Estimation of the mortality rate functions from time series field data in a stage-structured demographic model for Lobesia botrana
Estimation of the mortality rate functions from time series field data in a stage-structured demographic model for *Lobesia botrana*

Sara Pasquali, Cinzia Soresina
Corresponding author:

Sara Pasquali
Istituto di Matematica Applicata e Tecnologie Informatiche “E. Magenes” (IMATI-CNR)
Consiglio Nazionale delle Ricerche
Via Alfonso Corti 12, 20133 MILANO
E-mail: sara.pasquali@mi.imati.cnr.it

Cinzia Soresina
Department of Mathematics
Technical University of Munich
Boltzmannstr. 3, 85748 Garching bei München, Germany
E-mail: soresina@ma.tum.de
Abstract.

The estimation of the mortality rate function for a stage-structured population is obtained starting from time-series field data on the abundance of the species. The method is based on the formulation of the mortality as a combination of cubic splines and it is applied to the case of Lobesia botrana, the main pest in the European vineyards, with data collected in a location in the North of Italy. Mortality estimates are based on 3 years of data and are used to obtain the dynamics for two different years. These dynamics give a satisfactory fit of the phenology of the pest. The method presented allows to obtain more flexible shape for the mortality rate functions compared with previously methods applied for the same pest.

Keywords: Mortality rate function; stage-structured population; Lobesia botrana
Estimation of the mortality rate functions from time series field data in a stage-structured demographic model for *Lobesia botrana*

Sara Pasquali$^1$ and Cinzia Soresina$^2$

$^1$CNR-IMATI “Enrico Magenes”, via Alfonso Corti 12, 20133 Milano, Italy
e-mail: sara.pasquali@mi.imati.cnr.it

$^2$Department of Mathematics - Technical University of Munich
Boltzmannstr. 3, 85748 Garching bei München, Germany
e-mail: soresina@ma.tum.de

**Abstract**

The estimation of the mortality rate function for a stage-structured population is obtained starting from time-series field data on the abundance of the species. The method is based on the formulation of the mortality as a combination of cubic splines and it is applied to the case of *Lobesia botrana*, the main pest in the European vineyards, with data collected in a location in the North of Italy. Mortality estimates are based on 3 years of data and are used to obtain the dynamics for two different years. These dynamics give a satisfactory fit of the phenology of the pest.

The method presented allows to obtain more flexible shape for the mortality rate functions compared with previously methods applied for the same pest.

**Keywords:** Mortality rate function; stage-structured population; *Lobesia botrana*

1 Introduction

Population dynamics models play an important role in pest control. A good knowledge of the temporal dynamics of a pest population can help decision makers in the choice of the best strategy in terms of application of phytosanitary treatments. This is a fundamental task in light of the Directive 2009/128/EC on the sustainable use of pesticides in Europe.

To obtain a good description of the population dynamics it is necessary to take into account climatic factors, phenology of the plant and, in general, physical-biological characteristics of the environment in the site of interest. The population dynamics can be represented using both a phenological model that describes the percentage of individuals in the different stages and a demographic
model that accounts for the population abundance in time. Effects of mortality and fecundity rate functions in phenological models are discussed in Pasquali et al. [20]. Physiologically based demographic models allow to know the population abundance over time, taking into account the environmental variables influencing the dynamics of a species. These mechanistic models have been used since along time (see, for example, [7, 12, 21, 19, 18]). Often, it is useful to consider the population organized in stages because pests are dangerous for the crop when they are in a particular phase of their life. The model here considered describes an age-structured population and it is a particular case of a more general model presented in [3]. This kind of models have been used for various pests in the last years [13, 11, 16]. The model gives the abundance of the population in each stage in time and physiological age. It is based on biodemographic functions (development, mortality and fecundity) describing the biology of the species. Development, mortality and fecundity rate functions depend on environmental variables, mainly temperature.

It is important to have a good estimate of the biodemographic functions to obtain a reliable model. In general, these rate functions are estimated starting from literature data on the biology of the species (for example, duration in a stage for the development, number of eggs produced by an adult female for the fecundity). Starting from these data, a simple least square method allows to estimate the parameters of a biodemographic function of a given functional form. Unfortunately, for the mortality function data are often not available in literature. In this case, different methods of estimation have to be applied (see, e.g., [23] for a survey of mortality estimation methods). In case of absence of data on the mortality rate, mortality estimate can rely on the knowledge of time series data on population dynamics. Different methods to estimate mortality, starting from population dynamics time series data, have been proposed in the last years. Ellner et al. [9] proposed a non-parametric regression model, in [11] a method based on least squares is presented, while in [16] a Bayesian estimation method is proposed. In the two last approaches it is required a functional form for the mortality and the estimate concerns only parameters present in this function. This is restrictive, then to avoid heavy constraints on the mortality, we decided to follow the approach proposed by Wood [22] that does not require a functional form, but express the mortality as linear combination of elements of a suitable basis. The coefficients of the linear combination are estimated by minimizing a weighted least squares term that measures the “distance” between the simulated and the collected population abundance.

The mortality estimation method is applied to the case study of the grape berry moth *Lobesia botrana* which is considered the most dangerous pest in European vineyard. Data on population abundances were collected in Colognola ai Colli (Verona, Italy) in the period 2008-2012 for the cultivar Garganega. The method allows to know the behaviour of the mortality rates as function of the temperature and to forecast the population abundance in future periods.

The paper is organized as follows. In Section 2 the mathematical model describing the dynamics of the population is presented, in Section 3 the biodemographic functions for the grape berry moth are specified, in Section 4 is described
the mortality estimation method and results for the grape berry moth mortality are reported. Finally, Section 5 is devoted to discussion and concluding remarks.

2 The mathematical model

2.1 The stage-structured population model

The demographic model is based on a system of partial differential equations that allows to obtain the temporal dynamics of the stage-structured population and their distribution on physiological age within each stage. Let

\[ \phi^i(t, x) \, dx = \text{number of individuals in stage } i \text{ at time } t \]

with age in \((x, x + dx)\),

\(i = 1, 2, ..., s\), where \(s\) is the number of stages. Stages from 1 to \(s - 1\) are immature stages, and stage \(s\) represents the reproductive stage (adult individuals). Note that \(t\) denotes the chronological time while \(x\) is a developmental index which represents the physiological age indicating the development over time [3, 4, 5, 8].

Instead of a deterministic setting in which the population dynamics is described through von Foerster equations [3], we prefer to consider a stochastic approach which allows to take into account the variability of the development rate among the individuals [4, 5]. The dynamics is described in terms of the forward Kolmogorov equations [10, 6]

\[ \frac{\partial \phi^i}{\partial t} + \frac{\partial}{\partial x} \left[ v^i(t) \phi^i - \sigma^i \frac{\partial \phi^i}{\partial x} \right] + m^i(t) \phi^i = 0, \quad t > t_0, \quad x \in (0, 1), \]  

(1)

\[ \left. \left[ v^i(t) \phi^i(t, x) - \sigma^i \frac{\partial \phi^i}{\partial x} \right] \right|_{x=0} = F^i(t), \]  

(2)

\[ \left. -\sigma^i \frac{\partial \phi^i}{\partial x} \right|_{x=1} = 0, \]  

(3)

\[ \phi^i(t_0, x) = \hat{\phi}^i(x), \]  

(4)

where \(i = 1, 2, ..., s\), \(v^i(t)\) and \(m^i(t)\) are the specific development and mortality rates, respectively, assumed independent of the age \(x\), \(\hat{\phi}^i(x)\) are the initial distributions, while \(\sigma^i\) are the diffusion coefficients, assumed time independent. Moreover, the fluxes \(F^i(t)\) in the boundary condition (2) are evaluated as follows. The term \(F^1(t)\) is the egg production flux and is given by

\[ F^1(t) = v^s(t) \int_0^1 \beta(t, x) \phi^s(t, x) \, dx, \]  

(5)

where \(v^s(t)\beta(t, x)\) is the specific fertility rate. In particular, we consider
\[ v^*(t) \beta(t, x) = b(t) f(x) \quad \text{eggs/adults with age in } (x, x + dx)/\text{time unit}, \quad (6) \]

where \( b(t) \) takes into account the effect due to both diet and temperature, and \( f(x) \) is the fertility profile.

The other terms \( F^i(t) \), when \( i > 1 \), are the individual fluxes from stage \( i-1 \) to stage \( i \) and are given by

\[ F^i(t) = v^{i-1}(t) \phi^{i-1}(t, 1), \quad i > 1. \quad (7) \]

The boundary condition at \( x = 0 \) assigns the input flux into stage \( i \), while the boundary condition at \( x = 1 \) means that the output flux from stage \( i \) is due only to the advective component \( v^i(t) \phi^i(t, 1) \).

The functions \( \phi^i(t, x) \) allow to obtain the number of individuals in stage \( i \) at time \( t \):

\[ N^i(t) = \int_0^1 \phi^i(t, x) dx. \]

### 2.2 The biodemographic functions

System (1)-(4) requires an explicit formulation (depending on a certain number of parameters) of basic biodemographic rate functions, for development, fecundity and mortality for each stage. They model the physiological response of individuals to environmental forcing variables, which vary over the chronological time. For poikilotherm organisms, temperature is considered the most important driving variable; due to this reason, these biodemographic rate functions are commonly formulated in terms of temperature, which depends on the chronological time. Also the dependence on other environmental variables can be taken into account.

### 3 Structure of \textit{L. botrana} population

\textit{Lobesia botrana} has a stage structured population, generally considered composed by four stages: eggs, larvae, pupae and adults (\( s = 4 \)). Estimations of stage-specific biodemographic functions usually rely on bottom-up laboratory experimental data, while top-down field population data must be used to validate the model. Our approach is different: we use experimental data to estimate developmental and fecundity rates, but not mortality rates that are very difficult to measure. For these reasons, we use population time series data for the mortality estimations applying the method proposed by Wood [22] for formulate and fitting partially specified models.

In this section we define development and fecundity rate functions for \textit{L. botrana} that summarize our knowledge on the biology of the species. We suppose that development and mortality rate functions depend on time only through temperature, while the fecundity rate function depends also on the physiological age, as done in [11].
3.1 Development rate function

The development rate function $v(t)$, appearing in (1), describes the development response curve. Typically, there is no growth below a lower temperature threshold, while the developmental rate increases and reaches a maximum at an optimal temperature and then it declines rapidly approaching zero at a lethal temperature threshold. A lot of functional expressions have been proposed in literature to describe development [14]. Here, as in [11], we consider a Lactin function [15] to represent the development of all the stages:

$$v(t) = \delta^i \max \left\{ 0, e^{\alpha^i T} - e^{\alpha^i T_{m} - \beta^i T} - \gamma^i \right\}$$  \hspace{1cm} (8)

where $T_{m}$ is the lethal maximum temperature, $\alpha^i$ is the slope parameter describing the acceleration of the function from the low temperature threshold to the optimal temperature, $\beta^i$ is the width of the high temperature decline zone, $\gamma^i$ is the asymptote to which the function tends at low temperatures, and $\delta^i$ is a coefficient of amplification of the curve.

Parameters of the development rate functions are estimated by means of a least square method using the datasets in [1, 2] and are reported in Table 1 (see [11]).

3.1.1 Fecundity rate function

We assume that the eggs production depends on the physiological age of the adults, and on the chronological time through temperature and phenological stage of the host plant as environmental variables, as already supposed in (6). The oviposition profile $f(x)$, as function of the physiological age $x$, is assumed to be of the functional form

$$f(x) = ax^{b-1} \exp(-cx)$$

where $a$, $b$, $c$ are parameters to be estimated. This class of functions, reproducing the shape of a gamma distribution, is sufficiently general to allow the shift of the mode in all the values of the physiological age interval.

The term $b(t)$ takes into account the influence of environmental variables, temperature $T(t)$ and phenological stage of the plant $P(t)$, which vary on the
Table 2: Values of the step function $b_0(P)$, with steps in three plant phenological stages, following the BBCH-scale.

<table>
<thead>
<tr>
<th>Plant stage</th>
<th>$P$</th>
<th>$b_0(P)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflorescence</td>
<td>BBCH 53</td>
<td>0.31</td>
</tr>
<tr>
<td>Green Berries</td>
<td>BBCH 71</td>
<td>0.48</td>
</tr>
<tr>
<td>Maturing fruits</td>
<td>BBCH 81</td>
<td>1</td>
</tr>
</tbody>
</table>

The product $f(a_0(T)) = f(x)a_0(T)$ (eggs/female/day) on temperature ($^\circ$C) is illustrated in Figure 1. Function $b_0(P(t))$, which depends on the phenological age $P$ of the plant expressed in terms of BBCH-scale [17] is a step function with steps at the BBCH stages indicated in Table 2 [13].

### 3.1.2 Mortality rate function

Mortality rates are very difficult to measure, then the functional form of the mortality rate $m(t)$ cannot be easily determined as the development and the fecundity rate functions. Moreover, in [11, 16] the authors defined a mortality composed by two terms: an intrinsic temperature-dependent (abiotic) mortality depending on the development rate function, and a constant generation-dependent extrinsic mortality likely related to external natural control factors, to be estimated using time series field data on the population dynamics. In this paper we consider the mortality $m(t)$ as unknown and we apply an estimate
method proposed by Wood [22] for formulating and fitting partially specified models. With this approach, the obtained mortality carries out both extrinsic and intrinsic mortality factors. Since we want to estimate the mortality rates without assuming a specific functional form for them, we do not specify an analytical expression depending on a certain number of parameters. However, we assume some biologically meaningful hypothesis:

- the mortality rate depends on the chronological time through temperature;
- the mortality rate is a nonnegative continuous function of temperature;
- the mortality rate is strictly positive at two reference temperature.

These assumptions will constitute the constraints on the shape of the mortality rates in the sequel.

4 Estimation of the mortality rate function

Mortality rates are very difficult to measure and an estimation like those used for development and fecundity is not always possible. For this reason, we apply the method proposed by Wood [22] for formulate and fitting partially specified models. To apply this method we need a dataset of population dynamics. We consider the same dataset used in [11] relative to the dynamics of the grape berry moth in a vineyard of Garganega located in Colognola ai Colli, a hilly region in the North-East of Italy during the period 2008-2012. The experimental field was not treated with insecticides to avoid controls on the growth of the insect population. More precisely, to estimate the mortality rates functions we used the field data collected at Colognola ai Colli in the three years 2008, 2009 and 2011 (model calibration); the data for the other years 2010 and 2012 were used to test the model (validation), keeping all the other parameters of development and fecundity fixed.
4.1 The method

We consider system (1)-(4) in which the development functions and the fecundity function are chosen as in the previous, and the mortality rates are unspecified. We want to find the functions $m_i(t)$ that result the best fit of the model to field data of populations densities. Once the model functions are fixed, the system (1)-(4) can be numerically solved and produce a vector of model estimates $\mu$, representing the population abundances, corresponding to the observations $y$. The goodness of the fit can be quantitatively measured as a weighted least squared term

$$\sum_{i=1}^{d} w_i (y_i - \mu_i)^2,$$

where $d$ is the number of data and $w_i$ are the weights. Then, the best fitting functions $m_i$ are those which minimize this quantity.

The unknown functions $m_i$ can be expressed as linear combination of a suitable basis $\xi_{i,j}(t)$, $i = 1, \ldots, s$, $j = 1, \ldots, n_i$ (for instance, a polynomial or cubic spline basis)

$$m_i(t) = \sum_{j=1}^{n_i} p_{ij} \xi_{ij}(t).$$

Therefore, finding the best fitting functions $m_i$ is reduced to finding the best fitting parameters $p_{ij}$, $i = 1, \ldots, s$, $j = 1, \ldots, n_i$, collected into the vector $p = [p_{11}, \ldots, p_{1n_1}, \ldots, p_{sn_s}]^T$, which produces the model estimates $\mu(p)$ together with some constraints. The total number of parameters is denoted as $n_p = n_1 + \cdots + n_s$. Then our objective is to minimize

$$q(p) = \sum_{i=1}^{d} w_i (y_i - \mu_i)^2.$$

The procedure which leads to an estimates of the coefficients $p$ is the following.

- Given a guess of the model parameter vector $p$, the model equations are numerically solved and model estimates $\mu$ are obtained.
- By repeatedly solving the model with slight changes in parameters, we obtain an estimate of the $d \times n_p$ matrix $J$ where $J_{ij} = \partial \mu_i / \partial p_j$. We use

$$J_{ij} \approx \frac{\mu_i(P + \delta_j e_j) - \mu_i(P - \delta_j e_j)}{2\delta_j},$$

where $\delta_j$ is a small number and $e_j$ are vectors of the canonical basis.
- The quantity $\mu$ and $J$ are used to construct a quadratic model of the fitting objective as a functional of $p$

$$q(p) \sim (\hat{y} - Jp)^T W (\hat{y} - Jp),$$

where $\hat{y} = y - \mu + Jp$, and $W$ is a diagonal matrix with $W_{ii} = w_i$.  

8
We find a suggesting direction to modify $p$ in order to minimize the real fitting objective.

We iterate these steps to convergence.

It is worthwhile to note that this is a simple version of the method proposed in [22]. It can be improved choosing a different quadratic model of the fitting objective or taking into account an additional term in the objective which is a sum of the “wiggliness” measures for the model unknown functions. However, the version proposed in this paper, despite its simplicity, gives satisfactory results.

4.2 Model calibration

To estimate the mortality rates functions we used the field data collected in a Garganega vineyard located in Colognola ai Colli for the years 2008, 2009 and 2011. All the other parameters of development and fecundity are fixed and the values of the diffusion coefficients are set $\sigma_i = 0.0001$, $i = 1, 2, 3, 4$. We have to minimize the sum of the weighted squared differences between simulated dynamics and observations for all the three years considered in the estimation phase. More precisely, $d = d_{2008} + d_{2009} + d_{2011}$, where $d_Y$ is the number of observation in the year $Y$.

To run the model for every year, we must know population densities at the beginning of the season to drive the simulation during the entire growing season as no other information on the pest abundance is provided. In our study, the number of adults catches per trap per week recorded until the first larvae of the first generation are observed, were used as the initial condition of the model. Hourly temperature data, collected by a meteorological station close to the vineyard, are used as a driver environmental variable for the model simulation.

Furthermore, we chose the cubic B-spline basis to represent the mortality rate functions. The basis is built on the nodes $[0, 10, 20, 30, 40]$, which is a suitable interval of temperature, and hence it consists in seven polynomial $\xi_j(T)$, $j = 1, \ldots, 7$ defined on this interval. The mortality rates function are

$$m_i(t) = \sum_{j=1}^{7} p_{ij} \xi_j(t), \quad i = 1, \ldots, 4.$$  

Then, it is possible to express the constraints on the shape of the mortality rates through linear and nonlinear inequalities involving the parameters $p_{ij}$.

The weight matrix $W$ is set in the following way: $w_1 = w_3 = w_4, w_2 = 10w_1$. This choice gives much more importance to the larval stage with respect to the other stages. The practical reason is that field data are more reliable on the larval stage than the others, and furthermore, this stage is harmful for the plant, so the most important for biological control, consequently we focus on this stage.

The shape of the estimated mortality rate functions are reported in Figure 2, while the resulting population dynamics compared with the field data are shown in Figure 3.
Figure 2: Shape of the estimated mortality functions for all the stages of the grape berry moth: we used a cost functional in which we gave more relevance to the data of the larval stage.

We can observe that the mortality function shows an increasing behaviour for increasing temperatures greater than an upper threshold and for decreasing temperatures lesser than a lower threshold, while in the central part of the temperature interval mortality has lower values. This is in agreement with the mortality chosen in [11] where second order degree polynomials were chosen for low and high temperatures.

The simulated dynamics for the three years 2008, 2009 and 2011 (Figure 3) present a satisfactory fit of the phenology of the grape berry moth, obtained from the field observations. The simulations are comparable with those obtained in [11].

4.3 Model validation

The data for the years 2010 and 2012 recorded in Colognola ai Colli were used to test the model, keeping all the other parameters of development and fecundity fixed as in the model calibration. The simulated dynamics of all the stages, obtained using the estimated mortality in Figure 2 for the years 2010 and 2012, are represented in Figure 4.

Also in this case, a good representation of the phenology of the species, for the two years considered, is obtained.
Figure 3: Continuous lines: simulated population dynamics obtained with the estimated mortality functions of Figure 2. Asterisks: data collected in a Gar-ganega vineyard in Colognola ai Colli.
5 Discussion and concluding remarks

A realistic simulation of the population dynamics relies on a good knowledge of the biodemographic functions describing the biology of the species. Frequently, literature data on the mortality rate function are not available and the mortality cannot be easily estimated as the development and the fecundity rate functions. Other methods, based on the availability of population dynamics datasets, have been developed [9, 11, 22].

Here we consider the case of the grape berry moth, for which we dispose of 5 years of dynamics observations. We apply the method proposed by Wood [22] considering the data collected in three non-consecutive years. This method returns the estimation of the mortality rate functions corresponding to the best fit of the dynamics for all the three years, meant as the smaller sum of weighted squared differences between simulated dynamics and observations.

The estimation method considered in the present paper has some advantages with respect to the method proposed in [11] and [16] for the grape berry moth mortality estimation. In [11] and [16] the mortality was represented as sum of two terms: an intrinsic mortality due to abiotic factors and an extrinsic mortality due to biotic factors. The intrinsic mortality was estimated using literature data, while for the extrinsic mortality, estimation methods based on population dynamics observations were proposed. Here the mortality is considered as a whole and represented as linear combination of cubic splines. This assures a greater flexibility for the shape of the mortality.

When considering the sum of the weighted square differences between simulations and observations we give a higher weight to the larval stage because the measurements of larvae are more precise than for the other stages. Moreover,
to the end of pest control in vineyards, the larval stage is the most important since larvae, in particular of second generation, produce serious damages to the vineyards.

The estimated mortalities (Figure 2) have low values in the central part of the temperature interval and they, generally, increase for increasing large values of the temperatures and for decreasing small values of the temperatures. This behaviour is in agreement with the assumptions made in [11]. In fact, in that work the authors supposed that the mortality increases as a second order degree polynomial for large values and small values of the temperatures, obtaining a satisfactory fit of the population dynamics. Here we obtain an analogous behaviour of the mortality with again a good fit of the phenology of the grape berry moth in the years considered for the estimation phase. In some cases (for example for large and small temperature values for adults or for small values for eggs) the mortality does not increase quickly. This is not unexpected because very high and very small temperatures are achieved infrequently, then it is not easy to have a good estimate of the mortality in these temperature intervals. On the other hand a not very reliable mortality estimate for unlikely temperatures does not produce bad results in the simulation of the dynamics because these values are not reached and do not affect the dynamics.

Mortality estimation procedure has been performed considering three years of data on population dynamics. Then, the estimated mortalities have been used to simulate the dynamics for two further years. The satisfactory representation of the phenology in these years allows us to state that the mortality can be actually considered the same for different years (as the development and the fecundity rate functions). This is an important result because it allows to estimate the mortality once, considering a fixed number of years of observations and then to use the same mortalities for all the following years.

Acknowledgements

Support by INdAM-GNFM is gratefully acknowledged by CS.
The authors would like to thank Tommaso del Viscio for technical support.

References


Recent titles from the IMATI-REPORT Series:

2018


18-02: *A study on 3D shape interaction in virtual reality*, E. Cordeiro, F. Giannini, M. Monti, A. Ferreira.


18-09: *High order VEM on curved domains*, S. Bertoluzza, M. Pennacchio, D. Prada.

18-10: *Estimation of the mortality rate functions from time series field data in a stage-structured demographic model for Lobesia botrana*, S. Pasquali, C. Soresina.