, evaluates the corresponding rate $R[T_i]$. It is then inserted into the system (7.7), providing the corresponding daily density of population. Since the latter is a function of two variables, results are provided in different graphical forms, with the aim of relaying the widest scenario of the simulated situation. The first plot is a 3D-surface, in which independent variables are time (expressed in days) and physiological age. The dependent variable is the simulated density of population $N(t, x)$.

A second graphical result provides a summary of the temporal trends at fixed stage: in a single plot all the population densities corresponding to each life stage are superimposed.

7.4 Data collection and comparison between simulations and monitoring

Surveys on *L. botrana* populations were conducted in 2014, 2015 and 2016 in the same experimental vineyard (1 hectare) located in Castelli Romani area (Lazio Region, Central Italy). European grapevine moth presence was monitored using BIOGARD (BDT) pheromones Delta traps (S.r.l. Biogard Division, 20834 Nova Milanese (MB), Italy) with Selure-Lb dispenser (Shin-Etsu Chemical Co., Ltd., Tokyo). Three traps were deployed every year of experimentation from the end of April until the end of September and were checked weekly. Traps were hung in the plant canopy, along the diagonal of the vineyard. The number of traps used for the field survey was unaltered during the years, and also proportioned to the area involved in the survey [159] following the manufacturer instructions. Pheromone lures were changed every 2 weeks. According to the sampling time, it is possible to assume an uncertainty of ± 3.5 days, half of the 7 days between two inspections.

Regarding the measurement of daily average temperatures, a meteorological station from Vaisala company (Helsinki, Finland) was positioned in the aforementioned field. More specifically, a model AWS310 recorded one temperature measurement every 30 minutes, with a single daily value averaged from 48 samples.

The comparison between simulations and field data was conducted with a new tool included in “*fittasso*”. Mathematically, the nearness of two series of data can be evaluated using the definition of χ^2 :

$$\chi^2 = \sum_{n=1}^{n=N} \frac{(M_n - S_n)^2}{S_n} \quad (7.8)$$

The equation (7.8) compares N pairs of data points corresponding to the same time t (expressed in days), related to the simulations, S_n , and to the monitoring data, M_n . To support the expression (7.8), it is possible to also consider the R^2 -value, which indicates the extent to which the simulation is reliable in representing field data.

This operation is managed by C++ code which absorbs the values from two specific .txt files (one for simulation and one for validation). Thereafter, it prints χ^2 and R^2 -values on a computer console and on a dedicated .txt file.

7.5 Results

7.5.1 Parameter estimation from life tables

The grapevine moth life tables are available from the work of Moshtaghi Maleki et al. [55]. These data were inserted as input in the “*fittasso*” macro to obtain the parameters related to the linear, Brière and Logan development rate functions. Results of this first phase of the work are listed in Table 7.1, together with their standard errors, χ^2 and R^2 values. In addition, a visual support is provided in Fig. 7.1, where an overlap of the best fit development rate functions

is represented.

Table 7.1: Results from the linear and non-linear parameter estimation process through the use of “fittasso” macro. The parameters’ values for linear, Brière and Logan rate functions are reported with their corresponding standard errors. χ^2 -values, R^2 -values and the number of degrees of freedom (NDF) are also provided.

Rate function	Parameters ($\pm SE$)	χ^2 -value	R^2 -value	NDF (n)
Linear	$\frac{1}{S} = (1.324 \pm 0.007) \cdot 10^{-3}$ $\frac{T_L}{S} = -0.0126 \pm 0.0002$	0.0064	0.694	4
Brière	$a = (3.4 \pm 0.2) \cdot 10^{-5}$ $T_L = 6 \pm 1$ $T_M = 32.6 \pm 0.5$ $m = 3 \pm 1$	0.0036	0.867	2
Logan	$\psi = (7.0 \pm 0.5) \cdot 10^{-4}$ $\rho = 0.18 \pm 0.04$ $T_M = 34.1 \pm 0.8$ $\Delta T = 4 \pm 2$	0.0019	0.925	2

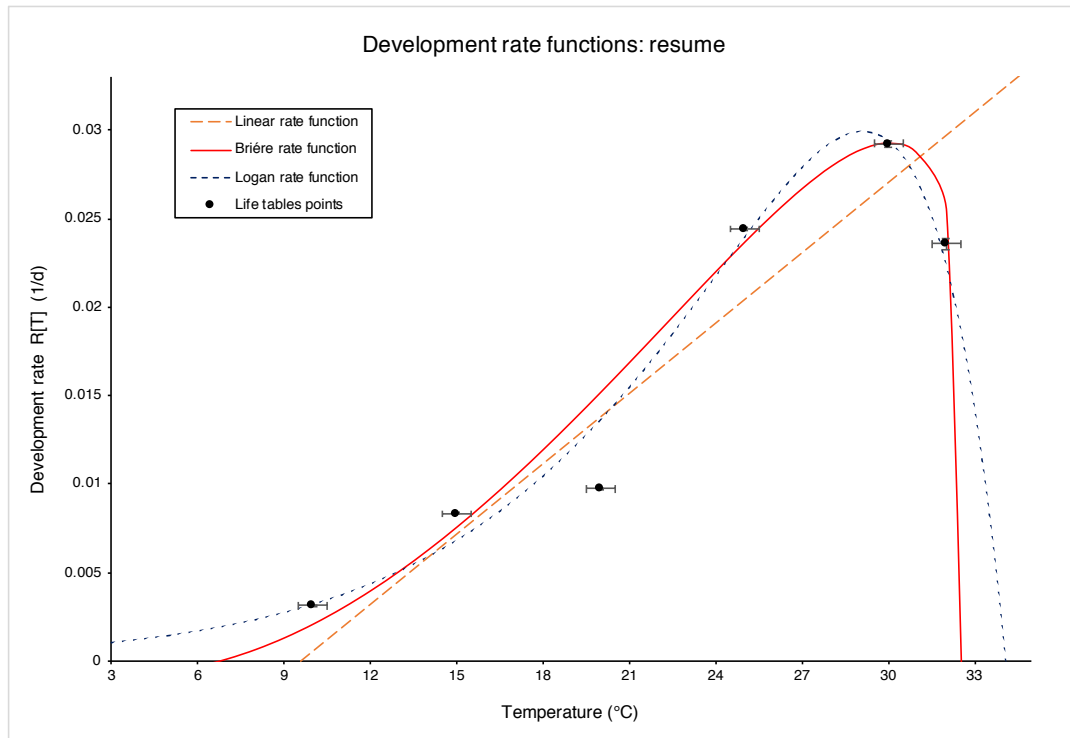


Fig. 7.1: Development rate functions describing the life tables of *Lobesia botrana*, estimated with “fittasso” macro. Linear, Brière and Logan functions are also compared with the experimental points. Moshtaghi Maleki et al. [55] provided the data for the calculation.

Examination of the χ^2 and R^2 values showed that the linear rate function was inappropriate for the simulation. Both Logan and Brière had a better fit. Concerning the optimal temperature for the development of the grapevine moth, T^* , a value of $(31 \pm 1)^\circ\text{C}$ was calculated with formula (7.4) for the Brière expression, while $(29 \pm 1)^\circ\text{C}$ was calculated with formula (7.5) for the Logan expression. Logan and Brière functions had 4 intersection points (Fig. 7.1). This led to a division of the temperature axis in 5 potential subranges in order to discern which was the best for representing development rate function for each of them (see supplementary material). The choice for the simulation was Brière for the ranges $(6.7 - 13]^\circ\text{C}$ and $[25 - 30)^\circ\text{C}$ and Logan for $(13 - 25)^\circ\text{C}$ and $[30 - 34)^\circ\text{C}$.

7.5.2 Simulations and validations

After the individuation of the specific development rate functions for each temperature range, the analysis provided a simulation of the second and third generation of *L. botrana* for the three years of field survey. For each year, the simulation inputs were the daily average temperatures provided by the meteorological station, and the first generation data provided by monitoring of the adult population.

Proceeding in order, Fig. 7.2 represents the results for the 2014 season. It shows a correspondence between the simulated population and the experimental points. Both simulated peaks were close to the monitoring data: for the second generation, model indicated a peak on 15th June, in accordance with field data on 16th June. A similar situation regards the third generation: forecasts indicate a maximum on 18th August, while monitoring reported this value on 19th August.

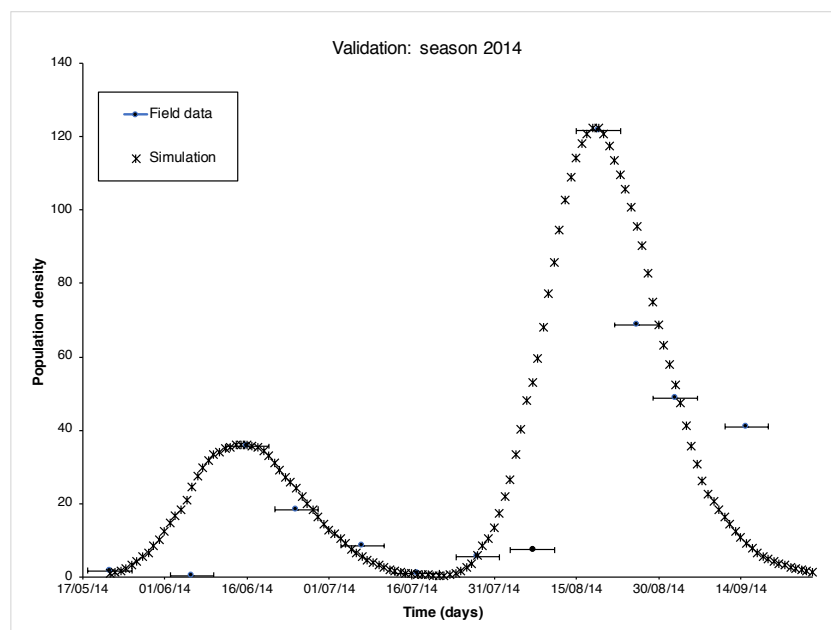


Fig. 7.2: 2014 Season. Comparison between simulated data (cross) and field data (circle) in the case of the second generation (left peak) and third generation (right peak).

Results for the 2015 season are shown in Fig. 7.3. Despite the lack of accordance between experimental points and simulation in the second generation, as shown towards the left-hand side of the time axis, the simulation predicted the peak of the real population reported on 16th June. Forecasts for the third generation report the highest number on 15th August, while the field survey indicated 19th August.

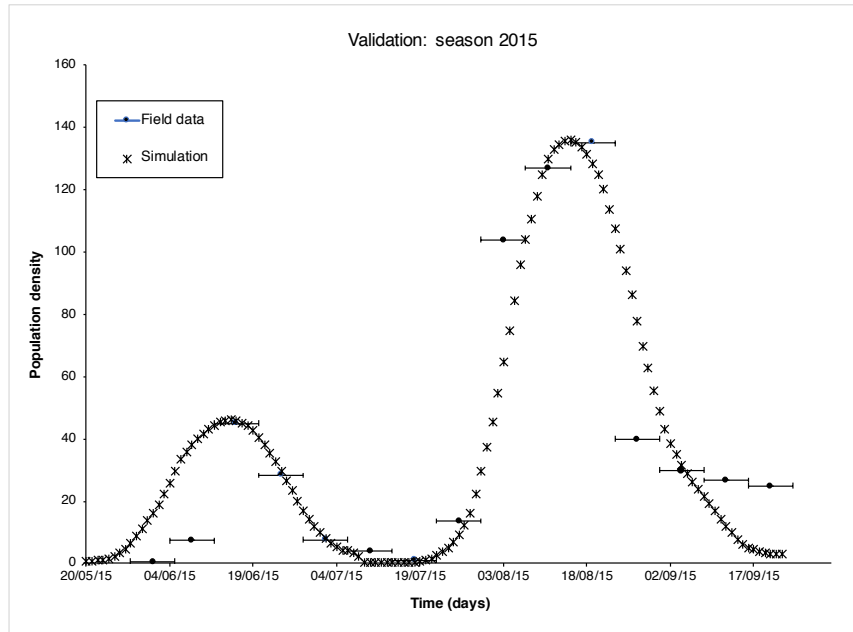


Fig. 7.3: 2015 Season. Comparison between simulated data (cross) and field data (circle) in the case of the second generation (left peak) and third generation (right peak).

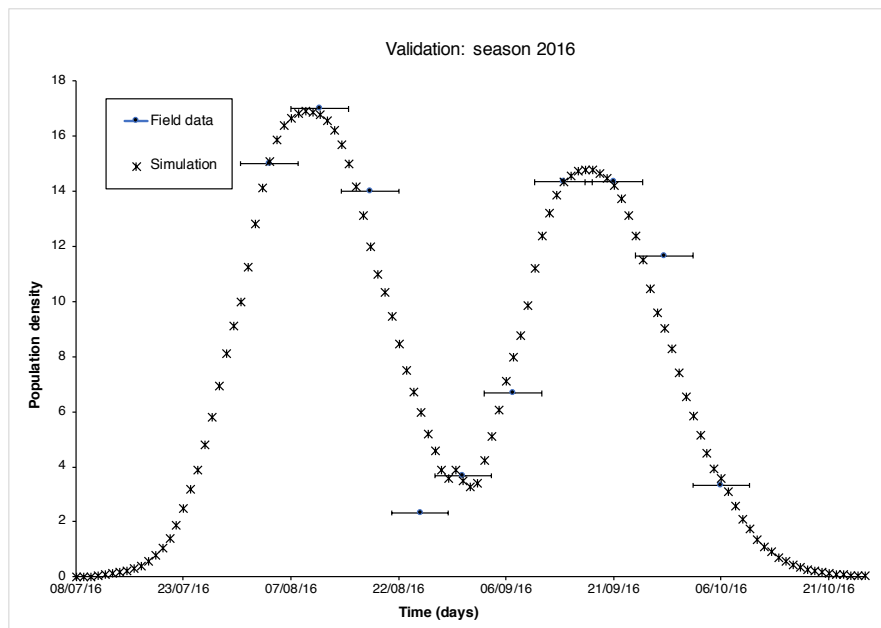


Fig. 7.4: 2016 Season. Comparison between simulated data (cross) and field data (circle) in the case of the second generation (left peak) and third generation (right peak).

Fig. 7.4 reports the results for the last year of the survey, 2016. In this case, the second generation recorded a maximum peak on 11th August, while simulation indicates the date of 10th August. For the third generation the situation was different, since it was not possible to individuate only one peak based on the field data: the highest values were recorded on the 14th and 21st September. Nevertheless, simulation indicated that the maximum peak of the population was on 18th September. The accuracy of the three simulations was also revealed using expression (7.8), which provided the respective χ^2 -values. R^2 value is also reported in order to underline the capability of the simulations to represent field data. These last results are listed in Table 7.2.

Table 7.2: : χ^2 and R^2 values which indicate the accordance between simulations and field data for each year of field survey.

Year of survey	χ^2 -value	R^2 -value
2014	162.10	0.736
2015	234.58	0.865
2016	3.58	0.932

7.6 Discussion and conclusions

The life tables analysed in the first part of the results provided satisfying values for both Brière and Logan development rate functions. Despite χ^2 and R^2 values listed in Table 7.1 highlighted small differences between the best fit development rate functions, the Logan expression resulted as the most appropriate for simulations. Nevertheless, one of the critical issues with Logan formula lies in the lack of a parameter which clearly indicates a lower temperature threshold for the development of the species. This aspect is usually not considered in the growing season, because of the daily average temperature fluctuations around values concentrated in the quasi-linear range of the plot. On the other hand, it becomes of crucial importance when there is an interest in simulating insects' populations during season changes. As a result, the compromise consisted of dividing the temperature axis in subranges and to assign a specific development rate function to each of them.

An alternative option could be to increase the number of points in the life tables to improve the reliability of fit results. In particular, the additional points should be concentrated in temperature ranges close to the biological limits, where a high mortality rate is usually reported [161]. Accordingly, it requires a conspicuous number of individuals in the initial cohort to obtain a reasonable result.

Overall, there was a good accordance between simulations and field data for all the three years examined: the R^2 -values listed in Table 7.2 indicated the high capability of *EntoSim* in describing the monitored population of *L. botrana*. The 2016 season was well represented by the simulation, but at the same time had lowest population density. Concerning the 2014 and 2015 seasons, the trend was similar, with a higher population in the third generation.

This work, in the light of results obtained, underlined that ROOT and C++ programming language represent a valid choice to manage data and simulations. Beyond the computer model, the aim of this work was also to show the main difficulties related to a physiologically based model application, and how a dedicated software could validate results. It is often difficult to obtain life tables with a high number of experimental points. In fact, this operation requires a series of rearing at different constant temperatures, which needs environmental chambers and high numbers of individuals. The possibility of limiting some parameters in the development rate functions, together with the insertion of initial values, allows one to obtain similarly realistic results.

A territorial monitoring and modelling activity could be helpful to identify areas susceptible to *L. botrana* infestation and to forecast when outbreaks will occur. This concept can be relevant particularly for vineyards characterized by different microclimatic and/or agronomical conditions where the maturation of the grapes can be delayed [137]. Insect movement is mediated by a large number of factors including food availability [162], which in an agricultural context can depend on harvest date [163]. For instance, a late grape harvest can cause dispersal of *L. botrana* from nearby fields, impacting their population density. Large-scale management with forecasting models can contribute to develop a strategy to reduce the population density and limit migration.

From a more general point of view, predictions can also act as decision support system tools to guide the restriction of the areas in which harmful alien species occur [59], [121]. For example, where an invasive species is detected for the first time, forecasts can provide an idea of its potential speed of development and about the suitability of the environment [115].

The early detection of invasive organisms results in a better chance of significantly reducing their diffusion, not only in agricultural fields but also in management of forests, urban and peri-urban parks and private gardens [164]. These represent a natural refuge for native species of plants and animals [165], and the use of chemical pesticides, not always allowed in urban areas by national regulations, can contribute to a reduction in biodiversity. Unexpected insect outbreaks can occur in such ecosystems, a phenomenon which is probably due to an environment simplification with negative interference of entomofauna [166]. Alternative control strategies have been developed both for agricultural [122] and forest [62] systems, but the common factor in their application is often the need for fast reactions to maximize efficiency [167].

Moreover, mathematical models in agriculture, forest and environmental sciences have recently been used to identify future areas at risk from climate change [35], [118], [168]. For instance, several results have been published showing how temperature increases can affect biodiversity [138], [169]. In this scenario, software as EntoSim could be helpful to predict which insect species can benefit of climate change in a particular area, supporting agronomists and technicians in the choice of the less susceptible cultivation [170].

Chapter 8

A second novel model: the integro-differential generalised Von Foerster equation

“Failure is not an option.”

- Gene Kranz

The previous articles [31], [41], [42] showed how the novel generalised Von Foerster equation is reliable in describing populations of different insects species in different living environments. However, the population density of the following generation strongly depends on the abundance of the previous one. This aspect is defined as “*memory effect*”, and it is typically verified in Ecology. The generalised Von Foerster equation considers reproduction as a boundary condition, but the master equation does not have a memory term. In this article [30] this issue has been solved, providing an additional brick to the mathematical theory showed in this thesis. Hence, this model represents the second main goal reached during the PhD work. Besides the mathematical details, a validation of this novel model is reported in the case study of *L. botrana*.

Article details:	
Title:	A novel version of the Von Foerster equation to describe poikilothermic organisms including physiological age and reproduction rate
Authors:	Luca Rossini, Mario Contarini and Stefano Speranza
Corresponding author:	Luca Rossini
Journal:	Ricerche di Matematica

A novel version of the Von Foerster equation to describe poikilothermic organisms including physiological age and reproduction rate

8.1 Abstract

The mathematical description of poikilothermic organisms' life cycle, of insects in particular, is a widely discussed argument, above all for its application in decision support systems. The increasing interest among agricultural industries in obtaining products with minor quantities of chemical inputs has led entomologists and model scientists to study in greater depth not only the biology and behaviour of the insects, but also the way to translate these mechanisms into mathematical language.

The aim of this work is to provide a new instrument to describe insect pests' population density. In particular, the study analyses insects' development through the life stages driven by environmental factors. This has led researchers to consider physiological age, instead of the more widely used chronological age. In addition to mortality and fertility rates, the possibility of improving the simulation by inserting as a boundary condition results from previous field monitoring or the links with diapause models has been considered. The work is validated in the case of the European grapevine moth *Lobesia botrana*.

8.2 Introduction

Since their first introduction, the McKendrick and Von Foerster equations [57], [109] aroused the interest of population dynamics scientists. This is due to the ability of these equations to describe analytically the temporal evolution of a population divided into age classes. Over the years, this model has been widely applied to several case studies, such as human populations, animals and cells [125], [171], [172]. In particular, its potentialities have been used to predict populations of insect pests [173], [174], to perform control strategies focused on the reduction of pesticide usage on crops [143], in line with the safeguarding of environmental and human health.

The main issue related to the formulation of a mathematical model in biology is the lack, in some cases, of theories describing the basic processes [1] and, accordingly, this makes a series of simplifications inevitable, in order to schematize the problem. In particular, the Von Foerster equation considers that the individual's growth increases linearly with time [131]. This aspect makes the application of the Von Foerster equation difficult in the case of poikilothermic organisms, which are strongly dependent on the environmental parameters, such as temperature [31]. However, to obtain a more realistic description of insect populations in their habitat, it is necessary to consider also *i*) the possibility to include data provided by previous field monitoring, *ii*) models to describe how insects deal with unfavourable environmental conditions (e.g. diapause), and *iii*) the initial population's distribution in age classes.

Hence, the aim of this work is to provide a new model as a decision support system in plant protection against insect pests. The model's efficiency has been tested using the case study of the

European grapevine moth *Lobesia botrana* (Denis & Schiffermüller) (Lepidoptera: Tortricidae).

8.3 Materials and Methods

8.3.1 The population density model

The mathematical description starts from the schematic representation showed in Fig. 8.1.

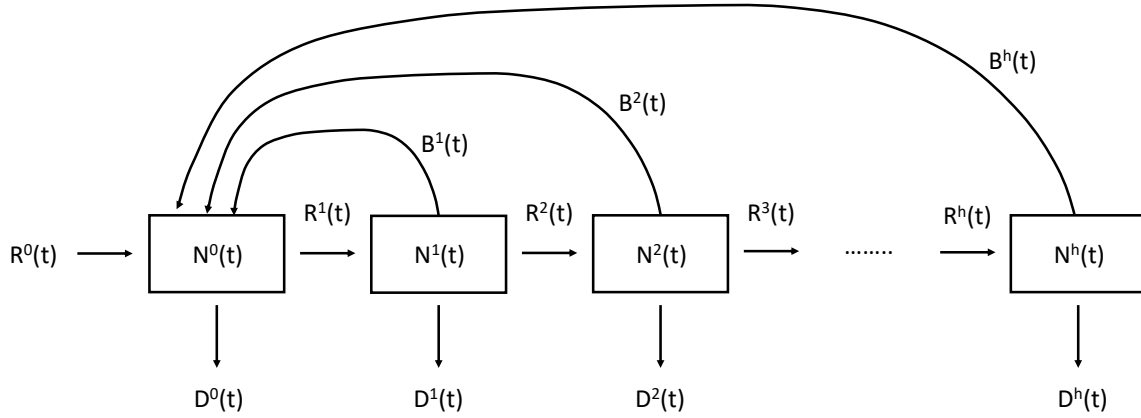


Fig. 8.1: Schematic representation of a life cycle considering reproduction rate. The life cycle is divided into age classes represented by a number $N^i(t)$ of individuals. On the whole, a cohort of individuals generates an input flux $R^{i-1}(t)$ and, after their maturation, an exit flux $R^i(t)$ toward the next stage. $D^i(t)$ indicates the flux of dead individuals, while $B^i(t)$ is the flux of new-born.

A life cycle divided into age classes can be represented with a series of chained boxes. The input flow in the first stage is composed of a cohort of individuals (e.g. eggs laid on the same day in the case of insects) and is indicated with the temporal function $R_0(t)$. Conceptually, each box represents an i -th age class, in which $N_i(t)$ individuals spend time for the process of maturation, dictated by environmental parameters and aging effects. When maturation is complete, an output flow of individuals passes to the next stage: this also represents the input flow for the following stage. This process is repeated until the last life stage (h -th stage), after which the development of the individual is no longer possible. Two additional flows specific to each stage are represented by progeny $B_i(t)$ and dead $D_i(t)$. The difference between $B_i(t)$ and $D_i(t)$ is that the first one represents the individuals which enter only in the first life stage, and the second one is similar to an exit flow, but it does not feed back into the maturation process of other life stages.

Let us focus on a single life stage, represented in Fig. 8.2. Introducing the population density function, $N(t, x)$, which describes the density of population over time t and physiological age x , it is possible to formulate a series of hypotheses at the basis of the model description.

Population starts from a time t , which is considered equal to zero, without losing the generality of the problem. At time $t = 0$, hence, there is a profile of individuals distributed into life stages, described by the following function:

$$N(0, x) = N^0(x). \quad (8.1)$$

The input flow into the i -th stage (Fig. 8.2) is indicated with $R_{x-\Delta x}(t)$ and it is ruled by a “generalized” development rate function which considers: *i*) the interaction between the species and the environmental parameters, usually time dependent, and *ii*) the effect of aging. This function in its general form is indicated with $G(t, x)$ and its introduction connects the entering flow to the population density function:

$$R_{x-\Delta x}(t) = G(t, x - \Delta x) \cdot N(t, x - \Delta x). \quad (8.2)$$

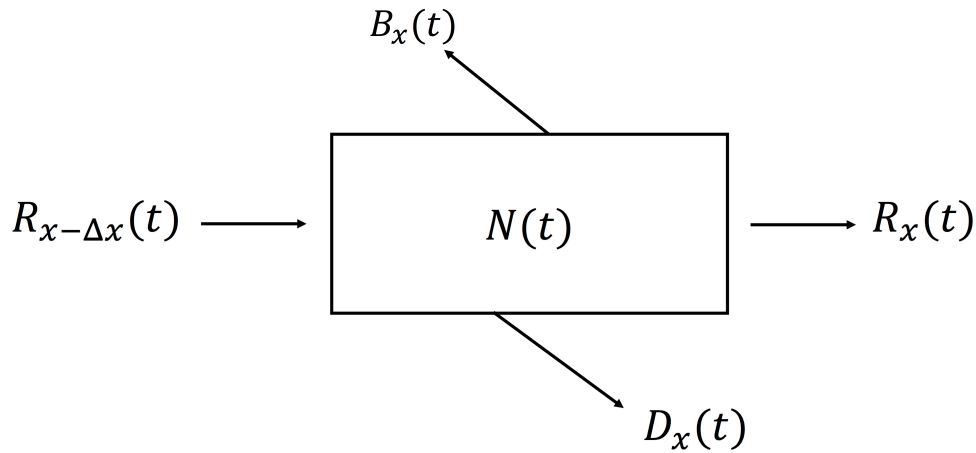


Fig. 8.2: Schematic representation of a single life stage. The infinitesimal input flow is represented by $R_{x-\Delta x}(t)$ function, $R_x(t)$ is the output flow of mature individuals which pass to the next stage, $D_x(t)$ is the flow of dead and $B_x(t)$ is the flow of progeny generated by the individuals in the stage, and which provides an input flow for the first stage of the life cycle. $N(t)$ is the count of the members present into the stage.

The development rate function, moreover, is defined as the temporal variation of the physiological age, that is:

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t} = G(t, x). \quad (8.3)$$

In other words, it is possible to assume that there is a relationship between age and time, described by the function $G(t, x)$. More specifically, the role of the development rate function is to describe the age increase in respect of time. Hence, with the definition (8.3), the development rate function assumes, also, the meaning of “speed of growth”.

Output flow from the i -th stage (Fig. 8.2) due to some individuals maturing is indicated with $R_x(t)$. In a similar way to the expression (8.2) it is possible to write:

$$R_x(t) = G(t, x) \cdot N(t, x). \quad (8.4)$$

Mortality flow from the i -th stage indicated with $D_x(t)$ is ruled by a mortality rate function $M(t, x)$. Its connection with the population density function is conceptually similar to the expressions (8.4) and (8.2):

$$D_x(t) = M(t, x) \cdot N(t, x). \quad (8.5)$$

The state function $N^i(t)$, which considers the number of individuals in the i -th stage, can be expressed integrating the population density function in the age interval $[x - \Delta x, x]$. Mathematically:

$$N^i(t) = \int_{x-\Delta x}^x N(t, x') dx'. \quad (8.6)$$

The flow of new-born $B_x(t)$ is ruled by a fertility rate function $\beta(t, x)$, time and age dependent.

Even if it is supposed that reproduction is possible in all of the life stages, the fertility rate function defines which of the age classes are fertile. In practical applications, the $\beta(t, x)$ function can be defined as a step function where the fertility coefficients, estimated from laboratory experiment, are reported. In a similar way to the expressions (8.2), (8.4) and (8.5):

$$B_x(t) = \beta(t, x) \cdot N(t, x). \quad (8.7)$$

Conceptually, it is possible to produce progeny at multiple stages, but the contribution deals with an entering flow only in the first life stage.

Now, let us focus on the temporal variation of $N^i(t)$. Using the hypotheses formulated above, it is possible to write a balance equation which rules all the input and output flows:

$$\frac{d}{dt} N^i(t) = R_{x-\Delta x} - R_x(t) - D_x(t) + \int_0^x B_x(t) dx'. \quad (8.8)$$

Expression (8.8) represents the balance between the input flow from the previous life stage, the output flow due to maturation of individuals, and the outlet flow of dead.

The last term on the right represents the contribution of the new-born and requires a more detailed discussion. Hypothetically, progeny can be produced by more life stages of the insect which are defined, in the most general case, by the fertility rate function $\beta(t, x)$. That means that each of these stages can produce a flow directed towards the first stage. Accordingly, the additional amount of individuals produced with reproduction is represented by the contributions derived from all the stages. Now, substituting the expressions (8.2), (8.4), (8.5), (8.6) and (8.7) the balance equation (8.8) is rewritten considering the density of population $N(t, x)$ and development, fertility and death rates:

$$\begin{aligned} \frac{d}{dt} \int_{x-\Delta x}^x N(t, x') dx' = & G(t, x - \Delta x) \cdot N(t, x - \Delta x) - G(t, x) \cdot N(t, x) + \\ & -M(t, x) \cdot N(t, x) + \int_0^x \beta(t, x') \cdot N(t, x') dx'. \end{aligned} \quad (8.9)$$

Dividing both members of the equation (98.) for Δx , and rewriting the terms:

$$\begin{aligned} \frac{d}{dt} \frac{1}{\Delta x} \int_{x-\Delta x}^x N(t, x') dx' &= \frac{G(t, x - \Delta x) \cdot N(t, x - \Delta x) - G(t, x) \cdot N(t, x)}{\Delta x} + \\ &\quad - \frac{M(t, x) \cdot N(t, x)}{\Delta x} + \frac{1}{\Delta x} \int_0^x \beta(t, x') \cdot N(t, x') dx'. \end{aligned} \quad (8.10)$$

At this point, each component of the expression (8.10) can be modified as follows. Considering the hypothesis of an infinitesimal stage, the mean value theorem for integrals can be applied to the left member integral:

$$\frac{1}{\Delta x} \int_{x-\Delta x}^x N(t, x') dx' = N(t, x). \quad (8.11)$$

The first term on the right member of expression (8.10) can be replaced considering the Taylor series expansion of the function $G(t, x) \cdot N(t, x)$ truncating on the first order [31]:

$$\frac{G(t, x - \Delta x) \cdot N(t, x - \Delta x) - G(t, x) \cdot N(t, x)}{\Delta x} \sim -\frac{\partial}{\partial x} [G(t, x) \cdot N(t, x)]. \quad (8.12)$$

Mortality term, divided by Δx , can be considered as a function which provides the probability of death at each life stage and indicated with $M^i(t, x)$ [31]:

$$\frac{M(t, x) \cdot N(t, x)}{\Delta x} = M^i(t, x) \cdot N(t, x). \quad (8.13)$$

Otherwise, it is not possible to apply the mean value theorem for integrals, as showed for the term (18.1), to the integral $\frac{1}{\Delta x} \int_0^x \beta(t, x') \cdot N(t, x') dx'$ on the right of equation (8.10), because of the width of the integration interval. To avoid this issue, it is possible to follow the same line of reasoning showed for the mortality term (8.13).

The function $\beta(t, x)$ divided by Δx represents the probability for individuals to generate progeny in each stage, synthetized with $\beta^i(t, x)$. In this way, the integral on the right of the equation (8.10) is rewritten as follows:

$$\frac{1}{\Delta x} \int_0^x \beta(t, x') \cdot N(t, x') dx' = \int_0^x \beta^i(t, x') \cdot N(t, x') dx'. \quad (8.14)$$

The new expressions (8.11), (8.12), (8.13) and (8.14) are now ready to be substituted in the equation (8.10). The Taylor series expansion (8.12) introduced the partial derivative in the x variable. For this reason, in the left member of the expression (8.10), the time derivative becomes $\frac{\partial}{\partial t} N(t, x)$. Hence, these assumptions lead to:

$$\frac{\partial}{\partial t} N(t, x) + \frac{\partial}{\partial x} [G(t, x) \cdot N(t, x)] = -M^i(t, x) \cdot N(t, x) + \int_0^x \beta^i(t, x') \cdot N(t, x') dx'. \quad (8.15)$$

The initial condition (8.1) described during the hypothesis' formulation can be associated

to the equation (8.15). Otherwise, in practical cases, the estimation of the initial population profile is difficult, above all in open field contexts.

Data related to insect monitoring activity mathematically represent a temporal evolution of the population in a specific life stage. In other words, this information represents a boundary condition for the equation (8.15), and in the most general form can be written as:

$$N(t, 0) = N^0(t). \quad (8.16)$$

The expression (8.15) with the initial condition (8.1) and the boundary condition (8.16) represents the complete version of the model.

8.3.2 Interaction between species and environment

At this point, it is important to draw attention to the development rate function $G(t, x)$. Even though the development is supposed to be dependent, in the most general case, on time and age, the only known dependence for insects is on the daily average temperature. Hence, there is a connection only with temporal variations of the development rate, and it is restricted only to the main driving environmental parameter.

In the surrounding literature, various expressions have been developed to describe the connection between the development rate and the temperature [44], [45], [128], usually written using the notation $R[T(t)]$. The choice for this work lies in the Brière [47] and Logan [46] expressions. The Brière expression is mathematically represented as follows:

$$R[T(t)] = a \cdot T(t) [T(t) - T_L] \cdot [T_M - T(t)]^{\frac{1}{m}}, \quad (8.17)$$

where $T(t)$ is the daily average temperature, a and m are empirical parameters, T_L and T_M are respectively the lower and upper thermal thresholds, below and above which life is theoretically not possible.

The Logan expression is mathematically represented as follows:

$$R[T(t)] = \psi \cdot \left[\exp(\rho T(t)) - \exp\left(\rho T_M - \frac{T_M - T(t)}{\Delta T}\right) \right], \quad (8.18)$$

where, $T(t)$ is the daily average temperature, ψ and ρ are empirical parameters, T_M is the upper thermal threshold, above which life is theoretically not possible, and ΔT is the temperature interval between the maximum peak of the function and the upper thermal threshold. The use of the expressions (8.17) and (8.18) as development rate functions implies to not consider the age dependence on the generalized development rate function $G(t, x)$ presented in the equation (8.15). Hence, to pursue the aim of this work it is now considered $G(t, x) = R[T(t)]$.

The parameters of the above expressions (8.17) and (8.18) are estimated through a non-linear fit operation with data from laboratory experiments which, in the case study presented in this article, are available from the work of Moshtaghi Maleki et al. [55].

A cross check between χ^2 -test and adjusted R^2 values [45] resulting from the non-linear regression will determine which is the best representing function between (8.17) and (8.18) for

the experimental data [31], [41], [42].

Even if expressions (8.17) and (8.18) have a similar shape, from an ecological point of view they provide different information. More specifically, the Logan function lacks of a parameter which indicates the lower thermal threshold for the development of the species [42], well defined by T_L in the Brière expression. On the other hand, the additional value provided by the Logan expression is ΔT , which can be interpreted as the tolerance of the species to high temperatures. Both the expressions (8.17) and (8.18) present a *quasi-linear* range, corresponding to the optimal thermal conditions usually assessed during grapevine's growing season. Accordingly, it is possible to choose, at the height of the growing season, the development rate function which represents better the experimental life tables. A shrewder choice is advised in areas featured by particularly low or particularly high temperatures. In this case, it is preferable to address the choice on the Brière expression at medium-low temperatures and on the Logan at medium-high temperatures. The experimental vineyard considered in this work belongs to an area not affected by particular climatic events, and for this way the choice of the best representing development rate function can be entrusted considering the χ^2 -test and the R^2 value, as already mentioned above.

8.3.3 Numerical solutions and model validation

The main problem related to the application of the model (8.15) concerns the lack of an analytical expression to describe the daily temperature variation throughout the year. More specifically, $T(t)$ is included into the expressions (8.17) and (8.18) and it refers to the temperature recorded in the experimental field, where insects develop. Accordingly, the only way to apply the model (8.15) in a varying temperature context is through numerical solutions.

A possible solution could be the application of the upwind scheme, the finite difference method and the rectangle method to the integro-differential equation (8.15) obtaining the following discrete form:

$$\begin{cases} N_h^{i+1} = \frac{i}{h} R[T_i] N_{h-1}^i + N_h^i \left[1 - \frac{i}{h} R[T_i] - i M_h^i \right] + i h \sum_{n=0}^{n=h} \beta_n^i N_n^i \\ N_0^i = N_0 \\ N_h^0 = N^0. \end{cases} \quad (8.19)$$

A different notation has been adopted to distinguish discrete and continuous variables: time-step is indicated with i and the age-step with h . Furthermore, in the expression (8.19) $R[T_i]$ indicates the discrete version of the expression (8.17) or (8.18).

Calculations have been conducted in a similar way, as described in the work of Rossini et al. [49], by the implementation of the numerical scheme into C++ software.

A comparison between field data and numerical solutions is possible applying, again, χ^2 formula:

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}, \quad (8.20)$$

where O_i are the observed data (monitored population) and E_i are their corresponding simulated data. In this case, expression (8.20) is used as an indicator of the distance between the numerical solution (8.19) and data provided by field monitoring. To support results provided by χ^2 formula, it is also possible to consider the R^2 , which shows how much simulation is capable of accurately representing experimental data.

8.3.4 The European grapevine moth *Lobesia botrana*

To test the model's efficacy, the case study of the European grapevine moth, *Lobesia botrana*, one of the most harmful insect pests affecting grapevines worldwide, has been chosen. Its dangerousness derives mainly from its capability to complete up to 4 generations during the year, and from its connection with secondary hosts such as *Botrytis cinerea*, a causal agent of fungal disease [175]–[178].

The life cycle has an average duration of 40 days in optimal conditions, and the yearly generations are synchronized with the main phenological stages of grape plants [112].

The first life stage is represented by the eggs, laid by females into flower buds during the first generation and laid into clusters and berries during the second and the third. After the eggs have hatched, five larval instars are largely responsible for the damages recorded in grape fields [179].

The total duration of the larval stage comprises a high percentage of the whole life cycle (around 60-65%) and it is usually the most difficult stage to control with either chemical or biological strategies [111]. The following stages are pupae and adults. The coupling activity starts 1-6 days after emergence [179] and females lay an average of 100 eggs [55] 1-2 days after coupling [179].

In accordance with this life cycle, this work considers a total number of 8 stages (Fig. 8.3) [55], [111], [179]. Since the first generation is considered the least relevant for damages reported in grapevine fields [178], information related to the monitoring operations can be inserted as a boundary condition into the model, in order to simulate the second and the third generations.

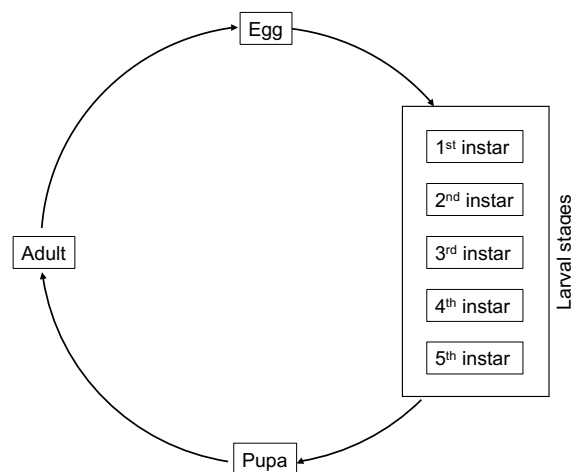


Fig. 8.3: Schematic representation of *L. botrana* life cycle.

8.3.5 Field data acquisition and analysis

Data were collected in an experimental vineyard (ca. 1 hectare) located in the Castelli Romani area (Lazio, Central Italy), a peri-urban area notable for high quality wine production [42], [180].

L. botrana population density was monitored by catching males with Biogard Delta Traps (BDT) (S.r.l. Biogard Division, 20834 Nova Milanese (MB), Italy), equipped with pheromone type Selure-Lb. Three traps were sufficient to cover the experimental area [159]. Data were collected in the 2015 growing season, during which three generations of European grapevine moth were recorded.

After their deployment (which occurred in the middle of April 2015), traps were checked once a week until the end of the season (late September 2015). The sampling range, moreover, provides an estimation of the uncertainty associated to each experimental point, fixed to ± 3.5 days.

A meteorological station from Vaisala company (Helsinki, Finland) (model AWS-310) was placed into the experimental field to record the daily average temperature. The mean values were obtained from 48 measurements every 24 hours and used as input into the numerical solution (8.19), once the best representative development rate function between (8.17) and (8.18) was defined.

8.4 Results and discussion

The first phase of model application concerns the development rate functions' parameters estimation through a non-linear fitting operation. Using data provided by Moshtaghi Maleki et al. [55], the results listed in Table 8.1 have been obtained. A visual output is represented in Fig. 8.4, where the best fit functions (8.17) and (8.18) are overlapped and compared with Moshtaghi Maleki et al. data.

At this point, the adjusted R^2 -values listed in Table 8.1 indicates that the appropriate function to use into numerical scheme (8.19) is the Logan, but this result needs to be confirmed with a χ^2 -test.

Considering a number of degrees of freedom $NDF = 2$, the P -values provided by χ^2 -test are $P_B = 0.9981$ for the Bri re expression (8.17) and $P_L = 0.9990$ for the Logan expression (8.18). According to this first phase of the analysis, the effects of temperature on *L. botrana* development are well described by the Logan formula (8.18). Hence, the best development rate function to use for simulations was chosen as the Logan development rate function. Other inputs are the daily average temperature values and field monitoring on the first generation of *L. botrana*.

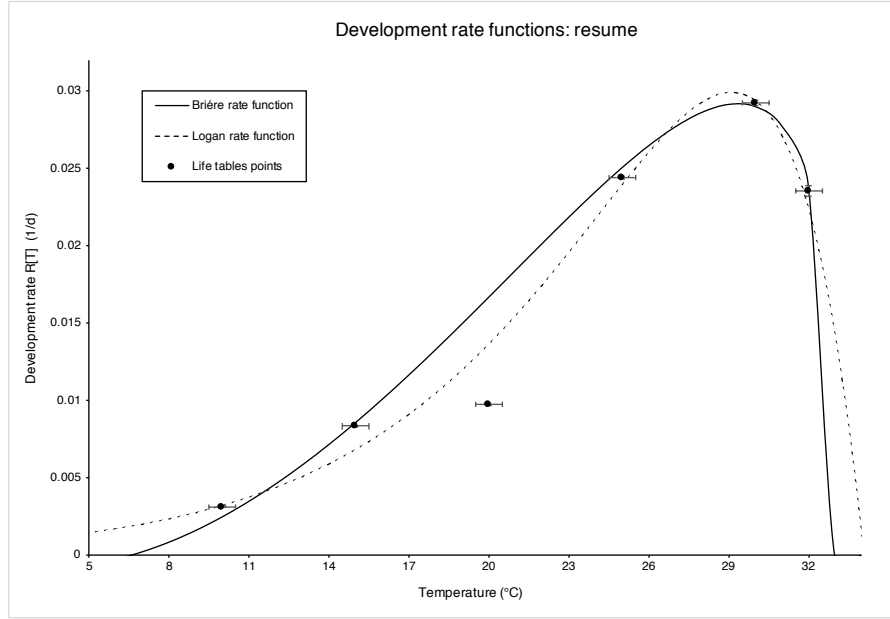


Fig. 8.4: Logan and Brière development rate functions estimated with experimental data from Moshtaghi Maleki et al. [55].

Table 8.1: Results from non-linear parameters estimation process through the use of “*fitasso*” macro [49]. The parameters’ values for Brière and Logan rate functions are reported with their corresponding errors. Additional information is provided by χ^2 -value, adjusted R^2 -value and the number of degrees of freedom (NDF).

Rate function	Parameters	χ^2 -value	Adjusted R^2 -value	NDF (n)
Brière	$a = (3.4 \pm 0.2) \cdot 10^{-5}$	$3.65 \cdot 10^{-3}$	0.867	2
	$T_L = 6 \pm 1$			
	$T_M = 32.6 \pm 0.2$			
	$m = 5 \pm 1$			
Logan	$\psi = (7.0 \pm 0.5) \cdot 10^{-4}$	$1.93 \cdot 10^{-3}$	0.925	2
	$\rho = 0.18 \pm 0.04$			
	$T_M = 34.1 \pm 0.8$			
	$\Delta T = 4 \pm 2$			

A graphical comparison between simulations and monitoring data is reported in Fig. 8.5. The peak on the left side represents adult flights of the second generation. The population’s profile is well represented, and a maximum is provided on 17th June by simulations, while the field data report this value on 16th June.

The third generation has been simulated considering the reproduction rate. Results are shown on the right part of Fig. 8.5. In this case, also, simulations reproduce the field data

profile well, but with less accordance than the second generation. The maximum peak for the third generation is on 14th August for simulations, and on 19th August for field data.

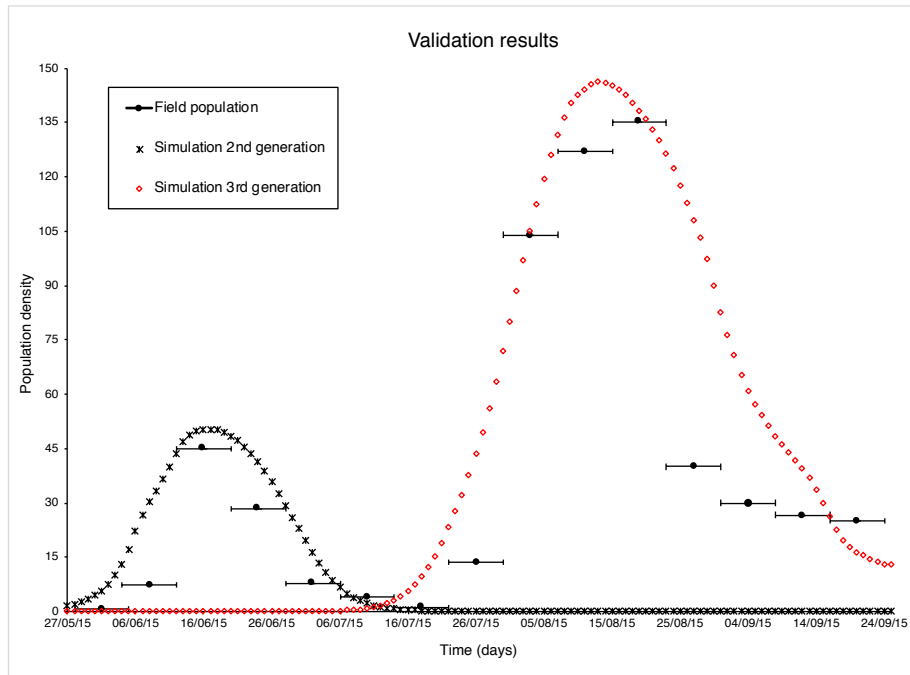


Fig. 8.5: Comparison between simulations of the first and second generation and field data.

The accuracy of simulations in representing field data has been investigated with the expression (8.20) and the R^2 -value. Results are listed in Table 8.2.

Table 8.2: χ^2 and R^2 values to compare simulations and field data in the case of the second and third generation.

Generation	χ^2 -value	R^2 -value
2 nd	21.02	0.841
3 rd	46.65	0.911

The second generation reported a lower χ^2 value, but at the same time the R^2 is lower than the third generation. Overall, these results confirm the reliability of the simulation in representing field data.

8.5 Conclusions

The comparison between simulations and field data shows that the model is appropriate to describe poikilothermic organisms' population density. In particular, it has demonstrated the importance of adjusting the model by providing additional information in the initial and boundary condition. For this case study, data about the first generation of the European grapevine

moth have been inserted as a boundary condition. There were two primary reasons for this choice.

The first generation usually has a low population density, making control action unnecessary [181]. At the same time, however, the damages inflicted on flower buds [54] and above all the crucial role that the first generation plays on the abundance of the following generations cannot be neglected [112], [176], [177], [179]. Accordingly, by inserting into the model the monitoring results on the first generation as a boundary condition, a more accurate simulation can be obtained, as shown in Fig. 8.5.

An additional advantage to having a boundary condition “*free of use*”, is that it allows the inclusion of diapause models, an argument which has aroused the interest of model scientists for a long time [182]–[184]. A mathematical representation of the diapause, in fact, is helpful for managing the initial population profile for the current year [185].

This is of extensive interest not only for the European grapevine moth, but also in relation to other aggressive phytophagous agents’ species such as the tomato leaf miner, *Tuta absoluta* [116], [117], or the pepper weevil, *Anthonomus eugenii* [41], [121].

The validation process reported as the conclusive phase of this research shows that the model (8.15) represents a good starting point to describe the poikilothermic organisms’ life cycle. Nevertheless, it provides evidence of the need to delve further into methods of translating into mathematical language the interaction between insects and environmental factors (to date limited only to the temperature), together with the effect of aging on the development rate, and fertility and mortality rate description.

Moreover, it has been shown that the model is able to describe multivoltine species. Nonetheless, the number of generations is expressed in terms of physiological age. In the case of *Lobesia botrana* the total number of life stages is 8, considering the eggs, 5 larval instars, pupa and adult (Fig. 8.3). In the results shown in Fig. 8.5, the second generation starts from the 9th stage, and is concluded by the 16th.

Cases in which the biology of the species is known will have a simple application, but more attention is required for insects that have not been completely explored by entomologists. Fortunately, this issue is overcome with more accurate studies on life tables [52], demonstrating again the importance of collecting biological information for the correct operativity of the model.

In addition to the fact that this kind of model is helpful for decision support systems at the service of agronomists, farmers and technicians, recently they are finding additional uses in other contexts. To cite some examples, a reliable mathematical description of the biological mechanisms in relationship with the environmental parameters, is crucial for studies related to climate change and forensic entomology [37]. In particular, several studies are focused on the possible diffusion of insect pests in an increasing temperature scenario [118], [135]–[137], [168], remarking on the need for progress in the modelling field.

Chapter 9

Conclusions

*“Oh, simple thing, where have you gone?
I’m getting old, and I need something to rely on
So tell me when you’re gonna let me in
I’m getting tired, and I need somewhere to begin”*

- Somewhere only we know, Keane

The study summarized in the previous chapters highlighted some of the problems affecting the modelling of ectotherms populations. The choice to divide the whole study in independent “blocks” led to analyse the strenghts and the weaknesses of each part of the theory. Surely, the main findings of this thesis are the generalised Von Foerster equation and the integro-differential Von Foerster equation. In fact, it has been shown that the chronological age and physiological time are not so adequate to describe ectotherm populations. Even if in the past the Distributed Delay Model has been widely applied, the first part of this study showed that not only it should be used in its “counts” version, but also that it needs a calibrations step. From an experimental point of view, calibration means to repeat an additional validation but “before the validation”.

In Biology, Ecology, Agricultural and Forest Sciences and all the related fields of study, data acquisition is usually the most expensive part. For this reason, the experimental part has to be supported as much as possible by a strong mathematical and statistical theory behind. In addition, the latter should be standardised and applied to all the models belonging to the same category. This phylosophy led to the formulation of the generalised and the integro-differential Von Foerster equations. The advantage to have preferred physiological age to chronological age or physiological time is the elimination of the calibration step. Accordingly, the validation of the model is easier, as well as its application as decision support system.

Other interesting aspects emerged from the study concerned the effective reliability of the *a priori* and *a posteriori* analysis. It has been demonstrated that an ideal model describes ectotherms population dynamics over time and through their life stages. Hence, the master equation is composed by a series of sub-models (e.g. the generalised development, mortality and feritility rate functions) containing biological information. The advantage is that the structure

of the master equation (either generalised or integro-differential Von Foerster equation) never change. Therefore sub-models can be developed independently. A practical example reported in the present work concerned the development rate functions: even if they represent a simplification (not considering ageing or other environmental parameters than temperature), they are developed independently from life tables. Accordingly, in the master equation is inserted a function already optimized and validated with other experimentations. The role of the *a priori* and *a posteriori* analysis for this aspect is fundamental. In fact, each sub-model describing development, fertility or mortality rates involve a function, as a solution, whose parameters need to be estimated through experimental data. When a linear or non-linear regression is involved, a cross check between χ^2 -test and R^2 -value is the strongest methodology to evaluate the goodness of the best fit function. But at the same time more than one expression can provide good values from the *a priori* analysis, as showed in the case study of the *A. eugenii*. Thereby, the *a posteriori* analysis can overcome the issues comparing the model's solution with field data varying the sub-models in input to the master equation. The *a posteriori* analysis was discussed and applied in the present study only in the case of the development rate functions (showing its reliability), but it can be extended also to fertility and mortality rate functions. Accordingly, this type of analysis can be easily extended to all the physiologically-based models based on ordinary differential equations (it has been showed that it worked on the Distributed Delay Model) and on partial differential equations (first and higher orders).

However, the glue between the mathematical theory (models, data analysis and validation) and the application is represented by *EntoSim*. The reasons behind the choice of C++ programming language and ROOT's software libraries have been discussed in its dedicated chapters, but some remarks are needed in this conclusion. Previously, the costs behind the experimentations were mentioned. Talking about costs the first thought falls on manpower (technicians, researchers etc.) and materials, but a great part of the costs are also related to software licenses. This cannot be a problem for the biggest research centres worldwide, but it can greatly affect small research centres or the ones located in developing countries. Hence, the use of open source software and the knowledge of programming languages are fundamental in model development in this field of research. A proof of that has been showed in this thesis, as well as the fact that the same field data can be helpful to validate different type of models. About this latter aspect, a demonstration was reported in the case of *L. botrana*: the same field data (the whole dataset or part of it) were used to validate the counts version of the Distributed Delay Model, and the generalised and integro-differential Von Foerster equations.

Even if the results of this research showed novel aspects and organised them as a theory, this is only the beginning of a long road. Several limitations, in fact, still affect the presented models. Besides the development of the various sub-models describing the biology of the species, two main issues affect the master equations of the generalised and of the integro-differential Von Foerster equation. The first one, is the biomass. In fact, it was not considered the food availability or the growth due to ectotherms' nutrition. Hence, to date it is not possible to consider the interaction between plants and insects, and also to provide an estimation of the losses. Biomass growth term, however, can be easily introduced into the master equation together with the "growth

rate”, its related sub-model. An approach of this type was already considered by Sinko and Streifer [102]: in future works the *principia* of their equation can be merged with the *principia* of the generalised and integro-differential Von Foerster equations. The second aspect still missing in the master equations presented in this work concerns the possibility of interaction between species. There are no terms, in fact, which provide the possibility to link a model describing a predator, and a model describing a prey, for example. An extension of this type surely can open the model applications to control strategies which provides for the use of natural enemies, or to control the equilibria of ecological niches. Even this extension can be easily considered in future works, merging the *principia* of the generalised and integro-differential Von Foerster equations with the Volterra-Lotka predator-prey models.

Hence, this is not the end of this work, this is just its beginning!

Prof.: “Dobbiamo prepararci per domani.”

Mignolo: “Perché? Cosa facciamo domani sera?”

*Prof.: “Quello che facciamo tutte le sere, Mignolo.
Tentare di conquistare il mondo!”*

- Mignolo e Prof.

Acknowledgements

Ed eccoci arrivati alla parte più difficile da scrivere.

Già, perché le emozioni sono state la benzina che ha alimentato tutto questo bellissimo percorso che ora, purtroppo, sta volgendo al tramonto. Pensiamo a domani sperando che il sole, nell'incertezza che in questo mondo accompagna noi giovani, sorga presto.

Noi siamo quelli che non mollano mai. Non siamo programmati per farlo.

Credo che sarei ingiusto se facessi un elenco di persone da ringraziare, perché sono sicuro che dimenticherei qualcuno. In cuor mio, ringrazio chi mi ha motivato, chi ha creduto in me soprattutto nei momenti più bui. Un grazie va a chi mi è sempre stato da esempio, a chi ha svegliato in me l'animo competitivo, a chi mi ha sempre donato sorrisi e amore e mi ha insegnato a renderlo. Grazie a chi c'era ieri, c'è oggi e ci sarà domani.

Un grazie, infine, va anche a chi puntualmente mi ha fatto riconoscere i miei limiti, perché con modi rudi mi avete insegnato a non darmi per vinto. Mai.

Senza di voi, tutto questo non sarebbe stato possibile.

*“Sono salito sulla cattedra per ricordare a me stesso
che dobbiamo sempre guardare le cose da angolazioni diverse.*

*E il mondo appare diverso da quassù.
Non vi ho convinti? Venite a vedere voi stessi.*

*Coraggio!
È proprio quando credete di sapere qualcosa
che dovete guardarla da un'altra prospettiva.”*

- L'attimo fuggente

Bibliography

- [1] M. Ageno, *Le radici della Biologia*, F. Editor, Ed. 1986, ISBN: 9788807100628.
- [2] L. Rossini, ‘Sviluppo stocastico di *Corcyra cephalonica* a diverse temperature’, Master’s thesis, Università degli Studi della Tuscia, 2016.
- [3] L. Rossini, M. Contarini, M. Severini and S. Speranza, ‘Reformulation of the Distributed Delay Model to describe insect pest populations using count variables’, *Ecological Modelling*, p. 109286, 2020. DOI: 10.1016/j.ecolmodel.2020.109286.
- [4] G. Gilioli, S. Pasquali and E. Marchesini, ‘A modelling framework for pest population dynamics and management: An application to the grape berry moth’, *Ecological Modelling*, vol. 320, 348–357, 2016, ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2015.10.018.
- [5] W. M. Getz and A. P. Gutierrez, ‘A perspective on systems analysis in crop production and insect pest management’, *Annual Review of Entomology*, vol. 27, no. 1, 447–466, 1982. DOI: 10.1146/annurev.en.27.010182.002311.
- [6] S. Orlandini, R. D. Magarey, E. W. Park, M. Spordler and J. Kroschel, ‘Methods of agroclimatology: Modeling approaches for pests and diseases’, *Agroclimatology*, pp. 453–488, 2018. DOI: 10.2134/agronmonogr60.2016.0027.
- [7] G. Bellocchi, M. Rivington, M. Donatelli and K. Matthews, ‘Validation of biophysical models: Issues and methodologies’, *Sustainable Agriculture*, vol. 2, pp. 577–603, 2011. DOI: 10.1007/978-94-007-0394-0_26.
- [8] I. Pertot, T. Caffi, V. Rossi, L. Mugnai, C. Hoffmann, M. Grando, C. Gary, D. Lafond, C. Duso, D. Thiery and et al., ‘A critical review of plant protection tools for reducing pesticide use on grapevine and new perspectives for the implementation of ipm in viticulture’, *Crop Protection*, vol. 97, 70–84, 2017, ISSN: 02612194. DOI: 10.1016/j.cropro.2016.11.025.
- [9] E. Han, W. E. Baethgen, A. V. Ines, F. Mer, J. S. Souza, M. Berterretche, G. Atunéz and C. Barreira, ‘SIMAGRI: An agro-climate decision support tool’, *Computers and Electronics in Agriculture*, vol. 161, no. June 2018, 241–251, 2019, ISSN: 01681699. DOI: 10.1016/j.compag.2018.06.034.
- [10] M. Amo-Salas, V. Ortega-Lopez, R. Harman and A. Alonso-Gonzalez, ‘A new model for predicting the flight activity of *Lobesia botrana* (Lepidoptera: Tortricidae)’, *Crop Protection*, vol. 30, no. 12, pp. 1586–1593, 2011. DOI: 10.1016/j.cropro.2011.09.003.

- [11] A. Cappio Borlino, G. Di Cola and G. Marras, 'Mathematical modelling of natural population dynamics', *Memorie dell'Istituto Italiano di Idrobiologia*, vol. 49, 127–162, 1991.
- [12] R. E. Plant and L. T. Wilson, 'Modelling and simulation in agricultural pest management', *SIAM Review*, vol. 29, pp. 247–262, 1987.
- [13] F. Rebaudo and V.-B. Rabhi, 'Modeling temperature-dependent development rate and phenology in insects: Review of major developments, challenges, and future directions', *Entomologia Experimentalis et Applicata*, vol. 166, no. 8, pp. 607–617, 2018. DOI: 10.1111/eea.12693.
- [14] E. Gilbert, J. A. Powell, J. A. Logan and B. J. Bentz, 'Comparison of three models predicting developmental milestones given environmental and individual variation', *Bulletin of Mathematical Biology*, vol. 66, no. 6, 1821–1850, 2004, ISSN: 00928240. DOI: 10.1016/j.bulm.2004.04.003.
- [15] T. J. Manetsch, 'Time-varying distributed delays and their use in aggregative models of large systems', *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-6, no. 8, 547–553, 1976, ISSN: 0018-9472. DOI: 10.1109/TSMC.1976.4309549.
- [16] J. Vansickle, 'Attrition in distributed delay models', *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 7, no. 9, 635–638, 1977, ISSN: 0018-9472. DOI: 10.1109/TSMC.1977.4309800.
- [17] B. K. Quinn, 'A critical review of the use and performance of different function types for modeling temperature-dependent development of arthropod larvae', *Journal of Thermal Biology*, vol. 63, no. November 2016, pp. 65–77, 2017, ISSN: 03064565. DOI: 10.1016/j.jtherbio.2016.11.013.
- [18] G. B. Schaalje and H. R. Van Der Vaart, 'Relationships among recent models for insect population dynamics with variable rates of development', *Journal of Mathematical Biology*, vol. 27, no. 4, 399–428, 1989, ISSN: 0303-6812. DOI: 10.1007/BF00290637.
- [19] A. P. Gutierrez, L. Ponti, M. L. Cooper, G. Gilioli, J. Baumgärtner and C. Duso, 'Prospective analysis of the invasive potential of the European grapevine moth *Lobesia botrana* (Den. & Schiff.) in california', *Agricultural and Forest Entomology*, vol. 14, no. 3, 225–238, 2012, ISSN: 14619555. DOI: 10.1111/j.1461-9563.2011.00566.x.
- [20] A. P. Gutierrez, L. Ponti, G. Gilioli and J. Baumgärtner, 'Climate warming effects on grape and grapevine moth (*Lobesia botrana*) in the palearctic region', *Agricultural and Forest Entomology*, 2017, ISSN: 14619555. DOI: 10.1111/afe.12256.
- [21] M. Severini, J. Baumgärtner and M. Ricci, 'Theory and practice of parameter estimation of distributed delay models for insect and plant phenologies', *"Meteorology and Environmental Sciences"*, 674–719, 1990.
- [22] L. Limonta, D. P. Locatelli, T. Broglia and J. Baumgärtner, 'Cohort development models for integrated *Corcyra cephalonica* (Stainton) population management', *Bollettino di Zoologia Agraria e di Bachicoltura*, vol. 3, 215–226, 2009.

- [23] A. P. Gutierrez and M. A. Pizzamiglio, 'A general distributed delay time varying life table plant population model: Cotton (*Gossypium hirsutum* L.) growth and development as an example', *Ecological Modelling*, vol. 26, 231–249, 1984.
- [24] G. Di Cola, G. Gilioli and J. Baumgärtner, 'Mathematical models for age-structured population dynamics', in. 1989, ch. 15.
- [25] J. Nance, R. T. Fryxell and S. Lenhart, 'Modeling a single season of *Aedes albopictus* populations based on host-seeking data in response to temperature and precipitation in eastern Tennessee', *Journal of Vector Ecology*, vol. 43, no. 1, 138–147, 2018, ISSN: 10811710. DOI: 10.1111/jvec.12293.
- [26] B. Ainseba and A. Sebastian, 'Age-dependent population dynamics with diffusion', *Abstract and Applied Analysis*, vol. 0, pp. 357–368, 2001.
- [27] V. Bellagamba, G. Di Cola and R. Cavalloro, 'Stochastic models in fruit-fly population dynamics', in *Proceedings of the CEC/IOBC International Symposium "Fruit Flies of Economic Importance 87"*, 1987, 91–98.
- [28] Y. He and B. Ainseba, 'Exact null controllability of the *Lobesia botrana* model with diffusion', *Journal of Mathematical Analysis and Applications*, vol. 409, no. 1, 530–543, 2014, ISSN: 0022247X. DOI: 10.1016/j.jmaa.2013.07.020.
- [29] S. Pasquali, C. Soresina and G. Gilioli, 'The effects of fecundity, mortality and distribution of the initial condition in phenological models', *Ecological Modelling*, vol. 402, pp. 45–58, 2019. DOI: 10.1016/j.ecolmodel.2019.03.019.
- [30] L. Rossini, M. Contarini and S. Speranza, 'A novel version of the Von Foerster equation to describe poikilothermic organisms including physiological age and reproduction rate', *Ricerche di Matematica*, 2020, ISSN: 0035-5038. DOI: 10.1007/s11587-020-00489-6.
- [31] L. Rossini, M. Severini, M. Contarini and S. Speranza, 'A novel modelling approach to describe an insect life cycle vis-à-vis plant protection: Description and application in the case study of *Tuta absoluta*', *Ecological Modelling*, vol. 409, no. July, p. 108778, 2019, ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2019.108778.
- [32] A. P. Gutierrez, F. Schulthess, L. T. Wilson, A. M. Villacorta, C. K. Ellis and J. Baumgärtner, 'Energy acquisition and allocation in plants and insects: A hypothesis for the possible role of hormones in insect feeding patterns', *The Canadian Entomologist*, vol. 119, 109–129, 1984.
- [33] B. Wermelinger, J. Baumgärtner and A. P. Gutierrez, 'A demographic model of assimilation and allocation of carbon and nitrogen in grapevines', *Ecological Modelling*, vol. 53, no. C, 1–26, 1991, ISSN: 03043800. DOI: 10.1016/0304-3800(91)90138-Q.
- [34] M. Severini, R. Alilla, S. Pesolillo and J. Baumgärtner, 'Fenologia della vite, e della *Lobesia botrana* (Lep. Tortricidae) nella zona dei Castelli Romani', *Rivista Italiana di Agrometeorologia*, vol. 39, no. 3, 34–39, 2005.

- [35] L. Schaub, S. Breitenmoser, J. Derron and B. Graf, 'Development and validation of a phenological model for the univoltine European corn borer', *Journal of Applied Entomology*, vol. 141, no. 6, 421–430, 2017, ISSN: 09312048. DOI: 10.1111/jen.12364.
- [36] J. Baumgärtner and M. Severini, 'Microclimate and arthropod phenologies: The leaf miner *Phyllonorycter blancardella* F. (Lep.) as an example', in *International Conference on Agrometeorology*, F. Prodi, F. Rossi and G. Cristoferi, Eds., Fondazione Cesena Agricoltura Publ., 1987, 225–243.
- [37] J. C. Allen and J. H. Byrd, *Computer modeling of insect growth and its application to forensic entomology*. 2000.
- [38] S. K. Spolia, 'Time-varying cascade model for flow forecasting', *Hydrological forecasting - Prévisions hydrologiques*, no. 129, 319–324, 1980.
- [39] G. G. Jensen, F. Uekermann and K. Sneppen, 'Multi stability and global bifurcations in epidemic model with distributed delay sirns-model', *The European Physical Journal*, vol. 92, no. 2, p. 28, 2019. DOI: 10.1140/epjb/e2018-90562-1.
- [40] M. K. Asplen, G. Anfora, A. Biondi, D.-S. Choi, D. Chu, K. M. Daane, P. Gibert, A. P. Gutierrez, K. A. Hoelmer, W. D. Hutchison, R. Isaacs, Z.-L. Jiang, Z. Kárpáti, M. T. Kimura, M. Pascual, C. R. Philips, C. Plantamp, L. Ponti, G. Véték, H. Vogt, V. M. Walton, Y. Yu, L. Zappalà and N. Desneux, 'Invasion biology of spotted wing drosophila (*drosophila suzukii*): A global perspective and future priorities', *Journal of Pest Science*, vol. 88, no. 3, pp. 469–494, 2015. DOI: 10.1007/s10340-015-0681-z.
- [41] L. Rossini, M. Contarini, M. Severini, D. Talano and S. Speranza, 'A modelling approach to describe the *Anthonomus eugenii* (Coleoptera: Curculionidae) life cycle in plant protection: *a priori* and *a posteriori* analysis', *Florida Entomologist*, vol. 103, no. 2, 259–263, 2020. DOI: 10.1653/024.103.0217.
- [42] L. Rossini, M. Severini, M. Contarini and S. Speranza, 'EntoSim, a ROOT-based simulator to forecast insects' life cycle: Description and application in the case of *Lobesia botrana*', *Crop Protection*, vol. 129, p. 105024, 2020, ISSN: 02612194. DOI: 10.1016/j.cropro.2019.105024.
- [43] A. P. Gutierrez and J. Baumgärtner, 'Multitrophic models of predator-prey energetics: II. A realistic model of plant-herbivore-parasitoid-predator interactions', *The Canadian Entomologist*, vol. 116, no. 07, 933–949, 1984, ISSN: 0008-347X. DOI: 10.4039/Ent116933-7.
- [44] P. Damos and M. Savopoulou-Soultani, 'Temperature-driven models for insect development and vital thermal requirements', *Psyche*, vol. 2012, 2012, ISSN: 00332615. DOI: 10.1155/2012/123405.
- [45] T. Ikemoto and K. Kiritani, 'Novel method of specifying low and high threshold temperatures using thermodynamic SSI model of insect development', *Environmental Entomology*, vol. 48, no. 3, 479–488, 2019, ISSN: 0046-225X. DOI: 10.1093/ee/nvz031.

- [46] J. A. Logan, D. J. Wollkind, S. C. Hoyt and L. K. Tanigoshi, 'An analytic model for description of temperature dependent rate phenomena in arthropods', *Environmental Entomology*, vol. 5, no. 6, 1133–1140, 1976, ISSN: 1938-2936. DOI: 10.1093/ee/5.6.1133.
- [47] J.-F. Briere, P. Pracros, A.-Y. Le Roux and J.-S. Pierre, 'A novel rate model of temperature-dependent development for arthropods', *Environmental Entomology*, vol. 28, no. 1, 22–29, 1999, ISSN: 0046-225X. DOI: 10.1093/ee/28.1.22.
- [48] D. G. Harcourt, 'Development and use of life tables in study of natural insect populations', *Annual Review of Entomology*, vol. 14, no. 1, p. 175, 1969, ISSN: 0066-4170. DOI: 10.1146/annurev.en.14.010169.001135.
- [49] L. Rossini, M. Severini, M. Contarini and S. Speranza, 'Use of ROOT to build a software optimized for parameter estimation and simulations with Distributed Delay Model', *Ecological Informatics*, vol. 50, no. January, 184–190, 2019, ISSN: 15749541. DOI: 10.1016/j.ecoinf.2019.02.002.
- [50] R. Alilla, M. Severini and S. Pesolillo, 'Modello a ritardo distribuito a temperatura variabile per la simulazione dello sviluppo ontogenetico in stadi giovanili di popolazioni peciloterme', *Rivista Italiana di Agrometeorologia*, vol. 3, 30–33, 2005.
- [51] L. Rossini, S. Speranza and M. Contarini, 'Distributed delay model and von foerster's equation: Different points of view to describe insects' life cycles with chronological age and physiological time', *Ecological Informatics*, vol. 59, p. 101 117, 2020. DOI: 10.1016/j.ecoinf.2020.101117.
- [52] M. A. Mirhosseini, Y. Fathipour and G. V. P. Reddy, 'Arthropod development's response to temperature: A review and new software for modeling', *Annals of the Entomological Society of America*, vol. 110, no. 6, 507–520, 2017, ISSN: 0013-8746. DOI: 10.1093/aesa/sax071.
- [53] M. A. Toapanta, D. J. Schuster and P. A. Stansly, 'Development and life history of *Anthonomus eugenii* (Coleoptera: Curculionidae) at constant temperatures', *Environmental Entomology*, vol. 34, no. 5, 999–1008, 2005, ISSN: 0046-225X. DOI: 10.1093/ee/34.5.999.
- [54] V. A. Vassiliou, 'Effectiveness of insecticides in controlling the first and second generations of the *Lobesia botrana* (Lepidoptera: Tortricidae) in table grapes', *Journal of Economic Entomology*, vol. 104, no. 2, 580–585, 2011, ISSN: 00220493. DOI: 10.1603/EC10343.
- [55] F. Moshtaghi Maleki, S. Iranipour, M. J. Hejazi and M. Saber, 'Temperature-dependent age-specific demography of grapevine moth (*Lobesia botrana*) (Lepidoptera: Tortricidae): Jackknife vs. bootstrap techniques', *Archives of Phytopathology and Plant Protection*, vol. 49, no. 11–12, 263–280, 2016, ISSN: 0323-5408. DOI: 10.1080/03235408.2016.1140566.
- [56] H. Colinet, B. J. Sinclair, P. Vernon and D. Renault, 'Insects in fluctuating thermal environments', *Annual Review of Entomology*, vol. 60, no. 1, 123–140, 2015, ISSN: 0066-4170. DOI: 10.1146/annurev-ento-010814-021017.

- [57] H. Von Foerster, 'Some remarks on changing populations', *In The Kinetics of Cellular Proliferation*, vol. 382–407, 1959.
- [58] R. E. Plant and L. T. Wilson, 'Models for age structured populations with distributed maturation rates', *Journal of Mathematical Biology*, vol. 23, no. 2, 247–262, 1986, ISSN: 0303-6812. DOI: 10.1007/BF00276960.
- [59] S. Speranza, R. Alilla, S. Pesolillo and M. Severini, 'Modello previsionale del rischio di attacco del punteruolo rosso delle palme (*Rynchophorus ferrugineus*) nel lazio: Ipotesi di lavoro', in *Quarte giornate studio su metodi numerici, statistici e informatici nella difesa delle colture agrarie e delle foreste. Ricerca ed applicazioni.*, 2007, 29–31.
- [60] J. L. Capinera, *Handbook of vegetable pests*, J. L. Capinera, Ed. Elsevier, 2001, ISBN: 978-0-12-158861-8. DOI: 10.1016/B978-0-12-158861-8.X5000-5.
- [61] D. Pilarska, M. McManus, P. Pilarski, G. Georgiev, P. Mirchev and A. Linde, 'Monitoring the establishment and prevalence of the fungal entomopathogen *Entomophaga maimaiga* in two *Lymantria dispar* L. populations in Bulgaria', *Journal of Pest Science*, vol. 79, no. 2, 63–67, 2006, ISSN: 1612-4758. DOI: 10.1007/s10340-005-0111-8.
- [62] M. Contarini, L. Ruiiu, D. Pilarska and P. Luciano, 'Different susceptibility of indigenous populations of *Lymantria dispar* to the exotic entomopathogen *Entomophaga maimaiga*', *Journal of Applied Entomology*, vol. 140, no. 4, 317–321, 2016, ISSN: 09312048. DOI: 10.1111/jen.12267.
- [63] G. Alsaedi, A. Ashouri and R. Talaei-Hassanloui, 'Evaluation of *Bacillus thuringiensis* to control *Tuta absoluta* (Meyrick) (Lepidoptera: Gelechiidae) under laboratory conditions', *Agricultural Sciences*, vol. 08, no. 07, 591–599, 2017, ISSN: 2156-8553. DOI: 10.4236/as.2017.87045.
- [64] J. Gonzalez-Cabrera, O. Molla, H. Montón and A. Urbaneja, 'Efficacy of *Bacillus thuringiensis* (Berliner) in controlling the tomato borer, *Tuta absoluta* (Meyrick) (Lepidoptera: Gelechiidae)', *BioControl*, vol. 56, no. 1, 71–80, 2011, ISSN: 13866141. DOI: 10.1007/s10526-010-9310-1.
- [65] S. Speranza, C. M. Melo, G. M. Luna and E. G. Virla, 'First record of *Zelus obscuridorsis* (Hemiptera: Reduviidae) as a predator of the south american tomato leafminer, *Tuta absoluta* (Lepidoptera: Gelechiidae)', *Florida Entomologist*, vol. 97, no. 1, 295–297, 2014, ISSN: 00154040. DOI: 10.1896/054.097.0144.
- [66] R. Alilla, S. Speranza, T. Perovic, S. Hrnčić, S. Pesolillo, C. Pucci and M. Severini, 'Modello a ritardo variabile per la simulazione della fenologia e della demografia della *Bactrocera oleae* (Gmel.)(Diptera: Tephritidae) in due diversi ambienti olivicoli e in condizioni di aumento della temperatura', in *Quarte giornate studio su metodi numerici, statistici e informatici nella difesa delle colture agrarie e delle foreste. Ricerca ed applicazioni.*, 2007, 48–50.

- [67] S. P. Blythe, R. M. Nisbet, W. S. C. Gurney and N. MacDonald, 'Stability switches in distributed delay models', *Journal of Mathematical Analysis and Applications*, vol. 109, no. 2, 388–396, 1985, ISSN: 10960813. DOI: 10.1016/0022-247X(85)90158-1.
- [68] L. Galeano-Vasco, M. Ceron-Munoz, D. Rodriguez and J. M. Cotes, 'Using the distributed-delay model to predict egg production in laying hens', *Revista Colombiana de Ciencias Pecuarias*, vol. 26, 270–279, 2013.
- [69] C. Pucci and A. F. Spanedda, 'Performance comparison between two forecasting models of infestation caused by olive fruit fly (*Bactrocera oleae* Rossi)', *Pomologia Croatia*, vol. 12, no. 3, 193–202, 2006.
- [70] M. Severini, *Modello previsionale delle conseguenze dell'aumento della temperatura e della irradianza ultravioletta sulla dinamica di popolazioni zooplanctoniche di interesse in acquacoltura*. 2004.
- [71] Y. Wang, A. P. Gutierrez, G. Oster and R. Daxl, 'A population model for plant growth and development: Coupling cotton-herbivore interaction', *The Canadian Entomologist*, vol. 109, no. 10, 1359–1374, 1977, ISSN: 0008-347X. DOI: 10.4039/Ent1091359-10.
- [72] L. Li, W.-Q. Qin, Z.-L. Ma, W. Yan, S.-C. Huang and Z.-Q. Peng, 'Effect of temperature on the population growth of *Rhynchophorus ferrugineus* (Coleoptera: Curculionidae) on sugarcane', *Environmental entomology*, vol. 39, no. 3, 999–1003, 2010, ISSN: 0046225X. DOI: 10.1603/EN09316.
- [73] N. B. Osman, V. F. Wright and R. B. Mills, 'The effects of rearing temperatures on certain aspects of the biology of *Corcyra cephalonica* (Stainton), the Rice Moth', 99–106, 1984. [Online]. Available: <http://spiru.cgahr.ksu.edu/proj/iwcspp/pdf2/3/99.pdf>.
- [74] M. S. Oztokce, A. Bayindir and I. Karaca, 'Temperature-dependent development of the tomato leaf miner, *Tuta absoluta* (Meyrick) (Lepidoptera: Gelechiidae) on tomato plant *Lycopersicon esculentum* Mill. (Solanaceae)', *Turkish Journal of Entomology*, vol. 40, no. 1, 51–59, 2016, ISSN: 1010-6960. DOI: 10.16970/ted.64743.
- [75] A. Golizadeh, K. Kamali, Y. Fathipour and H. Abbasipour, 'Life table and temperature-dependent development of *Diadegma anurum* (Hymenoptera: Ichneumonidae) on its host *Plutella xylostella* (Lepidoptera: Plutellidae).', *Environmental entomology*, vol. 37, no. 1, 38–44, 2008, ISSN: 0046-225X. DOI: 10.1603/0046-225X(2008)37.
- [76] A. L. Nielsen, G. C. Hamilton and D. Matadha, 'Developmental rate estimation and life table analysis for *Halyomorpha halys* (Hemiptera: Pentatomidae)', *Environmental Entomology*, vol. 37, no. 2, 348–355, 2008.
- [77] S.-S. Liu, G.-M. Zhang and J. Zhu, 'Influence of temperature variations on rate of development in insects: Analysis of case studies from entomological literature', *Annals of the Entomological Society of America*, vol. 88, no. 2, 107–119, 1995, ISSN: 1938-2901.
- [78] M. Severini and G. Gilioli, 'Storia e filosofia dei modelli di simulazione nella difesa delle colture agrarie', *Notiziario sulla protezione delle piante*, 2002.

- [79] R. Brun and F. Rademakers, ‘ROOT - An object oriented data analysis framework’, *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 389, no. 1–2, 81–86, 1997, ISSN: 01689002. DOI: 10.1016/S0168-9002(97)00048-X.
- [80] G. L. Curry and R. M. Feldman, *Mathematical foundations of population dynamics*, First. Texas A&M University press college station, 1987, ISBN: 0-89096-256-1.
- [81] P. J. Sharpe and D. W. DeMichele, ‘Reaction kinetics of poikilotherm development’, *Journal of Theoretical Biology*, vol. 64, no. 4, 649–670, 1977, ISSN: 00225193. DOI: 10.1016/0022-5193(77)90265-X.
- [82] G. Di Cola and G. Gilioli, ‘Mathematical models for age-structured population dynamics: An overview’, in *20th International congress of Entomology, Florence*, 1996, 45–61.
- [83] J. Vansickle, ‘Analysis of a distributed-parameter population model based on physiological age’, *Journal of Theoretical Biology*, vol. 64, no. 3, 571–586, 1977, ISSN: 00225193. DOI: 10.1016/0022-5193(77)90289-2.
- [84] S. Pesolillo, M. Severini and R. Alilla, ‘Il modello simulatore di ritardo “delay simulator” versione 2.0: Un modello di simulazione su foglio di calcolo elettronico della demografia e della fenologia delle popolazioni naturali’, in *Metodi numerici, statistici e informatici nella difesa delle colture agrarie e delle foreste: ricerca e applicazioni*, 2004.
- [85] M. Jorgensen and M. Shepperd, ‘A systematic review of software development cost estimation studies’, *IEEE Transactions on Software Engineering*, vol. 33, no. 1, 33–53, 2007, ISSN: 0098-5589. DOI: 10.1109/TSE.2007.256943.
- [86] J. Grudin, ‘Interactive systems: Bridging the gaps between developers and users’, *Computer*, vol. 24, no. 4, 59–69, 1991, ISSN: 0018-9162. DOI: 10.1109/2.76263.
- [87] E. Von Hippel, ‘Innovation by user communities: Learning from open-source software’, *MIT Sloan Management Review*, vol. 42, no. 4, 82–86, 2001.
- [88] G. Câmara and F. Fonseca, ‘Information policies and open source software in developing countries’, *Journal of the American Society for Information Science and Technology*, vol. 58, no. 1, 121–132, 2007, ISSN: 15322882. DOI: 10.1002/asi.20444.
- [89] J. F. Liénard, D. Gravel and N. S. Strigul, ‘Data-intensive modeling of forest dynamics’, *Environmental Modelling & Software*, vol. 67, 138–148, 2015, ISSN: 13648152. DOI: 10.1016/j.envsoft.2015.01.010.
- [90] N. S. Strigul, I. Florescu, A. R. Welden and F. Michalczewski, ‘Modelling of forest stand dynamics using Markov chains’, *Environmental Modelling & Software*, vol. 31, 64–75, 2012, ISSN: 13648152. DOI: 10.1016/j.envsoft.2011.12.004.
- [91] R. Brun and F. Rademakers, *ROOT user’s guide*, 2007. [Online]. Available: <https://root.cern/root/htmldoc/guides/users-guide/ROOTUsersGuide.html>.

- [92] A. P. Gutierrez, L. Ponti and Q. A. Cossu, 'Effects of climate warming on olive and olive fly (*Bactrocera oleae* (gmelin)) in california and italy', *Climatic Change*, vol. 95, no. 1–2, 195–217, 2009, ISSN: 0165-0009. DOI: 10.1007/s10584-008-9528-4.
- [93] S. A. Estay, M. Lima and F. Bozinovic, 'The role of temperature variability on insect performance and population dynamics in a warming world', *Oikos*, vol. 123, no. 2, 131–140, 2014, ISSN: 00301299. DOI: 10.1111/j.1600-0706.2013.00607.x.
- [94] P. E. Hulme, 'Trade, transport and trouble: Managing invasive species pathways in an era of globalization', *Journal of Applied Ecology*, vol. 46, no. 1, 10–18, 2009, ISSN: 00218901. DOI: 10.1111/j.1365-2664.2008.01600.x.
- [95] C. Robinet, H. Kehlenbeck, D. J. Kriticos, R. H. A. Baker, A. Battisti, S. Brunel, M. Dupin, D. Eyre, M. Faccoli, Z. Ilieva and et al., 'A suite of models to support the quantitative assessment of spread in pest risk analysis', vol. 7, no. 10, 2012. DOI: 10.1371/journal.pone.0043366.
- [96] G.-r. Walther, A. Roques, P. E. Hulme, M. T. Sykes, P. Pyšek, C. Robinet and V. Semchenko, 'Alien species in a warmer world: Risks and opportunities', *Trends in Ecology and Evolution*, vol. 24, no. 12, 686–693, 2009. DOI: 10.1016/j.tree.2009.06.008.
- [97] M. P. Hill, S. Clusella-Trullas, S. Terblanche John and D. M. Richardson, 'Drivers, impacts, mechanisms and adaptation in insect invasions', *Biological Invasions*, no. 2002, 2016, ISSN: 1573-1464. DOI: 10.1007/s10530-016-1088-3.
- [98] A. Roques, M.-A. Auger-Rozenberg, T. M. Blackburn, J. Garnas, P. Pyšek, W. Rabitsch, D. M. Richardson, M. J. Wingfield, A. M. Liebhold and R. P. Duncan, 'Temporal and interspecific variation in rates of spread for insect species invading europe during the last 200 years', *Biological Invasions*, vol. 18, no. 4, 907–920, 2016, ISSN: 1387-3547. DOI: 10.1007/s10530-016-1080-y.
- [99] H. A. Mooney and E. E. Cleland, 'The evolutionary impact of invasive species', in *Proceedings of the National Academy of Sciences of the USA 98*, vol. 98, 2001, 5446–5451.
- [100] D. Pimentel, *Biological invasions: economic and environmental costs of alien plant, animal and microbe species*, 2nd. Boca Raton, FL: CRC Press, 2011, 2011.
- [101] D. DeAngelis and M. Huston, 'Effects of growth rates in models of size distribution formation in plants and animals', *Ecological Modelling*, vol. 36, no. 1–2, 119–137, 1987, ISSN: 03043800. DOI: 10.1016/0304-3800(87)90062-7.
- [102] J. W. Sinko and W. Streifer, 'A new model for age-size structure of a population', *Ecology*, vol. 48, no. 6, 910–918, 1967, ISSN: 00129658. DOI: 10.2307/1934533.
- [103] A. P. Gutierrez, L. A. Falcon, W. Loew, P. A. Leipzig and R. Van Den Bosch, 'An analysis of cotton production in California: A model for Acala cotton and the effects of defoliators on its yields', *Environmental Entomology*, vol. 4, no. 1, 125–136, 1975.

- [104] A. P. Gutierrez, D. W. Williams and H. Kido, 'A model of grape growth and development: The mathematical structure and biological considerations', *Crop Science*, vol. 25, no. 5, p. 721, 1985, ISSN: 0011-183X. DOI: 10.2135/cropsci1985.0011183X002500050001x.
- [105] A. Cappio Borlino, G. Di Cola and G. Marras, 'I modelli compartimentali nello studio della dinamica delle popolazioni naturali', *Bollettino della società sarda di scienze naturali*, vol. 27, 77–114, 1990.
- [106] A. M. Lutambi, M. A. Penny, T. Smith and N. Chitnis, 'Mathematical modelling of mosquito dispersal in a heterogeneous environment', *Mathematical Biosciences*, vol. 241, no. 2, 198–216, 2013, ISSN: 00255564.
- [107] T. J. Manetsch, 'Bilateral distributed delays and their use in modeling classes of distributed parameter processes', *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 10, no. 2, 61–67, 1980, ISSN: 0018-9472. DOI: 10.1109/TSMC.1980.4308434.
- [108] M. Severini, J. Baumgärtner and L. Limonta, 'Parameter estimation for distributed delay based population models from laboratory data: Egg hatching of *Oulema duftschmiedi* Redthenbacher (Coleoptera: Chrysomelidae) as an example', *Ecological Modelling*, vol. 167, no. 3, 233–246, 2003, ISSN: 03043800. DOI: 10.1016/S0304-3800(03)00188-1.
- [109] A. G. McKendrick, 'Applications of mathematics to medical problems', in *Edinburgh Mathematical Society*, 1926, 98–130, ISBN: 1464-3839. DOI: 10.1017/S0013091500034428.
- [110] G. Buffoni, G. Di Cola and A. Ugolini, 'Numerical methods for the solution of PDE describing the stochastic development of an age-structured population', *Computer science and mathematical methods in plant protection - Proceeding of the International Workshop*, 1990.
- [111] J.-F. Briere and P. Pracros, 'Comparison of temperature-dependent growth models with the development of *Lobesia botrana* (Lepidoptera: Tortricidae)', *Environmental Entomology*, vol. 27, no. 1, 94–101, 1998, ISSN: 1938-2936. DOI: 10.1093/ee/27.1.94.
- [112] C. R. Preto, D. E. Bellamy, S. S. Walse and F. G. Zalom, 'Predicting larval stage distribution of *Lobesia botrana* (Lepidoptera: Tortricidae) at three constant temperatures', *Journal of Economic Entomology*, vol. 112, no. 2, 585–590, 2019, ISSN: 1938291X. DOI: 10.1093/jee/toy374.
- [113] G. M. Luna, N. E. Sánchez, P. C. Pereyra, E. Nieves, V. Savino, E. Luft, E. G. Virla and S. Speranza, 'Biological control of *Tuta absoluta* in Argentina and Italy: Evaluation of indigenous insects as natural enemies', *EPPO Bulletin*, vol. 42, no. 2, 260–267, 2012, ISSN: 02508052. DOI: 10.1111/epp.2564.
- [114] EPPO, EPPO web page, 2020. [Online]. Available: <https://gd.eppo.int/taxon/>.
- [115] A. Biondi, R. N. C. Guedes, F.-H. Wan and N. Desneux, 'Ecology, worldwide spread, and management of the invasive south american tomato pinworm, *Tuta absoluta*: Past, present, and future', *Annual Review of Entomology*, vol. 63, no. 1, 239–258, 2018, ISSN: 0066-4170. DOI: 10.1146/annurev-ento-031616-034933.

- [116] S. Speranza and L. Sannino, 'The current status of *Tuta absoluta* in Italy', *EPPO Bulletin*, vol. 42, no. 2, 328–332, 2012, ISSN: 02508052. DOI: 10.1111/epp.2579.
- [117] A. Cocco, S. Deliperi, A. Lentini, R. Mannu and G. Delrio, 'Seasonal phenology of *Tuta absoluta* (Lepidoptera: Gelechiidae) in protected and open-field crops under Mediterranean climatic conditions', *Phytoparasitica*, vol. 43, no. 5, 713–724, 2015, ISSN: 18767184. DOI: 10.1007/s12600-015-0486-x.
- [118] H. E. Z. Tonnang, S. F. Mohamed, F. Khamis and S. Ekesi, 'Identification and risk assessment for worldwide invasion and spread of *Tuta absoluta* with a focus on sub-Saharan Africa: Implications for phytosanitary measures and management', *PLOS ONE*, vol. 10, no. 8, R. N. C. Guedes, Ed., 1–19, 2015, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0135283.
- [119] C. A. Marchioro, F. S. Krechemer and L. A. Foerster, 'Estimating the development rate of the tomato leaf miner, *Tuta absoluta* (Lepidoptera: Gelechiidae), using linear and non-linear models.', *Pest management science*, vol. 73, no. 7, 1486–1493, 2017, ISSN: 1526-4998. DOI: 10.1002/ps.4484.
- [120] L. Sannino, 'Inquadramento sistematico e aspetti morfo-biologici di *Tuta absoluta* (Meyrick) (Lepidoptera, Gelechiidae)', *Atti Accademia Nazionale Italiana di Entomologia Anno 2012*, vol. LX, 67–75, 2012.
- [121] S. Speranza, E. Colonnelli, A. P. Garonna and S. Laudonia, 'First record of *Anthonomus eugenii* (Coleoptera: Curculionidae) in Italy', *Florida Entomologist*, vol. 97, no. 2, 844–845, 2014, ISSN: 0015-4040. DOI: 10.1653/024.097.0275.
- [122] L. Ruiu, 'Microbial biopesticides in agroecosystems', *Agronomy*, vol. 8, no. 11, p. 235, 2018, ISSN: 2073-4395. DOI: 10.3390/agronomy8110235.
- [123] R. Mannu, G. Gilioli and P. Luciano, 'Occupancy of the territory by *Lymantria dispar* (L.) (Lepidoptera Erebiidae) egg masses as a predictive index of damage', *Redia*, vol. 100, no. April, 157–165, 2017, ISSN: 03704327. DOI: 10.19263/REDIA-100.17.20.
- [124] M. Severini, R. Alilla and S. Pesolillo, *Modello previsionale delle conseguenze dell'aumento della temperatura e dell'irradianza ultravioletta sulla dinamica di popolazioni zooplanctoniche di interesse in acquacoltura*. 2006.
- [125] E. Trucco, 'Mathematical models for cellular systems. the Von Foerster equation. Part I', *The Bulletin of Mathematical Biophysics*, vol. 27, no. 3, 449–471, 1965, ISSN: 00074985. DOI: 10.1007/BF02476849.
- [126] Cern, ROOT *cern web page*, 2020. [Online]. Available: <http://root.cern.ch>.
- [127] J. A. Logan and D. W. Hilbert, 'Modeling the effects of temperature on arthropod population systems', in *Developments in Environmental Modelling*. 1983, 113–122. DOI: 10.1016/B978-0-444-42179-1.50017-1.

- [128] R. M. Schoolfield, P. J. Sharpe and C. E. Magnuson, 'Non-linear regression of biological temperature-dependent rate models based on absolute reaction-rate theory.', *Journal of theoretical biology*, vol. 88, no. 4, 719–31, 1981, ISSN: 0022-5193. DOI: 10.1016/0022-5193(81)90246-0.
- [129] R. Sipahi, A. M. Fatican and S. I. Niculescu, 'Stability of traffic flow behavior with distributed delays modeling the memory effects of the drivers', *Society for Industrial and Applied Mathematics*, vol. 68, no. 3, 738–759, 2007. DOI: 10.1137/0606738.
- [130] B. Keyfitz and N. Keyfitz, 'The McKendrick partial differential equation and its uses in epidemiology and population study', *Mathematical and Computer Modelling*, vol. 26, no. 6, 1–9, 1997, ISSN: 08957177. DOI: 10.1016/S0895-7177(97)00165-9.
- [131] O. Diekmann, H. A. Lauwerier, T. Aldenberg and J. A. J. Metz, 'Growth, fission and the stable size distribution', *Journal of Mathematical Biology*, vol. 18, no. 2, 135–148, 1983, ISSN: 03036812. DOI: 10.1007/BF00280662.
- [132] D. S. Saunders and R. D. Lewis, 'A damped circadian oscillator model of an insect photoperiodic clock. I. Description of the model based on a feedback control system', *Journal of Theoretical Biology*, vol. 128, no. 1, 47–59, 1987, ISSN: 10958541. DOI: 10.1016/S0022-5193(87)80030-9.
- [133] D. B. Silva, V. H. Bueno, J. C. Lins and J. C. Van Lenteren, 'Life history data and population growth of *Tuta absoluta* at constant and alternating temperatures on two tomato lines', *Bulletin of Insectology*, vol. 68, no. 2, 223–232, 2015, ISSN: 17218861.
- [134] B. Ainseba, D. Picart and D. Thiéry, 'An innovative multistage, physiologically structured, population model to understand the European grapevine moth dynamics', *Journal of Mathematical Analysis and Applications*, vol. 382, no. 1, 34–46, 2011, ISSN: 0022247X. DOI: 10.1016/j.jmaa.2011.04.021.
- [135] V. Castex, M. Beniston, P. Calanca, D. Fleury and J. Moreau, 'Pest management under climate change: The importance of understanding tritrophic relations', *Science of The Total Environment*, vol. 616–617, 397–407, 2018, ISSN: 00489697. DOI: 10.1016/j.scitotenv.2017.11.027.
- [136] R. R. Nemani, M. A. White, D. R. Cayan, G. V. Jones, S. W. Running, J. C. Coughlan and D. L. Peterson, 'Asymmetric warming over coastal california and its impact on the premium wine industry', *Climate Research*, vol. 19, no. 1, 25–34, 2001, ISSN: 0936577X. DOI: 10.3354/cr019025.
- [137] F. Pavan, C. Floreani, P. Barro, P. Zandigiacomo and L. Dalla Montà, 'Occurrence of two different development patterns in *Lobesia botrana* (Lepidoptera: Tortricidae) larvae during the second generation', *Agricultural and Forest Entomology*, vol. 15, no. 4, 398–406, 2013, ISSN: 14619555. DOI: 10.1111/afe.12027.
- [138] D. S. Pureswaran, A. Roques and A. Battisti, 'Forest insects and climate change', *Current Forestry Reports*, vol. 4, no. 2, 35–50, 2018, ISSN: 2198-6436. DOI: 10.1007/s40725-018-0075-6.

- [139] K. M. Addesso, H. J. McAuslane, P. A. Stansly and D. J. Schuster, 'Host-marking by female pepper weevils, *Anthonomus eugenii*', *Entomologia Experimentalis et Applicata*, vol. 125, no. 3, 269–276, 2007, ISSN: 00138703. DOI: 10.1111/j.1570-7458.2007.00626.x.
- [140] R. J. Patrock and D. J. Schuster, 'Feeding, oviposition and development of the pepper weevil, (*Anthonomus eugenii* Cano), on selected species of Solanaceae', *Tropical Pest Management*, vol. 38, no. 1, 65–69, 1992, ISSN: 01436147. DOI: 10.1080/09670879209371648.
- [141] J. L. Capinera, 'Pepper weevil, *Anthonomus eugenii* Cano (Insecta: Coleoptera: Curculionidae)', *UF IFAS Extension University of Florida*, 1–6, 2011.
- [142] P. Wu, M. Haseeb, W. Diedrick, H. Ouyang, R. Zhang, L. H. B. Kanga and J. C. Legaspi, 'Influence of plant direction, layer, and spacing on the infestation levels of *Anthonomus eugenii* (Coleoptera: Curculionidae) in open jalapeño pepper fields in North Florida', *Florida Entomologist*, vol. 102, no. 3, p. 501, 2019, ISSN: 0015-4040. DOI: 10.1653/024.102.0319.
- [143] R. Rupnik, M. Kukar, P. Vračar, D. Košir, D. Pevec and Z. Bosnić, 'Agrodss: A decision support system for agriculture and farming', *Computers and Electronics in Agriculture*, vol. 161, no. November 2017, 260–271, 2019, ISSN: 01681699. DOI: 10.1016/j.compag.2018.04.001.
- [144] P. S. Carberry, Z. Hochman, R. L. McCown, N. P. Dalgliesh, M. A. Foale, P. L. Poulton, J. N. G. Hargreaves, D. M. G. Hargreaves, S. Cawthray, N. Hillcoat and et al., 'The farmscape approach to decision support: Farmers', advisers', researchers', monitoring, simulation, communication and performance evaluation', *Agricultural Systems*, vol. 74, 141–177, 2002.
- [145] S. Tang and R. A. Cheke, 'Models for integrated pest control and their biological implications', *Mathematical Biosciences*, vol. 215, no. 1, 115–125, 2008, ISSN: 00255564. DOI: 10.1016/j.mbs.2008.06.008.
- [146] M. V. Chiriaco, C. Belli, T. Chiti, C. Trotta and S. Sabbatini, 'The potential carbon neutrality of sustainable viticulture showed through a comprehensive assessment of the greenhouse gas (GHG) budget of wine production', *Journal of Cleaner Production*, vol. 225, 435–450, 2019, ISSN: 09596526. DOI: 10.1016/j.jclepro.2019.03.192.
- [147] E. Schimmenti, G. Migliore, C. P. Di Franco and V. Borsellino, 'Is there sustainable entrepreneurship in the wine industry? Exploring Sicilian wineries participating in the SOSustain program', *Wine Economics and Policy*, vol. 5, no. 1, 14–23, 2016, ISSN: 22129774. DOI: 10.1016/j.wep.2016.05.001.
- [148] A. Valero, J. Howarter and J. Sutherland, 'Sustainable wine scoring system (SWSS): A life cycle assessment (LCA) multivariable approach', *BIO Web of Conferences*, vol. 12, J.-M. Aurand, Ed., p. 03016, 2019, ISSN: 2117-4458. DOI: 10.1051/bioconf/20191203016.

- [149] H. Steinitz, A. Sadeh, A. Kliot and A. Harari, 'Effects of radiation on inherited sterility in the European grapevine moth (*Lobesia botrana*)', *Pest Management Science*, vol. 71, no. 1, 24–31, 2015, ISSN: 1526498X. DOI: 10.1002/ps.3797.
- [150] V. Partel, S. Charan Kakarla and Y. Ampatzidis, 'Development and evaluation of a low-cost and smart technology for precision weed management utilizing artificial intelligence', *Computers and Electronics in Agriculture*, vol. 157, no. November 2018, 339–350, 2019, ISSN: 01681699. DOI: 10.1016/j.compag.2018.12.048.
- [151] N. Holst, 'A universal simulator for ecological models', *Ecological Informatics*, vol. 13, 70–76, 2013, ISSN: 15749541. DOI: 10.1016/j.ecoinf.2012.11.001.
- [152] R. Criollo, V. Velasco, A. Nardi, L. Manuel de Vries, C. Riera, L. Scheiber, A. Jurado, S. Brouyère, E. Pujades, R. Rossetto and et al., 'AkvaGIS: An open source tool for water quantity and quality management', *Computers & Geosciences*, vol. 127, no. November 2018, 123–132, 2019, ISSN: 00983004. DOI: 10.1016/j.cageo.2018.10.012.
- [153] D. Kpienbaareh, M. Kansanga and I. Luginaah, 'Examining the potential of open source remote sensing for building effective decision support systems for precision agriculture in resource-poor settings', *GeoJournal*, vol. 0123456789, 2018, ISSN: 0343-2521. DOI: 10.1007/s10708-018-9932-x.
- [154] R. Rossetto, G. De Filippis, I. Borsi, L. Foglia, M. Cannata, R. Criollo and E. Vázquez-Suñé, 'Integrating free and open source tools and distributed modelling codes in GIS environment for data-based groundwater management', *Environmental Modelling & Software*, vol. 107, no. May, 210–230, 2018, ISSN: 13648152. DOI: 10.1016/j.envsoft.2018.06.007.
- [155] M. Bange, S. Deutscher, D. Larsen, D. Linsley and S. Whiteside, 'A handheld decision support system to facilitate improved insect pest management in Australian cotton systems', *Computers and Electronics in Agriculture*, vol. 43, no. 2, 131–147, 2004, ISSN: 01681699. DOI: 10.1016/j.compag.2003.12.003.
- [156] P. Flinn, D. Hagstrum, C. Reed and T. Phillips, 'Stored grain advisor pro: Decision support system for insect management in commercial grain elevators', *Journal of Stored Products Research*, vol. 43, no. 4, 375–383, 2007, ISSN: 0022474X. DOI: 10.1016/j.jspr.2006.09.004.
- [157] P. Cravedi and E. Mazzoni, 'Somme termiche per la previsione degli sfarfallamenti di *Lobesia botrana* Den. & Schiff. (Lepidoptera, Tortricidae).', *Redia*, vol. 73, no. 1, 17–40, 1990.
- [158] A. Ferrara, L. Salvati, A. Sateriano, M. Carlucci, I. Gitas and R. Biasi, 'Unraveling the 'stable' landscape: A multi-factor analysis of unchanged agricultural and forest land (1987–2007) in a rapidly-expanding urban region', *Urban Ecosystems*, vol. 19, no. 2, 835–848, 2016, ISSN: 15731642. DOI: 10.1007/s11252-015-0509-x.

- [159] A. Lucchi, P. L. Scaramozzino, G. Michl, A. Loni and C. Hoffmann, 'The first record in Italy of *Trichogramma cordubense* Vargas & Cabello 1985 (Hymenoptera: Trichogrammatidae) emerging from the eggs of *Lobesia botrana* (Denis & Schiffermüller, 1775) (Lepidoptera: Tortricidae)', *Vitis - Journal of Grapevine Research*, vol. 55, no. 4, 161–164, 2016, ISSN: 00427500. DOI: 10.5073/vitis.2016.55.161-164.
- [160] C. Ioriatti, B. Bagnoli, A. Lucchi and V. Veronelli, 'Vine moths control by mating disruption in Italy: Results and future prospects', *Redia*, vol. 87 (2004), no. January, 117–128, 2005, ISSN: 0370-4327.
- [161] L. M. Thompson, T. M. Faske, N. Banahene, D. Grim, S. J. Agosta, D. Parry, P. C. Tobin, D. M. Johnson and K. L. Grayson, 'Variation in growth and developmental responses to supraoptimal temperatures near latitudinal range limits of gypsy moth *Lymantria dispar* (L.), an expanding invasive species', *Physiological Entomology*, vol. 42, no. 2, 181–190, 2017, ISSN: 03076962. DOI: 10.1111/phen.12190.
- [162] D. Mazzi and S. Dorn, 'Movement of insect pests in agricultural landscapes', *Annals of Applied Biology*, vol. 160, no. 2, 97–113, 2012, ISSN: 00034746. DOI: 10.1111/j.1744-7348.2012.00533.x.
- [163] G. G. Kennedy and N. P. Storer, 'Life systems of polyphagous arthropod pests in temporally unstable cropping systems', *Annual Review of Entomology*, vol. 45, no. 1, 467–493, 2000, ISSN: 0066-4170. DOI: 10.1146/annurev.ento.45.1.467.
- [164] L. Rossini, M. Severini, M. Contarini and S. Speranza, 'Distributed Delay Models: A proposal of application in urban context to forecast pest insects' life cycle', in *Environmental and territorial modelling for planning and design*, A. Leone and C. Gargiulo, Eds. Naples: FedOAPress, 2018, 169–178, ISBN: 9788868870485. DOI: 10.6093/978-88-6887-048-5.
- [165] A. Guglielmino, F. Modola, E. Scarici, S. Speranza and C. Bückle, 'The auchenorrhyncha fauna (Insecta: Hemiptera) of Villa Lante, Bagnaia (Italy): A study of an urban ecosystem', *Bulletin of Insectology*, vol. 68, no. 2, 239–253, 2015, ISSN: 17218861.
- [166] M. J. Raupp, J. J. Holmes, C. Sadof, P. Shrewsbury and J. A. Davidson, 'Effects of cover sprays and residual pesticides on scale insects and natural enemies in urban forests', *Journal of Arboriculture*, vol. 27, no. 4, 203–214, 2001, ISSN: 02785226.
- [167] B. Bagnoli and A. Lucchi, 'Parasitoids of *Lobesia botrana* (Den. & Schiff.) in Tuscany.', *IOBC/WPRS Bulletin*, vol. 29, no. 11, 139–142, 2006.
- [168] H. Fraga, J. A. Santos, J. Moutinho-Pereira, C. Carlos, J. Silvestre, J. Eiras-Dias, T. Mota and A. C. Malheiro, 'Statistical modelling of grapevine phenology in portuguese wine regions: Observed trends and climate change projections', *Journal of Agricultural Science*, vol. 154, no. 5, 795–811, 2016, ISSN: 14695146. DOI: 10.1017/S0021859615000933.
- [169] C. Bellard, C. Bertelsmeier, P. Leadley, W. Thuiller and F. Courchamp, 'Impacts of climate change on the future of biodiversity', *Ecology Letters*, vol. 15, no. 4, 365–377, 2012, ISSN: 1461023X. DOI: 10.1111/j.1461-0248.2011.01736.x.

- [170] F. Pavan, G. Stefanelli, A. Villani and E. Cargnus, ‘Influence of grapevine cultivar on the second generations of *Lobesia botrana* and *Eupoecilia ambiguella*’, *Insects*, vol. 9, no. 1, p. 8, 2018, ISSN: 2075-4450. DOI: 10.3390/insects9010008.
- [171] T. Chou and C. D. Greenman, ‘A hierarchical kinetic theory of birth, death and fission in age-structured interacting populations’, *Journal of Statistical Physics*, vol. 164, 49–76, 2016, ISSN: 0022-4715. DOI: 10.1007/s10955-016-1524-x.
- [172] E. Trucco, ‘Mathematical models for cellular systems. The Von Foerster equation. Part I’, *The Bulletin of Mathematical Biophysics*, vol. 27, no. 3, 449–471, 1965, ISSN: 00074985. DOI: 10.1007/BF02476849.
- [173] R. M. Feldman and G. L. Curry, ‘A PDE formulation and numerical solution for a boll weevil-cotton crop model’, *Computers and Mathematics with Applications*, vol. 9, no. 3, 393–402, 1983, ISSN: 08981221. DOI: 10.1016/0898-1221(83)90054-8.
- [174] F. Rojano, L. A. Ibarra-Juárez, A. Lira-Noriega, F. Escobar-Sarria and D. González-Tokman, ‘Application of a modified McKendrick-Von Foerster equation to predict beetle population dynamics (*Xyleborus affinis*) under artificial medium in growth chambers’, *AgEng Conference*, no. July, 1–8, 2018.
- [175] C. Iltis, G. Martel, D. Thiéry, J. Moreau and P. Louâpre, ‘When warmer means weaker: High temperatures reduce behavioural and immune defences of the larvae of a major grapevine pest’, *Journal of Pest Science*, vol. 91, no. 4, 1315–1326, 2018, ISSN: 1612-4758. DOI: 10.1007/s10340-018-0992-y.
- [176] L. Delbac and D. Thiéry, ‘*Lobesia botrana* (Lepidoptera: Tortricidae) larval population assessment by damage to grape flowers: Could empty larval nests monitoring be useful?’, in *IOBC-WPRS Bulletin*, 2017, 37–44.
- [177] D. Thiéry, K. Monceau and J. Moreau, ‘Larval intraspecific competition for food in the European grapevine moth *Lobesia botrana*’, *Bulletin of Entomological Research*, vol. 104, no. 4, 517–524, 2014, ISSN: 0007-4853. DOI: 10.1017/S0007485314000273.
- [178] A. Verpy, F. Gil, S. Mary, L. Delbac and D. Thiéry, ‘Temporal differences in *Lobesia botrana*’s lifecycle at local scale, the example of the saint emilion vineyard’, in *IOBC-WPRS Bulletin*, vol. 105, 2014, 197–204.
- [179] L. G. Varela, R. J. Smith, M. L. Cooper and R. W. Hoenisch, ‘European grapevine moth, *Lobesia botrana*, in Napa Valley vineyards’, *Practical Winery & Vineyard*, no. March/April, 1–5, 2010.
- [180] E. Brunori, R. Farina and R. Biasi, ‘Sustainable viticulture: The carbon-sink function of the vineyard agro-ecosystem’, *Agriculture, Ecosystems and Environment*, vol. 223, 10–21, 2016, ISSN: 01678809. DOI: 10.1016/j.agee.2016.02.012.
- [181] T. Pennington, J. M. Reiff, K. Theiss, M. H. Entling and C. Hoffmann, ‘Reduced fungicide applications improve insect pest control in grapevine’, *BioControl*, vol. 63, no. 5, 687–695, 2018, ISSN: 1386-6141. DOI: 10.1007/s10526-018-9896-2.

- [182] D. S. Saunders, ‘Insect photoperiodism: Effects of temperature on the induction of insect diapause and diverse roles for the circadian system in the photoperiodic response’, *Entomological Science*, vol. 17, no. 1, 25–40, 2014, ISSN: 13438786. DOI: 10.1111/ens.12059.
- [183] E. Tauber and B. P. Kyriacou, ‘Insect photoperiodism and circadian clocks: Models and mechanisms’, *Journal of Biological Rhythms*, vol. 16, no. 4, 381–390, 2001, ISSN: 07487304. DOI: 10.1177/074873001129002088.
- [184] S. W. Gillanders and D. S. Saunders, ‘A coupled pacemaker-slave model for the insect photoperiodic clock: Interpretation of ovarian diapause data in *Drosophila melanogaster*’, *Biological Cybernetics*, vol. 49, 149–154, 1984, ISSN: 19326203. DOI: 10.1371/journal.pone.0046143.
- [185] J. Baumgärtner, A. P. Gutierrez, S. Pesolillo and M. Severini, ‘A model for the overwintering process of European grapevine moth *Lobesia botrana* (Denis & Schiffermüller) (Lepidoptera: Tortricidae) populations’, *Journal of Entomological and Acarological Research*, vol. 44, no. 1, p. 2, 2012, ISSN: 2279-7084. DOI: 10.4081/jear.2012.e2.