Life tables data collection in entomology: an overview on the differential and the integral representation and proposal for a standard electronic file

Luca Rossini[1,*], Arthur Lots[1], Grégoire Noël[2], Arnaud Segers[2], Serhan Mermer[3,4], Mario Contarini[5], Stefano Speranza[5,6], Vaughn Walton[3], Frédéric Francis[2], Emanuele Garone[1]

[1] *Service d'Automatique et d'Analyse des Systèmes, Université Libre de Bruxelles, Brussels, Belgium*

[2] *Functional and Evolutionary Entomology, Gembloux Agro-Bio-Tech, University of Liège, Gembloux, Belgium*

[3] *Department of Environmental and Molecular Toxicology, Oregon State University, Corvallis, Oregon, United States of America*

[4] *Department of Horticulture, Oregon State University, Corvallis, Oregon, United States of America*

[5] *Dipartimento di Scienze Agrarie e Forestali, Università degli Studi della Tuscia, Viterbo, Italy*

[6] *Centro de Estudios Parasitológicos y de Vectores (CEPAVE, CONICET-UNLP), La Plata, Argentina*

[*] Corresponding author: luca.rossini@ulb.be (LR)

Abstract

Life tables are the most common tool to explore the biology of insects and of many terrestrial arthropods, as well as to understand their response to external factors (i.e., temperature, relative humidity, diet, etc.). The data collection process has been partially standardised over the years and life tables results are presented mainly depending on the purpose of the study. In fact, two different representations of the data can be obtained from the raw dataset: a differential and an integral representation. The two representations provide relevant biological information, but they lead to a loss of information (e.g., the stage-duration of each specimen) with respect to the raw dataset. To date, a conceptual explanation of how the two representations can be obtained from the raw data, and of their respective properties, is still missing; moreover, providing the raw dataset as

supporting information of the published papers is still not customary by the authors. This paper highlights three main points: *i)* how the two representations are obtained by the life tables raw dataset, *ii)* without raw data it is not possible to switch between the two representations, with a subsequent loss of information, *iii)* why there is the need for a data collection standard. In addition to the conceptual explanation, this work also provides a proposal for a standard data collection spreadsheet that could support data collection, data sharing, and that automatically transform the data in the two representations. We believe that this study is a first step towards a more efficient diffusion of the information among the scientific community, maximising the efforts made by scholars during the experimental and data analysis process.

Keywords: key 1

1. Introduction

Measurements, data collection, and data representation standards are essential for scientific research, as they ensure the consistency of collected data; this is in turn fundamental to allow the reproducibility of experiments, and coherent data sharing (Mari et al., 2012; Zhan and Xie, 2020). Indeed, there is a high value in collecting and organising data using standardised protocols, allowing for comparability and re-usability of the data (Mons et al., 2017; Stall et al., 2019). Such aspect is particularly important whenever data sharing between research groups is foreseen. It is thus not surprising that all scientific communities that develop and use shared databanks (e.g. NCBI for genetic sequencing https://www.ncbi.nlm.nih.gov/) have defined strict standards on how to carry out the experiments and report data (NCBI Resource Coordinators, 2012; Sherry, 2001). In entomology and ecology, the transition towards more quantitative approaches is highlighting the need for defining clear and standard data measurement and representation protocols.

Life tables are one of the most recurring laboratory experiments in entomology to explore the biology of the species (of insects and terrestrial arthropods at large) and their reaction to external factors (i.e., temperature, relative humidity, diet, etc.) (Rossini et al., 2024a, 2025). Such factors consist of rearing individuals under controlled conditions and sampling their development at fixed time intervals (Bellows, 1992; Harcourt, 1969; Hsieh, 1991). The resulting dataset from the subject species provides developmental information on a series

of important qualitative (e.g., morphological traits) and quantitative (e.g., stage duration, minimum development time, mortality, fertility/egg production) features of the species (Chi and Liu, 1985). Studies are typically carried out under different rearing conditions which may include factors as temperature, humidity, CO_2, nitrogen, diet, or pesticide treatment (Alves et al., 2013; Korpelainen, 1986; Liu et al., 2017; Rossini et al., 2024b; Sánchez-Ramos et al., 2007; Vantornhout et al., 2005; Wen et al., 2020). Even if conditions and goals of the experiments may differ, the procedure is always the same: all experiments determine dead or alive life stages at each inspection and measure qualitative and/or quantitative insect traits depending on the research question (Chi et al., 2023, 2020). This allows researchers to track the life history of the population and define development time, mortality, and fertility quantitatively at different rearing conditions (Chi et al., 2023, 2020; Rossini et al., 2024a). Typically, researchers are interested in analysing differences among various rearing conditions, to validate the hypotheses behind the experiment. However, to reach this goal there is a need for a clear organisation of the dataset that possibly warrants reproducibility of results and application of standard procedures for data analysis (Rossini et al., 2024a).

The starting point of this paper is the observation that for this crucial type of experiment, the data collection and data reporting is a process that, to date, is still not univocally defined among entomologists. In fact, if we look at the specific literature, we do not find either the raw data (except in very few cases), or a standardised method for acquisition and reporting of the data (Kareithi et al., 2019; Rossini et al., 2025, 2024a). Currently there are two main representations of the data that have been used in the literature, and whose analysis will be the starting point for this study.

The first representation of life tables data records daily changes. This data typically includes recordings of newly developed life stages, dead individuals, and new offspring (eggs) per female (Bieri et al., 1983; De Campos et al., 2020; Genç and Nation, 2008; Moshtaghi Maleki et al., 2016; Nielsen et al., 2008; Samayoa et al., 2018; Southwood, 1978; Tochen et al., 2014; Van Nieuwenhove et al., 2016). The representation leads to a distribution of the development times of each life stage, datasets from which we can obtain descriptive statistics parameters such as the mean value, standard deviation, median, mode, kurtosis, or skewness (Rossini et al., 2024a). Such values are then listed in specific tables as a result and are widely employed in the parameterisation of mathematical models that describe the population dynamics over time (Rossini et al., 2019a, 2020a, 2021a, 2021b, 2025).

The second representation was introduced by Chi et al. (Chi, 1988; Chi and Liu, 1985), and consists of developmental life stage, mortality, and fertility matrices containing life stages and age. This framework is inspired by the Leslie matrix (Leslie, 1948, 1945), which however is not suitable to describe insect populations because of the different age-stage division (Chi and Liu, 1985). Elements contained in the matrices depend on their purpose and should contain count of individuals, survival rates, mortality rates, and fecundity. Such data allows scholars to compute some key parameters that characterise the population, such as the intrinsic rate of increase, mean generation time, and age-specific fecundity (Bellows, 1992; Chi et al., 2020, 2023; Rossini et al., 2024a).

The two representations can be seen as a "*differential*" representation (the first) and an "*integral*" representation (the second). In this paper we show that these two approaches have benefits and disadvantages and that they are not equivalent: from the differential representation is not possible to recover the information of the integral one and *vice versa*, unless the raw data is available. However, although Chi (Chi, 1988) provides a valuable example of how raw data should be collected (see Table 1 of their study), raw data is not commonly provided in entomology.

It is worth remarking that the generation of life tables' data is characterised by high costs in terms of time and personnel (Kareithi et al., 2019). Given the large effort to generate the data, the fact that the two main representations used in the literature are not comparable nor mergeable is a major issue. This work aims at clarifying the relationship between the common "differential" and "integral" representation presented in the literature and to suggest a standardised data collection and documentation procedure. This methodology will allow subsequent studies to more effectively determine both population dynamics and cohort representations. In addition, we also propose a standardised spreadsheet template that can support the scientific community to plan experiments and that can be used as supplementary material to first share data during paper submissions, and then after publication. This tool will allow us for a simple and versatile solution that can facilitate open data while not requiring high programming skills, and that may also constitute a basis for the future development of a common database, as already done in other fields of research.

2. Materials and Methods

To fully understand the problem of the data collection, the study was structured in two parts. The first one is more conceptual and highlights the *pros* and *cons* of the most common representations of data for life tables, and the connections with the raw dataset. The second part introduces our proposal for a standard as well as its implementation through an electronic spreadsheet file. It is worth pointing out that this study will focus only on development and mortality, but not on fertility. The two representations, in fact, do not affect the information on fertility, as this experiment can be conceptually considered as independent. Fertility has been already well discussed by Chi et al. (Chi, 1988; Chi et al., 2023, 2020; Chi and Liu, 1985) and it is fully inserted in the framework discussed hereafter.

2.1. Background: how do we collect raw data?

Data collection is the starting point of the process, and it is directly related to the experimental trials. The eggs are placed in climatic chambers at day zero and are usually separated and labelled to maintain their identity. According to the suggestions of Chi (Chi, 1988), the experimenter should build a table where the lines indicate the number of the specimen reared (so that the individuality can be maintained) and the columns represent the sampling time. For the sake of exposition, and with no loss of generality, in this work we consider a sampling time of one day. The life stages composing the life cycle can be abbreviated with acronyms or letters, as for instance “E” for “egg”, “L1” or “N1” for first instar larvae or nymphs, until the death, that may be indicated as “D”. Those acronyms will be the elements of the table, indicating the stage in which the individual “*i*” was at time *t*: an idea of how this kind of tables should look like is provided in Table 1.

Table 1: Example of table to report the raw data of life tables experiments. Rows indicate the number of the specimens, while columns the days. Each label composing the element of the table indicates the life stage: in this example “E” means *egg*, “L1”, “L2”, and “L3” indicate the *first*, *second*, and *third larval stage*, respectively, and “D” indicates the *dead*.

	Time (day)												
	1	2	3	4	5	6	7	8	9	10	11	12	…
Specimen													
SP-1	E	E	E	L1	L1	L1	L2	L2	L2	L2	L3	L3	…

SP-2	E	E	L1	L1	L1	L2	L2	L2	L2	L3	L3	L3	…
SP-3	E	E	E	L1	L1	L1	L1	L2	L2	L2	L2	L3	…
SP-4	E	E	L1	L1	L1	L1	L2	L2	L2	L2	L3	L3	…
SP-5	E	E	E	D	--	--	--	--	--	--	--	--	--
SP-6	E	L1	L1	L1	L2	L2	L2	L2	L3	L3	L3	L3	…
SP-7	E	E	E	L1	L1	L1	L2	L2	L2	L3	L3	L3	…
SP-8	E	E	E	L1	L1	L1	L1	L2	L2	L2	L2	L3	…
SP-9	E	E	E	E	L1	D	--	--	--	--	--	--	--
SP-10	E	E	L1	L1	L1	L2	L2	L2	L2	L3	L3	L3	…
SP-11	E	E	E	L1	L1	L1	L1	L2	L2	L2	L3	L3	…
SP-12	E	E	L1	L1	L1	L1	L2	L2	L2	L2	D	--	--
SP-13	E	E	L1	L1	L1	L1	L2	L2	L2	L3	L3	L3	…
SP-14	E	E	L1	L1	L1	L1	L2	L2	L2	L2	L3	L3	…
SP-15	E	E	L1	L1	L1	L1	L2	L2	L2	L2	L2	L3	…
SP-16	E	E	E	L1	L1	L1	L2	L2	L2	D	--	--	--
…	…	…	…	…	…	…	…	…	…	…	…	…	…

Typically, these raw data are not published, and they are further analysed or processed. The ways this operation is usually carried out leads to the two different representations that we discuss in Sections 2.2 and 2.3. To fix the ideas, the differential and the cohort representation can be obtained by manipulating the rows and the columns of the Table 1, respectively.

2.2. Distribution of the development times for each stage and differential representation

A typical first step to extract information from the raw data is to report the life stage duration of each specimen, i.e. counting how many elements with the same stage label are along each line of Table 1. The values hereby computed are reported in two tables: one accounting for individuals who successfully developed to the next stage, hereafter denoted as *stage-development table* (Table 2), and one accounting for died individuals, hereafter denoted as *stage-mortality table* (Table 3). In other words, this step converts Table 1 into two quantitative tables that can be further processed.

Table 2 is an example of stage-development table, where the rows contain the life histories of the individuals, the columns indicate the life stages, and the elements represent the duration of the stage between life stages. Table 3 is an example of stage-mortality table, conceptually similar to Table 2 except for the meaning of the elements, that hereby indicate the duration from the last metamorphosis to death. Adult survival (or longevity) is part of the stage-mortality table (Table 3) to be coherent with the definition of development, as "survival" indicates the times between the adult emergence and the death.

Table 2: First elaboration of the dataset collected in Table 1 – *Stage-development table*. The rows indicate the individuals of the population, the columns the life stages, and the elements the duration in days (or other time sub-units). This table should be read as follows: "*Specimen SP-1 spent 3 days to pass from egg to L1*", or, as an alternative "*The duration of the egg stage of the specimen SP-1 was 3 days*". LN denotes the last preimaginal stage, while the double dash "--" indicates that the specimen died in the latest stage having numeric values in the corresponding cell.

	Stages					
Specimen	E	L1	L2	L3	…	LN
SP-1	3	3	4	4	--	13
SP-2	2	3	4	5	…	12
SP-3	3	4	4	5	…	12
SP-4	2	4	4	4	…	12
SP-5	--	--	--	--	…	--
SP-6	1	3	4	5	…	13
SP-7	3	3	3	5	…	13
SP-8	3	4	4	4	…	13
SP-9	4	--	--	--	…	--
SP-10	2	3	4	4	…	13
SP-11	3	4	3	4	…	12
SP-12	2	4	--	--	…	--
SP-13	2	4	3	5	…	12
SP-14	2	4	4	5	…	12

SP-15	2	4	5	4	…	13
SP-16	3	3	--	--	…	--
…	…	…	…	…	…	…

Table 3: First elaboration of the dataset collected in Table 1 – *Stage-mortality table*. Rows indicate the individuals of the population, columns the life stages, and the elements the days from the last metamorphosis to death (D). This table should be read as follows: "*Specimen SP-5 died 4 days after the egg oviposition*", or, as an alternative "*Specimen SP-9 died 2 days after turning L1*". The double dash "--" indicates that the specimen successfully developed in the previous stages having numeric values or died before, while AF and AM denote adult males and females, respectively.

	Stages						
Specimen	E-D	L1-D	L2-D	…	…	AF-D	AM-D
SP-1	--	--	--	--	…	--	74
SP-2	--	--	--	--	…	--	78
SP-3	--	--	--	--	…	79	--
SP-4	--	--	--	--	…	--	72
SP-5	4	--	--	--	…	--	--
SP-6	--	--	--	--	…	72	--
SP-7	--	--	--	--	…	80	--
SP-8	--	--	--	--	…	--	64
SP-9	--	2	--	--	…	--	--
SP-10	--	--	--	--	…	63	--
SP-11	--	--	--	--	…	--	81
SP-12	--	--	5	--	…	--	--
SP-13	--	--	--	--	…	60	--
SP-14	--	--	--	--	…	74	--
SP-15	--	--	--	--	…	--	70
SP-16	--	--	4	--	…	--	--
…	…	…	…	…	…	…	…

It is worth remarking that although this first step of the data analysis does not introduce any loss of information with respects to the raw data (while making them much more readable), it is not customary to report these tables in scientific papers (Kareithi et al., 2019).

Indeed, the information listed in the stage-development table (e.g. Table 2) and in the stage-death table (e.g. Table 3) is mostly used as an intermediate step (usually not published) to reach a *differential representation* of life tables data, also known as "*population dynamics representation*". The differential representation is usually reported as a table or, equivalently, as a family of frequency histograms (an example of the table and of the associated histograms are provided in Table 4 and Fig. 1, respectively) describing the distribution of the stage-development and mortality times of each stage and using bins size coinciding with the sampling time of the experiment. As detailed by Rossini et al. (Rossini et al., 2024a), the stage-distributions of development and mortality can be interpreted as a stage-population dynamics, namely the set of newly emerged (or died) individuals per stage and per sampling unit (e.g., per day). From a mathematical point of view, we can summarise the information of the j-th stage histograms as the column vectors

$$G_D^j = \left(g_{1j}^D, g_{2j}^D, \dots, g_{ij}^D\right)^T \quad (1)$$

and

$$D_D^j = \left(d_{1j}^D, d_{2j}^D, \dots, d_{ij}^D\right)^T, \quad (2)$$

where the elements g_{ij}^D and d_{ij}^D denote the number of individuals that spent exactly i days in stage j before passing to stage $j + 1$ or dying.

Table 4: Example of table data typically resulting from the life tables *differential representation*. The table was made for the sake of exposition using a small portion of unpublished data on *Drosophila suzukii* life tables data at 20 °C. E indicates the egg stage, L1, L2, and L3 indicate the first, the second, and the third larval instars, respectively, and D indicates the death.

Commenté [LR1]: If accepted, cite Baser et al. 2025

Time	Development			Mortality		
(days)	E to L1	L1 to L2	L2 to L3	E to D	L1 to D	L2 to D
1	0	0	0	0	0	0

2	0	0	0	0	0	0
3	3	0	0	0	0	
4	4	0	0	1	0	0
5	18	9	0	0	0	0
6	20	23	0	1	0	0
7	0	6	7	1	1	0
8	0	0	10	0	5	0
9	0	0	14	0	1	0
10	0	0	5	0	0	0
11	0	0	0	0	0	1
12	0	0	0	0	0	0

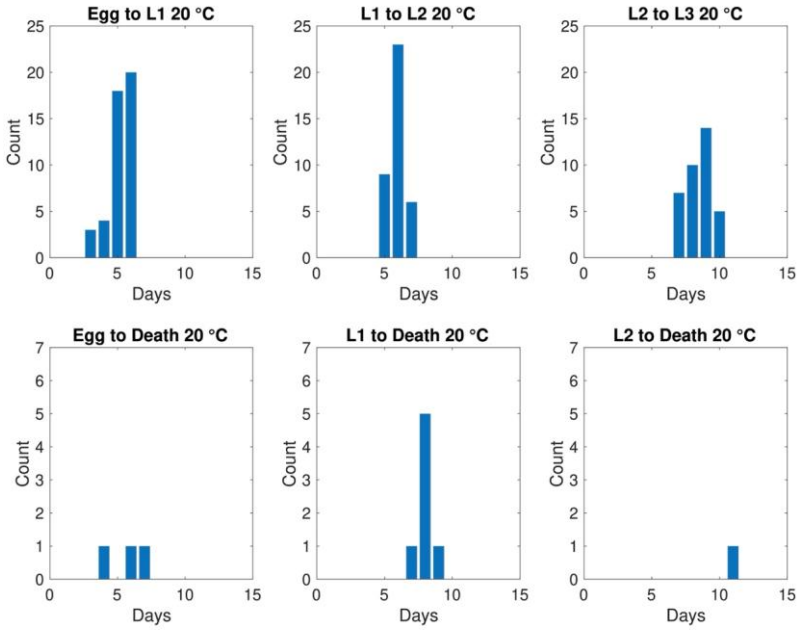

Figure 1: Example of histogram of the data typically resulting from the life tables *differential representation*. These plots were made for the sake of exposition using a small portion of unpublished data on *Drosophila suzukii* life tables data at 20 °C.

The set of vectors (1) and (2) are organised to form the stage-development and the stage-mortality matrices G_D and D_D, respectively, that assume the following forms:

$$G_D = \begin{pmatrix} g_{11}^D & g_{12}^D & g_{13}^D & \cdots \\ g_{21}^D & g_{22}^D & g_{23}^D & \cdots \\ g_{31}^D & g_{32}^D & g_{33}^D & \cdots \\ g_{41}^D & g_{42}^D & g_{43}^D & \cdots \\ g_{51}^D & g_{52}^D & g_{53}^D & \cdots \\ g_{61}^D & g_{62}^D & g_{63}^D & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (3)$$

and

$$D_D = \begin{pmatrix} d_{11}^D & d_{12}^D & d_{13}^D & \cdots \\ d_{21}^D & d_{22}^D & d_{23}^D & \cdots \\ d_{31}^D & d_{32}^D & d_{33}^D & \cdots \\ d_{41}^D & d_{42}^D & d_{43}^D & \cdots \\ d_{51}^D & d_{52}^D & d_{53}^D & \cdots \\ d_{61}^D & d_{62}^D & d_{63}^D & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (4)$$

An interesting property of G_D and D_D is the mass conservation along the columns: by summing all the values on the columns, we obtain the total number of individuals alive (matrix G_D) or dead (matrix D_D). The matrices G_D and D_D can also be normalised by dividing each element g_{ij} and d_{ij} by the size of the initial cohort. Accordingly, the columns can be seen as the impulse response (of the alive or of the dead) as described by Rossini et al. (Rossini et al., 2024a), which is a convenient way to describe the stage-distributions of individuals over time. A biological interpretation of the impulse response is the following: the elements of the matrices indicate the stage-probability that the individuals have of developing (G_D) or dying (D_D) over time.

Matrices G_D and D_D are extremely useful to understand the patterns of development of the population and to extract synthetic values such as the mean duration of the stage (or from the metamorphosis to mortality), its standard deviation or standard error, and other parameters that depict the shape of the distribution, namely median, mode, skewness, and kurtosis. Other useful information that can be extracted from the differential representation is the minimum and the maximum time before and after which, respectively, no individuals

should be expected to develop. This is a biologically meaningful information that is of great importance in many fields, as for instance in pest control. Knowing the minimum amount of time required by an organism growing under given conditions (i.e., the *minimum development time*) is fundamental for planning effective control strategies that possibly target the peak of the population of the life stage more susceptible to the type of treatment chosen. The differential representation, accordingly, is a very useful tool to analyse the effect of the rearing conditions on the different life stages.

Despite the advantages that the differential representation provides, it is worth remarking that it introduces a loss of information with respect to the dataset of the stage development and stage mortality matrices (such as, e.g. Table 2 and 3). Indeed, in this representation the individual story of each specimen is lost, and time is reset at each metamorphosis. This loss of information is not a minor aspect, and it might hide some important information about how the development in one stage influences the next one. Furthermore, it might prevent the detection of anomalies (e.g. to identify sub-populations with different development patterns). To better understand this fact, let us consider the following example.

Example 1. This example aims to show, in a simple way, that the differential representation introduces a loss of information w.r.t. the original *stage-development table* and *stage-mortality table* (and thus to the raw data), and that this might even hide possible problems in the experiments, hindering the comparability between results obtained independently. Let us consider a population with only three stages (E, L, A) and a population of 180 individuals at day 0. The corresponding matrix G_D (3) is reported in Table 5. This example focuses on a differential representation where the development from E to L is in 1 day for ⅓ of the population and in 2 days for the rest of the population, and the development from L to A is in 1 day for ⅓ of the population and in 2 days for the rest of the population.

Table 5: Differential representation of the dataset discussed in the *Example 1*.

Time	Development	
(days)	E to L	L to A
1	60	60
2	120	120

Looking at Table 5, it is possible to see that from this matrix we cannot reconstruct the information contained in the stage-development table (Table 2): in fact, there are multiple datasets that can give rise to such a G_D. In Table 6 we report two different *stage-development tables* which both generate Table 5. This example highlights one of the risks in using only the differential representation to describe life tables experiments. In fact, if we look at the two datasets, the first one shows a homogenous population where the development in the first stage does not influence the development in the second stage. *Vice versa*, in the second dataset, there is a full correlation between the development from E to L and the development from L to A (if they did the first transformation in 1 day, the second will be in 1 day, if they did the second transformation in 2 days, the second will be in 2 days). This important information (which might suggest either the presence of some biological insights or the fact that part of the population had a genetic mutation) is completely lost by only using the differential approach.

Table 6: Stage-development tables (a and b) that can be obtained by the differential representation listed in Table 5.

	Stage-development table a	
Specimen	E	L1
SP-1	1	1
⋮	⋮	⋮
SP-20	1	1
SP-31	1	2
⋮	⋮	⋮
SP-60	1	2
SP-61	2	1
⋮	⋮	⋮
SP-100	2	1
SP-101	2	2

	Stage-development table b	
Specimen	E	L1
SP-1	1	1
⋮	⋮	⋮
SP-20	⋮	⋮
SP-31	⋮	⋮
⋮	⋮	⋮
SP-60	1	1
SP-61	2	2
⋮	⋮	⋮
SP-100	⋮	⋮
SP-101	⋮	⋮

⋮	⋮	⋮	⋮	⋮	⋮
SP-180	2	2	SP-180	2	2

2.3. Cohort (integral) representation

The second representation that can be obtained by the raw dataset in Table 1 is the *integral*, or cohort, form introduced by Chi et al. (Chi, 1988; Chi and Liu, 1985). While the differential representation observes the newly individuals that over time are developing towards the next stages, the integral representation provides a picture of the number of individuals into each stage over time. In this case, accordingly, instead of a density we obtain the cumulative evolution of the population in each stage.

Integral representation datasets can be obtained by the raw data in Table 1 starting from some considerations. According to its structure, Table 1 lists along the columns the labels (and consequently the life stage) of the individuals over time. The number of alive individuals in each life stage at time t can be hence obtained by counting how many times a given label occurs in each column. An example of how the integral representation looks like is numerically provided in Table 7, while Fig. 2 shows the usual way to plot results. An analogous process can be carried out for the dead: in this case the process is slightly different, as over time we count the cumulative number of dead per stage. In case an individual in the stage i dies at time t, the number of dead corresponding to i will be increased of one unit.

Table 7: Example of table data typically resulting from the life tables *integral representation*. The table was made for the sake of exposition using a small portion of unpublished data on *Drosophila suzukii* life tables data at 20 °C. E indicates the egg stage, L1 and L2, indicate the first and the second larval instars, respectively, and D indicates the death.

Commenté [LR2]: If accepted, cite Baser et al. 2025

Time	Development			Mortality		
(days)	E	L1	L2	E to D	L1 to D	L2 to D
1	50	0	0	0	0	0
2	50	0	0	1	0	0
3	46	3	0	1	0	0
4	42	7	0	2	0	0

5	23	25	0	3	1	0
6	0	46	0	3	6	0
7	0	41	3	3	7	0
8	0	37	9	3	7	0
9	0	31	24	3	7	1
10	0	15	37	3	7	1
11	0	0	39	3	7	1
12	0	0	39	3	7	1
13	0	0	38	3	7	1
14	0	0	38	3	7	1
15	0	0	34	3	7	1
16	0	0	24	3	7	1
17	0	0	6	3	7	1
18	0	0	1	3	7	2
19	0	0	0	3	7	2
20	0	0	0	3	7	2

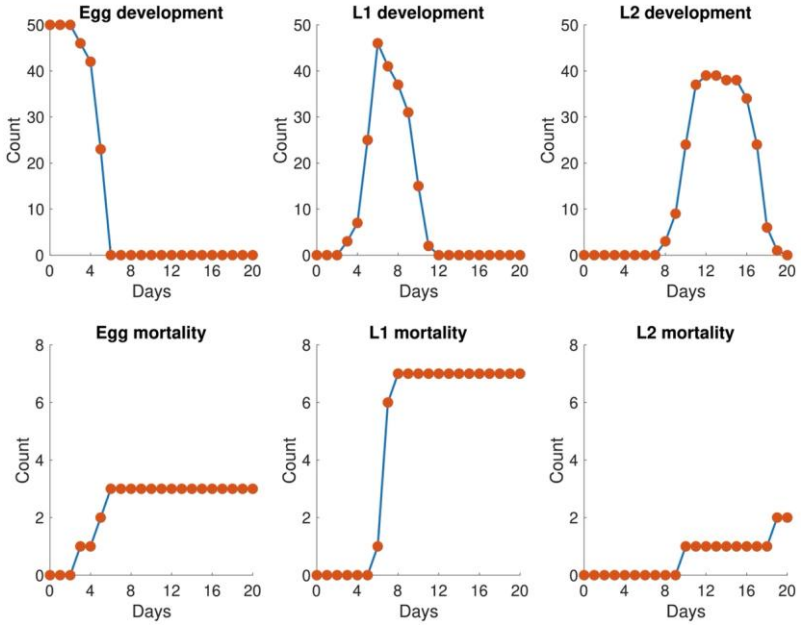

Figure 2: Example of plot of the data typically resulting from the life tables *integral representation*. These plots were made for the sake of exposition using a small portion of unpublished data on *Drosophila suzukii* life tables data at 20 °C.

This procedure led us to obtaining the matrices G_C and P_C already described by Chi et al. (Chi, 1988; Chi and Liu, 1985) mathematically defined as follows:

$$G_C = \begin{pmatrix} g_{11}^C & g_{12}^C & g_{13}^C & g_{14}^C & \cdots \\ g_{21}^C & g_{22}^C & g_{23}^C & g_{24}^C & \cdots \\ g_{31}^C & g_{32}^C & g_{33}^C & g_{34}^C & \cdots \\ g_{41}^C & g_{42}^C & g_{43}^C & g_{44}^C & \cdots \\ g_{51}^C & g_{52}^C & g_{53}^C & g_{54}^C & \cdots \\ g_{61}^C & g_{62}^C & g_{63}^C & g_{64}^C & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (5)$$

and

$$P_C = \begin{pmatrix} p_{11}^C & p_{12}^C & p_{13}^C & p_{14}^C & \cdots \\ p_{21}^C & p_{22}^C & p_{23}^C & p_{24}^C & \cdots \\ p_{31}^C & p_{32}^C & p_{33}^C & p_{34}^C & \cdots \\ p_{41}^C & p_{42}^C & p_{43}^C & p_{44}^C & \cdots \\ p_{51}^C & p_{52}^C & p_{53}^C & p_{54}^C & \cdots \\ p_{61}^C & p_{62}^C & p_{63}^C & p_{64}^C & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (6)$$

The matrices (5) and (6) are composed of the column vectors

$$G_C^{\,j} = \left(g_{1j}^C, g_{2j}^C, \dots, g_{ij}^C\right)^T \quad (7)$$

and

$$P_C^{\,j} = \left(p_{1j}^C, p_{2j}^C, \dots, p_{ij}^C\right)^T, \quad (8)$$

where the elements g_{ij}^D and p_{ij}^D indicates the number of individuals (dead or alive) in the life stage j at time i. Unlike the differential representation, where the sum of the elements on the columns of the matrices (3) and (4) provides the total number of individuals into each life stage, in the matrix (5) and (6) the sum of the elements on the rows provides the total number of individuals alive (5) or dead (6) at time i. This property is further underlying the different focus that the two representations have: the differential representation is always centred on the stage, while the integral representation on the whole life cycle.

Before going ahead, it is worth remarking two important aspects. In this paper we are not going to recall the overall methodology of Chi et al. (Chi, 1988; Chi and Liu, 1985) as it has well explained in several works over the years (Amir-Maafi et al., 2022; Chi et al., 2023, 2020), but instead we are providing a connection between the differential and the integral representation. Accordingly, we have maintained the same notation of Chi et al. (Chi, 1988; Chi and Liu, 1985) for the matrices (5) and (6), even if the mortality matrix P_C has been indicated in a slightly different way. Chi et al. (Chi, 1988; Chi and Liu, 1985) in fact, do not cumulate the daily values of the dead, as carried out for the stage-development in the matrix (5), but they just report the number of new died individuals per day. In this study we are conceptually separating the differential from the integral representation, this is the reason why for the sake of coherence it would be better to express the matrix (6) as daily cumulative values. This however does not limit further analyses, as the datasets hereby represented provide the necessary information to fully apply the framework of Chi et al. (Amir-Maafi et al., 2022; Chi, 1988; Chi et al., 2023, 2020; Chi and Liu, 1985).

Unlike the differential representation, the vectors (7) and (8) of the integral representation report the information in absolute time, namely by setting the time reference on the oviposition day for all the stages. On the other hand, there is a loss of information with respect to Table 2 and 3, as already showed for the differential representation. As conclusion of this Section, we discuss the *Example 2*, that shows the consequences of the loss of information in case of a dataset collected in integral form.

Example 2. This example aims to show that the integral representation leads to a loss of information with respect to Table 2 and 3. For the sake of simplicity, let us consider a population composed of three individuals, whose eggs are laid at time $t = 0$. Moreover, let us focus on the egg (E) and larva (L) stage only. By observing the stage development on the following days, we can write the dataset as shown in Table 1. From the raw data (Table 8), we assess that: *i)* the specimen SP-1 dies on day two, *ii)* the specimen SP-2 develops to larva on day three, and *iii)* the specimen SP-3 is still in the egg stage on day four. Accordingly, from Table 8 we can know who the individuals involved in the aforementioned scenario are.

Table 8: Raw dataset discussed in the *Example 2*. Rows indicate the number of the specimens, while columns the days. "E" indicates *egg*, "L" *larvae*, and "D" the *dead*.

Specimen	Time (day)					
	1	2	3	4	5	…
SP-1	E	E	D	--	--	…
SP-2	E	E	E	L	L	…
SP-3	E	E	E	E	E	…

From Table 8, we can obtain the integral representation by counting how many individuals are in each stage on day two, three, and four (Table 9): *i)* on day two we have 1 individual dead and 2 individuals in the egg stage, *ii)* on day three we have one egg and one larva, *iii)* on day four we have one egg and one larva.

Table 9: Integral representation of the dataset discussed in the *Example 2*.

Time	Development		Mortality	
(days)	E	L	E	L
1	3	0	0	0
2	2	0	1	0
3	1	1	1	0
4	1	1	1	0
⋮	⋮	⋮	⋮	⋮

Table 9 shows that we are losing, with respect to Table 8, the following information: *i)* the specimen one died on day two, *ii)* the duration of the egg stage of the individual two was three days, *iii)* the duration of the egg stage of the individual four is 4 days. As already discussed for the differential representation in the *Example 1*, there might be different combinations of Table 9 that leads to the original configuration shown in Table 8.

2.4. Spreadsheet file for standard of data collection

The two representations described in Section 2.2 and 2.3 are complementary to each other and allow an overall understanding of the biology of the species. During this study, we analysed the whole process of data collection and transformation, highlighting the importance of data sharing to maximise the information that life table studies provide to the scientific community. To pursue this aim, we hereby propose a simple solution, namely an electronic spreadsheet file where scholars can insert data and automatically obtain the differential and the integral representation as a result. The idea behind this choice comes from our experience: experimenters are used to collecting life tables data directly on electronic spreadsheets, often organised according to their personal needs. As a consequence, it is extremely difficult, for third people, to understand the logic behind the data collection, as well as to retrieve the main information needed to fully reproduce the results of the experiments. Further calculations to reach the differential or the integral representation are then carried out by transforming and/or reshaping the data, a very time-consuming operation that further slows down the overall experimental process.

350 The main goal of the spreadsheet we are proposing is either to standardise the data collection or to guide
351 scholars in set up their experiments such that the reproducibility of the results and the reusability of the data is
352 warranted. Data elaboration is not the only disadvantage of the experimental process, as data are usually
353 collected and analysed at the end of the trial. An online processing of the data, instead, may be helpful to
354 understand if there are anomalies or adjustments needed in the experimentation. As a last motivation of our
355 choice, we can say that it is not worthy, at this stage, to build an *ad hoc* software without having a well-defined
356 standard, forcing the potential users to install and get familiar with it. An electronic spreadsheet can further be
357 uploaded on online platforms that extrapolate the data and store them as already common in gene banks.
358 The development of the spreadsheet hereby proposed is based on the following logic process. The file should
359 be self-explanatory for external readers; accordingly, the first tab should contain a detailed resume of the data
360 about the species, the experimental conditions explored, and the acronyms that denote the life stages.
361 Moreover, it would be extremely helpful if the researchers report additional details, such as the main aim of
362 the experiment and the rearing protocol, besides citations to helpful literature. From our experience, a clear
363 summary of all the information in a single file will speed up the manuscript writing process as well, as the
364 spreadsheet can be considered, *de facto*, an electronic logbook. The overview tab is also a first important step
365 to give credit to the research group.
366 The second part of the file concerns the data acquisition. We may use a single spreadsheet for each constant
367 condition experiment; however, this is not practical for data sharing and further analysis. This concept., can be
368 explained with a practical example: let us consider that we aim to explore the thermal response of a given
369 insect species; to do that, we should repeat the rearing at different constant temperatures. If the dataset
370 corresponding to the different experiments is included in a single spreadsheet, it is possible to easily compare
371 the overall results. For this reason, the spreadsheet proposed consider 15 tabs, corresponding to 15 independent
372 rearing carried out under different conditions. The number of tabs is just a compromise, as in the average the
373 conditions explored (e.g., temperature, humidities, diets, treatments) is around 6–10 and, in the best of our
374 knowledge, there are no cases that exceed 15. Each tab allows the experimenter to collect life history data of
375 up to 1000 individuals: data obtained from multiple experiments under the same conditions can be grouped in
376 the same tab and analysed together. Moreover, we foreseen a textbox in each tab, as it can be helpful either to
377 take personal notes or observations to share with third people.

The third part of the file, instead, should automatically calculate the differential and the integral representation of each dataset, so that they are ready to be copy/pasted for further analyses that depends on the aim of the research. Note that since the calculations are automated, the spreadsheet is uploaded as soon as a new data is inserted somewhere in a tab, an operation that allows to double check if the ongoing experiment shows anomalies and/or unexpected behaviours of the population that need further investigation. The two representations can be obtained by transforming the raw data according to the procedure detailed in Section 2.2 and 2.3.

The fourth part of the file is the life tables statistics and plots, reported in two corresponding tabs. The life tables statistics are computed from the differential representation and provide a summary of the stage durations under the different rearing conditions. More specifically, the mean and the standard deviation of the stage duration and the moments of the distribution (median, mode, skewness, and kurtosis) are automatically computed for each rearing condition. The plot tab, instead, plots the distributions coming from the different rearing conditions, grouped by stage, for quick and intuitive comparison of the results.

3. Results

The theoretical considerations discussed in this study lead us to organise the spreadsheet for life tables data collection such that the information is standardised and complete of the details needed to reproduce the experiments. The datasheet can be downloaded at the following GitHub page https://github.com/lucaros1190/LifeTables-StandardSpreadsheet, where it is also possible to notify further improvements or report bugs. Two versions of the datasheet are available: a "user" version, protected to avoid accidental modifications of the formulas in the different cells, and a "developer" version, free and ready to be modified or improved according to special needs. The web page contains a detailed guide of the datasheet, and will be enriched with some examples based on real experiments carried out over the years by our research group and collaborators, as well. Let us now discuss the details of each tab, to show how the theoretical aspects have been implemented.

405 *DatasetOverview*: The first page of the file contains details about the species, the credits to the research group
406 that carried out the experiment, the definitions that will be further provided as input to the other tabs, and a
407 text box where the rearing protocols and the aim of the experiment should be reported. For the sake of
408 generality, the rearing conditions are indicated by "Cond-1, …, Cond-15", the same name indicated on the raw
409 data tabs. In this first page, it is possible to assign names/acronyms to the different "Cond-", that will be
410 automatically propagated along the spreadsheet.
411 A second general trait concerns the number of life stages and the sex division. We are aware that the number
412 of stages can vary among the species, and this is the reason why they have been exceeded on purpose: in case
413 the species under study has less preimaginal stages, the exceeding ones can be skipped by indicating "--" as a
414 label. The same reasoning is applied to parthenogenetic species: if there are only females in the populations,
415 the adult males' cell can be skipped by indicating "--", as well.
416
417 *LifeTables-Statistics:* This tab provides a summary of the data inserted, by exploiting the differential
418 representation from the *Differential-Rep* tab. Here, users can either have real-time overview/synthesis of the
419 experimental data, helpful during the trials, or the life tables values ready to be used elsewhere (e.g., in a
420 manuscript/report). This tab provides life tables statistics of the different life stages and experimental
421 conditions explored, according to the indications of Rossini et al. (Rossini et al., 2024a). This tab provides a
422 textbox as well, helpful to write notes for third people or for the experimenter themself.
423
424 *Plots:* After the numerical synthesis is carried out in the *LifeTables-Statistics* tab, a graphical summary follows
425 as well. This tab provides a plot of the differential representation, coming from the *Differential-Rep* tab. The
426 stage-populations corresponding to the different rearing conditions are grouped together in the same plot, so
427 that users can easily see differences and insights of their studies. For the sake of completeness, plots are
428 extended to mortality tables as well.
429
430 *Cond-1, …, Cond-15:* These tabs are the core of the datasheet file, as the data collected for each rearing
431 condition should be directly inserted here to produce the Table 1 of this study. Even if the sampling time is
432 usually set to one day, this is not generally true, as some species develop more rapidly (above all at optimal

433 temperatures), requiring multiple inspections per day. For this reason, each tab provides for a cell where the
434 sampling time (expressed in hours) should be indicated. Once defined, the times corresponding to that
435 experimental conditions are automatically converted in days (or fraction of days) in the overall spreadsheet.
436 For the users' convenience, a list indicating the labels defined in the *DatasetOverview* is provided on top of
437 the tab.
438 The table, instead, is organised in columns indicating the time step of each sampling, and in rows indicating
439 the individuals: note that in this way it is possible to uniquely identify the individuals, taking track of the life
440 history of up to 1000 specimens. It may be possible that, to reach a statistically significant number of reared
441 individuals, the experimenter should rear many cohorts under the same conditions. In this case, the data can
442 be pooled together in the same tab, as they are independent repetitions of the same experiment, and the different
443 rearing sessions can be indicated in the textbox. An example at hand can be: "Specimens 1-50 correspond to
444 the rearing session 1, Specimens 51-100 correspond to the rearing session 2".
445
446 *Individual-LifeHistory:* This tab contains the first summary of the data, more precisely it transforms each raw
447 dataset in the stage-development table (Table 2), according to the procedure described in Section 2.2. For each
448 individual, in fact, this tab counts how many cells (corresponding to a single time step increase) contain the
449 same label. The same is carried out for mortality, obtaining the stage-mortality table (Table 3). The stage-
450 development table contains only the preimaginal stages, as adult survival is included in the stage-mortality
451 table.
452
453 *Differential-Rep:* This tab contains the second summary of the data, namely the differential representation.
454 Calculations are carried out according to the procedure described in Section 2.2, organising the dataset as in
455 Table 4. Since the differential representation is providing us the number of individuals that *complete* the stage
456 (or that exit from the stage because they die) over time, adult males' and females' survival should belong to
457 mortality table. This is the reason why the differential representation tables related to development include
458 only the preimaginal stages.
459

460 *Integral-Rep:* This tab contains the third summary of the raw data, namely the integral representation. The
461 tables report, for each experimental condition, the number of individuals that over time are into the different
462 life stages (Table 7), this is the reason why this time the adults are present in both development and mortality
463 tables.
464
465 **4. Discussion**
466
467 Standardising data collection is fundamental to endorse open science and reproducibility of experiments and
468 results. We contributed to this goal in the case of life table experiments in entomology, widely used to explore
469 the biology and the response of insect populations to different external conditions. Besides providing a possible
470 solution, this study wanted to also revise the theoretical background behind life tables experiments, to better
471 clarify which are the procedures that lead to the two different representations of the data.
472 The differential and the integral representation have been used since long time by the community of
473 entomologists working in life tables experiments but, to the best of our knowledge, there are no in-depth studies
474 that discuss their connection and their advantages. The differential representation, for instance, provides the
475 distribution of the development times of the individuals that develop under different external conditions
476 (Carey, 2001; Harcourt, 1969; Rossini et al., 2024a; Southwood, 1978). From those distributions it is possible
477 to carry out different analyses: the most common one is the calculation of the synthetic values listed in
478 *LifeTables-Statistics*, as they provide a direct idea of the impact of the conditions on the different life stages of
479 the species. More accurate analyses of the differences between the experimental datasets, instead, can be
480 carried out through statistical tests, such as Generalised Linear Models (GLM), classical parametric tests (e.g.,
481 analysis of the variance ANOVA) or non-parametric tests (e.g., Kruskal-Wallis test). The common process of
482 data analysis provides for an intermediate step of data preparation, usually by organising the data in .csv or
483 .xlxs files, that is regardless of the software used. By experience, this phase is extremely time consuming and
484 slows down the overall analysis and the further manuscript preparation process. The datasheet provided with
485 this study automatically prepares the dataset for the most common software utilised by the community (e.g.,
486 R, SAS, custom Python scripts, or Matlab), endorsing the researchers to easily get their results. Future studies

will focus on the development of scripts that automatically analyse the data, providing an additional benefit for the scientific community.

There is the same advantage for the integral representation, even if the data analysis in this case has been better discussed and organised by Chi et al. (Chi et al., 2023, 2020). While the differential representation focuses more on how the rearing conditions affect the single stage populations, the integral one provides an overview of the overall life cycle. This aspect is also underlined by the different time scales considered: as discussed in Section 2.2 and 2.3, the absolute time scale (the time zero set to the oviposition day for each life stage population) is typical of the integral representation, while the differential one entrusts on a relative time scale (the time zero is set at the beginning of each life stage). For this reason, the integral representation is more suitable to calculate parameters such as the age-stage specific survival rate, distribution of the mortality, age-specific survival rate, net reproductive rate, intrinsic rate of increase, finite rate of increase, age-stage life expectancy, and reproductive value (Chi, 1988; Chi and Liu, 1985). Given the higher complexity of this analysis with respect to the differential representation, Chi et al. introduced an *ad hoc* software, TWOSEX-MSChart (http://140.120.197.173/Ecology/prod02.htm). TWOSEX-MSChart provides for an input that is close to the *Individual-LifeHistory* tables: even in this case our datasheet shows its utility to speed up the data analysis process, besides supporting the data sharing.

The suitability of the data for third parties' software is not the only advantage of the spreadsheet introduced with this study, as it can be useful as a guide for researchers approaching to life tables experiments. The data, in fact, should be inserted in the same way, independently from the species or from the rearing conditions: the logic behind the datasheet and its structure provides instructions on how to carry out the experiments. Additionally, as every dataset contains information in the same format and units, it can endorse the constitution of a common database, as in the case of gene banks. Our wish for the future is that the datasheet starts to be used to collecting the data, and further uploaded on web databases (with specific identifiers). Even if this hypothesis is still remote, the advantage would be great, as datasets obtained under the same experimental conditions can be combined and available to the scientific community with the observations and notes reports as well.

Open science practices are going to be supported and, in some cases imposed, by always more institutions, journals, and funding agencies worldwide (Mons et al., 2017). This willingness led to the constitution of the

Findable, Accessible, Interoperable, and Reusable (FAIR) policies (Stall et al., 2019), adopted for instance by the European Union, where "*open science is at centre of European research policy*" (https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en). FAIR policy endorses the transparency of the experiments while offering the right protection of the data (they might be openly accessible, or before proper requests) and credit for the authors (Mons et al., 2017). This tendency should further motivate the standardisation of the data collection (even if FAIR policies do not impose standards) and analysis, so as everyone from the scientific community can easily understand what the data and the results mean without going through further explanations contained in *readme* files or published papers. On the other hand, the scientific literature associated with the data will be the reference only for the objectives, rearing conditions, and results.

However, we should clarify that standardisation of the data collection is not a fast process, as it needs to be discussed by the scientific community before being accepted. What this study introduces is a proposal that can be tested and further improved by the scientific community, as it still contains limitations. For instance, we focused on the different representations and to an easy way to collect the data about insect (and terrestrial arthropods at large) life cycles, discarding reproduction aspects. We are aware that this is an important aspect, but we should also recognise that the method introduced by Chi et al. (Chi, 1988; Chi and Liu, 1985) is already suitable to deal with egg production. What we suggest to the researchers focusing on this specific part of the life tables study, is to recall the cited literature, as including this aspect in our framework would be just a repetition of something that already exist and that has been widely accepted by the community. Under this point of view, our proposal is an extension and a clarification of the works of Chi et al. (Chi, 1988; Chi et al., 2023, 2020; Chi and Liu, 1985) above all concerning the theoretical aspects on the differential representation, that to date has never been discussed properly.

Raw data of life tables experiments are extremely important, above all in the current age of the entomological research. Besides the biological information and their implication in many fields as, for instance, insect pest control or edible insect industry, they are also the basics of parameter estimation for mathematical models (Buffoni and Pasquali, 2007; Castex et al., 2018; Damos and Savopoulou-Soultani, 2012; Gutierrez et al., 1981, 2012; Ikemoto and Kiritani, 2019; Quinn, 2017; Rossini et al., 2019a, 2019b, 2020a). Modelling insect populations is one of the hot topics of the modern research, as they can potentially provide predictions on

543 future population trends (Grimm et al., 2020; Rossi et al., 2019; Rupnik et al., 2019; Zhai et al., 2020). There
544 are many proposals, to date, based on different mathematical frameworks, but the common starting point is
545 laboratory data that allows the parameter estimation (Rossini et al., 2025, 2020b). Therefore, the unavailability
546 of data can be a strong limitation to this field of research, as not all the researchers involved in this kind of
547 studies have access to experimental data, or connections with groups specialised in life tables experiments.
548 Additional aspect, not negligible in many cases, is the budget constrain: climatic chamber experiments require
549 specialised laboratory with proper equipment and experienced personnel, and this is one of the main limitations
550 that affect several research groups worldwide (Kareithi et al., 2019). Multidisciplinary research can be
551 endorsed by open data and by collecting data according to standards, as the interpretation of the dataset can be
552 easier for researchers coming from different fields.
553
554 **5. Conclusion**
555
556 This work dealt with the conceptual clarification and standardisation of the data collection process of life tables
557 in entomology. We have provided a conceptual explanation of the different representations that can be obtained
558 from the raw data, underlying why only sharing the experimental dataset provides the complete information to
559 the scientific community. We are aware that this is just a first step that needs further adjustments and feedback
560 from the insiders, but we strongly believe that this can be carried out in the very next future. Our hope is that
561 the scientists working with life tables experiments will adopt our data collection protocol and shares the data
562 in public repositories, enriching the knowledge on many topics that are relevant for the current and future
563 challenges of research.
564
565 **Acknowledgements**
566 L.R. is funded by the European Commission under the Marie Sklodowska Curie Actions Postdoctoral
567 Fellowship (MSCA-PF-2022) project “PestFinder” Grant n. 101102281.
568
569 **References**
570

Alves, F., Edlmann, M., Costa, J., Costa, P., Macedo, P., Leal Da Costa, P., Symington, C., 2013. Heat requirements and lenght of phenological stages. Effects of rootstock on red grape varieties at Douro region. 18th International Symposium GIESCO, Porto, Portugal, 7-11, July 2013. 5p.

Amir-Maafi, M., Chi, H., Chen, Z.-Z., Xu, Y.-Y., 2022. Innovative bootstrap-match technique for life table set up. Entomologia Generalis 42, 597–609. https://doi.org/10.1127/entomologia/2022/1334

Bellows, T.S., 1992. Life-table construction and analysis in the evaluation of natural enemies. Annual Review of Entomology 37, 587–614.

Bieri, M., Baumgärtner, J., Bianchi, G., Delucchi, V., Arx, R., 1983. Development and fecundity of pea aphid (Acyrthosiphon pisum Harris) as affected by constant temperatures and by pea varieties. Mitteilungen der Schweizerischen Entomologischen Gesellschaft 56, 163–171.

Buffoni, G., Pasquali, S., 2007. Structured population dynamics: Continuous size and discontinuous stage structures, Journal of Mathematical Biology. https://doi.org/10.1007/s00285-006-0058-2

Carey, J.R., 2001. Insect biodemography. Annual Review of Entomology 46, 79–110.

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

Chi, H., 1988. Life-table analysis incorporating both sexes and variable development rates among individuals. Environmental Entomology 17, 26–34. https://doi.org/10.1093/ee/17.1.26

Chi, H., Kavousi, A., Gharekhani, G., Atlihan, R., Salih Özgökçe, M., Güncan, A., Gökçe, A., Smith, C.L., Benelli, G., Guedes, R.N.C., Amir-Maafi, M., Shirazi, J., Taghizadeh, R., Maroufpoor, M., Xu, Y.-Y., Zheng, F.-Q., Ye, B.-H., Chen, Z.-Z., You, M.-S., Fu, J.-W., Li, J.-Y., Shi, M.-Z., Hu, Z.-Q., Zheng, C.-Y., Luo, L., Yuan, Z.-L., Zang, L.-S., Chen, Y.-M., Tuan, S.-J., Lin, Y.-Y., Wang, H.-H., Gotoh, T., Shaef Ullah, M., Botto-Mahan, C., De Bona, S., Bussaman, P., Gabre, R.M., Saska, P., Schneider, M.I., Ullah, F., Desneux, N., 2023. Advances in theory, data analysis, and application of the age-stage, two-sex life table for demographic research, biological control, and pest management. Entomologia Generalis 43, 705–732. https://doi.org/10.1127/entomologia/2023/2048

Chi, H., Liu, H., 1985. Two new methods for the study of insect population ecology. Bull. Inst. Zool 24, 225–240.

599 Chi, H., You, M., Atlıhan, R., Smith, C.L., Kavousi, A., Özgökçe, M.S., Güncan, A., Tuan, S.-J., Fu, J.-W.,
600 Xu, Y.-Y., Zheng, F.-Q., Ye, B.-H., Chu, D., Yu, Y., Gharekhani, G., Saska, P., Gotoh, T., Schneider,
601 M.I., Bussaman, P., Gökçe, A., Liu, T.-X., 2020. Age-Stage, two-sex life table: an introduction to
602 theory, data analysis, and application. Entomologia Generalis 40, 103–124.
603 https://doi.org/10.1127/entomologia/2020/0936
604 Damos, P., Savopoulou-Soultani, M., 2012. Temperature-driven models for insect development and vital
605 thermal requirements. Psyche 2012. https://doi.org/10.1155/2012/123405
606 De Campos, M.R., Béarez, P., Amiens-Desneux, E., Ponti, L., Gutierrez, A.P., Biondi, A., Adiga, A., Desneux,
607 N., 2020. Thermal biology of Tuta absoluta: demographic parameters and facultative diapause. Journal
608 of Pest Science. https://doi.org/10.1007/s10340-020-01286-8
609 Genç, H., Nation, J.L., 2008. Survival and development of Bactrocera oleae Gmelin (Diptera:Tephritidae)
610 immature stages at four temperatures in the laboratory. African Journal of Biotechnology 7, 2495–
611 2500. https://doi.org/10.5897/AJB08.395
612 Grimm, V., Johnston, A.S.A., Thulke, H.-H., Forbes, V.E., Thorbek, P., 2020. Three questions to ask before
613 using model outputs for decision support. Nature Communications 11, 4959.
614 https://doi.org/10.1038/s41467-020-17785-2
615 Gutierrez, A.P., Baumgärtner, J., Hagen, K.S., 1981. A conceptual model for growth, development, and
616 reproduction in the ladybird beetle, Hippodamia convergens (Coleoptera: Coccinellidae). The
617 Canadian Entomologist 113, 21–33. https://doi.org/10.4039/Ent11321-1
618 Gutierrez, A.P., Ponti, L., Cooper, M.L., Gilioli, G., Baumgärtner, J., Duso, C., 2012. Prospective analysis of
619 the invasive potential of the European grapevine moth Lobesia botrana (Den. & Schiff.) in
620 California. Agricultural and Forest Entomology 14, 225–238. https://doi.org/10.1111/j.1461-
621 9563.2011.00566.x
622 Harcourt, D.G., 1969. Development and use of life tables in study of natural insect populations. Annual Review
623 of Entomology 14, 175. https://doi.org/10.1146/annurev.en.14.010169.001135
624 Hsieh, J., 1991. A general theory of life table construction and a precise abridged life table method. Biomedical
625 Journal 33, 143–162.

626 Ikemoto, T., Kiritani, K., 2019. Novel Method of Specifying Low and High Threshold Temperatures Using
627 Thermodynamic SSI Model of Insect Development. Environmental Entomology 48, 479–488.
628 https://doi.org/10.1093/ee/nvz031
629 Kareithi, D.N., Salifu, D., Owuor, N., Subramanian, S., Tonnang, E.Z.H., 2019. An algorithm for data
630 reconstruction from published articles – Application on insect life tables. Cogent Mathematics &
631 Statistics 6, 1701377. https://doi.org/10.1080/25742558.2019.1701377
632 Korpelainen, H., 1986. The effects of temperature and photoperiod on life history parameters of *Daphnia*
633 *magna* (Crustacea: Cladocera). Freshwater Biology 16, 615–620. https://doi.org/10.1111/j.1365-
634 2427.1986.tb01004.x
635 Leslie, P.H., 1948. Some further notes on the use of matrices in population mathematics. Biometrika 35, 213.
636 https://doi.org/10.2307/2332342
637 Leslie, P.H., 1945. On the Use of Matrices in Certain Population Mathematics. Biometrika 33, 183.
638 https://doi.org/10.2307/2332297
639 Liu, J., Huang, W., Chi, H., Wang, C., Hua, H., Wu, G., 2017. Effects of elevated CO2 on the fitness and
640 potential population damage of Helicoverpa armigera based on two-sex life table. Sci Rep 7, 1119.
641 https://doi.org/10.1038/s41598-017-01257-7
642 Mari, L., Carbone, P., Petri, D., 2012. Measurement fundamentals: A pragmatic view. IEEE Transactions on
643 Instrumentation and Measurement 61, 2107–2115. https://doi.org/10.1109/TIM.2012.2193693
644 Mons, B., Neylon, C., Velterop, J., Dumontier, M., Da Silva Santos, L.O.B., Wilkinson, M.D., 2017. Cloudy,
645 increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud.
646 ISU 37, 49–56. https://doi.org/10.3233/ISU-170824
647 Moshtaghi Maleki, F., Iranipour, S., Hejazi, M.J., Saber, M., 2016. Temperature-dependent age-specific
648 demography of grapevine moth (Lobesia botrana) (Lepidoptera: Tortricidae): jackknife vs. bootstrap
649 techniques. Archives of Phytopathology and Plant Protection 49, 263–280.
650 https://doi.org/10.1080/03235408.2016.1140566
651 NCBI Resource Coordinators, 2012. Database resources of the National Center for Biotechnology Information.
652 Nucleic Acids Research 41, D8–D20. https://doi.org/10.1093/nar/gks1189

Nielsen, A.L., Hamilton, G.C., Matadha, D., 2008. Developmental Rate Estimation and Life Table Analysis for Halyomorpha halys (Hemiptera : Pentatomidae) Developmental Rate Estimation and Life Table Analysis for Halyomorpha halys (Hemiptera : Pentatomidae). Environmental Entomology 37, 348–355.

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

Rossi, V., Sperandio, G., Caffi, tito, Simonetto, A., Gilioli, G., 2019. Critical success factors for the adoption of decision tools in IPM. Agronomy 9, 710. https://doi.org/10.3390/agronomy9110710

Rossini, L., Bono Rosselló, N., Benhamouche, O., Contarini, M., Speranza, S., Garone, E., 2025. A general DDE framework to describe insect populations: Why delays are so important? Ecological Modelling 499, 110937. https://doi.org/10.1016/j.ecolmodel.2024.110937

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

Rossini, L., Contarini, M., Severini, M., Speranza, S., 2020a. Reformulation of the Distributed Delay Model to describe insect pest populations using count variables. Ecological Modelling 436, 109286. https://doi.org/10.1016/j.ecolmodel.2020.109286

Rossini, L., Contarini, M., Speranza, S., 2021b. A novel version of the Von Foerster equation to describe poikilothermic organisms including physiological age and reproduction rate. Ricerche di Matematica 70, 489–503. https://doi.org/10.1007/s11587-020-00489-6

Rossini, L., Contarini, M., Speranza, S., Mermer, S., Walton, V., Francis, F., Garone, E., 2024a. Life tables in entomology: A discussion on tables' parameters and the importance of raw data. PLoS ONE 19, e0299598. https://doi.org/10.1371/journal.pone.0299598

Rossini, L., Locatelli, D.P., Limonta, L., 2024b. Development of Idaea inquinata (Lepidoptera Geometridae) at different constant temperatures and relative humidities under controlled conditions. Journal of Stored Products Research 109, 102466. https://doi.org/10.1016/j.jspr.2024.102466

680 Rossini, L., Severini, M., Contarini, M., Speranza, S., 2020b. EntoSim, a ROOT-based simulator to forecast
681 insects' life cycle: Description and application in the case of Lobesia botrana. Crop Protection 129,
682 105024. https://doi.org/10.1016/j.cropro.2019.105024
683 Rossini, L., Severini, M., Contarini, M., Speranza, S., 2019a. A novel modelling approach to describe an insect
684 life cycle vis-à-vis plant protection: description and application in the case study of Tuta absoluta.
685 Ecological Modelling 409, 108778. https://doi.org/10.1016/j.ecolmodel.2019.108778
686 Rossini, L., Severini, M., Contarini, M., Speranza, S., 2019b. Use of ROOT to build a software optimized for
687 parameter estimation and simulations with Distributed Delay Model. Ecological Informatics 50, 184–
688 190. https://doi.org/10.1016/j.ecoinf.2019.02.002
689 Rupnik, R., Kukar, M., Vračar, P., Košir, D., Pevec, D., Bosnić, Z., 2019. AgroDSS: A decision support system
690 for agriculture and farming. Computers and Electronics in Agriculture 161, 260–271.
691 https://doi.org/10.1016/j.compag.2018.04.001
692 Samayoa, A.C., Choi, K.S., Wang, Y.-S., Hwang, S.-Y., Huang, Y.-B., Ahn, J.J., 2018. Thermal effects on the
693 development of Bactrocera dorsalis (Hendel) (Diptera: Tephritidae) and model validation in Taiwan.
694 Phytoparasitica 46, 365–376. https://doi.org/10.1007/s12600-018-0674-6
695 Sánchez-Ramos, I., Álvarez-Alfageme, F., Castañera, P., 2007. Effects of relative humidity on development,
696 fecundity and survival of three storage mites. Exp Appl Acarol 41, 87–100.
697 https://doi.org/10.1007/s10493-007-9052-7
698 Sherry, S.T., 2001. dbSNP: the NCBI database of genetic variation. Nucleic Acids Research 29, 308–311.
699 https://doi.org/10.1093/nar/29.1.308
700 Southwood, T.R.E., 1978. The construction, description and analysis of age-specific life-tables, in: Ecological
701 Methods.
702 Stall, S., Yarmey, L., Cutcher-Gershenfeld, J., Hanson, B., Kerstin, L., Nosek, B., 2019. Make all scientific
703 data FAIR. Nature 570, 27–29.
704 Tochen, S., Dalton, D.T., Wiman, N., Hamm, C., Shearer, P.W., Walton, V.M., 2014. Temperature-related
705 development and population parameters for *Drosophila suzukii* (Diptera: Drosophilidae) on cherry and
706 blueberry. Environ Entomol 43, 501–510. https://doi.org/10.1603/EN13200

Van Nieuwenhove, G.A., Frías, E.A., Virla, E.G., 2016. Effects of temperature on the development, performance and fitness of the corn leafhopper Dalbulus maidis (DeLong) (Hemiptera: Cicadellidae): implications on its distribution under climate change. Agricultural and Forest Entomology 18, 1–10. https://doi.org/10.1111/afe.12118

Vantornhout, I., Minnaert, H.L., Tirry, L., De Clercq, P., 2005. Influence of diet on life table parameters of Iphiseius degenerans (Acari: Phytoseiidae). Exp Appl Acarol 35, 183–195. https://doi.org/10.1007/s10493-004-3940-x

Wen, D., Liu, J., Fan, S., Zhang, Z., Wu, G., 2020. Evaluation on the fitness and population projection of *Nilaparvata lugens* in response to elevated CO_2 using two-sex life table. International Journal of Pest Management 66, 368–377. https://doi.org/10.1080/09670874.2019.1654146

Zhai, Z., Martínez, J.F., Beltran, V., Martínez, N.L., 2020. Decision support systems for agriculture 4.0: Survey and challenges. Computers and Electronics in Agriculture 170, 105256. https://doi.org/10.1016/j.compag.2020.105256

Zhan, M., Xie, X., 2020. Precision measurement physics: physics that precision matters. National Science Review 7, 1795–1795. https://doi.org/10.1093/nsr/nwaa271