

Modelling

Elsevier Editorial System(tm) for Ecological

Manuscript Draft

Manuscript Number: ECOMOD-18-711R1

Title: INSTAR: An Agent-Based Model that integrates existing knowledge to simulate the biological cycle of a forest pest

Article Type: VSI: Modelling forest systems

Keywords: Thaumetopoea pityocampa; Agent-Based Model (ABM); forest pest; biological cycle; pattern-oriented

Corresponding Author: Dr. Francisco Bonet,

Corresponding Author's Institution:

First Author: María Suárez-Muñoz

Order of Authors: María Suárez-Muñoz; Francisco Bonet; José A Hódar; Javier Herrero; Mihai Tanase; Lucía Torres-Muros



Professor Brian D. Fath, Ph.D.  
Editor-in-Chief, *Ecological Modelling*

Ref. Manuscript ECOMOD-18-711

“INSTAR: an Agent-Based model linking climate and the biological cycle of forest pests in Mediterranean ecosystems.” by María Suárez-Muñoz, Francisco Bonet-García, José A. Hódar, Javier Herrero, Mihai Tanase and Lucía Torres-Muros.

Dear Editor and Guest Editors,

Thank you for your letter of 18 February 2019, in which you include your and the reviewers’ comments of our paper. We are thankful for the opportunity to resubmit our manuscript. We greatly appreciate the input received and the opportunity to address your concerns. We have carefully taken the comments into consideration in preparing our revision, which has resulted in a paper that is clearer, broader and more compelling. We provide our detailed responses to your and the reviewer’s criticisms on the pages below, with explanations of the changes we made and their locations in the text.

Reviews provided many useful comments and suggestions. The major concern, shared by editor and reviewers was the ambition on the work. We were definitely too ambitious by referring to the long-term goal of our project instead of presenting current state of the model and its capabilities. As the revised text reflects, *INSTAR* is not a predictive model but rather an open product which allows to identify gaps in our knowledge and is flexible enough to incorporate new knowledge in the future. This point has been thoroughly incorporated into the text in order to portrait a realistic perspective of the model state. Specifically, model purpose (line 227 and followings) has been fully rewritten to adjust it to this idea and avoid giving incorrect expectations of the model potentials. Moreover, the experiments are now introduced as an initial consistency test, which corresponds better to what we aimed with these tests, as reviewer #2 pointed out. The conclusions section now highlights which type of information can be extracted from *INSTAR*, and which are the future steps to follow. Thus, we considered that the test is now better contextualised and readers will have a better perspective of the model capabilities.

The other major concern was related to the model assumptions and parameter estimation as there still exist big gaps of knowledge regarding the species biology. We agree that the level of knowledge regarding *T. pityocampa* is rather low or imperfect. We are aware that a lot of the available information about the species is not optimal for modelling purposes, as it is poorly reported and could be outdated (such as data from Dèmolin). However, we considered that as this is the available information, we would rather use it to build a functioning model which contains several interlinked processes than not building the model at all. Therefore, some of the parameters are given estimated values based on this information, but would require further parameterisation in the future. That is why the idea of *INSTAR* as an open product has been thoroughly highlighted in the new version of the manuscript. With these changes, we would like to avoid the readers to extract the wrong impression of *INSTAR* being a ready-to-use model. By building this model, we precisely would like to provide a tool which helps to identify gaps of knowledge. This idea was not properly reflected in the previous version of this manuscript, we hope that it is clearer now.

Apart from focusing on each of the editors and reviewers' comments, the whole text has been revised and minor changes have been made to improve the writing, understanding and flow: programming language has been avoided, unnecessary comments have been deleted, and model purpose, abstract and conclusions have been completely rewritten in order to adjust to the new focus.

We are grateful by the effort and thought that the reviewers have put into this manuscript, which definitely made a great improvement to our manuscript.

We hope you will now find our revised text suitable for publication in *Ecological Modelling* within the special issue entitled "Modelling forest ecosystems".

Yours sincerely,



Francisco Bonet-García

## Revision notes

Response to reviewer #2:

Comment 1:

This study is very ambitious since it aims at developing a model to describe the phenology of the pine processionary moth (PPM) and its spatial distribution at local scale. Due to the amount of data needed to parametrize such a model, a lot of parameters are "self-defined" or based on historical reports in the 1960s ignoring the current changes.

Since the aim of this study is to provide a support for decision making related to pest outbreaks, the model should rigorously describe the processes and it should be carefully parametrized. Unfortunately, a lot of data is still missing and it would be necessary to collect biological data first to be able to parametrize correctly this model then.

Answer 1:

- The high number of parameters is a common characteristic of ABM (Kelly et al., 2013, mentioned in the text in line 382). As pointed out by the review, a lot of parameters are based on historical reports. We agree that this may be a problem as they may ignore current changes. We also agree with the reviewer that some parameter estimates do not have the quality to provide a model which can be directly used for decision making. However, at this stage our aim here is to present a fully functioning model which simulates several interlinked and complex processes rather than a fully parameterised model. Moreover, the model structure allows an easy integration of new updated information as this becomes available, as referred in the text (lines 57, 389, 508-509). Besides, the lack of quality information has been highlighted in the conclusions to encourage further research (line 900 and followings). Nevertheless, parameters have been better contextualised and the reasons for parameter estimate values are better explained now (lines 391-398 and Table 2).
- The model purpose has been revised and it is now in sound with the current model state, which aims to deepen our knowledge regarding *T. pityocampa* rather than assisting decision-making as this objective is far ahead the current situation (line 227 and followings).

In addition to this main concern, I have several questions or comments:

Comment 2:

Why is it a model for forest pests in Mediterranean ecosystems? I mean: 1) on one side, what is the validity beyond the PPM? but also 2) on the other side, what does restrict the application to forest pest in Mediterranean ecosystems only? In fact, the model is applied only to a small study area in Spain (not clearly defined). It seems very difficult to have data for the PPM, so it seems

likely impossible for less-documented forest pests. The title does not match the study.

Answer 2:

- The title of the manuscript has been modified from “INSTAR: an Agent-Based model linking climate and the biological cycle of forest pests in Mediterranean ecosystems” to “INSTAR: An Agent-Based Model that integrates existing knowledge to simulate the biological cycle of a forest pest”, which is more accurate.
- As pointed out, there is no restriction in the application to the Mediterranean ecosystems and therefore such references have been eliminated.
- Regarding the of the model validity beyond *T. pityocampa*, this aspect was highlighted several times in the previous version of this manuscript, while now it is only mentioned as a possibility, as it is not the main goal of the model. Nevertheless, “Model description” section has been revised for clarification (lines 492-496: “Moreover, it could constitute a framework to build models simulating other forest pests in a wide variety of ecological systems, since it summarises the key biological processes that take place in them. As an example, movement submodels could be adjusted to simulate the dispersal and establishment of mistletoe, another common pest in pine plantations, by following Mellado and Zamora (2016)”).
- Besides this, explanation about the study area has been revised (line 631 and followings: “A virtual experimental area was created using a combination of real and built-in datasets...”), and the description figure has been improved (Appendix E).

Comment 3:

The authors claim to reproduce three patterns: phenology, spatial distribution and response to temperature. This presentation is quite strange since temperature can affect both species phenology and distribution, and thus response to temperature cannot be seen as a 3<sup>rd</sup> "pattern".

Answer 3:

- Indeed, naming it as a third pattern is not fully correct. Phenology and spatial distribution are now the two identified patterns, while the temperature experiment is presented as what it was meant to be, a stress test (lines 37, 41, 56, 616-617, 731, 894-895, among others).

Comment 4:

Explanation about different estimates are too vague. For instance, line 169: "hatching occur after a month under optimal conditions". What are these conditions? This is just an example among others.

#### Answer 4:

- It should be noted that section “Ecological background” aims to give an overview of the pest biology for those unfamiliar with this species, and not an exhaustive literature review. For an in-depth review we refer to Roques (2015) (lines 152-154: “A complete outline on the biology of *Thaumetopoea* species in general, and of *T. pityocampa* in particular, can be found in Roques (2015)”). Nevertheless, vague or undefined expressions have been reviewed, modified and referenced as much as possible (e.g. lines 179-181: “higher survival rates are observed for pine trees not affected by defoliation during previous years” (Hóðar et al. 2004), lines 206-207: “The conditions inducing extended diapause are still uncertain (but see Salman et al. 2019)”).
- “Model description” section has been reviewed as well, as explained in answer 1.

#### Comment 5:

Models describing phenology cannot be based on a given duration. Phenology is closely linked to temperature. The authors calculate a given time (days) above a threshold of development (based on historical data) (page 31 lines 546-559), but do not accumulate degrees. This is a very basic approach that could hardly capture the real life and ongoing changes with climate warming. Why not considering more classical and sophisticated phenological models?

#### Answer 5:

- The duration of the cycle is given a fixed length since *T. pityocampa* is known to keep a univoltine cycle by controlling its pupae phase in order to emerge when environmental conditions are appropriate (Berardi et al., 2015). Although, indeed, a classical phenology model based on degree accumulation would be the ideal solution, such knowledge is not (yet) available. Therefore, we opted for a simpler approach that, nevertheless, is able to capture the complex relationship between development and temperature. By setting minimum length for each phase ( $Egg_{min}$ ,  $L1_{min}$ ,  $L2_{min}$ ) and the temperature thresholds controlling how these phases are lengthened ( $T_{egg\_devel}$ ,  $T_{larvae\_devel\_high}$ ,  $T_{larvae\_devel\_internal}$ ,  $T_{larvae\_devel\_low}$ ), a minimum degree accumulation is implicitly considered for each phase (e.g. in order to complete the egg phase, a colony needs 30 days reaching temperatures above  $T_{egg\_devel}$ ). It is not, obviously, a strict degree accumulation model, since thresholds are contrasted with daily minimum and maximum temperatures. However, this approach has other advantages:
  - It is based on easily obtained available behaviour observations in the field, such as temperature at which larvae leave the nests to feed or temperatures at which they look for shelter (both high and low temperatures), while data for degree accumulation models are difficult to obtain and sensitive to experimental conditions;
  - It can be calibrated for each study area, as we present in sections 4.1.1 and 4.2.1, as populations are expected to be adapted to local conditions;

- It requires less detailed input data (minimum and maximum daily temperature versus degree accumulation), decreasing significantly the computing power required to execute the model. Nevertheless, the architecture of the model does not impede, but rather it is open to incorporate this kind of calculations if data and functions become available and computing power is enough.

Comment 6:

*Pinus* spp is important for the PPM phenology and the pest density (PPM develops faster on some *Pinus* species; and PPM has preferred host trees although it may attack less preferred species when PPM level is high). The authors only refer to the pine species (*Pinus halepensis*) on line 430 regarding another study. The authors should mention on which *Pinus* their model is applied. It is all the more important that the model is dedicated to be used for decision making in pine plantations.

Answer 6:

- *INSTAR* refers to *Pinus* spp. regarding the carrying capacity ( $cc$ ) of the hosts and the cohort division of needles ( $QI_{threshold}$ ). The carrying capacity is calculated based on data for *P. halepensis*, as explained in section 3.3.1, and it could be adapted if the model is executed on other pine species (lines 429-430: “This is the most important model assumption regarding *Pinus* species, and therefore in this study *INSTAR* is considered to apply on *P. halepensis* stands (see section 4.1). Nevertheless, the model could be applied to other pine stands, provided a carrying capacity equation for the corresponding species”). This is the most important model assumption regarding pine species. Besides this, the parameter  $QI_{threshold}$  is taken from Muukkonen (2005), based on data from *P. sylvestris*. Although ideally data from *P. halepensis* would be used, in the lack of data for *P. halepensis* we decided that it was better to use the data from Muukkonen instead of establishing an arbitrary threshold. In any case, as it has been highlighted through the text and mentioned above, the model has been designed to incorporate information and therefore allow the simulation of different *Pinus* sp., provided there is available information (lines 57, 389, 431-432, 508-509).

Comment 7:

Some figures are very difficult to read and/or understand (less realistic but more informative figures would be better). Improving the figures and the explanations in the legends is necessary. Besides, the names of the parameters look like names used in programming but not in mathematical modeling nor scientific articles.

Answer 7:

- Figures 1 and 2 and their corresponding legends have been modified in order to improve their readability.
- Parameter names, as well as state variables names, have been modified (see tables 1 and 2). Moreover, text has been fully revised and programming expressions has been rewritten for a more natural reading.

Comment 8:

Some information is also missing in the appendices. For instance, "procession distance": line 60-61, it would be good to provide the value of the average and maximum displacement. Line 117, where do the values (direction coefficients) come from? And so on...

Answer 8:

- Certainly, it would be good to provide the value of the average and maximum displacement, but we do not provide because in the reference (Robredo 1963) these data are not available. This is a common situation when working on *T. pityocampa*: a lot of classic papers give the impression that everything is known about the species, while the reality is that the level of knowledge is rather low or imperfect. This is also one of the reasons why we built the model in such a way that works even with rather deficient or incomplete data while it still allows identifying these knowledge gaps and incorporating new knowledge as this becomes available. Similar reasons apply to the moths' preference for oviposition on different tree orientations. Corresponding sections in Appendix A and B have been modified for clarification.

Comment 9:

The authors explore climate change considering  $\pm 10^{\circ}\text{C}$  (page 38 lines 707-720), which is clearly unrealistic. IPCC scenarios predict a warming of 0.3 to 4.8 $^{\circ}\text{C}$  by the end of the century, and a warming of +1.5 $^{\circ}\text{C}$  could already have large effects on forest ecosystems following a recent report of IPCC (2018). Why considering so strong changes?

Answer 9:

- This experiment is not meant to be a climate scenario to observe what would happen under the expected climate change, but rather a stress test. The importance of stress tests has been highlighted through the text (lines 624-626: "Stress tests often help to identify errors in the model that would not be noted under normal conditions and therefore extreme unrealistic inputs are used in these tests (Railsback and Grimm 2011)"), the name of this section has been modified and more explanation has been included in the revised version of the text to avoid confusion (lines 732-733: "By exposing the model to extreme climate, we aim to test the model consistency in relation to temperature influence in development. [...], extreme unrealistic inputs are used in these tests in order to observe divergent behaviours in the model.").



Response to reviewer #5:

Comment 10:

This work is a magnificent example of how to mobilize relevant, ecological sound information into a modelling framework that allows to gain significant insight into the study object at question (in this case, the Processionary moth). The authors have made an enormous effort to synthesize and use available data and evidence about this forest pest and show their far-reaching knowledge of its dynamics and ecology. This is definitely a must and a pioneer study that should make its way to the public.

My main concern is actually related to the ambition of the work. The model is justified but to my opinion too ambitious, specially in some practical aspects. Although built with elegance, the different elements included in the model may lead to huge amounts of uncertainty that it will not be easy to quantify or assess. There is also a kind of divide between the model introduction and description and the actual test conducted. It looks to me that the ambition of the model (its scope and submodules) is actually much bigger than the actual test conducted over a very small spatial scale).

My suggestion is that the authors should better describe their approach and either they restrict the description of the model to the level of detail over it was tested or that this current mismatch is better explained.

Answer 10:

- Thank you very much for your nice words about our model. Indeed, we have made a big effort to integrate and mobilize the available information about *T. pityocampa*.
- As the review points out, our work showed a great ambition, which we have tried to reduce in the current version of the manuscript. This definitely came from a wrong explanation about the model capabilities and the misunderstanding between our long-term project goal and the current model state. In our opinion, one of the main interests of building such models is that it improves our understanding of the processes that result in a system behaving the way it does. This idea was not clearly stated before, but it has been underlined now. Thus, the importance of the model in deepening our knowledge regarding this species has been highlighted against the predictive objective, which does not apply at the current model stage. The model purpose has been specially revised and it is now in sound with the current state (line 227 and followings: “The overarching purpose of *INSTAR* is to generate a deeper understanding of the population dynamics of *T. pityocampa*...”).
- Moreover, the experiments are now presented highlighting the fact that they constitute an initial consistency test, which will definitely require further work (lines 38-41, 45-47, 143-147, 603, 882-883). We hope that our approach is clearer now and expresses the model state and capabilities in a realistic way.

- Regarding model uncertainty, we agree with the review that *INSTAR* contains a high level of uncertainty which needs to be addressed. As Kelly et al. (2013) highlight, “it is still very difficult to address uncertainty in most ABMs and their simulation outputs”. Nevertheless, it is possible and therefore we highlight this in our text as future analysis (lines 919-920).

Comment 11:

Second, the applicability of the model should be much better communicated and contextualized (there is no even an explicit discussion section now in the manuscript), because the predictive ability of the model is likely to be low for common practices (i.e. pest control) unless lots of contextual information, parametrization and initialization information is actually available and known with a level of detail unlikely for this study model (I think here the example of fire spread models and decision making may be good to mention).

Answer 11:

- We agree with this point and we have revised the text accordingly. We still have not included a discussion section as the manuscript is already relatively long, but rather we have reworked the conclusions section to make it more explicit and highlight the potential of the model (i.e. knowledge gaps identification). Moreover, the predictive aspect of the model has been discarded as, as the review pointed out, at the moment the predictive ability of the model is very low.

Comment 12:

Finally, the study claims that to generate a model which can be used in decision-making, it has to be able to reproduce high level patterns observed in the real system. However, I am not sure (or I do not understand) how the authors achieve this goal. It is not clear when they use real data (i.e. phenology) to validate some patterns emerging from the model, when the paper aims at calibration of the parameters to achieve sound values and when experiments are run as a kind of sensitivity analyses to get a sense of the potential of the model capabilities. Current, the results part reads more like a general consistency test of the submodules assemblage than a real model application. I think the authors should be more explicit about this and frame it in the context of either decision-making applications or further exploration of the pest dynamics to generate hypothesis about their dynamics that may be later tested in the field... and these two objectives are very different.

Answer 12:

- At this stage, the model helps us to explore pest dynamics, indeed, while it is our hope that it will have decision-making applications in the future. This idea is translated in the text as mentioned in answer 10 (lines 227 and followings), as well as in the conclusions section.
- Results are now presented as a consistency test of different submodels (Section 4 “Model consistency test”, line 603).

- Table 3 has been included, as well as explanation through the text (lines 658-662) in order to clarify which data have been used for phenology calibration. Moreover, “Input data” and “Details” sections have been revised.

Minor comments:

Comment 13:

Why authors target pupae spatial distribution and not nest distribution in space (which is probably more likely to be detailed information obtained from the field)?

Answer 13:

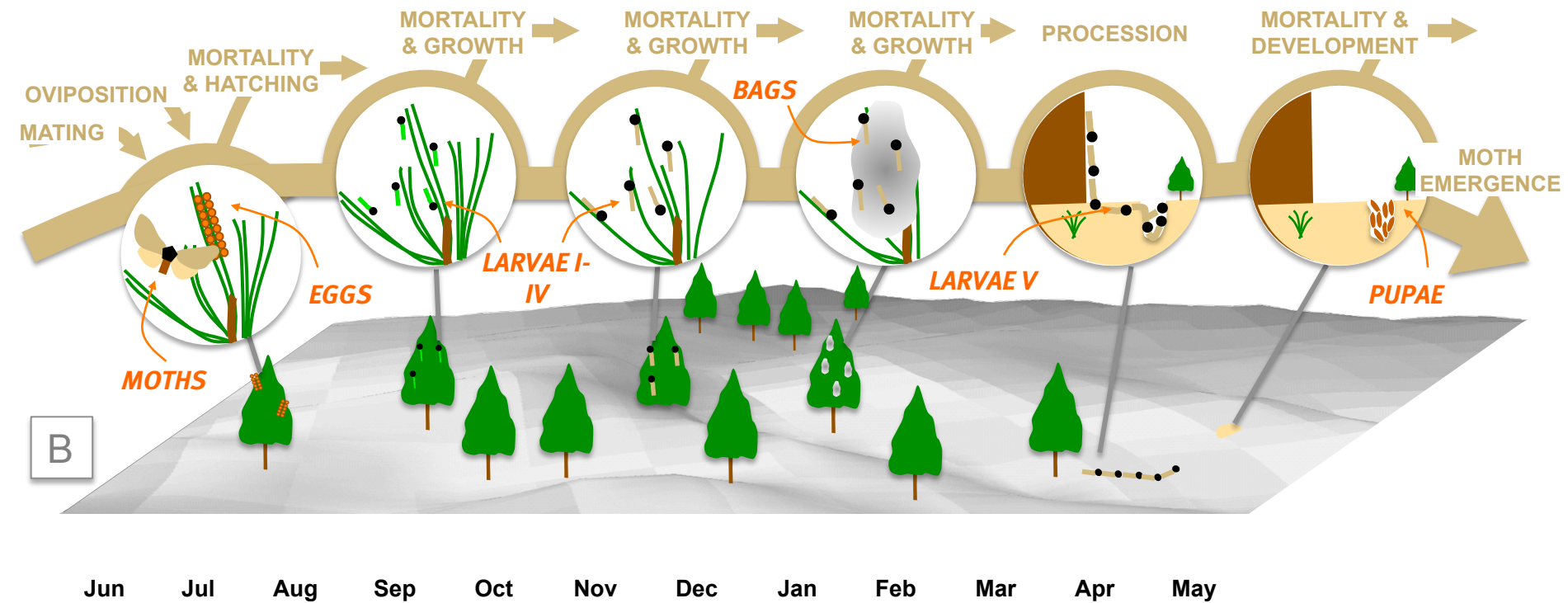
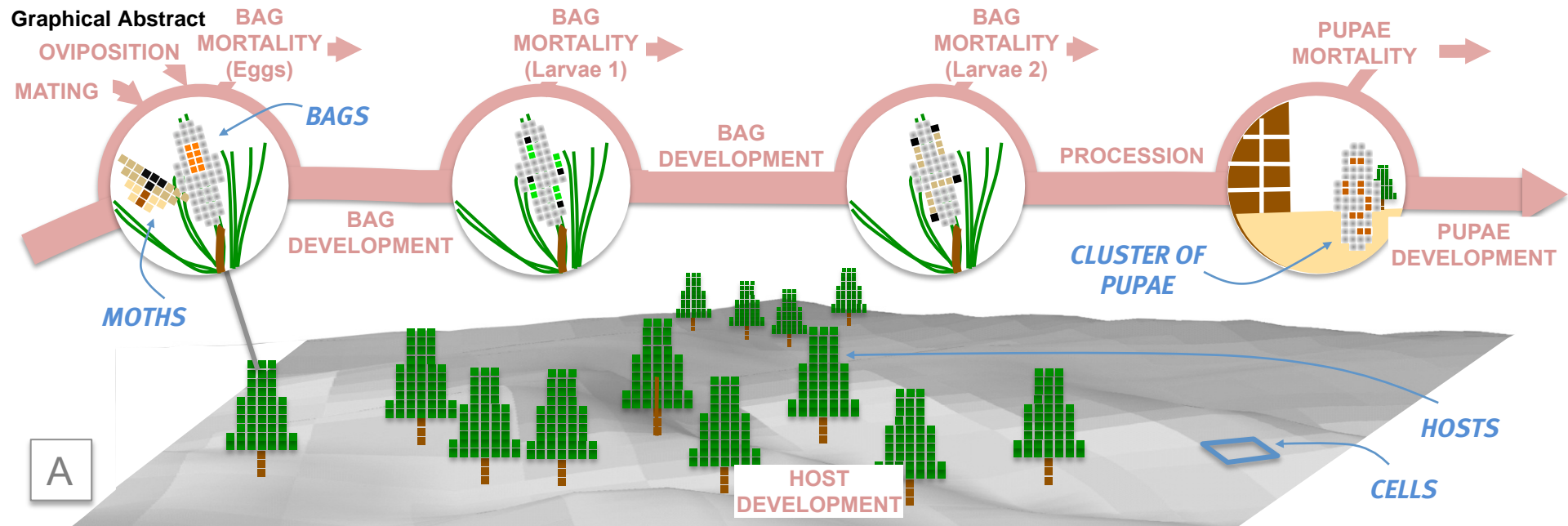
- The purpose of this experiment is to conduct a robustness analysis of the procession submodel, rather than a validation based on field observations, as we could do with nests observations. With this experiment, we aimed to answer the question: Is the submodel robust? In other words, if we deconstruct the submodel by neglecting its most important assumption (i.e. selection of burial spots based on minimum pine density), do we get the expected pattern? Nevertheless, a robustness analysis of the oviposition submodel could also be conducted. But, as this movement submodel is applied just after mating, which also implies movement, deconstruction of this submodel is not as direct and a validation based on observations would be more appropriate. The aim of the manuscript is to present the model and the initial testing performed and therefore no validation experiments have been conducted yet.

Comment 14:

Line 820. Extreme climate scenarios. This section is a good example of the mixture of calibration, validation, sensitivity analyses and future test evaluation (this section) in the paper. I think that the model is extremely powerful, but the different stages of model preparation and use should be better conveyed to the reader.

Answer 14:

- Indeed, it was not well explained. The name of the experiment has been modified and lines 732-735 (“By exposing the model to extreme climate, we aim to test the model consistency in relation to temperature influence in development. [...], extreme unrealistic inputs are used in these tests in order to observe divergent behaviours in the model”) have been added for clarification. We hope that now it is clear that the purpose of this experiment was just a stress test to check that the population collapses under unrealistic extreme temperatures, as it is expected.



### Highlights

- Agent-Based Models are a useful approach to simulate forest pests' dynamics
- *INSTAR* simulates the biological cycle of *Thaumetopoea pityocampa* in pine plantations
- Key patterns are reproduced: phenology and spatial distribution
- *INSTAR* has been designed to easily incorporate new information about the pest biology
- *INSTAR* has been designed in a modular way to allow an easy upgrade and reusability

1  
2  
3 1 **INSTAR: An Agent-Based Model that integrates existing**  
4  
5  
6  
7 2 **knowledge to simulate the biological cycle of a forest**  
8  
9  
10  
11 3 **pest**

12  
13  
14  
15 4 Suárez-Muñoz, María (1); Bonet-García, Francisco (2); Hódar, José A. (3); Herrero, Javier (4);  
16  
17 5 Tanase, Mihai (5); Torres-Muros, Lucía (6);

18  
19  
20  
21  
22  
23 6 **Affiliations**

24  
25  
26  
27  
28 7 (1) iecolab, Grupo de Ecología Terrestre, Andalusian Institute for Earth System Research  
29  
30 8 (IISTA), University of Granada, Spain

31  
32  
33  
34 9 (2) iecolab, Grupo de Ecología Terrestre, Andalusian Institute for Earth System Research  
35  
36 10 (IISTA), University of Córdoba, Spain

37  
38  
39 11 (3) Facultad de Ciencias, Departamento de Ecología, University of Granada, Spain

40  
41  
42 12 (4) Fluvial Dynamics and Hydrology Research Group, Andalusian Institute for Earth System  
43  
44 13 Research (IISTA), University of Córdoba, Spain

45  
46  
47 14 (5) Department of Geology, Geography and Environment, University of Alcalá, C. Colegios 2,  
48  
49 15 Alcalá de Henares 28801, Spain; Center for the Study of the Biosphere from Space,  
50  
51 16 University of Paul Sabatier, 18 av. Edouard Belin, Toulouse 2801, France

52  
53  
54 17 (6) Universidad Nacional de Educación del Ecuador (UNAE, Chuquipata, Ecuador)

## 18 Corresponding author

19 Francisco Bonet-García (fjbonet@gmail.com). Departamento de Botánica, Ecología y Fisiología  
20 Vegetal, Área de Ecología, Universidad de Córdoba, Edificio Celestino Mutis (C-4), 14014  
21 Córdoba, Spain

## 22 Licence

23 Creative Commons Attribution (CC BY)

## 24 Abstract

25 Pine plantations, very common in the Mediterranean basin, are recurrently affected by forest  
26 pests due to intrinsic characteristics (high density, low spatial heterogeneity) and external  
27 factors (consistent trend towards a warmer and drier climate). *INSTAR* is an Agent-Based  
28 Model designed to generate a deeper understanding of the population dynamics of the  
29 *Thaumetopoea pityocampa* forest pest.

30 The model has been designed using a modular approach: several interconnected modules  
31 (submodels) facilitate the incorporation of new knowledge about the pest biology and can  
32 serve as template for the design of other similar models. The model is spatially and temporally  
33 explicit and allows its implementation under different climate and land use scenarios. *INSTAR*  
34 is described in detail in this manuscript using the standardized ODD (Overview, Design  
35 concepts, and details) protocol.

36 In order to be coherent and structurally realistic, *INSTAR* reproduces two important patterns:  
37 the observed timing of the different life stages, and the spatial distribution of the pupae within

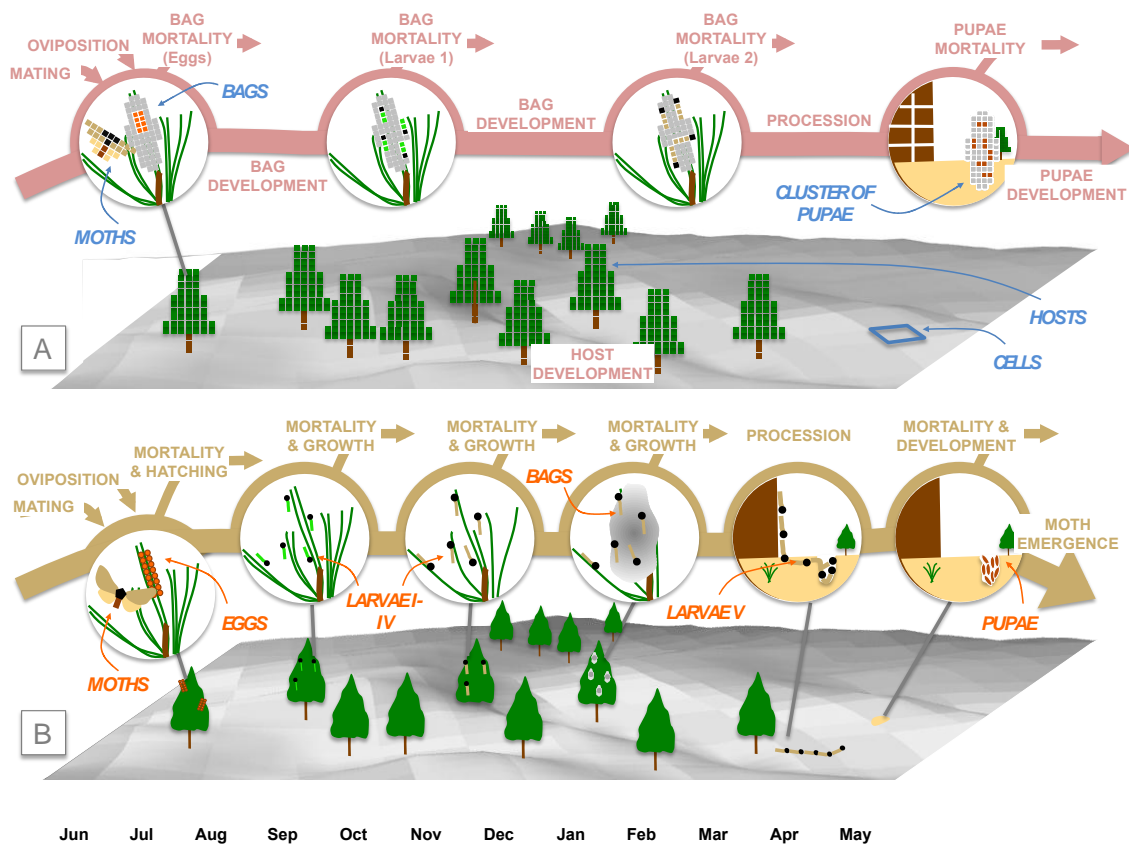
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38 the landscape. These requirements have been assessed here through a consistency test of the  
39 model and its submodules. This assessment is constituted by a calibration analysis of the pest  
40 phenology and a robustness analysis of the processes involved in the spatial distribution of the  
41 pest. Besides this, a stress test was performed by exposing the model to extreme climate  
42 inputs. The model successfully reproduces the phenology of the species in the study area.  
43 Moreover, the spatial distribution of the pest concurs with the literature, showing a higher  
44 emergence of moths in unshaded areas. The stress test confirmed that the model behaves as  
45 expected when exposed to extreme input values. The experiments presented in this  
46 manuscript constitute a first internal validation of the model, which can now be fully calibrated  
47 and parameterised.

## 48 **Key words**

49 *Thaumetopoea pityocampa*; Agent-Based Model (ABM); forest pest; biological cycle; pattern-  
50 oriented; scenario

## 51 **Graphical Abstract**





## 53 Highlights

- 54 • Agent-Based Models are a useful approach to simulate forest pests' dynamics
- 55 • *INSTAR* simulates the biological cycle of *Thaumetopoea pityocampa* in pine plantations
- 56 • Key patterns are reproduced: phenology and spatial distribution
- 57 • *INSTAR* has been designed to easily incorporate new information about the pest biology
- 58 • *INSTAR* has been designed in a modular way to allow an easy upgrade and reusability

# 1. Introduction

Pine plantations are very common in the Mediterranean basin, where they were used in the early decades of the 20th century as management tools to halt soil loss. In Spain alone, 2.5 million hectares were planted between 1940 and 1980 (Ministerio de Medio Ambiente 2002). Currently, their intrinsic characteristics (mainly low diversity and high stand density), together with external factors (i.e. climate change) cause pests to be a recurrent problem within these forests.

Forest pests receive great attention due to their relevance in ecosystems functioning and the economic impacts they cause on managed forests (see Wingfield et al. 2015). Numerous articles are published every year providing insights on species physiology, reporting outbreaks and first records of species and showing up-to-date occurrences (Higham et al. 2017, Valenta et al. 2017, Zanetti et al. 2017, Tanase et al. 2018). Researchers also highlight the potential impacts of forest pests under climate change scenarios, which are expected to be more acute in coniferous forests and the boreal biome (Seidl et al. 2017). In this regard, current knowledge confirms the lower resistance and thus higher vulnerability of single species forests to disturbances such as forest pests, among others (Fares et al. 2015).

The lack of integrative and synthetic knowledge makes it difficult to forecast pest behaviour under specific abiotic scenarios and landscape configurations. Although experiments are possible on a small scale (individuals, small areas), obvious restrictions exist when the goal is to understand the dynamics of large-scale systems, such as extensive plantations and/or long time periods. This situation reveals an important drawback when trying to design more effective and feasible management strategies that use up-to-date knowledge. Therefore, there is a need to generate tools that can help foresters to avoid or minimise pest damage and ensure sustainability among forests masses (Fares et al. 2015).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

83 Modelling has demonstrated its usefulness when carrying out *in silico* experiments  
84 (Chowdhury and Stauffer 2005) and synthesising information from different scales, thus  
85 helping with acquiring more comprehensive and holistic knowledge. Our aim is to improve the  
86 general understanding of forest pest dynamics by studying specific cases, as suggested by  
87 Fares et al. (2015). Thus, our approach is to create a model which explicitly simulates the  
88 concatenated linkages between the abiotic configuration of the landscape (topography,  
89 climate, meteorology, etc.) and the performance of the ecosystem components (pine  
90 plantations and forest pests) in a global change scenario. The pine processionary moth  
91 (*Thaumetopoea pityocampa*) is especially suited for modelling purposes for several reasons:

92 i) The pine processionary moth lifecycle allows for simplification without neglecting crucial  
93 processes: its characteristic gregarious behaviour allows modelisation into groups of  
94 individuals that behave as fully functional units which constitutes an advantage in terms  
95 of computing power needs. Nevertheless, the lifecycle of this species is influenced by  
96 interacting abiotic and biotic factors, and therefore modelling such a complex system  
97 constitutes an interesting challenge. Winter climate and habitat (Démolin 1969a, Masutti  
98 and Battisti 1990, Hódar et al. 2002) are among the most important abiotic factors  
99 influencing these pest dynamics, while biotic factors such as parasitism or amount and  
100 quality of available pine biomass also play a key role in modulating *T. pityocampa*  
101 populations (Battisti 1988, Hódar et al. 2002, 2004, Charbonnier et al. 2014, Torres Muros  
102 2015). The linkage between abiotic and biotic factors also presents an opportunity to  
103 couple climate and hydrologic models with ecological ones.

104 ii) The pine processionary moth lifecycle is representative of many forest pests: long  
105 defoliating larval stage followed by pupation and short non-feeding moth stage (Dajoz  
106 2000, Netherer and Schopf 2010, Barbosa et al. 2012), thus simulating its biological cycle  
107 can serve as inspiration to model the behaviour of other forest pests. Similar to other  
108 species, *T. pityocampa* is experiencing an expansion process fostered by the spread of the

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

109 above-mentioned pine plantations in which it feeds as well as climate change (Hódar and  
110 Zamora 2004, 2004, Battisti et al. 2005, Petrucco Toffolo et al. 2006, 2006). Modelling the  
111 biological cycle of this species can help in understanding the potential impact of global  
112 change on the functioning of pine plantations in the Mediterranean area. This is of high  
113 importance given the extent of such forests and since their potential growth decline due  
114 to the interaction between defoliation and drought stress (Linares et al. 2014, Sangüesa-  
115 Barreda et al. 2014). Besides, modelling the biological cycle of this species will also allow  
116 the study of the interaction between pests and land uses, therefore helping to improve  
117 management strategies in the long term (Régolini et al. 2014, Castagneyrol et al. 2014b).

118 iii) The pine processionary moth constitutes a forestry, ecological and public health problem  
119 nowadays since affected trees reduce their growth and are more sensitive to other pests  
120 (Markalas 1998, Kanat et al. 2005, Arnaldo et al. 2010, Jacquet et al. 2012). It also causes  
121 stinging and allergic reactions to humans and animals due to its urticating hairs (Rebollo  
122 et al. 2002, Vega et al. 2003, Battisti et al. 2017). Thus, simulating the population  
123 dynamics of *T. pityocampa* could help manage its outbreaks to minimise health and  
124 ecological impacts.

125 Some authors have modelled the distribution of *T. pityocampa* by means of equations-based  
126 models (Robinet et al. 2014) or statistical analysis (Seixas Arnaldo and Monteiro Torres 2005,  
127 Régolini et al. 2014). Most authors interested in this species highlight the importance of  
128 climate on the impact of the pest, and some have even attempted to model the effects of  
129 climate change on it (Robinet et al. 2007, 2014, Robinet et al. 2015, Seixas Arnaldo et al. 2011).  
130 However, no Agent-Based Model (ABM) model is available for *T. pityocampa*.

131 In this article, a spatial and temporally explicit ABM called *INSTAR* is presented. *INSTAR* has  
132 been designed to provide a deeper understanding of the population dynamics of *T.*  
133 *pityocampa*. Specifically, it will help us to test whether our current knowledge on the species

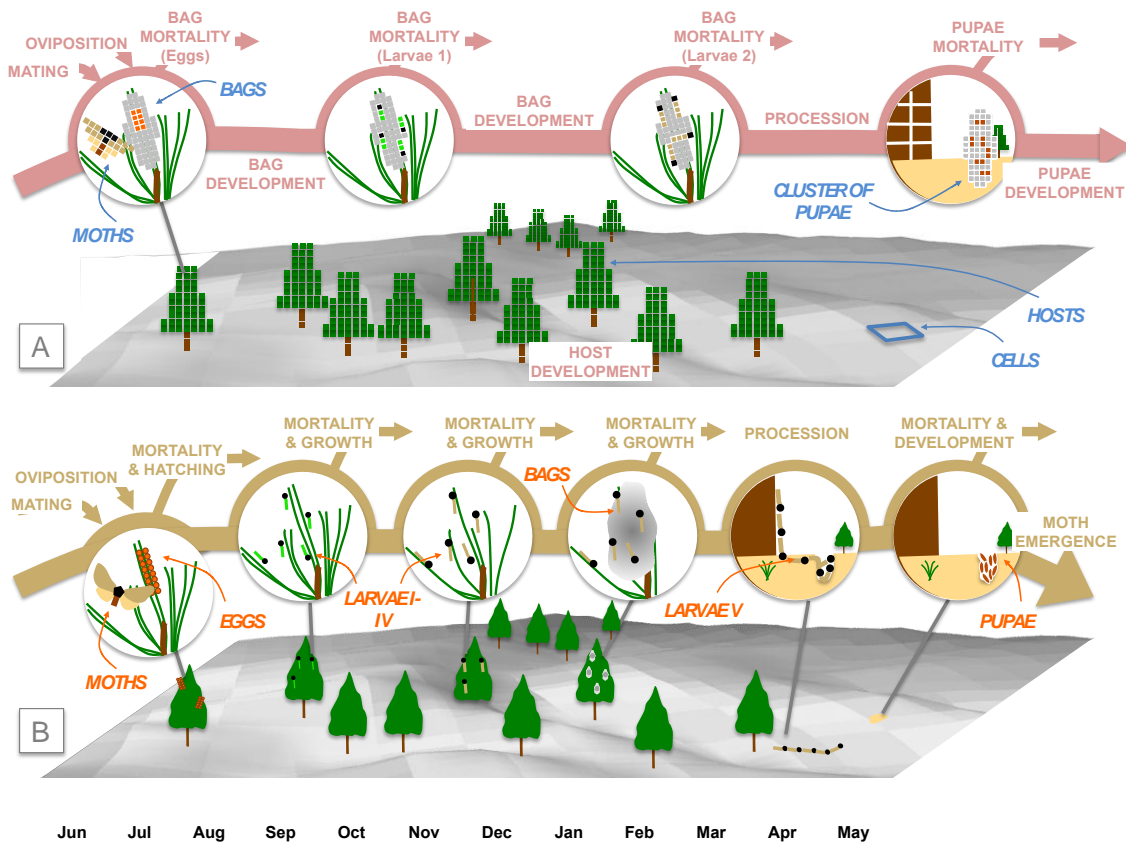
134 biology and individuals' behaviour explain the population dynamics patterns observed in  
135 selected forests. It has been conceived as an open product which allows the incorporation of  
136 new knowledge and further development. Once the model is fully calibrated and validated,  
137 simulations can be considered as *in silico* performed experiments that are not feasible in the  
138 real world. It will also allow us to forecast the probability of occurrence and intensity of the  
139 pest outbreaks on pine plantations under different climate and land use scenarios. Therefore,  
140 it will help improve planning and management activities on *T. pityocampa* outbreaks. This  
141 manuscript firstly provides an overview about the ecological background of the target species.  
142 The structure and functioning of *INSTAR* are then described in detail according to the ODD  
143 protocol ("Overview, Design concepts and Details") proposed by Grimm et al. (2010). Finally,  
144 we present a consistency test of the model and its submodules, constituted by a calibration  
145 analysis of the most important processes simulated in *INSTAR*, a robustness analysis of the  
146 processes involved in the spatial distribution of the pest and a stress test on the influence of  
147 extreme temperature on the model.

## 148 2. Ecological background of the species

149 *T. pityocampa* is present in the entire Mediterranean Basin, from North Africa to the South of  
150 Europe and from the Atlantic Coast to the western part of Turkey, and it is especially present in  
151 afforested areas. It is a univoltine oligophagous species feeding on *Pinus* sp., and more rarely  
152 on *Cedrus* sp. and planted *Pseudotsuga* sp. (Battisti et al. 2015). A complete outline on the  
153 biology of *Thaumetopoea* species in general, and of *T. pityocampa* in particular, can be found  
154 in Roques (2015). However, the detailed features of the *T. pityocampa* cycle widely vary  
155 between sites and between years.

156 The biological cycle of *T. pityocampa* is shown in **Figure 1B**. After mating, female moths lay  
157 between 150-350 eggs around pine needles or, more rarely, twigs. Eggs are exposed to  
158 predation and parasitism mainly caused by *Ooencyrtus pitocampae* and *Baryscapus servadeii*  
159 (**Figure 1B**). Parasites are temperature-limited, and therefore the parasitism rate shows an  
160 altitude gradient: the higher the altitude, the lower the parasitism rate (Torres Muros 2015).  
161 Hatching occurs in early summer in northern (and higher) locations and late summer in  
162 southern (lower) ones, thus avoiding high temperatures which are lethal for the larvae (**Figure**  
163 **1B**) (Démolin 1969b, Seixas Arnaldo et al. 2011).  
164 Larvae are mainly characterised by their gregarious behaviour. Just after hatching, larvae build  
165 silk nests, to protect against low temperatures and predators. Since *T. pityocampa* is a  
166 poikilothermic species, air temperature plays a key role during larval development (Démolin  
167 1969b, Huchon and Demolin 1970, Robinet et al. 2007, Hoch et al. 2009), as well as in later  
168 stages. Moderately high temperatures (below 30 °C according to Démolin (1969b)) accelerate  
169 larval growth, while cold temperatures delay development and can cause mortality due to  
170 freezing or starvation. Thus, *T. pityocampa* has a higher affinity for trees situated at the edge  
171 of the stands or around clear areas, since nests there receive more light and therefore are in a  
172 more advantageous situation in terms of temperature (Démolin 1969b). Moreover, larvae  
173 colonies can move within the tree to find the most exposed areas of the pines at each moment  
174 (Démolin 1969b, Hoch et al. 2009, Pimentel et al. 2011). Normally, larvae feed during the night  
175 and digest during the day inside their nests. In contrast to other defoliator species, the larval  
176 phase of *T. pityocampa* takes place during winter. Therefore, increasing temperatures due to  
177 climate change are expected to favour this species (Netherer and Schopf 2010).  
178 Larval development requires five larval stages or "instars" (**Figure 1B**). Growth and survival of  
179 early development stages depend on food quality (Hódar et al. 2002) and higher survival rates  
180 are observed for pine trees not affected by defoliation during previous years (Hódar et al.  
181 2004). On the other hand, late stages are more limited by the amount of food available

182 (Démolin 1969b, Huchon and Demolin 1970, Hóðar et al. 2002, Robinet et al. 2007, Hoch et al.  
 183 2009). Full larval development cannot last less than 4 months under optimal conditions. For all  
 184 stages, it is reported that optimal daily temperature is between 20-25 °C and mortality quickly  
 185 reaches 100 % at 32 °C (Démolin 1969b). Apart from temperature, larval survival is also  
 186 affected by predators (Barbaro and Battisti 2011) and parasites (Battisti et al. 2015), with the  
 187 incidence of these factors being quite variable among studies.



188  
 189 **Figure 1.** Diagram showing the simplified version of the cycle implemented in *INSTAR* (A), as  
 190 well as the biological cycle of the processionary moth (B). The modelled entities are shown in  
 191 blue, while the real entities are shown orange. The biological cycle of the species is simulated  
 192 in *INSTAR* by three types of submodels: development (bag, pupae and host development),  
 193 mortality (bag and pupae mortality) and movement (procession, mating and oviposition).  
 194 When larvae complete their development, usually at the end of winter or beginning of spring,  
 195 they constitute the characteristic processions and search for a place to bury (Figure 1B). Once

196 the leading larva chooses a suitable spot, the whole group buries and pupal development  
197 starts. Pupation success rate depends on shading and parasitism and therefore optimal burial  
198 spots are characterised by high radiation and reduced vegetation cover (Robredo 1963,  
199 Dulaurent et al. 2011, Dulaurent et al. 2012, Torres-Muros et al. 2017). Moths select trees on  
200 the edge of the stand when laying their eggs. These behaviours generate a characteristic  
201 spatial pattern: the pest concentrates in low density stands or stand edges, and pupae are  
202 especially present in clear areas between trees.

203 Like all other *Thaumetopoea* sp., *T. pityocampa* is a univoltine species, but pupae can engage  
204 in extended diapause. In Mediterranean environments only 2.5 % of pupae show extended  
205 diapause (i.e. more than two years, Torres-Muros et al. (2017)), but in other environments the  
206 proportion of latent pupal population can be much higher (Aimi et al. 2006). The conditions  
207 inducing extended diapause are still uncertain (but see Salman et al. 2019). When pupation  
208 process ends, moths emerge and live between 1 and 2 days, during which time they mate and  
209 lay their eggs on a nearby pine (Zhang and Paiva 1998) ("Mating" and "Oviposition" in **Figure**  
210 **1B**).

### 211 3. Model description

212 According to Gertseva and Gertseva (2004), *INSTAR* can be classified as a homomorphic (all  
213 components of the model have analogous components in reality, but not *vice versa*), time-  
214 dependent, continuous (it represents continuous changes of an object over time) and  
215 deterministic Agent-Based Model (ABM). ABM has been chosen as the modelling technique,  
216 since this approach allows an easy integration of multiscale information, which is specially  
217 abundant in the case of *T. pityocampa*: physiological data based on observations of individuals  
218 under laboratory conditions (Hoch et al. 2009, Berardi et al. 2015), influence of environmental



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

219 factors based on observations of populations and field experiments (Démolin 1969a, Samalens  
220 and Rossi 2011, Castagneyrol et al. 2014a, Cayuela et al. 2014, Linares et al. 2014, Torres-  
221 Muros et al. 2017), and population history based on observations of outbreaks on a regional  
222 scale (Hódar et al. 2012, Li et al. 2015). The following sections describe *INSTAR* according to  
223 the "Overview, Design Concepts, and Details" (ODD) protocol proposed by Grimm et al. (2010),  
224 which is a very useful tool to characterise ABMs using a standard structure.

## 225 3.1 Overview

### 226 3.1.1 Purpose

227 The overarching purpose of *INSTAR* is to generate a deeper understanding of the population  
228 dynamics of *Thaumetopoea pityocampa*. In the short term, it aims to identify caveats and gaps  
229 in our knowledge about the species biology, and promote research aiming to fill these gaps. In  
230 the long term, it aims to forecast the probability of occurrence and intensity of the pest  
231 outbreaks at a landscape scale under different climate and land use scenarios, this way aiding  
232 environmental decision making in pine plantations affected by this pest. The first step in this  
233 process is the construction, documentation and initial testing of the model presented in this  
234 manuscript.

235 More specifically, and following Jakeman et al. (2006), the context of *INSTAR* refers to the  
236 following topics which help clarify our purpose. Firstly, it addresses questions such as: How  
237 well can *INSTAR* simulate the functioning of the processionary moth forest pest in a spatially  
238 and temporally explicit way? Is it possible to obtain landscape level metrics (defoliation rate  
239 per year and watershed) modelling the behaviour of individual agents? Is *INSTAR* able to  
240 reproduce natural patterns of the *T. pityocampa* cycle such as density-dependent restrictions?

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

241 How well does the coupling between a hydrometeorological model and an ecological ABM  
242 (*INSTAR*) work?

243 Moreover, *INSTAR* has been designed to satisfy the needs of two types of interest groups or  
244 users. On the one hand, scientists interested in simulating ecological processes and,  
245 specifically, *T. pityocampa* population dynamics, who would find in *INSTAR* a modular tool to  
246 conduct *in silico* experiments. On the other hand, and in the long term, *INSTAR* could be useful  
247 for the environmental decision makers in charge of managing the large area covered by pine  
248 plantations affected by this pest.

249 The model has two main forcing variables or drivers: climate and land use. Since this species is  
250 poikilotherm, the air temperature is very important to determine its performance. *INSTAR* uses  
251 daily temperature maps as inputs to simulate the biological cycle of *T. pityocampa*. Regarding  
252 land use, the distribution and density of pine plantation stands has been demonstrated to be a  
253 key factor explaining the population dynamics of the processionary moth (Dulaurent et al.  
254 2011, Li et al. 2015, Torres-Muros et al. 2017).

### 255 3.1.2 Entities, state variables and scales

256 An overview of entities and state variables can be found in **Table 1**. *INSTAR* contains a  
257 simplified version of the biological cycle of *T. pityocampa*, represented by three entities:  
258 **colonies**, **cluster of pupae** and **moths**; the pine trees on which larvae feed are represented by  
259 the entity **hosts**; and the environment is constituted by **cells**.

260 Each **colony** represents all surviving individuals from the same egg laying and its *stage* defines  
261 the phase of the lifecycle in which the **colony** is, namely: "egg", larvae 1 ("L1", representing the  
262 first two instars: I and II) or larvae 2 ("L2", representing the last three instars: III, IV and V). This  
263 entity is also defined by the state variables *number of individuals* (i.e. the number of eggs,  
264 larvae 1 or larvae 2 contained in the **colony**), *days as egg*, *days as L1* and *days as L2* (days

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

265 remaining for the *stage* to be completed), and *host* (**colonies** develop on pine crowns, and  
266 therefore their location within the environment is described by the **host** they are linked to).  
267 Except for the host they are linked to, all **colony** variables values change throughout the  
268 simulation, i.e. they are dynamic state variables.

269 The entity **cluster of pupae** refers to all individuals from an egg laying that have reached the  
270 end of the larval stage and have successfully buried on a **cell**. This entity is used to model the  
271 development of pupae and is defined by similar state variables as the **colonies**: *number of*  
272 *individuals*, *days as pupae* and the static variable *location*.

273 Finally, the entity **moth** represents an individual moth in the real world. It has a static assigned  
274 *sex* and can be *mated* or not. Thus, in the transformation from **cluster of pupae** into **moths**  
275 there is an individualisation process, which is explained in detail in section 3.3.3.

276 **Host** entity represents trees (*Pinus* sp.) on which the larvae develop. **Hosts** are characterised  
277 by their static variables *height* (as surrogate of tree size), which defines its carrying capacity,  
278 i.e. how many **colonies** can feed on it before it gets defoliated, and *location*. Moreover, they  
279 also have the dynamic state variables *quantity* of available biomass, and *quality*, which  
280 indicates if the tree was defoliated in the previous season or not. *INSTAR* can be applied to any  
281 *Pinus* spp. affected by *T. pityocampa* by parameterising the state variables and submodels  
282 according to each species (see section 3.3).

283 *INSTAR* works with a 30 m spatial resolution at a temporal resolution of 1 day. The world in the  
284 model is constituted by **cells**, which represent 30 x 30 m, and have *elevation* associated as a  
285 static variable. **Cells** also have dynamic state variables referring to weather conditions, which  
286 change in every time step (1 day). These are the *minimum* and the *maximum air temperature*.  
287 Moreover, **cells** also have an assigned *integrated temperature*, an internal variable referring to  
288 the temperature of a **colony** which is exposed to the **cell** air temperature and the number of  
289 insolation hours per day, according to Démolin (1969b). This variable is calculated as:

290

$$\text{integrated temperature} = \text{maximum temperature} + (1.5 \times \text{insolation hours})$$
291 **Table 1.** Summary of entities and state variables within *INSTAR*.

ENTITY	STATE VARIABLE	DESCRIPTION	VALUES (units)
<b>Colony</b>	<i>stage</i>	Indicates in which phase the colony is.	"Egg", "L1", "L2"
	<i>number of individuals</i>	Number of individuals within the colony	0 - ~ 200 (ind.)
	<i>days as egg</i>	Number of days remaining before the colony in instar "Egg" will change to instar "L1".	0 - 30 (days)
	<i>days as L1</i>	Number of days remaining before the colony in instar "L1" will change to instar "L2".	0 - 30 (days)
	<i>days as L2</i>	Number of days remaining before the colony in instar "L2" will transform into a cluster of pupae.	0 - 90 (days)
	<i>host</i>	Pine on which the colony is located, to which it is linked.	Host ID
<b>Cluster of pupae</b>	<i>number of individuals</i>	Number of individuals within each cluster.	0 - ~ 200 (ind.)
	<i>days as pupae</i>	Number of days remaining before the cluster of pupae will become moths.	0 - 215 (days)
	<i>location</i>	Cell where it is located.	Cell coordinates
<b>Moth</b>	<i>sex</i>	Sex of the moth.	"Female", "Male"
	<i>mated</i>	Describes whether the moth is mated or not.	Boolean (true/false)
	<i>location</i>	Cell where it is located.	Cell coordinates
<b>Host</b>	<i>height</i>	Pine height.	1 - 20 (m)
	<i>quantity</i>	Relative amount of available biomass for larvae: 0% corresponds to a completely defoliated tree while 100 % corresponds to a tree in the best possible conditions.	0 - 100 (%)
	<i>quality</i>	Indicates whether a tree was defoliated the previous year or not.	"defoliated", "not defoliated"
	<i>location</i>	Position of the host in the world.	x and y coordinates
<b>Cells</b>	<i>elevation</i>	Elevation above sea level.	0 - 8000 (m.a.s.l.)
	<i>probability of emergence</i>	Probability of emergence which defines the number of individuals from a cluster of pupae those survive and become moths (see "Initialisation").	0 - 12 (%)
	<i>maximum air</i>	Maximum air temperature each day.	-50 - +50 (°C)

---

*temperature*

*minimum air*

Minimum air temperature each day.

-50 - +50 (°C)

---

*temperature*

---

### 292 3.1.3 Process overview and scheduling

293 **Figure 1A** depicts a schematic overview of the submodels that take place in each time step as  
294 well as the entities which perform them. After initialisation (see section 3.3.1.), every time step  
295 consists of:

- 296 • Environment reading: **cells** *minimum* and *maximum temperature* are updated to the  
297 current date.
- 298 • **Colonies** update: *number of individuals* is updated following the "colony mortality"  
299 submodel and *stage* is updated following the "colony development". Then, "procession"  
300 submodel is applied to **colonies** which have completed their development and will  
301 become a **cluster of pupae**.
- 302 • **Clusters of pupae** update: *number of individuals* is updated following the "pupae  
303 mortality" and the length of this phase is adjusted through the "pupae development"  
304 submodel. Once a **cluster** completes its development it results in a number of **moths**.
- 305 • **Moths** update: **moths** become *mated* following the "mating" submodel and choose a  
306 **host** on which to lay their eggs through the "oviposition" submodel.
- 307 • **Hosts** update: the available biomass, *quantity*, is updated every day based on growth and  
308 larvae consumption through the "host development" submodel.

309 These submodels are explained in detail in section 3.3.3. Besides that, from the observer  
310 perspective, in each time step several state variables values are stored. For a detailed  
311 explanation of how model outputs are structured see details regarding observation in the  
312 following section.

## 313 3.2 Design concepts

1  
2  
3  
4 314 This section explains how some important concepts are implemented in *INSTAR*. These basic  
5  
6 315 concepts describe characteristics of ABMs that are not common in other modelling  
7  
8 316 approaches. Describing these concepts in a standard structured format can help transmit the  
9  
10  
11 317 key aspects of the model to others, either ABM experts or those not familiar with this  
12  
13 318 approach (Railsback and Grimm 2011).

14  
15  
16 319 The basic principle of *INSTAR* model is that *T. pityocampa* population dynamics are strongly  
17  
18 320 influenced by the environmental (mainly climate) conditions and therefore these define the  
19  
20  
21 321 impact of the forest pest, the timing of its outbreaks and the areas where it has the highest  
22  
23 322 incidence.

24  
25  
26 323 The spatial distribution of the pest as well as the forest state (amount and quality of available  
27  
28 324 biomass) emerges from the simulation as a result of an individual's behaviour. On the one  
29  
30 325 hand, the selection of burial spots and the pines on which moths lay their eggs define the  
31  
32 326 spatial location of the newly created agents, and thus their impact on the pines. On the other  
33  
34  
35 327 hand, timing and scale of the pest outbreaks is regulated through mortality and development.  
36  
37  
38 328 Literature shows a higher incidence of the pest on areas with low pine density (Régolini et al.  
39  
40 329 2014) and shifts in phenology due to climate changes (Hódar and Zamora 2004).

41  
42  
43 330 **Colonies** in *stage* "L2" adapt to increase their success probability by selecting the burial site  
44  
45 331 considering the number of **hosts** on the **cells** around them and choosing the one which  
46  
47 332 optimises the probability of survival.

48  
49  
50 333 *T. pityocampa* development is dependent on environmental conditions and adjusts its  
51  
52 334 biological cycle to keep it univoltine (Démolin 1969b, Berardi et al. 2015). *INSTAR* implements  
53  
54 335 this by indirect objective-seeking, constraining the total duration of the biological cycle to one  
55  
56 336 year by shortening and lengthening the **colony** and **cluster of pupae** phases depending on the  
57  
58 337 environmental conditions.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

338 **Colonies** can sense the state variables of their **host** and the **cell** where they are located.

339 Moreover, **colonies** are also able to sense the number of **hosts** within a certain radius to select

340 a spot where to bury (see "procession" submodel in section 3.3.3.3.). Similarly to **colonies**, the

341 **clusters of pupae** can also sense the state variables of the **cells** where they are, since it affects

342 their mortality. Finally, at the end of the cycle, female **moths** can sense if there is a male **moth**

343 within a certain distance and where the surrounding pines are (see "mating" and "oviposition"

344 submodels in section 3.3.3.3). No prediction or learning mechanisms are implemented in

345 *INSTAR*.

346 Agents interact with each other in an indirect way. **Colonies** within the same **host** compete

347 with each other for their food. When the number of **colonies** on a **host** equals its carrying

348 capacity, the biomass consumption will slowly lead to complete defoliation of the tree, and

349 therefore to the death of larvae present on it (see "colony mortality" submodel in section

350 3.3.3.1.). Moreover, there are intra- and intergenerational interactions. The intragenerational

351 interaction appears when **colonies** turn into **clusters of pupae** and these into **moths** as

352 development takes place, inheriting state variables values. The *number of individuals* of the

353 **cluster of pupae** is inherited from the *number of individuals* of the **colony** it comes from, and

354 the number of **moths** that emerge from the burial site is defined by the *number of individuals*

355 within the parent **cluster of pupae**. Moreover, the location of each agent is defined based on

356 the location of its predecessor. On the other hand, the intergenerational interaction comes

357 from the defoliation of one season defining the *quality* state of the **host** in the coming season

358 and thus affecting the probability of survival of the coming generation of larvae.

359 To simulate the intrinsic variability of natural processes, several stochastic processes are

360 included during *INSTAR* initialisation: assigning *quantity* to **hosts**; assigning *number of*

361 *individuals* to **colonies**; assigning *days as egg* to **colonies**; and selecting which are the initial

362 infected **hosts**. Moreover, pseudorandom numbers are also used during the model running to:

363 assign **moth sex** (1:1 ratio); decide whether a female **moth** becomes *mated* or not once it finds

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

364 a male **moth** and choose the **hosts** where **moths** lay their eggs, following a weighted random  
365 selection based on distance to the **moth**. These functions are explained in details in the  
366 corresponding submodels descriptions (section 3.3.3).

367 The representation of *T. pityocampa* biological cycle in three agent types, responds to a  
368 necessity to simulate the behaviour and characteristics of each development phase. **Colonies**  
369 and **clusters of pupae** are, in a way, collectives since they represent a group of individuals. This  
370 assumption does not constitute an oversimplification of the real system considering the  
371 characteristic gregarious behaviour of this species, where **colonies** behave as fully functional  
372 units (Démolin 1969b). Besides, defining them as collectives is an advantage since it reduces  
373 computing power needs, thus allowing the simulation of larger areas.

374 To observe the pest evolution and **hosts** state, *INSTAR* provides several outputs. At system-  
375 level, average **hosts quantity**, percentage of infected pines and number of individuals of *T.*  
376 *pityocampa* at each life stage are stored for each day. Moreover, monthly averages of all these  
377 state variables are also stored at **cell** level. This allows the evaluation of the forest pest  
378 development over time for a whole landscape unit (i.e. a specific forest, a basin, etc.) as well as  
379 within the simulated area (stand edges, distribution and movement of each phase, most  
380 affected **hosts**, etc.).

### 381 **3.3 Details**

382 ABMs often require a high number of parameters and detailed information is needed to fully  
383 parameterise the model (Kelly et al. 2013). Thus, as any other ABM, *INSTAR* parameterisation  
384 requires an exhaustive review of literature as well as testing. Our model is based on local data  
385 from Sierra Nevada (SE Spain) whenever possible. When unavailable, data were taken from  
386 scientific literature elsewhere. The parameters used by *INSTAR* are listed in **Table 2** and  
387 decisions for parameter estimation are fully explained in **Appendix A**. It should be noted that



388 *INSTAR* was built as a customisable model in which parameter values can be changed to fit  
389 biological information to local conditions as well as to incorporate new information when it  
390 becomes available, constituting one of its potential.

391 Some of *INSTAR* parameter values are well referenced pest traits ( $CS_{mean}$ ,  $CS_{sd}$ ,  
392  $Procession_{max\_distance}$ ,  $Mating_{prob}$ ) or have been based on best available knowledge considering  
393 certain assumptions ( $Egg_{min}$ ,  $L1_{min}$ ,  $L2_{min}$ ,  $Pupae_{max}$ ,  $NPP$ ,  $QI_{threshold}$ , see section 3.3.3.2 and  
394 **Appendix A**). Other parameters require calibration for each study area ( $T_{egg\_devel}$ ,  $T_{larvae\_devel\_high}$ ,  
395  $T_{larvae\_devel\_internal}$ ,  $T_{larvae\_devel\_low}$ , as presented in this manuscript). Finally, another set of  
396 parameters are required for model functioning and can serve in the future for purposes such  
397 as scenario setting ( $Qt_{mean}$ ,  $Qt_{sd}$ ,  $L1_{mort}$ ,  $L2_{mort}$ ,  $L2_{mort\_threshold}$ ) or pest expansion experiments  
398 ( $Mating_{max\_distance}$ ,  $Flight_{max\_distance}$ ,  $T_{lethal\_max}$ ,  $T_{lethal\_min}$ ).

399 **Table 2.** *INSTAR* parameters

PARAMETER	VALUE (units)	DESCRIPTION AND REFERENCES	More details in
$Qt_{mean}$ , $Qt_{sd}$	50, 20 (%)	Mean and standard deviation used to assign initial values of <i>quantity</i> to the <b>hosts</b> , following a normal distribution with such parameters. Reference: This study, assuming an intermediate health state of the simulated pine stand.	3.3.1
$CS_{mean}$ , $CS_{sd}$	193.2, 4.5 (ind.)	Mean and standard deviation used to assign initial values of <i>number of individuals</i> to the <b>colonies</b> , following a normal distribution with such parameters. Reference: Torres Muros (2015).	3.3.1
$QI_{threshold}$	57.16 (%)	Threshold used to define the <i>quality</i> of the <b>host</b> based on their <i>quantity</i> at the end of the defoliating season. Reference: Muukkonen (2005).	3.3.1
$Egg_{min}$	30 (days)	Minimum time needed for <b>colonies</b> at <i>stage</i> "egg" to become "L1". Reference: Démolin (1969b).	3.3.1
$L1_{min}$	30 (days)	Minimum time needed for <b>colonies</b> at <i>stage</i> "L1" to become "L2". Reference: Démolin (1969b).	3.3.1
$L2_{min}$	90	Minimum time needed for <b>colonies</b> at <i>stage</i> "L2"	3.3.1

	(days)	to become <b>cluster of pupae</b> . Reference: Démolin (1969b).	
<b>L1<sub>mort</sub></b>	10 (%)	Percentage of individuals which die in a <b>colony</b> at <i>stage</i> "L1" if its <i>host</i> was defoliated the previous season. Reference: This study, to be parameterised with real data when available.	3.3.3.1
<b>L2<sub>mort</sub></b>	2 (%)	Percentage of individuals which die daily in a <b>colony</b> at <i>stage</i> "L2" if its <b>host</b> has a <i>quantity</i> below <b>L2<sub>mort_threshold</sub></b> . Reference: This study, to be parameterised with real data when available.	3.3.3.1
<b>L2<sub>mort_threshold</sub></b>	20 (%)	Threshold defining the <i>quantity</i> of <b>hosts</b> below which individuals of <b>colonies</b> at <i>stage</i> "L2" die due to food scarcity. Reference: This study, to be parameterised with real data when available.	3.3.3.1
<b>T<sub>lethal_max</sub></b>	32 °C	Maximum temperature above which <b>colonies</b> die. Reference: Démolin (1969b).	3.3.3.1
<b>T<sub>lethal_min</sub></b>	-12 °C	Minimum temperature below which <b>colony</b> die. Reference: Démolin (1969b).	3.3.3.1
<b>NPP</b>	28 (%)	Net Primary Production rate defining the daily increase of <i>quantity</i> in the <b>hosts</b> during the growing season. Reference: self-defined based on <b>QI<sub>threshold</sub></b> , assuming constant growth rate and according to the growing season described by Alcaraz-Segura et al. (2016).	3.3.3.2
<b>T<sub>egg_devel</sub></b>	20 °C *	Maximum temperature above which <b>colonies</b> at <i>stage</i> "egg" develop. Reference: based on Démolin (1969b).	3.3.3.2
<b>T<sub>larvae_devel_high</sub></b>	25 °C *	Maximum temperature below which <b>colonies</b> at <i>stage</i> "L1" or "L2" develop. This parameter is used in combination with <b>T<sub>larvae_devel_internal</sub></b> and <b>T<sub>larvae_devel_low</sub></b> . Reference: based on Démolin (1969b).	3.3.3.2
<b>T<sub>larvae_devel_internal</sub></b>	20 °C *	Minimum integrated temperature above which <b>colonies</b> at <i>stage</i> "L1" or "L2" develop. This parameter is used in combination with <b>T<sub>larvae_devel_high</sub></b> and <b>T<sub>larvae_devel_low</sub></b> . Reference: based on Démolin (1969b).	3.3.3.2
<b>T<sub>larvae_devel_low</sub></b>	-10 °C *	Minimum temperature below which <b>colonies</b> at <i>stage</i> "L1" or "L2" stop their activity. This parameter is used in combination with <b>T<sub>larvae_devel_high</sub></b> and <b>T<sub>larvae_devel_internal</sub></b> . Reference: based on Démolin (1969b).	3.3.3.2
<b>Pupae<sub>max_</sub></b>	215 (days)	Maximum time needed for <b>cluster of pupae</b> to become <b>moths</b> . Reference: Démolin (1969b).	3.3.3.3

<b>Procession</b> <sub>max_distance</sub>	1.23 (cells)	Maximum distance that <b>colonies</b> can processionate to find a suitable place to bury. Reference: Robredo (1963).	3.3.3.3
<b>Mating</b> <sub>max_distance</sub>	3 (cells)	Maximum distance that <b>moths</b> can fly to find a suitable partner to mate. Reference: This study, see <b>Appendix A</b> for a detailed explanation.	3.3.3.3
<b>Mating</b> <sub>prob</sub>	95 (%)	Probability for a "female" <b>moth</b> to become <i>mated</i> when a "male" <b>moth</b> is found within <b>Mating</b> <sub>max_distance</sub> . Reference: Rhainds (2010).	3.3.3.3
<b>Flight</b> <sub>max_distance</sub>	3 (cells)	Maximum distance that <i>mated</i> <b>moths</b> can fly to find a suitable <b>host</b> to lay their eggs. Reference: This study, see <b>Appendix A</b> for a detailed explanation.	3.3.3.3

400 \*Parameter values before calibration analysis presented in section 4.

### 401 3.3.1 Initialisation

402 *INSTAR* is initialised by creating **cells** and **hosts** and assigning their state variables values for  
403 the simulated area. *Elevation*, **host location** and *height* are loaded as input data (see section  
404 3.3.2.). **Moth probability of emergence** is then calculated for each **cell** based on number of  
405 **hosts** on it. Below 3 **hosts** per **cell probability of emergence** remains constant at 0.12 and  
406 above 10 it equals 0. Between 3 and 10 **hosts** per **cell probability of emergence** is calculated  
407 based on the following linear regression:

$$408 \text{ probability of emergence} = 0.1636 - 0.0169 \times (\text{number of } \mathbf{hosts} \text{ in } \mathbf{cell})$$

409 This function assumes that the number of **hosts** per **cell** is a good surrogate of shaded surface,  
410 which affects soil moisture and temperature. These seem to be the main factors explaining the  
411 **cluster of pupae** emergence rate according to Torres-Muros et al. (2017), who reported  
412 mortality and emergence rates of **cluster of pupae** based on vegetation cover and soil  
413 moisture.

414 After this, *quantity* and *quality* values of **hosts** are assigned. The *quantity* assigned to each **host**  
415 **is** based on a normal distribution with mean  $Qt_{mean}$  and standard deviation  $Qt_{sd}$ , therefore

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

416 allowing the user to setup the initial state of the pine stand. *Quality* is defined as "defoliated"  
417 or "not defoliated" according to a threshold value  $QI_{threshold}$ . Three extra internal variables are  
418 calculated for **hosts**:

419 • The carrying capacity (*cc*) of each **host**: the number of **colonies** that a pine can host. It is  
420 calculated based on observations relating to number of **colonies** on highly defoliated  
421 trees (below 10% of biomass remaining) and tree *height* (unpublished data from J.A.  
422 Hódar). Briefly, within the framework of a different study on *T. pityocampa*, 20 trees of *P.*  
423 *halepensis* were checked during two consecutive winters in Sierra Nevada Protected  
424 Area. The number of winter colonies in the trees suffering a severe (~90% defoliation)  
425 were related to the height of the tree. In this way, we obtained an estimate of how many  
426 colonies are needed to completely defoliate a tree depending on its size (*height*):

$$cc = 11.63 \times \ln(\text{height}) - 4.60 \quad (R^2 = 0.65)$$

428 This is the most important model assumption regarding *Pinus* species, and therefore in  
429 this study *INSTAR* is considered to apply on *P. halepensis* stands (see section 4.1).  
430 Nevertheless, the model could be applied to other pine stands, provided a carrying  
431 capacity equation for the corresponding species.

432 • The daily consumption per colony (*consumption*): the relative amount of biomass  
433 consumed per day by one **colony** in "L2" *stage*, given the carrying capacity (*cc*) of the  
434 **host** on which it is feeding. It is calculated by assuming that all **colonies** consume the  
435 same amount of biomass per day, independently of the number of individuals in it; it also  
436 assumes that **colonies** feed during the minimum number of days needed for a **colony** in  
437 "L2" *stage* to develop into a **cluster of pupae** (i.e. 90 days) and considering full  
438 consumption a defoliation of 90%, in agreement with observations used for the carrying  
439 capacity calculations which are as follows:

$$consumption = 90 / ((cc \times 90)) = 1 / cc$$

441 • and moth suitability ( $\mu$ ) of each **host**, which is calculated as:

$$\mu = \sum(\text{count} \times \text{coefficient})_i$$

443 Where  $i$  corresponds to the cardinal and intercardinal directions (i.e. north, northeast, east,  
444 etc.), the count is the number of trees in each direction  $i$  at a radius of 1 **cell** (30 m) and each  
445 coefficient provides a weighting (see **Appendix B**) for the influence on each direction of the  
446 suitability of a tree for the **moth**. This weighting creates an initial distribution of **colonies** in  
447 which there are more colonies on more exposed trees, which is what happens in the real world  
448 since burials occur in clear spots and moths choose a host based on the distance to their  
449 emergence point.

450 Afterwards, **colonies** are initialised. The number of **colonies** to be created is defined by the  
451 percentage of infected pines, which can be set by the user as a model input. In this study, real  
452 data from a defoliation monitoring programme within the Andalusia region (southern Spain)  
453 are used. This programme assigns a defoliation value from 0 (no defoliation) to 5 (complete  
454 defoliation) to each forest plot at the end of every defoliating season. These data are used to  
455 define the percentage of infected pines and thus the number of **colonies** present in the  
456 landscape. **Colonies** are then distributed among the **hosts** by weighted random selection  
457 based on moth suitability: the higher the moth suitability, the less probability for the **host**  
458 to be infected (i.e. to have colonies on it). **Colonies** are initialised as "egg" stage with *days as*  
459 *egg* corresponding to a random number between 0 and  $Egg_{min}$ .

### 460 3.3.2 Input data

461 *INSTAR* uses a set of input data as follows:

462 • Map of elevation for the simulated area, obtained from the Andalusian Network of  
463 Environmental Information (REDIAM) in a raster format and 30 x 30 m resolution.

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32
- 464 • Daily maps of meteorological variables (maximum and minimum temperature) for the  
465 simulated area. These maps are raster files generated by a hydrometeorological model  
466 (WiMMed, (Herrero et al. 2009) at 30 x 30 m resolution (see **Appendix C**).
  - 467 • Map of the pines present in the simulated forest. This map is a vector file derived from  
468 Airborne Laser Scanning data, which contains the location and height of trees (more  
469 details about this processing in **Appendix D**). All identified trees are considered to be  
470 pines (i.e. **hosts**).
  - 471 • Initial percentage of infested pines. In this study, this percentage was based on  
472 defoliation data from Andalusian defoliation monitoring programme. This dataset  
473 provides an infestation rate based on observed defoliation for each plot within the  
474 region, on a 0-5 scale. The initial percentage of infested pines is calculated based on the  
475 defoliation assigned by the monitoring programme to the plot where the simulated area  
476 is located. This dataset is fully documented by Ros Candeira et al. (forthcoming).

### 477 3.3.3 Submodels

33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

478 *INSTAR* uses a series of submodels to simulate the development of the pest, as well as its  
479 interactions with the **hosts** and the **cells** (**Figure 1A**). These submodels make use of best  
480 available current knowledge about *T. pityocampa*, but have been built in such a way that will  
481 allow an easy integration of new information as knowledge on this pest increases. As an  
482 example, if new information on pupae distribution is found, the "procession" submodel could  
483 be adjusted to follow such distribution by including a kernel function. These submodels can be  
484 classified into three types: mortality, development and movement. The classification of  
485 submodels into a reduced number of classes has some important advantages. ABMs have been  
486 criticised due to the difficulties of reproducing them, standardising structures and defining  
487 guidelines to design, calibrate and evaluate them (DeAngelis and Grimm 2014). The modular

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

488 structure of *INSTAR* eases the exchange of information among modellers as well as the  
489 comparison between models. Thus, such structure improves the understanding of the model  
490 and facilitates communication with non-modellers. Moreover, it could constitute a framework  
491 to build models simulating other forest pests in a wide variety of ecological systems, since it  
492 summarises the key biological processes that take place in them. As an example, movement  
493 submodels could be adjusted to simulate the dispersal and establishment of mistletoe,  
494 another common pest in pine plantations, by following Mellado and Zamora (2016).

### 495 3.3.3.1 Mortality submodels

496 Mortality submodels simulate the death of organisms due to abiotic or biotic conditions. Since  
497 some of *INSTAR* agents represent groups of organisms, mortality submodels may apply to the  
498 agent as a whole (i.e. death of the **colony**), or just to a portion of the individuals that the agent  
499 represents (i.e. reduction in the *number of individuals*). *INSTAR* applies mortality submodels to  
500 **colonies** and **clusters of pupae**, while **moth** mortality is already included in pupae mortality  
501 and **hosts** do not have an associated mortality submodel since *T. pityocampa* rarely causes  
502 pine death.

503 The submodel "colony mortality" affects the variable *number of individuals* of the **colonies** or  
504 kills whole **colonies** to simulate mortality rates due to different abiotic and biotic factors. Such  
505 mortality rates differ between each *stage*. At the beginning of *stage* "egg", when the **colony** is  
506 created, a mortality factor due to parasitism is applied. This mortality is calculated based on  
507 Torres Muros (2015) and causes the original *number of individuals* to be reduced as follows:

$$508 \text{ number of individuals}_{\text{after mortality}} = \text{number of individuals}_{\text{original}} * p_i$$

509 where

$$510 p_i = e^{n_i} / (1 + e^{n_i})$$

511 and

$$n_i = 0.2082 + (0.0006 * \textit{elevation})$$

At stage "L1", two mortality factors are applied. The first one is related to *quality* of the **host** and is applied at the beginning of this *stage*, when "eggs" become "L1". If *quality* is "defoliated", *number of individuals* is reduced by a certain percentage  $L1_{mort}$ . The second factor is due to lethal temperatures and is applied every day: a **colony** dies if *maximum temperature* is higher than  $T_{lethal\_max}$  or *minimum temperature* is below  $T_{lethal\_min}$ .

At stage "L2", two mortality factors are applied every day. The first one is related to scarcity of food: if *quantity* of the **host** is below a certain threshold  $L2_{mort\_threshold}$ , *number of individuals* is reduced by a certain percentage  $L2_{mort}$ . The second factor is the same as for stage "L1": a **colony** dies if *maximum temperature* is higher than  $T_{lethal\_max}$  or *minimum temperature* is below  $T_{lethal\_min}$ .

The submodel "pupae mortality" affects the variable *number of individuals* of the **cluster of pupae** to simulate mortality rates due to shading (Torres-Muros et al. 2017), which is assumed to be related to the number of **hosts** on a **cell**. A mortality factor is applied at the beginning of this phase, when **colonies** in stage "L2" become a **cluster of pupae**. The state variable *probability of emergence* defines the percentage of *number of individuals* which remain alive and results in the number of **moths** which are created (see "pupae development" submodel, section 3.3.3.2.).

No mortality submodel is applied to **moths'** agents, since it has been included in the *probability of emergence* of "pupae mortality" submodel.

### 3.3.3.2 Development submodels

Development submodels are used to simulate how an organism increases its biomass over time or how it changes to the next stage of its life-cycle. These submodels are depicted in **Figure 1A**. In *INSTAR*, **hosts**, **colonies** and **clusters of pupae** have an associated development



1  
2  
3  
4  
536 submodel, while a **moths'** development submodel is not required based on the assumption  
537 that moths live one day.

538 The submodel "hosts development" defines how *quantity* and *quality* state variables change  
539 over time. The *quantity* of a **host** on a given day is defined as:

540 
$$quantity_{(t)} = quantity_{(t-1)} - \text{consumption} + \text{growth}, \text{ where}$$

541 
$$\text{consumption} = \text{daily consumption per colony} * \text{number of "L2" colonies on the host}$$

542 
$$\text{growth} = \text{NPP between April and August and 0 \% the rest of the year}$$

543 Regarding *quality*, this variable is assessed before the growing season starts (i.e. on the 31st  
544 March). *Quality* will be defined as "defoliated" if *quantity* at that moment is below  $QI_{\text{threshold}}$ ,  
545 and "not defoliated" otherwise.

546 The submodel "colonies development" sets the time that is needed for a **colony** to change its  
547 *stage*. **Colonies** in *stage* "egg" will become "L1" and *days as L1* will be set to  $L1_{\text{min}}$  when *days as*  
548 *egg* become 0. Similarly, **colonies** in *stage* "L1" will become "L2" and *days as L2* will be set to  
549  $L2_{\text{min}}$  when *days as L1* become 0. Finally, **colonies** in *stage* "L2" will perform the "procession"  
550 submodel when *days as L2* become 0. As the time that a **colony** needs to change between  
551 stages increases, the length of the following pupae phase is decreased (see "pupae  
552 development" submodel, section 3.3.3.2) in order to maintain a univoltine cycle as literature  
553 refers (Berardi et al. 2015). *Days as egg*, *days as L1* and *days as L2* are adjusted by applying the  
554 following rules:

- 555 • *days as egg* is decreased by 1 each day if *maximum temperature* is above  $T_{\text{egg\_devel}}$ .  
556 Otherwise *days as egg* remains the same and the following pupae phase will be increased  
557 by 1 day. Thus, *INSTAR* halts the development of eggs on those days where temperature  
558 is lower than the threshold ( $T_{\text{egg\_devel}}$ ).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

- 559 • *days as L1* is decreased by 1 each day if *maximum temperature* is below  $T_{larvae\_devel\_high}$   
560 and the integrated temperature is above  $T_{larvae\_devel\_internal}$ . When these conditions are not  
561 met, if *minimum temperature* is above  $T_{larvae\_devel\_low}$ , *days as L1* remains the same and the  
562 following pupae phase increases by 1 day. Otherwise, *days as L1* increases by 1 and the  
563 following pupae phase increases by 2 days.
- 564 • *days as L2* follow the same rules as *days as L1*.

565 The submodel "pupae development" sets the length of the phase **cluster of pupae** within the  
566 lifecycle of the processionary by adjusting the *days as pupae* state variable. *Days as pupae* are  
567 decreased by 1 in every time step. When it becomes 0, the **cluster of pupae** disappears and  
568 becomes a certain number of **moths**. The number of **moths** which are then created is defined  
569 by *number of individuals*. For the sake of simplicity, *INSTAR* does not consider extended  
570 diapause, as our study area is located in the Mediterranean area, where *T. pityocampa* rarely  
571 engage in extended diapause (see Section 2). Nevertheless, a submodel implementing this  
572 phenomenon could be easily integrated.

### 573 3.3.3.3 Movement submodels

574 Movement submodels are used to describe processes where an agent selects a place to move  
575 to based on information from its surroundings. Three movement submodels are defined in  
576 *INSTAR*: "procession", which simulates the search for a burial spot, performed by **colonies** in  
577 "L2" stage; "mating", which simulates the search of a male **moth**, performed by female **moths**;  
578 and "oviposition", which simulates the selection of a **host** to lay the eggs on (i.e. creation of a  
579 new **colony**), and is performed by mated "female" **moths**.

580 The "procession" submodel defines the place where **colonies** at stage "L2" and *days as L2*  
581 equal to 0 become a **cluster of pupae**. Agents performing this submodel search among the  
582 surrounding **cells** within  $Procession_{max\_distance}$  and select the ones with the smallest pine density  
583 (i.e. number of pines in **cell**). If several **cells** have the smallest pine density, they choose the

584 closest one among them. A **cluster of pupae** is created in that **cell** and the **colony** disappears.

585 The state variables of the newly created **cluster of pupae** are inherited from the parent

586 **colony**: *number of individuals* corresponds to the *number of individuals* of the parent **colony**

587 and *days as pupae* equals  $Pupae_{max}$  minus the number of days that the parent **colony** was

588 extended (see "colony development" submodel description).

589 The "mating" submodel defines which **moths** will become *mated* and where they will be

590 located. If there is a "male" **moth** within  $Mating_{max\_distance}$ , "female" **moths** become mated with

591 a probability  $Mating_{prob}$ . When this happens, *mated moths* move halfway towards their

592 partner.

593 The submodel "oviposition" selects the **host** on which a new **colony** will be created (i.e. the

594 **host** to which the new **colony** will be linked to). After "mating", *mated moths* choose a **host** to

595 lay their eggs within  $Flight_{max\_distance}$ . Weighted random selection is used to define which will be

596 the selected **host**. This selection gives a higher probability of being chosen the closer the **host**

597 is to the **moth**. If no **host** is found, no **colony** is created. The state variables of the newly

598 created **colony** are defined as: *number of individuals* corresponds to a random number within

599 a normal distribution with mean  $CS_{mean}$  and standard deviation  $CS_{sd}$ , *stage* is "egg" and *days as*

600 *egg* equals  $Egg_{min}$ . The other state variables are set to 0.

## 4. Model consistency test

602 *INSTAR* has been written in NetLogo language (Wilensky 1999) and reviewed in an iterative

603 process by a multidisciplinary team consisting of a computer scientist, and various modellers

604 and ecologists. An independent implementation of each individual submodel has been used to

605 check that the model agrees with the conceptual model initially proposed by Torres Muros

606 (2015) or, in other words, for model verification in the sense of Findeisen et al. (1978). The

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

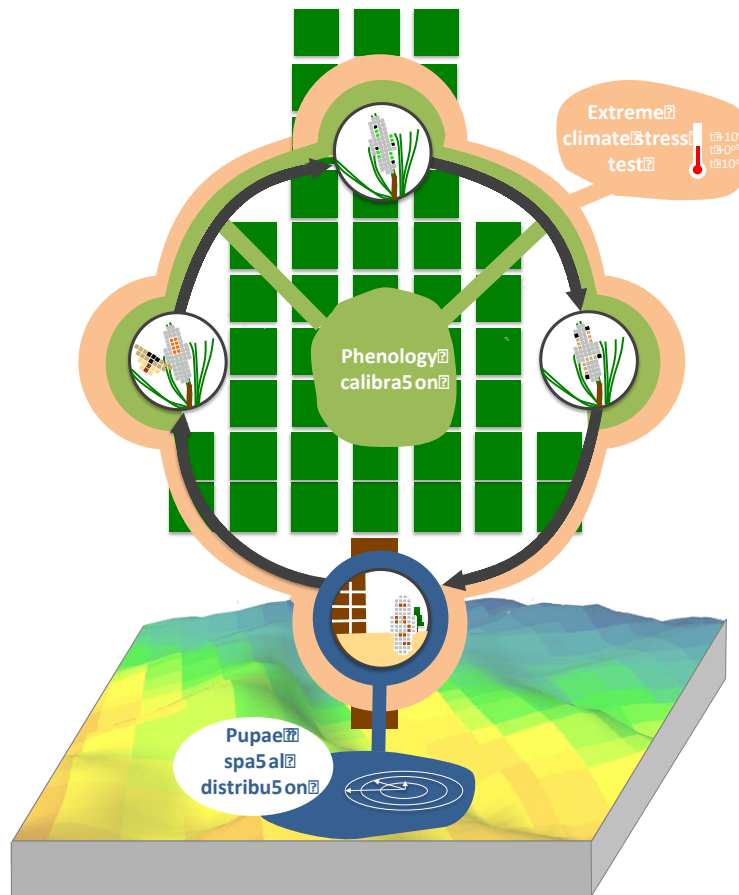
607 process of writing and testing has been documented using a version control tool (Github). This  
608 allows for easy and systematic access to different versions of the code, as well as tracking of  
609 changes. The last *INSTAR* stable version can be freely downloaded from GitHub  
610 ([http://sl.ugr.es/github\\_instar](http://sl.ugr.es/github_instar)) and run under a NetLogo free licence  
611 (<https://ccl.northwestern.edu/netlogo>). Besides this, a demo execution of *INSTAR* can be run  
612 at <https://vlab.geodab.eu>, a virtual laboratory promoted by ECOPOTENTIAL project.

## 613 4.1 Methodology

614 Pattern-oriented modelling led *INSTAR* design and helped decide which processes should be  
615 included in the model, as recommended by Wiegand et al. (2003). Two main patterns of the  
616 pine processionary population dynamics were considered: phenology and spatial distribution.  
617 In order to be coherent and structurally realistic, *INSTAR* should reproduce the observed  
618 timing of the different life stages in the study area, as well as the spatial distribution of the  
619 pupae within the landscape. Therefore, development and movement submodels are key  
620 processes. Considering these requirements, a calibration of the parameters involved in the  
621 development submodels and a robustness analysis of the procession submodel were  
622 performed. Besides this, a stress test was performed by exposing the model to extreme  
623 climate scenarios. Stress tests often help to identify errors in the model that would not be  
624 noted under normal conditions and therefore extreme unrealistic inputs are used in these  
625 tests (Railsback and Grimm 2011). A summary of the experiments and tests is shown in **Figure**  
626 **2**. These experiments constitute a first internal validation, since they can help with confirming  
627 hypothesis and observing unexpected behaviours which cannot be detected when the model is  
628 run under normal conditions. After these tests, parameterisation and sensitivity analysis are  
629 needed before the model can be fully applied.

630 A virtual experimental area was created using a combination of real and built-in datasets (see  
631 Appendix E for more details). The area has been created for the sake of the consistency

632 analysis, since this is not a simulation experiment. This virtual area is 15 x 12 cells large,  
633 covering an area of 450 x 360 m<sup>2</sup> (16.2 ha). The location is real: Sierra Nevada Protected Area  
634 (south-eastern Spain) at an elevation of 1706 +/- 34 m.a.s.l. and contains 1128 pine trees,  
635 which have been extracted using an existing LIDAR dataset (see section 3.3.2). In this study, we  
636 assumed that the pine species is *Pinus halepensis*, since it is the species for which we have a  
637 carrying capacity equation. Finally, the climatic data needed to conduct the consistency  
638 analysis were generated by WiMMED hydrological model (see section 3.3.2).



639

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

640 **Figure 2.** Diagram illustrating the consistency test presented in this manuscript, constituted by  
641 three experiments: phenology calibration, which consisted of two phases: egg development  
642 calibration through parameters  $T_{egg\_devel}$  and larvae development calibration through  
643 parameters  $T_{larvae\_devel\_internal}$ ,  $T_{larvae\_devel\_high}$  and  $T_{larvae\_devel\_low}$ ; pupae spatial distribution:  
644 analysed the effect of parameter **Procession**<sub>max\_distance</sub> by executing the model with three  
645 different values; and extreme climate stress test: tested the effect of temperature change on  
646 model performance by comparing  $\pm 10$  °C scenarios against the normal one.

#### 647 **4.1.1 Phenology calibration**

648 Each phase in the life-cycle of *T. pityocampa* shows a characteristic phenology, which shifts  
649 each year within certain limits depending on climate conditions. This process is mainly  
650 controlled in *INSTAR* by the "colonies development" submodel. In order to produce realistic  
651 outcomes where the phenology of the model reproduces field observations, a calibration  
652 analysis was performed. The parameters involved in the "colonies development" submodel are  
653 the temperature thresholds regulating the length of each phase and the minimum lengths of  
654 each stage under optimal conditions (see **Table 2**). The aim of the calibration was to find the  
655 values of the parameters  $T_{egg\_devel}$ ,  $T_{larvae\_devel\_high}$ ,  $T_{larvae\_devel\_low}$  and  $T_{larvae\_devel\_internal}$  that best  
656 reproduce *T. pityocampa* phenology in the study area.

657 For this calibration, field observations were used. They were provided by the Sierra Nevada  
658 Global Change Observatory (Bollullos and Aspizua 2014), and consist on a dataset including  
659 hatching, procession and egg laying dates recorded for several plots within the Sierra Nevada  
660 mountain range during 2009, 2010 and 2011. A summary of this dataset can be observed in

661 **Table 3.**

662 **Table 3.** Summary of hatching, procession and egg laying dates (in Julian dates) for  
 663 *Thaumetopoea pityocampa* observed in Sierra Nevada. Data from Sierra Nevada Global Change  
 664 Observatory (Bollullos and Aspizua 2014).

	EARLIEST DATE	EARLIEST DATE FOR 95% OF OBSERVATIONS	LATEST DATE FOR 95% OF OBSERVATIONS	LATEST DATE
<b>Hatching date</b>	204	211	256	288
<b>Procession date</b>	331 (year <sub>x</sub> )	331 (year <sub>x</sub> )	124 (year <sub>x+1</sub> )	139 (year <sub>x+1</sub> )
<b>Egg laying date</b>	99	167	261	275

665 Calibration was performed in a stepwise approach: egg development calibration followed by  
 666 larvae development calibration. In the first step, the parameter  $T_{egg\_devel}$ , which affects how  
 667 long the *stage* "egg" takes by adjusting *days as egg*, was calibrated in order to obtain  
 668 simulated hatching dates within the range of the above-mentioned observations. For this  
 669 analysis, a systematic execution of the model was performed with values of  $T_{egg\_devel}$  ranging  
 670 from 15-25 °C in 1 °C steps. Every run was initialised at Julian day 210 and lasted 4 months,  
 671 since tests indicated that such a period was long enough for all eggs within the modelled world  
 672 to hatch. Input data consisted of 13 datasets, corresponding to years 2001-2013 and runs were  
 673 executed in triplicates. For each simulation, the hatching date of each **colony** was recorded.

674 Once the parameter  $T_{egg\_devel}$  was calibrated, the second step of the calibration took place. The  
 675 parameters  $T_{larvae\_devel\_high}$ ,  $T_{larvae\_devel\_low}$  and  $T_{larvae\_devel\_internal}$  were calibrated to obtain  
 676 procession dates within the range of the above-mentioned observations. These parameters  
 677 affect how long the *stages* "L1" and "L2" take by adjusting *days as L1* and *days as L2* (see  
 678 section 3.3.3.2). For this analysis, a new systematic execution of the model was performed. A  
 679 range of values for each parameter was tested: 20-30 °C for  $T_{larvae\_devel\_high}$ , 15-25 °C for  
 680  $T_{larvae\_devel\_internal}$  and -15--5 °C for  $T_{larvae\_devel\_low}$ , in 1 °C step. Every execution was initialised  
 681 with **colonies** at *stage* "L1", to avoid noise coming from the previous *stage*. Thus, executions  
 682 started at Julian day 226 of the corresponding year (mean of the hatching dates derived from  
 683 previous experiment for all  $T_{egg\_devel}$  values). Executions continued for 9 months and, as in the

684 previous experiment, input data consisted of 13 datasets, corresponding to years 2001-2013.

685 For every combination of parameter values and input dataset, five executions were run.

686 In this case, the procession date of each **colony** was recorded in each simulation. Simulations

687 in which at least one execution did not cause all **colonies** to processionate were discarded for

688 the analysis. To avoid noise coming from other submodels, mortality submodels based on **host**

689 *quality* and *quantity* were disabled and the procession submodel was unconstrained (**colonies**

690 could bury in all **cells** around them). Since several parameters were calibrated in this step, an

691 error metric "procession criteria" (PC) was used. The model was run for each combination of

692 parameters until the error was minimised, given the objective function defined as:

693  $PC = \sum(y_i)$

694  $y_i = \left\{ \begin{array}{ll} 0 & \text{if } 1 \leq x_i \leq 124 \end{array} \right.$

695  $\left. \begin{array}{ll} 0 & \text{if } 331 \leq x_i \leq 366 \end{array} \right\}$

696  $\left. \begin{array}{ll} \left( \frac{124+139/2-x_i}{124+139/2} \right)^2 & \text{if } 124 \leq x_i \leq 139 \end{array} \right\}$

697  $\left. \begin{array}{ll} \left( \frac{139+331/2-x_i}{139+331/2} \right)^4 & \text{if } 139 \leq x_i \leq 331 \end{array} \right\}$

698  $\left. \begin{array}{l} \end{array} \right\}$

699 Where  $x_i$  corresponds to the procession date of **colony**  $i$  and therefore  $y_i$  measures the

700 deviation of the procession date from the field observations for each simulation. Note that this

701 deviation is increased if the procession date falls outside the extreme values observed in the

702 field. PC therefore increases as more procession dates within a simulation fall outside the

703 ranges of the field observations.

#### 704 4.1.2 Pupae spatial distribution



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

705 To study how *INSTAR* reproduces the pattern of pupae distribution, the "procession" submodel  
706 was deconstructed to conduct a robustness analysis by neglecting its most important  
707 assumption. The experiment consisted in testing different designs of the submodel to check  
708 how such modifications affect the spatial distribution of the pupae over the landscape. Thus,  
709 the "procession" submodel was modified to disable the constriction of burial spots due to pine  
710 density (i.e. selection was set as random among all **cells** within *Procession*<sub>max\_distance</sub>, rather  
711 than selecting the cell with the smallest pine density). With this modification, we aimed to  
712 verify that such constriction was responsible for the observed spatial distribution. In addition,  
713 two values for *Procession*<sub>max\_distance</sub> were tested (6.15 and 12.3 **cells**) against the normal value  
714 of 1.23 cells. The reason why three values of *Procession*<sub>max\_distance</sub> were tested is based on the  
715 hypothesis that as *Procession*<sub>max\_distance</sub> increases, the spatial distribution of the pupae  
716 becomes more explained by the *probability of emergence* and **host** location loses importance.  
717 One should note that since mortality due to *probability of emergence* is applied immediately  
718 after burial, the distribution of pupae over the landscape refers to the number of pupae that  
719 will actually emerge as **moths** when their development is completed (see section 3.3.1 and  
720 3.3.3 for more details). Therefore, as *Procession*<sub>max\_distance</sub> increases, the selection of a **cell** to  
721 bury becomes more random and the pupae distribution will increasingly resemble the  
722 *probability of emergence* map and not a random distribution. Thus, we would expect a high  
723 correlation between *probability of emergence* map and modelled pupae distribution when  
724 *Procession*<sub>max\_distance</sub> values are high.

725 For each *Procession*<sub>max\_distance</sub> value, triplicates were run for a 5-years period and the average  
726 number of **cluster of pupae per cell** was recorded each month. Maps showing the average  
727 number of pupae were calculated for each *Procession*<sub>max\_distance</sub> value, considering only the  
728 months when a **cluster of pupae** were present. These maps were then compared with the map  
729 of *probability of emergence* through a correlation analysis.

### 730 4.1.3 Extreme climate stress test

731 By exposing the model to extreme climate, we aim to test the model consistency in relation to  
732 temperature influence in development. As mentioned at the beginning of this section, extreme  
733 unrealistic inputs are used in these tests in order to observe divergent behaviours in the  
734 model. The following results are expected from this experiment. Firstly, larvae mortality rates  
735 will increase, even causing pest annihilation, due to the higher occurrence of extreme  
736 temperatures above or below the species survivability thresholds. Secondly, larval  
737 development will be accelerated if temperatures are higher and delayed if they are lower.

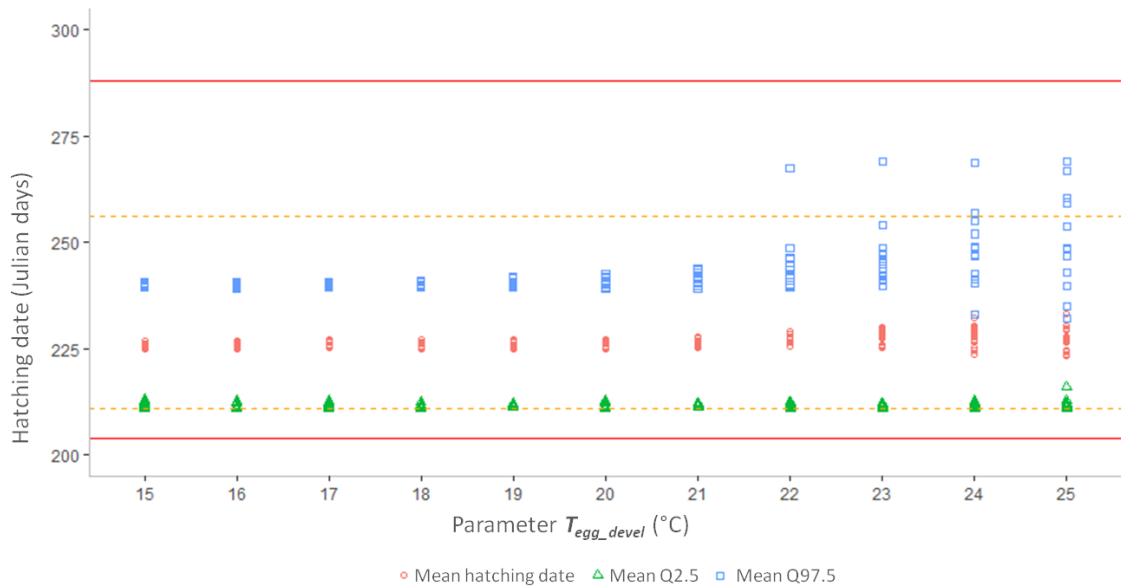
738 Three scenarios were used in this experiment: NORMAL, corresponding to the measured  
739 temperatures, MINUS10 and PLUS10. Scenarios MINUS10 and PLUS10 were generated by  
740 respectively subtracting and adding an offset (10 °C) to the measured temperatures, to reach  
741 lethal temperatures for the processionary moth and therefore test whether *INSTAR* simulates  
742 mortality rates as expected. The offset was applied equally to maximum and minimum  
743 temperatures, therefore maintaining the thermal range. Simulations were run in triplicate for  
744 the period 2001-2006.

## 745 4.2 Results and discussion

### 746 4.2.1 Phenology calibration

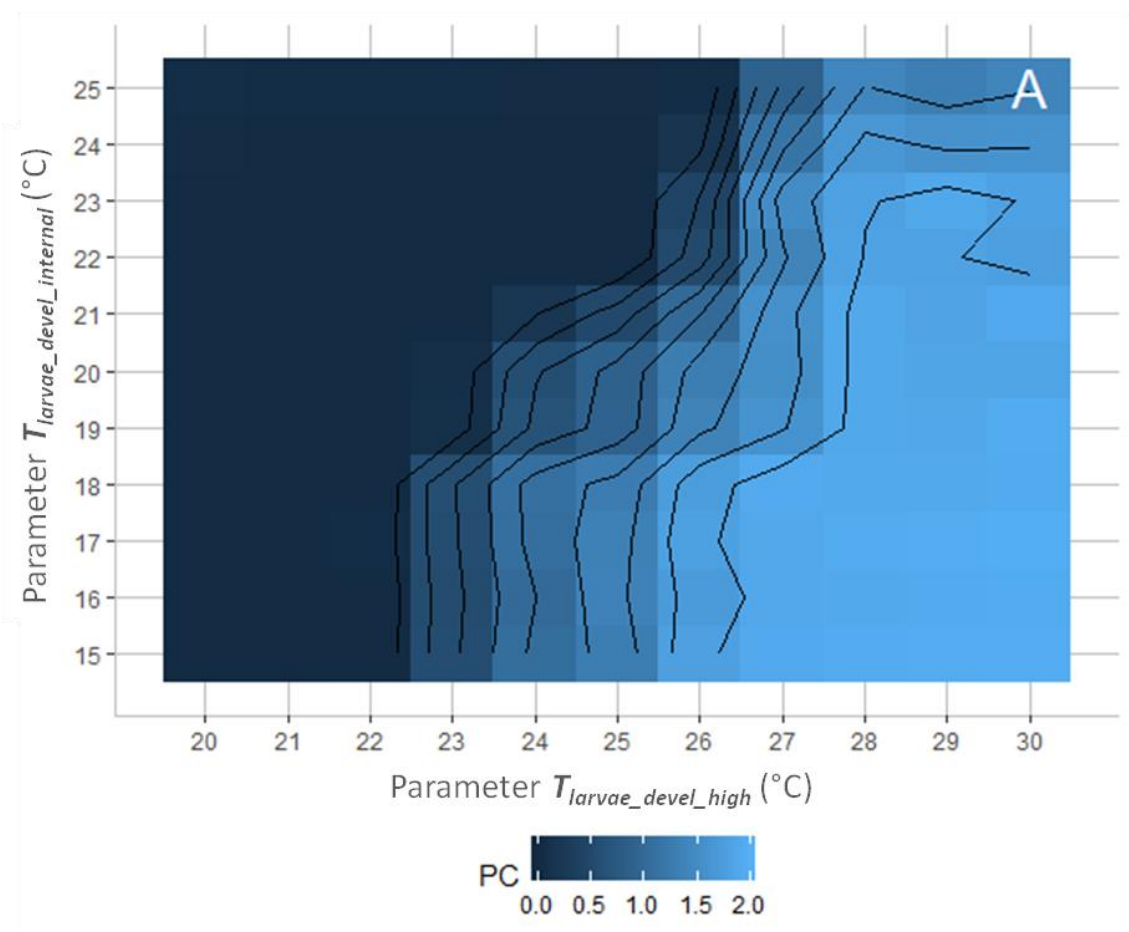
747 Results from the egg development calibration can be observed in **Figure 3**, which shows the  
748 mean hatching date, as well as the mean 2.5 % and 97.5 % quantile of the hatching dates for  
749 each tested value of  $T_{egg\_devel}$  and input dataset. As long as  $T_{egg\_devel}$  remains below 22 °C,  
750 hatching dates remain within field observation ranges. It can also be observed that, as  $T_{egg\_devel}$   
751 increases, the mean hatching date and the 97.5% quantile show increased spreading while the  
752 2.5 % quantile remains almost constant for all simulations. The reason is that as  $T_{egg\_devel}$

753 increases its value, the probability of a **colony** delaying its development also increases.  
 754 Therefore, the mean hatching date and the 97.5% quantile increase. The 2.5 % quantile is  
 755 defined by the **colonies** which were initialised with *days as egg* equal to 1 and therefore  
 756 cannot be further reduced. Considering these results, the value of  $T_{egg\_devel}$  was set to 20 °C,  
 757 since no clear differences between values below 22 °C were observed.

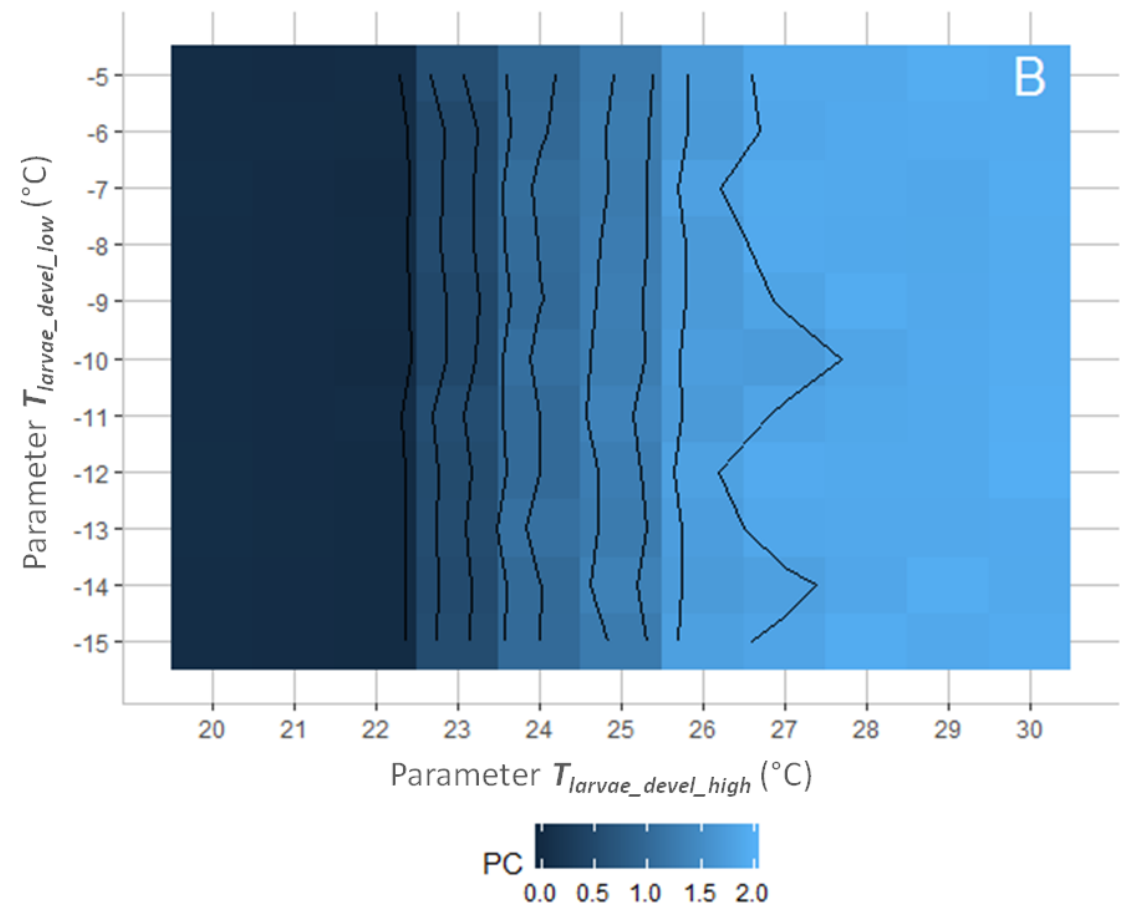


758 **Figure 3.** Plot showing the mean, 2.5 % and 97.5 % quantile of the hatching dates for each  
 759 simulation alongside the tested gradient in  $T_{egg\_devel}$  values. Simulations were run in triplicate  
 760 for each value of  $T_{egg\_devel}$  and input dataset. Hatching dates observed in the field are plotted as  
 761 horizontal lines: yellow dashed lines indicate the 2.5 % and 97.5 % quantile and red lines  
 762 indicate the earliest and latest observed hatching date. Note that axes do not cross in the 0,0  
 763 coordinate.  
 764 coordinate.

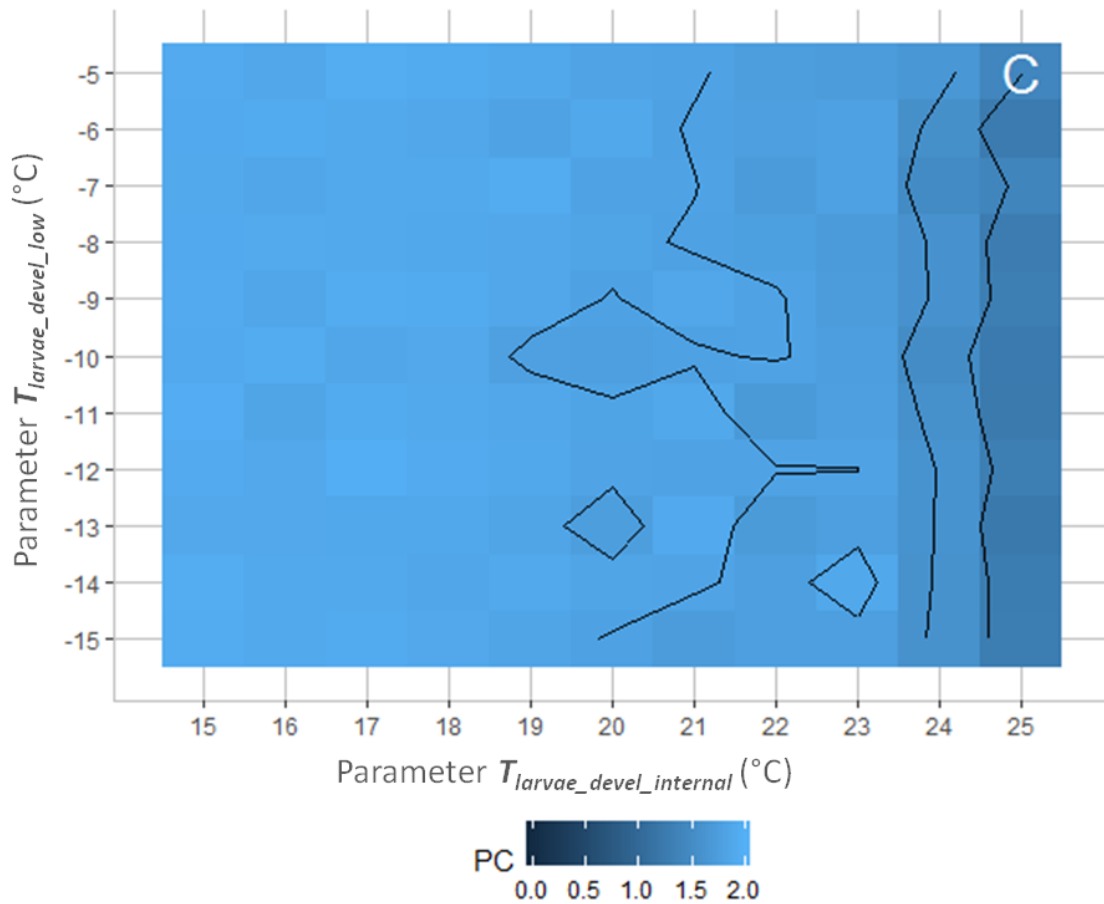
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65



765



766



767

768 **Figure 4.** Plot showing the maximum value of the procession criteria (PC) for each combination  
 769 of parameter values:  $T_{larvae\_devel\_internal}$  and  $T_{larvae\_devel\_high}$  values (A);  $T_{larvae\_devel\_low}$  and  
 770  $T_{larvae\_devel\_high}$  (B) and  $T_{larvae\_devel\_low}$  and  $T_{larvae\_devel\_internal}$  values (C).

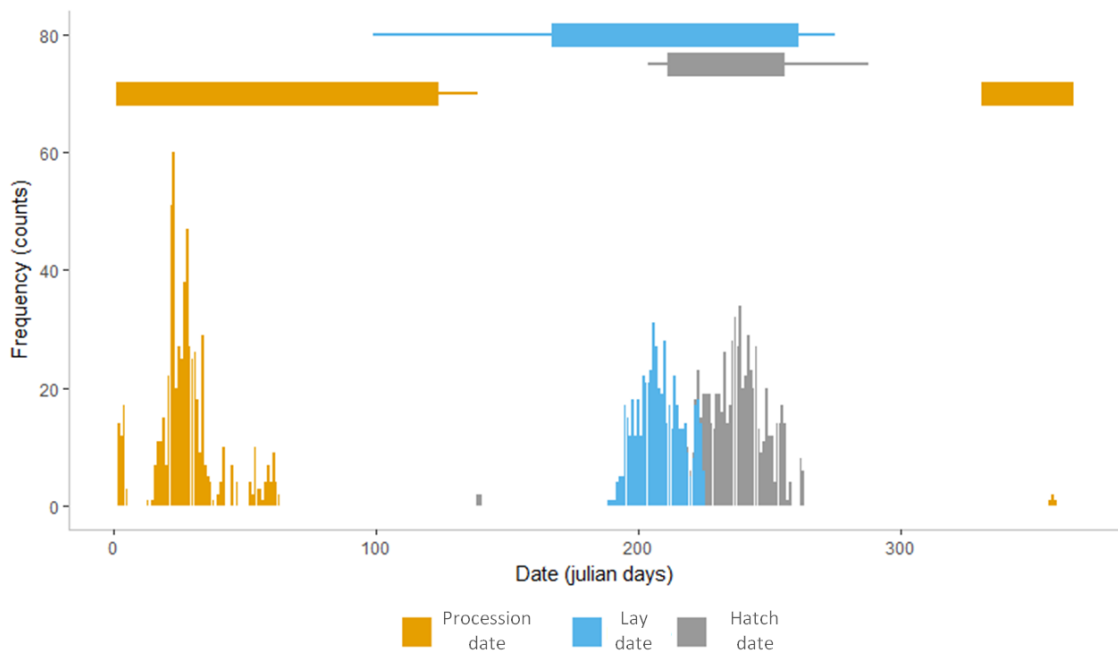
771 Once the egg phase was calibrated, larvae development calibration was carried out. **Figures**  
 772 **4A-C** depict the maximum value of the PC error metric in relation to two out of the three  
 773 tested parameters in each plot. Zero PC values mean that all simulations resulted in  
 774 processions dates falling within the range of the field observations. Similar figures depicting  
 775 the mean and minimum value of PC can be found in **Appendix F**.

776 When analysing the relationship between  $T_{larvae\_devel\_internal}$  and  $T_{larvae\_devel\_high}$  (**Figure 4A**), it can  
 777 be observed that the whole range of  $T_{larvae\_devel\_internal}$  values can result in PC equal to 0, as long  
 778 as  $T_{larvae\_devel\_high}$  value is 20 °C or 21 °C. Above this value, PC gradually increases as  
 779  $T_{larvae\_devel\_internal}$  decreases and  $T_{larvae\_devel\_high}$  increases. Regarding  $T_{larvae\_devel\_low}$ , it can be

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

780 observed that the whole range of tested values can result in a maximum PC equal to 0 if  
781  $T_{larvae\_devel\_high}$  remains between 20 °C and 22 °C (**Figure 4B**) although no  $T_{larvae\_devel\_internal}$  value  
782 fulfils this condition (**Figure 4C**). In other words, no combination of parameter values ensures  
783 that all modelled procession dates will fall within 95% of the field observations for every year.  
784 Nevertheless, several combinations of parameter values result in at least some of the  
785 simulations reproducing the field observations.  
786 Therefore, the parameter values were selected considering the combination which results in  
787 more simulations reproducing the observed *in situ* patterns and the species physiology as  
788 described in the literature. The selected values are 22 °C for  $T_{larvae\_devel\_high}$ , 20 °C for  
789  $T_{larvae\_devel\_internal}$  and -10 °C for  $T_{larvae\_devel\_low}$ . These values seem in agreement with those  
790 previously reported in the literature. Démolin (1969b) reported normal larvae activity if  
791 maximum temperature remains below 25 °C and integrated temperature above 20 °C and  
792 colony survival if minimum temperature remains above -10 °C. Moreover, Hoch et al. (2009)  
793 established the supercooling point of individual larvae at -7 °C. Naturally, discrepancies are  
794 expected between the literature and our results. Démolin observations took place in France  
795 and correspond to thresholds defining the behaviour of larvae and colonies, and not the  
796 accumulative effect of such temperatures. Similarly, Hoch results correspond to survival  
797 thresholds, while our parameter  $T_{larvae\_devel\_low}$  regulates the temperature below which larvae  
798 development stops, but is restored as temperatures raise again (note the difference between  
799 the parameters  $T_{larvae\_devel\_low}$  and  $T_{lethal\_min}$ ). The stepwise calibration was based on hatching  
800 and procession dates within 1-cycle runs. Once the parameter values have been calibrated, it  
801 becomes necessary to test whether these parameters reproduce the right pattern over the  
802 years, i.e. if the model reproduces faithfully the hatching and procession dates as the  
803 simulation time passes or it shifts towards a non-realistic phenology. Moreover, egg laying  
804 dates should also fall within the observed range. To test these criteria, *INSTAR* was run for the  
805 period 2001-2014 with the calibrated value for each parameter (i.e. 20 °C for  $T_{egg\_devel}$ , 22 °C

806 for  $T_{larvae\_devel\_high}$ , 20 °C for  $T_{larvae\_devel\_internal}$  and -10 °C for  $T_{larvae\_devel\_low}$ . The modelled  
 807 phenology shows a pattern similar to the one observed in the field (**Figure 5**). Modelled  
 808 hatchings fall outside of the observations range in 0.6% of the data, while processions and egg  
 809 laying dates fall within 95% of the field observations range in all cases. Thus, the stepwise  
 810 calibration seems to provide realistic values for the parameters involved in the development  
 811 submodel which reproduced the phenology pattern observed in the real system.



812  
 813 **Figure 5.** Histogram showing the frequency of procession (orange), egg laying (blue) and  
 814 hatching (grey) dates for each day of the year. Data obtained from a simulation run for the  
 815 period 2001-2014. Field observations are also depicted in the plot, differentiating between the  
 816 central 95% of the observations (thick horizontal segments) and the extreme observations  
 817 (thin horizontal lines).

#### 818 4.2.2 Pupae spatial distribution

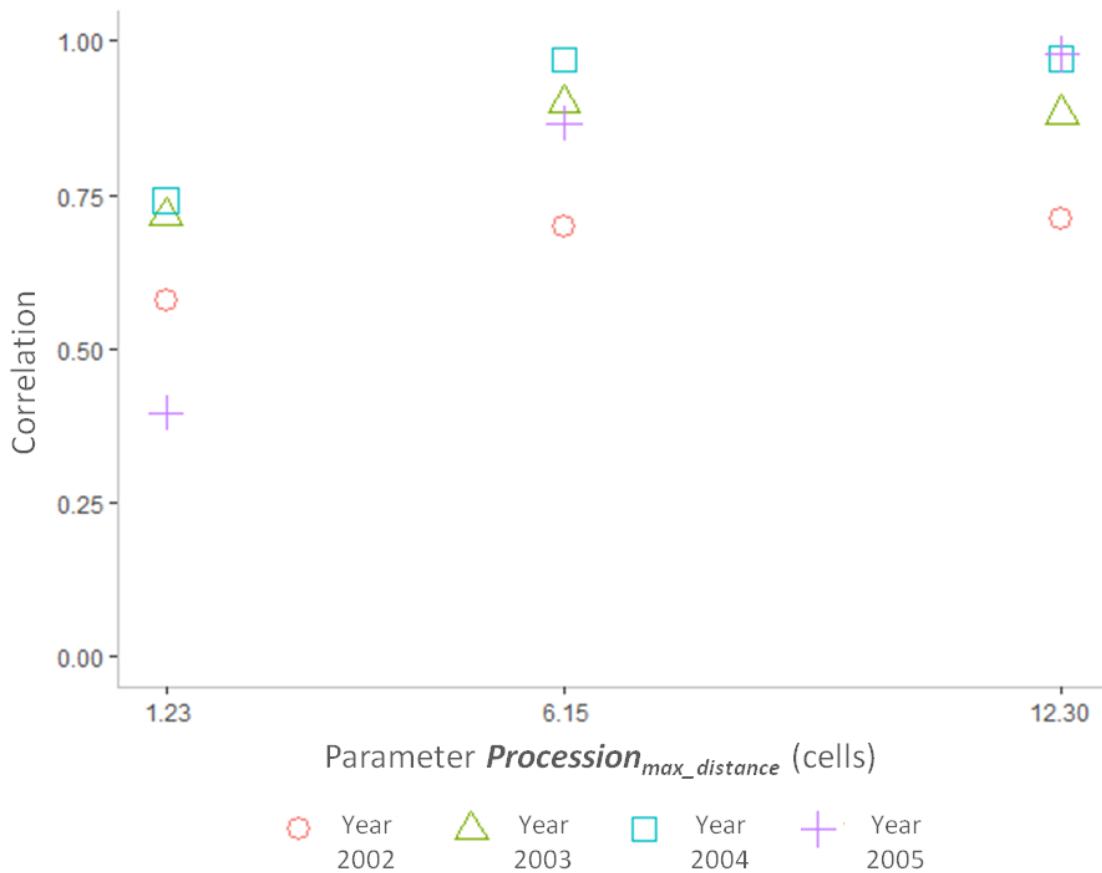
819 Regarding the distribution of the pupae, maps show a high correlation with the *probability of*  
 820 *emergence* map. This indicates the relevance of this parameter in explaining the spatial

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

821 pattern. One should note that *probability of emergence* is calculated based on the number of  
822 **hosts per cell**, which constitutes a proxy for vegetation cover and, consequently, surface soil  
823 moisture (Zribi et al. 2010). Since the presence of shady habitats is considered the most  
824 important factor explaining pupae mortality, the relevance of this parameter within the model  
825 agrees with our current understanding of the system. Moreover, the correlation values  
826 increase as *Procession<sub>max\_distance</sub>* increases, supporting our initial hypothesis. In other words,  
827 the *probability of emergence* map shows the highest similarity with the pupae distribution at  
828 higher *Procession<sub>max\_distance</sub>* values (**Figure 6**).

829 This simulation experiment has confirmed that the mechanism generating the spatial  
830 distribution of the pupae corresponds to the **cell** selection within the "procession" submodel,  
831 as we intended. Moreover, the importance of the *probability of emergence* map has been  
832 highlighted as an important aspect to consider when choosing areas to be simulated, since  
833 pupae spatial distribution is highly influenced by stand density, which is responsible for the  
834 generation of this map. Further work in this regard could be done in the form of a sensitivity  
835 analysis to quantify the influence of the "procession" submodel in the model outputs.





836

837 **Figure 6.** Correlation values between the emergence probability map and the average pupae  
 838 density map for each tested value of *Procession*<sub>max\_distance</sub>.

839 **4.2.3 Extreme climate stress test**

840 Pest dynamics for each scenario are shown in **Figure 7**. As expected, high larvae mortality rates  
 841 resulted in a reduced number of individuals in both MINUS10 and PLUS10 scenarios compared  
 842 to the NORMAL scenario. In the MINUS10 scenario, the population was greatly limited by low  
 843 temperatures (**Figure 7A**). The fact that egg mortality is only due to parasitism explains the  
 844 presence of eggs throughout the whole simulation. No references have been found reporting  
 845 egg mortality due to extreme temperatures and therefore egg mortality within *INSTAR* is not  
 846 influenced by this factor. However, parasitism probability could be assumed to increase as  
 847 time passes since the probability of an egg getting infected by a parasite increases over time.

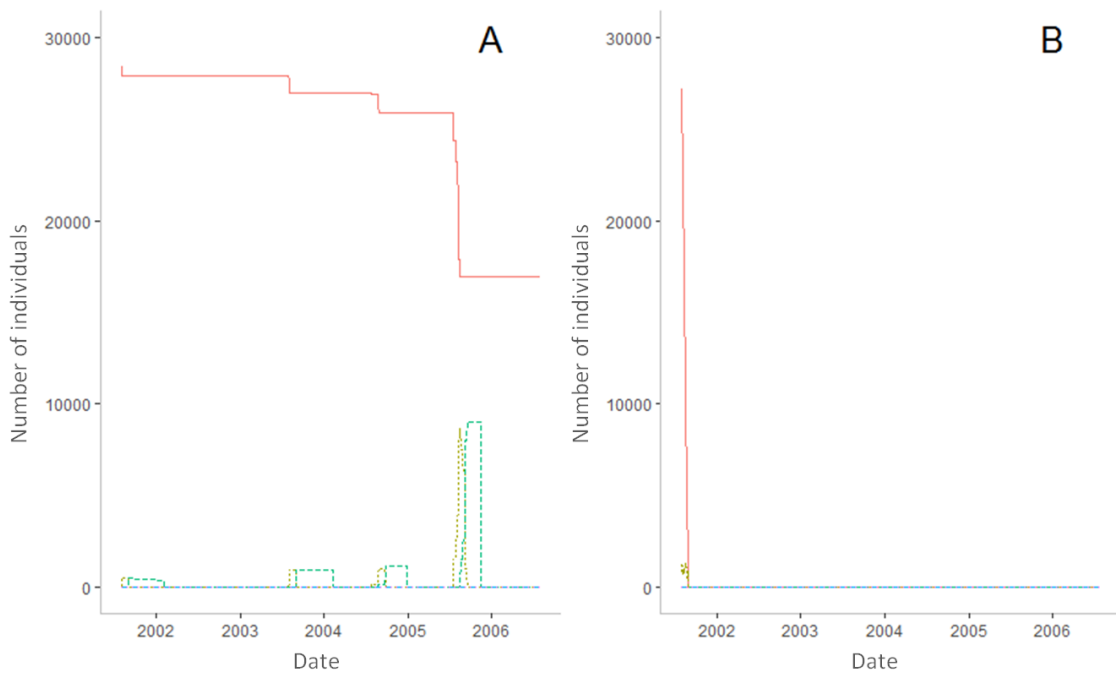
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

848 Thus, the model could be improved in further versions by applying an egg mortality submodel  
849 relating survival rate and time.

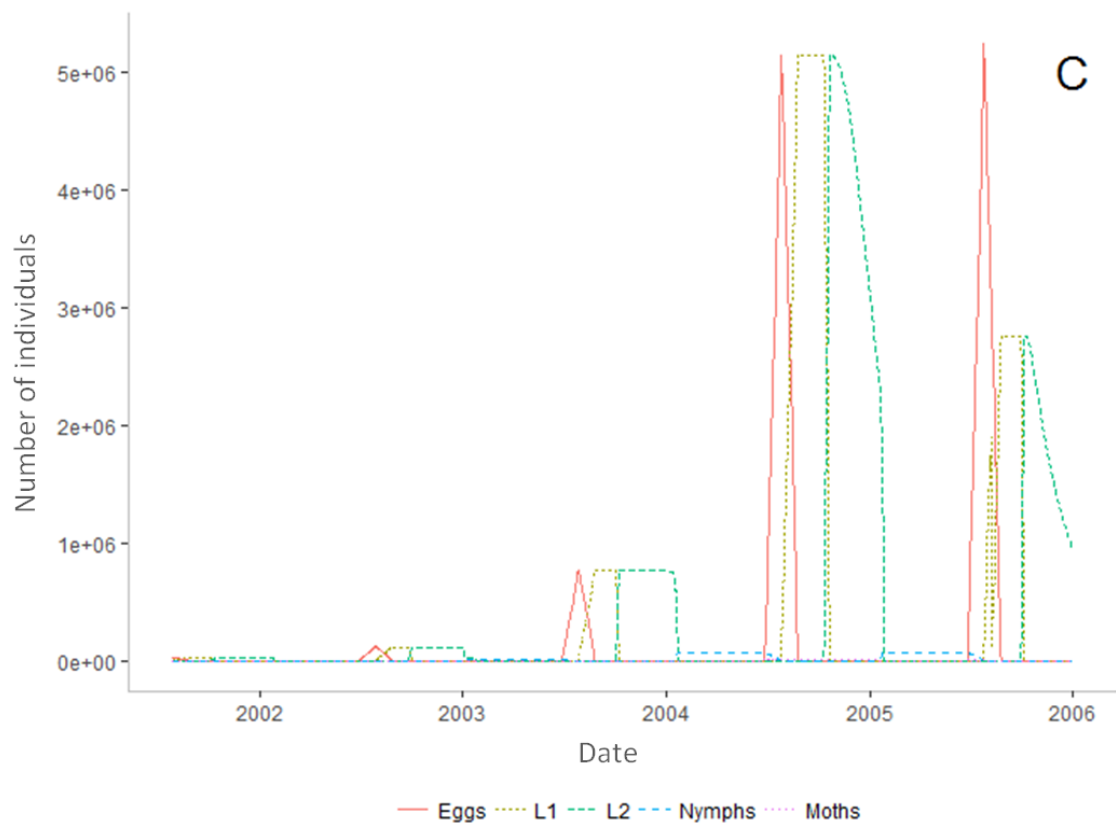
850 Under the PLUS10 scenario, however, larvae mortality due to high temperatures was as high as  
851 to cause full annihilation of the pest before the first cycle was completed (**Figure 7B**). Our  
852 simulated area represents a warm region of the distribution area of the pest. Therefore, the  
853 population is probably closer to its higher temperature threshold than to its lower one. Thus,  
854 an extreme increase in temperature causes most cells in the landscape to reach  $T_{lethal\_max}$  (32  
855 °C). Conversely, a change towards lower temperatures does not cause many cells to reach the  
856  $T_{lethal\_min}$  (-12 °C). Nevertheless, these conclusions should be taken cautiously. The parameter  
857 values have been taken from literature, they refer to air temperatures and represent the  
858 general knowledge of *T. pityocampa*. The actual lethal values may differ for different  
859 populations and been influenced by other factors (e.g. wind speed, shading) which can  
860 modulate the actual temperature to which larvae are exposed.

861 Another reason for the difference between PLUS10 and MINUS10 scenarios can be that at  
862 higher temperatures, eggs develop fast and hatching is anticipated, causing the larvae to be  
863 exposed to the high summer temperatures. Literature suggests that *T. pityocampa* adapts to  
864 areas with warm summers by delaying the egg phase to late summer, thus avoiding lethal  
865 temperatures (Chapter 2 in Roques (2015)). At the moment, *INSTAR* does not possess a  
866 mechanism simulating this process, since initialisation is fixed at a certain date as it is  
867 commonly done in processes that deal with seasonal and annual patterns (such as hydrological  
868 models). This limitation, however, can be overcome by calibrating the initialisation date for the  
869 area to be simulated based on field data. The temperature stress test confirmed that the  
870 model behaves as expected when exposed to extreme input values.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65



871



872

873 **Figure 7.** Plots showing the pest development as number of eggs (red continuous line), larvae 1  
874 (light brown pointed line), larvae 2 (green short-dashed line), pupae (blue long-dashed line)

1  
2 875 and moths (pink pointed line) under three different scenarios: MINUS10 (A), PLUS10 (B) and  
3 876 NORMAL (C), for the period 2001-2006.  
4  
5  
6  
7

## 8 877 5. Conclusions 9

10  
11  
12 878 To the best of our knowledge, *INSTAR* is the first ABM simulating the biological cycle of the  
13  
14 879 processionary moth. The model aims to improve our understanding of the species population  
15  
16 880 dynamics and identify gaps in our knowledge. Our model constitutes an effort to integrate and  
17  
18 881 make use of the available knowledge about a forest pest. In this paper we present the model  
19  
20 882 structure as well as a consistency test which constitutes a first internal validation of the model.  
21  
22 883 The experiments presented here assess how the model reproduces the most important  
23  
24 884 patterns of the pest: phenology and spatial distribution. As a general conclusion, *INSTAR* has  
25  
26 885 demonstrated to be robust and internally coherent, since it reproduces the most important  
27  
28 886 phases of the target species' biological cycle. More specifically, the main conclusions of this  
29  
30 887 work are:  
31  
32  
33  
34  
35

- 36 888
- 37 • The simulated phenology of the species agrees with field observations for a given area.  
38  
39 889 This means that the submodels responsible for the phenology within *INSTAR* are well  
40  
41 890 designed and implemented.
  - 42  
43 891 • Pupae spatial distribution is in accordance to what it is expected based on literature.  
44  
45 892 Thus, the submodel responsible for simulating the pupae distribution reproduces  
46  
47 893 reasonably well the pattern described from the field.
  - 48  
49 894 • The model behaves as expected when exposed to extreme input values through a  
50  
51 895 climate stress test, i.e. extreme high temperatures cause full annihilation of the pest  
52  
53 896 before the first cycle is completed and extreme low temperatures greatly limit the pest  
54  
55 897 population.  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

898       • *INSTAR* has been built with a modular structure to allow an easy integration of new  
899       knowledge.

900 Besides, *INSTAR* design, implementation and testing, helps in identifying gaps of knowledge  
901 which can promote further advances in the research focused on *T. pityocampa*. In this regard  
902 we have identified the following gaps:

903       • It appears that there is a lack of theoretical knowledge about larval biomass  
904 consumption which has avoided a more realistic modelisation of the host-pest  
905 interaction. Forest pests are commonly monitored by assessing their impact as  
906 defoliation or other stress signals on the host. In the case of the pine processionary  
907 moth, the number of colonies is commonly reported, but data at a lower level (i.e.  
908 number of larvae per colony) are scarce. Such information would be particularly  
909 valuable to improve host-pest interaction modelling and to calibrate the mortality  
910 submodels.

911       • Factors affecting egg, larvae and pupae development also appear as an important  
912 aspect to improve our understanding of *T. pityocampa* population dynamics. Although  
913 there is literature about the pest development, the level of detail of this information is  
914 low, and not enough knowledge is available to understand how *T. pityocampa*  
915 regulates its life cycle.

916 Due to the above-mentioned issues and the inherent characteristics of ABMs, there is a high  
917 degree of uncertainty in the model. At this moment, our aim was to present a fully functioning  
918 model which simulates several interlinked and complex processes and therefore no  
919 uncertainty analysis has been performed so far. Further work should focus on model  
920 calibration, sensitivity analysis and quantification of model uncertainty. Addressing these  
921 topics will help us to create a predictive tool able to forecast the behaviour of the target forest  
922 pest.

## 923 6. Author contributions

924 María Suárez-Muñoz led the manuscript writing process as well as the *INSTAR* code review  
 925 and testing. She also designed the experiments showed in the text. Francisco Bonet-García  
 926 contributed to the initial conception of the model and to write the manuscript. He supervised  
 927 the whole process and contributed to the funding acquisition. José A. Hódar contributed to the  
 928 conceptual model and provided basic data and knowledge regarding pine processionary moth  
 929 natural history and ecology. He also participated in the model implementation phase, writing  
 930 of the manuscript and the funding acquisition. Javier Herrero designed and executed the  
 931 hydrometeorological model and contributed to the manuscript. Mihai Tanase pre-processed  
 932 the LiDAR dataset and produced the LiDAR based metrics used in this study. He also  
 933 contributed to the manuscript with the description of the LiDAR data analysis. Lucía Torres-  
 934 Muros developed the initial conceptual model as well as to the collection of data needed to  
 935 calibrate some *INSTAR* functions. The table shown below describes with more detail the  
 936 specific contributions of each author.

Author name	Study conception	Methods	Computation	Formal analysis	Perform the experiments	Data collection	Data curation	Writing the manuscript	Data presentation	Supervision	Funding acquisition / project admin
María Suárez-Muñoz	3	3	4	3	4	1	5	5	4	0	0
Francisco J. Bonet-García	4	2	2	3	2	2	3	4	4	5	5
José Antonio Hódar	4	2	0	2	3	3	1	3	1	2	2
Javier Herrero-Lantarón	0	2	2	2	1	0	2	1	2	1	0
Mihai Tanase	0	2	2	2	1	0	1	1	2	0	0
Lucía Torres Muros	4	2	0	2	2	3	3	1	0	0	0

## 937 7. Acknowledgements

938 This work has been carried out within Sierra Nevada Global Change Observatory (LTER  
 939 platform, South Spain). This research infrastructure has provided datasets to calibrate the

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

940 model as well as hardware/software support for the authors. Besides, the tasks described in  
941 this work are outcomes of the H2020 project "ECOPOTENTIAL: Improving Future Ecosystem  
942 Benefits Through Earth Observations" (<http://www.ecopotential-project.eu>). The project has  
943 received funding from the European Union's Horizon 2020 research and innovation  
944 programme under grant agreement No 641762.

945 María Suárez-Muñoz has been funded by the European Social Fund through the Youth  
946 Employment Initiative and the *Ministerio de Educación y Formación Profesional* under the  
947 framework of *Plan Estatal de Investigación científica y Técnica y de Innovación 2013-2016*.

948 The authors thank Andalusian Environmental Information Network REDIAM, the National  
949 Agency of Meteorology AEMET and the National Parks Network for the meteorological data  
950 provided.

951 We also want to acknowledge the role played by Ramón Pérez-Pérez, who programmed the  
952 first version of *INSTAR*. His ability to translate natural processes into programming code is only  
953 comparable to his shyness to participate as co-author of this manuscript.

954 Finally, we thank the anonymous reviewers for their useful comments and suggestions which  
955 serve to improve earlier versions of this manuscript.

## 956 **7. Declaration of interest**

957 None

## 958 8. References

- 959 Aimi, A., A. Zocca, S. Minerbi, K. Hellrigl, P. Gatto, and A. Battisti. 2006. The outbreak of the  
960 pine processionary moth in Venosta/Vinschgau: ecological and economic aspects. *Forest*  
961 *Observer* 2/3:69–80.
- 962 Alcaraz-Segura, D., A. Reyes, and J. Cabello. 2016. Changes in vegetation productivity according  
963 to teledetection. Pages 142–145 *in* *Impacts of global change in sierra nevada: Challenges for*  
964 *conservation*. Consejería de Medio Ambiente y Ordenación del Territorio. Junta de Andalucía.
- 965 Arnaldo, P., S. Chacim, and D. Lopes. 2010. Effects of defoliation by the pine processionary  
966 moth *Thaumetopoea pityocampa* on biomass growth of young stands of *Pinus pinaster* in  
967 northern Portugal. *iForest - Biogeosciences and Forestry* 3:159–162.
- 968 Barbaro, L., and A. Battisti. 2011. Birds as Predators of the Pine Processionary Moth  
969 (*Lepidoptera: Notodontidae*). *Biological Control* 56:107–114.
- 970 Barbosa, P., D. K. Letourneau, and A. A. Agrawal, editors. 2012. *Insect Outbreaks Revisited*.  
971 John Wiley & Sons Inc (Sea) Pte Ltd.
- 972 Battisti, A. 1988. Host-plant relationships and population dynamics of the Pine Processionary  
973 Caterpillar *Thaumetopoea pityocampa* (Denis & Schiffermuller). *Journal of Applied Entomology*  
974 105:393–402.
- 975 Battisti, A., M. Avci, D. N. Avtzis, M. L. Ben Jamaa, L. Berardi, W. Berretima, M. Branco, G.  
976 Chakali, M. A. El Alaoui El Fels, B. Frérot, J. A. Hódar, I. Ionescu-Malancus, K. Ipekdal, S. Larsson,  
977 T. Manole, Z. Mendel, N. Meurisse, P. Mirchev, N. Nemer, M.-R. Paiva, J. Pino, A. Protasov, N.  
978 Rahim, J. Rousselet, H. Santos, D. Sauvard, A. Schopf, M. Simonato, A. Yart, and M. Zamoum.  
979 2015. Natural History of the Processionary Moths (*Thaumetopoea* spp.): New Insights in  
980 Relation to Climate Change. Pages 15–79 *in* R. A., editor. *Processionary moths and climate*  
981 *change: An update*. Springer, Dordrecht.



- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65
- 982 Battisti, A., S. Larsson, and A. Roques. 2017. Processionary Moths and Associated Urtication  
983 Risk: Global Change-Driven Effects. Pages 323–342 *in* M. Berenbaum, editor. Annual review of  
984 entomology.
- 985 Battisti, A., M. Stastny, S. Netherer, C. Robinet, A. Schopf, A. Roques, and S. Larsson. 2005.  
986 Expansion of geographic range in the pine processionary moth caused by increased winter  
987 temperatures. *Ecological Applications* 15:2084–2096.
- 988 Berardi, L., M. Branco, M. R. Paiva, H. Santos, and A. Battisti. 2015. Development time plasticity  
989 of the pine processionary moth (*Thaumetopoea pityocampa*) populations under laboratory  
990 conditions. *Entomologia* 3.
- 991 Bollullos, C., and R. Aspizua. 2014. Pine processionary moth. Pages 94–95 *in* R. Aspizua, J.  
992 Barea-Azcón, A. Bonet, F.J., Pérez-Luque, and R. Zamora, editors. Sierra nevada global change  
993 observatory. monitoring methodologies. Consejería de Medio Ambiente, Junta de Andalucía.
- 994 Castagneyrol, B., H. Jactel, Y. Charbonnier, L. Barbaro, and A.-M. Dulaurent-Mercadal. 2014a.  
995 Egg mortality in the pine processionary moth: habitat diversity, microclimate and predation  
996 effects. *Agricultural and Forest Entomology* 16:284–292.
- 997 Castagneyrol, B., M. Régolini, and H. Jactel. 2014b. Tree species composition rather than  
998 diversity triggers associational resistance to the pine processionary moth. *Basic and Applied*  
999 *Ecology* 15:516–523.
- 1000 Cayuela, L., R. Hernández, J. A. Hódar, G. Sánchez, and R. Zamora. 2014. Tree damage and  
1001 population density relationships for the pine processionary moth: Prospects for ecological  
1002 research and pest management. *Forest Ecology and Management* 328:319–325.
- 1003 Charbonnier, Y., L. Barbaro, A. Theillout, and H. Jactel. 2014. Numerical and functional  
1004 responses of forest bats to a major insect pest in pine plantations. *PLoS One* 9:e109488.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1005 Chowdhury, D., and D. Stauffer. 2005. Evolutionary ecology in silico: Does mathematical  
1006 modelling help in understanding 'generic' trends? *Journal of Biosciences* 30:277–287.  
1007 Dajoz, R. 2000. *Insects and Forests: The Role and Diversity of Insects in the Forest*  
1008 *Environment*. Intercept Limited.  
1009 DeAngelis, D. L., and V. Grimm. 2014. Individual-based models in ecology after four decades.  
1010 *F1000Prime Reports* 6:39.  
1011 Démolin, G. 1969a. Comportement des adultes de *Thaumetopoea pityocampa* Schiff.  
1012 Dispersion spatiale, importance écologique. *Annales des Sciences Forestières* 26:81–102.  
1013 Démolin, G. 1969b. Bioecología de la “procesionaria del pino” *Thaumetopoea pityocampa*  
1014 Schiff. Incidencia de los factores climáticos. *Boletín del Servicio de Plagas Forestales*, ISSN:  
1015 0584-7885 12:9–24.  
1016 Dulaurent, A., A. Porté, I. van Halder, F. Vetillard, P. Menassieu, and H. Jactel. 2012. Hide and  
1017 seek in forests: colonization by the pine processionary moth is impeded by the presence of  
1018 nonhost trees. *Agricultural and Forest Meteorology*:19–27.  
1019 Dulaurent, A.-M., A. J. Porté, I. van Halder, F. Vétillard, P. Menassieu, and H. Jactel. 2011. A  
1020 case of habitat complementation in forest pests: Pine processionary moth pupae survive  
1021 better in open areas. *Forest Ecology and Management* 261:1069–1076.  
1022 Fares, S., G. S. Mugnozza, P. Corona, and M. Palahí. 2015. Sustainability: Five steps for  
1023 managing Europe’s forests. *Nature* 519:407–409.  
1024 Findeisen, W., A. Lastebro, R. Lande, J. Lindsay, M. Pearson, and E. Quade. 1978. A sample  
1025 glossary of systems analysis. International Institute for Applied Systems Analysis, Laxenburg,  
1026 Austria.  
1027 Gertseva, V., and V. Gertseva. 2004. Classification of mathematical models in ecology.  
1028 *Ecological Modelling* 178:329–334.

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65
- 1029 Grimm, V., U. Berger, D. L. DeAngelis, J. G. Polhill, J. Giske, and S. F. Railsback. 2010. The ODD  
1030 protocol: A review and first update. *Ecological Modelling* 221:2760–2768.
- 1031 Herrero, J., M. J. Polo, A. Moñino, and M. A. Losada. 2009. An energy balance snowmelt model  
1032 in a Mediterranean site. *Journal of Hydrology* 371:98–107.
- 1033 Higham, M., B. M. Hoven, D. L. Gorchov, and K. S. Knight. 2017. Patterns of Coarse Woody  
1034 Debris in Hardwood Forests across a Chronosequence of Ash Mortality Due to the Emerald Ash  
1035 Borer (*Agrilus planipennis*). *Natural Areas Journal* 37:406–411.
- 1036 Hoch, G., E. P. Toffolo, S. Netherer, A. Battisti, and A. Schopf. 2009. Survival at low  
1037 temperature of larvae of the pine processionary moth *Thaumetopoea pityocampa* from an  
1038 area of range expansion. *Agricultural and Forest Entomology* 11:313–320.
- 1039 Hódar, J. A., and R. Zamora. 2004. Herbivory and climatic warming: a Mediterranean  
1040 outbreaking caterpillar attacks a relict, boreal pine species. *Biodiversity & Conservation*  
1041 13:493–500.
- 1042 Hódar, J. A., R. Zamora, and J. Castro. 2002. Host utilisation by moth and larval survival of pine  
1043 processionary caterpillar *Thaumetopoea pityocampa* in relation to food quality in three *Pinus*  
1044 species. *Ecological Entomology* 27:292–301.
- 1045 Hódar, J. A., R. Zamora, and L. Cayuela. 2012. Climate change and the incidence of a forest pest  
1046 in Mediterranean ecosystems: can the North Atlantic Oscillation be used as a predictor?  
1047 *Climatic Change* 113:699–711.
- 1048 Hódar, J. A., R. Zamora, J. Castro, and E. Baraza. 2004. Feast and famine: previous defoliation  
1049 limiting survival of pine processionary caterpillar *Thaumetopoea pityocampa* in Scots pine  
1050 *Pinus sylvestris*. *Acta Oecologica* 26:203–210.
- 1051 Huchon, H., and G. Demolin. 1970. La bioécologie de la processionnaire du pin. Dispersion  
1052 potentielle, dispersion actuelle. *Revue Forestière Française (sp.)*:220–234.

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65
- 1053 Jacquet, J.-S., C. Orazio, and H. Jactel. 2012. Defoliation by processionary moth significantly  
1054 reduces tree growth: a quantitative review. *Annals of Forest Science* 69:857–866.
- 1055 Jakeman, A. J., R. A. Letcher, and J. P. Norton. 2006. Ten iterative steps in development and  
1056 evaluation of environmental models. *Environmental Modelling & Software* 21:602–614.
- 1057 Kanat, M., M. H. Alma, and F. Sivrikaya. 2005. Effect of defoliation by *Thaumetopoea*  
1058 *pityocampa* (Den. & Schiff.) (Lepidoptera: Thaumetopoeidae) on annual diameter increment of  
1059 *Pinus brutia* Ten. in Turkey. *Annals of Forest Science* 62:91–94.
- 1060 Kelly (Letcher), R. A., A. J. Jakeman, O. Barreteau, M. E. Borsuk, S. ElSawah, S. H. Hamilton, H. J.  
1061 Henriksen, S. Kuikka, H. R. Maier, A. E. Rizzoli, H. van Delden, and A. A. Voinov. 2013. Selecting  
1062 among five common modelling approaches for integrated environmental assessment and  
1063 management. *Environmental Modelling and Software* 47:159–181.
- 1064 Li, S., J. J. Daudin, D. Piou, C. Robinet, and H. Jactel. 2015. Periodicity and synchrony of pine  
1065 processionary moth outbreaks in France. *Forest Ecology and Management* 354:309–317.
- 1066 Linares, J. C., K. Senhadji, A. Herrero, and J. A. Hódar. 2014. Growth patterns at the southern  
1067 range edge of Scots pine: Disentangling the effects of drought and defoliation by the pine  
1068 processionary caterpillar. *Forest Ecology and Management* 315:129–137.
- 1069 Markalas, S. 1998. Biomass Production of *Pinus pinaster* after Defoliation by the Pine  
1070 Processionary Moth (*Thaumetopoea pityocampa* Schiff.). Pages 292–302 in M. McManus and  
1071 A. Liebhold, editors. *Proceedings: Population dynamics, impacts, and integrated management*  
1072 of forest defoliating insects. usda forest service general technical report ne-247.
- 1073 Masutti, L., and A. Battisti. 1990. *Thaumetopoea pityocampa* (Den. & Schiff.) in Italy Bionomics  
1074 and perspectives of integrated control. *Journal of Applied Entomology* 110:229–234.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1075 Mellado, A., and R. Zamora. 2016. Spatial heterogeneity of a parasitic plant drives the seed-  
1076 dispersal pattern of a zoochorous plant community in a generalist dispersal system. *Functional*  
1077 *Ecology* 30:459–467.

1078 Ministerio de Medio Ambiente. 2002. Plan Forestal Español. Secretaria General de Medio  
1079 Ambiente. Dirección General de Conservación de la Naturaleza. Ministerio de Medio  
1080 Ambiente.

1081 Muukkonen, P. 2005. Needle biomass turnover rates of Scots pine (*Pinus sylvestris* L.) derived  
1082 from the needle-shed dynamics. *Trees* 19:273–279.

1083 Netherer, S., and A. Schopf. 2010. Potential effects of climate change on insect herbivores in  
1084 European forests-General aspects and the pine processionary moth as specific example. *Forest*  
1085 *Ecology and Management* 259:831–838.

1086 Petrucco Toffolo, E., I. Bernardinelli, F. Stergulc, and A. Battisti. 2006. Climate change and  
1087 expansion of the Pine Processionary Moth, *Thaumetopoea pityocampa*, in Northern Italy.  
1088 Pages 331–340 in IUFRO working party 7.03.10 proceedings of the workshop. Gmunden,  
1089 Austria.

1090 Pimentel, C., T. Calvao, and M. P. Ayres. 2011. Impact of climatic variation on populations of  
1091 pine processionary moth *Thaumetopoea pityocampa* in a core area of its distribution.  
1092 *Agricultural and Forest Entomology* 13:273–281.

1093 Railsback, S. F., and V. Grimm. 2011. Agent-Based and Individual-Based Modeling: A Practical  
1094 Introduction. Page 352. Princeton University Press.

1095 Rebollo, S., I. Moneo, J. M. Vega, I. Herrera, and M. L. Caballero. 2002. Pine processionary  
1096 caterpillar allergenicity increases during larval development. *International Archives of Allergy*  
1097 *and Immunology* 128:310–314.

1098 Régolini, M., B. Castagneyrol, A.-M. Dulaurent-Mercadal, D. Piou, J.-C. Samalens, and H. Jactel.  
1099 2014. Effect of host tree density and apparency on the probability of attack by the pine  
1100 processionary moth. *Forest Ecology and Management* 334:185–192.

1101 Rhainds, M. 2010. Female mating failures in insects. *Entomologia Experimentalis et Applicata*  
1102 136:211–226.

1103 Robinet, C., P. Baier, J. Pennerstorfer, A. Schopf, and A. Roques. 2007. Modelling the effects of  
1104 climate change on the potential feeding activity of *Thaumetopoea pityocampa* (Den. & Schiff.)  
1105 (Lep., Notodontidae) in France. *Global Ecology and Biogeography* 16:460–471.

1106 Robinet, C., M. Laparie, and J. Rousselet. 2015. Looking Beyond the Large Scale Effects of  
1107 GlobalChange: Local Phenologies Can Result in Critical Heterogeneity in the Pine Processionary  
1108 Moth. *Frontiers in Physiology* 6:334.

1109 Robinet, C., J. Rousselet, and A. Roques. 2014. Potential spread of the pine processionary moth  
1110 in France: preliminary results from a simulation model and future challenges. *Annals of Forest*  
1111 *Science* 71:149–160.

1112 Robredo, F. 1963. Las procesiones de crisalidación de *Thaumetopoea pityocampa* Schiff.  
1113 *Boletín del Servicio de Plagas Forestales* 6:122–130.

1114 Roques, A. 2015. *Processionary Moths and Climate Change: An Update.* (A. Roques, Ed.).  
1115 Springer Netherlands.

1116 Ros Candeira A, S.-M. M., Pérez-Luque AJ, and O. D. E. (forthcoming). (n.d.). Dataset of  
1117 occurrence and incidence of pine processionary moths in Andalusia, south Spain. Zookeys.

1118 Salman, B., MHR, A. El Alaoui El Fels, F. Giomi, J. Hodar, M. Laparie, L. Marini, C. Merel, M.  
1119 Zalucki, M. Zamoum, and A. Battisti. 2019. Winter temperature predicts prolonged diapause in  
1120 pine processionary moth species across their geographic range. *PeerJ* 7.

1 1121 Samalens, J.-C., and J.-P. Rossi. 2011. Does landscape composition alter the spatiotemporal  
2 1122 distribution of the pine processionary moth in a pine plantation forest? *Population Ecology*  
3  
4 1123 53:287–296.  
5  
6  
7 1124 Sangüesa-Barreda, G., J. J. Amarero, A. García-Martín, R. Hernández, and J. de la Riva. 2014.  
8  
9 1125 Remote-sensing and tree-ring based characterization of forest defoliation and growth loss due  
10  
11 1126 to the Mediterranean pine processionary moth. *Forest Ecology and Management* 320:171–  
12  
13 1127 181.  
14  
15  
16  
17 1128 Seidl, R., D. Thom, M. Kautz, D. Martin-Benito, M. Peltoniemi, G. Vacchiano, J. Wild, D. Ascoli,  
18  
19 1129 M. Petr, J. Honkaniemi, M. J. Lexer, V. Trotsiuk, P. Mairota, M. Svoboda, M. Fabrika, T. A.  
20  
21 1130 Nagel, and C. P. O. Reyer. 2017. Forest disturbances under climate change. *Nature Climate*  
22  
23 1131 *Change* 7:395–402.  
24  
25  
26  
27 1132 Seixas Arnaldo, P., and L. Monteiro Torres. 2005. Spatial distribution and sampling of  
28  
29 1133 *Thaumetopoea pityocampa* (Den. & Schiff.) (Lep. Thaumetopoeidea) populations on *Pinus*  
30  
31 1134 *pinaster* Ait. in Montesinho, N. Portugal. *Forest Ecology and Management* 210:1–7.  
32  
33  
34  
35 1135 Seixas Arnaldo, P., I. Oliveira, J. Santos, and S. Leite. 2011. Climate change and forest plagues:  
36  
37 1136 the case of the pine processionary moth in Northeastern Portugal. *Forest Systems* 20:508–515.  
38  
39  
40 1137 Tanase, M. A., C. Aponte, S. Mermoz, A. Bouvet, T. Le Toan, and M. Heurich. 2018. Detection of  
41  
42 1138 windthrows and insect outbreaks by L-band SAR: A case study in the Bavarian Forest National  
43  
44 1139 Park. *Remote Sensing of Environment* 209:700–711.  
45  
46  
47  
48 1140 Torres Muros, L. 2015. Estudio de la dinámica poblacional de la procesionaria del pino  
49  
50 1141 (*Thaumetopoea pityocampa*) en el Espacio Natural de Sierra Nevada, dentro del marco del  
51  
52 1142 cambio global. Análisis aplicado para una gestión adaptativa. PhD thesis, Universidad de  
53  
54 1143 Granada.  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65
- 1144 Torres-Muros, L., J. A. Hódar, and R. Zamora. 2017. Effect of habitat type and soil moisture on  
1145 pupal stage of a Mediterranean forest pest (*Thaumetopoea pityocampa*). *Agricultural and*  
1146 *Forest Entomology* 19:130–138.
- 1147 Valenta, V., D. Moser, S. Kapeller, and F. Essl. 2017. A new forest pest in Europe: a review of  
1148 Emerald ash borer (*Agrilus planipennis*) invasion. *Journal of Applied Entomology* 141:507–526.
- 1149 Vega, M. L., J. Vega, J. M. Vega, I. Moneo, E. Sanchez, and A. Miranda. 2003. Cutaneous  
1150 reactions to pine processionary caterpillar (*Thaumetopoea pityocampa*) in pediatric  
1151 population. *Pediatric Allergy and Immunology* 14:482–486.
- 1152 Wiegand, T., F. Jeltsch, I. Hanski, and V. Grimm. 2003. Using pattern-oriented modeling for  
1153 revealing hidden information: a key for reconciling ecological theory and application. *Oikos*  
1154 100.
- 1155 Wilensky, U. 1999. NetLogo. Center for Connected Learning and Computer-Based Modeling,  
1156 Northwestern University, Evanston, IL.
- 1157 Wingfield, M. J., E. G. Brockerhoff, B. D. Wingfield, and B. Slippers. 2015. Planted forest health:  
1158 The need for a global strategy. *Science* 349:832–836.
- 1159 Zanetti, R., C. S. Abreu, S. H. P. Silveira, and E. D. Andrade. 2017. First report of *Hypsipyla*  
1160 *grandella* (Lepidoptera: Pyralidae) on African mahogany *Khaya ivorensis*. *Scientia Agricola*  
1161 74:492–494.
- 1162 Zhang, Q.-H., and M. R. Paiva. 1998. Female calling behaviour and male response to the sex  
1163 pheromone in *Thaumetopoea pityocampa* (Den. & Schiff.) (Lep., Thaumetopoeidae). *Journal of*  
1164 *Applied Entomology* 122:353–360.
- 1165 Zribi, M., T. Paris Anguela, B. Duchemin, Z. Lili, W. Wagner, S. Hasenauer, and A. Chehbouni.  
1166 2010. Relationship between soil moisture and vegetation in the Kairouan plain region of  
1167 Tunisia using low spatial resolution satellite data. *Water Resources Research* 46:2221.



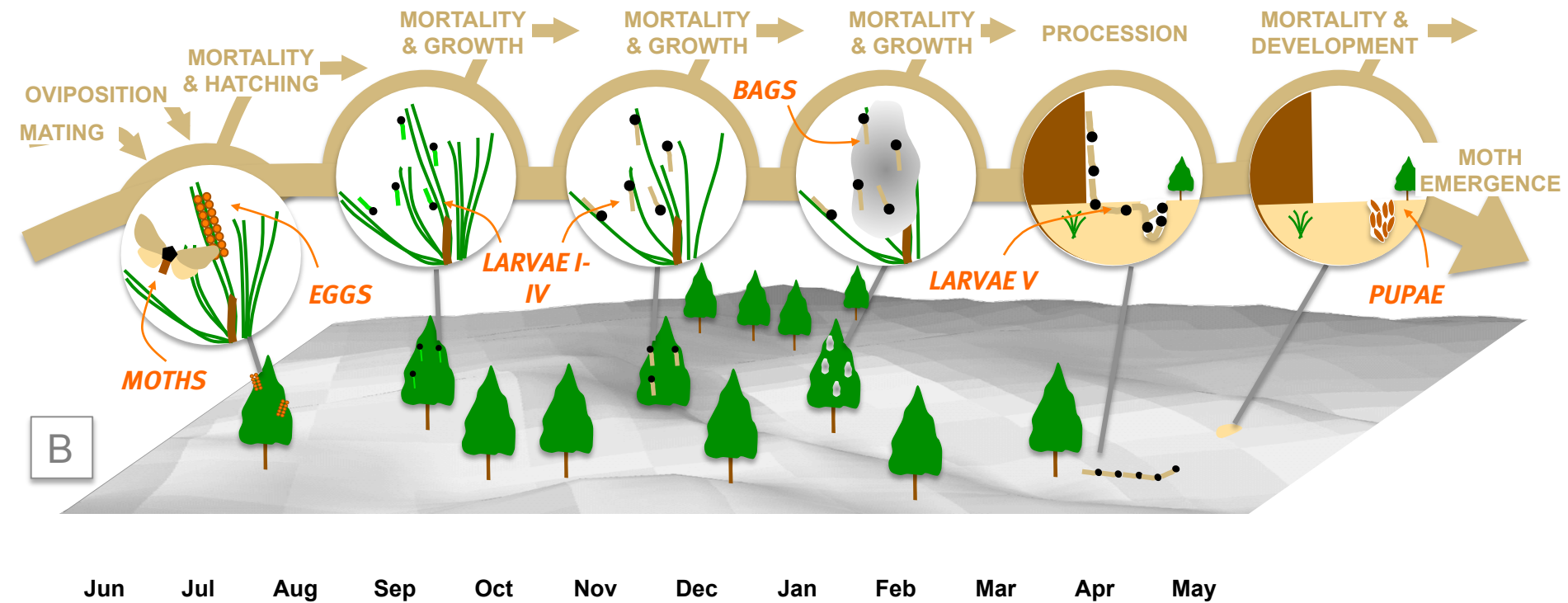
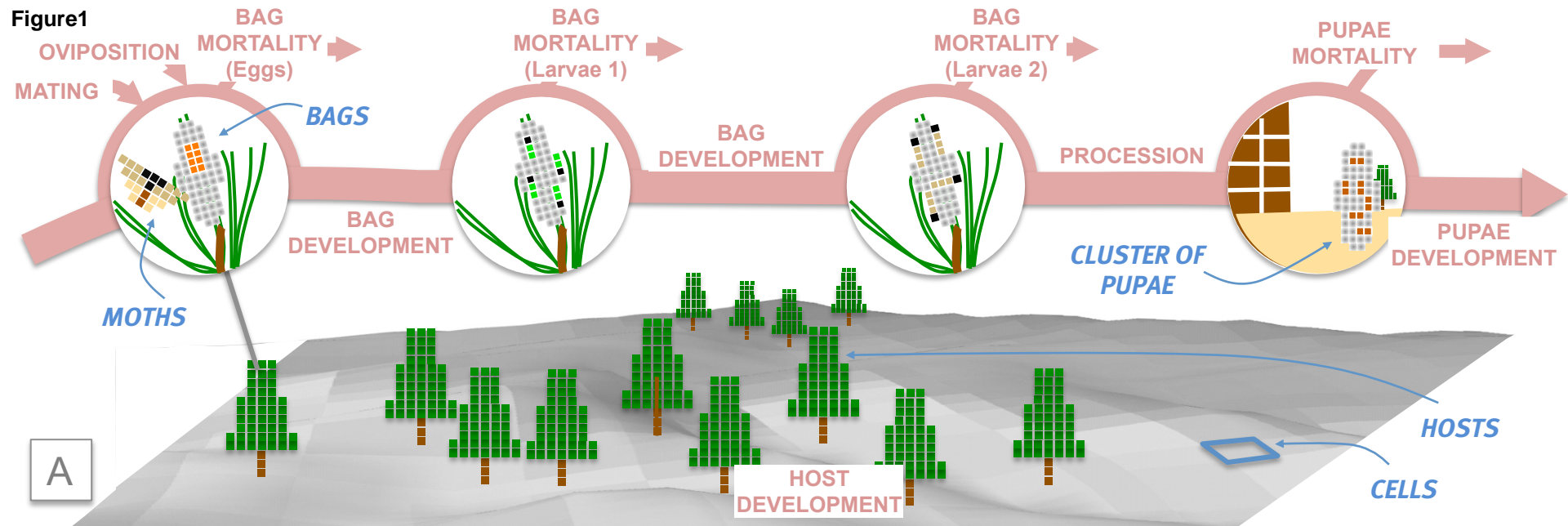


Figure2

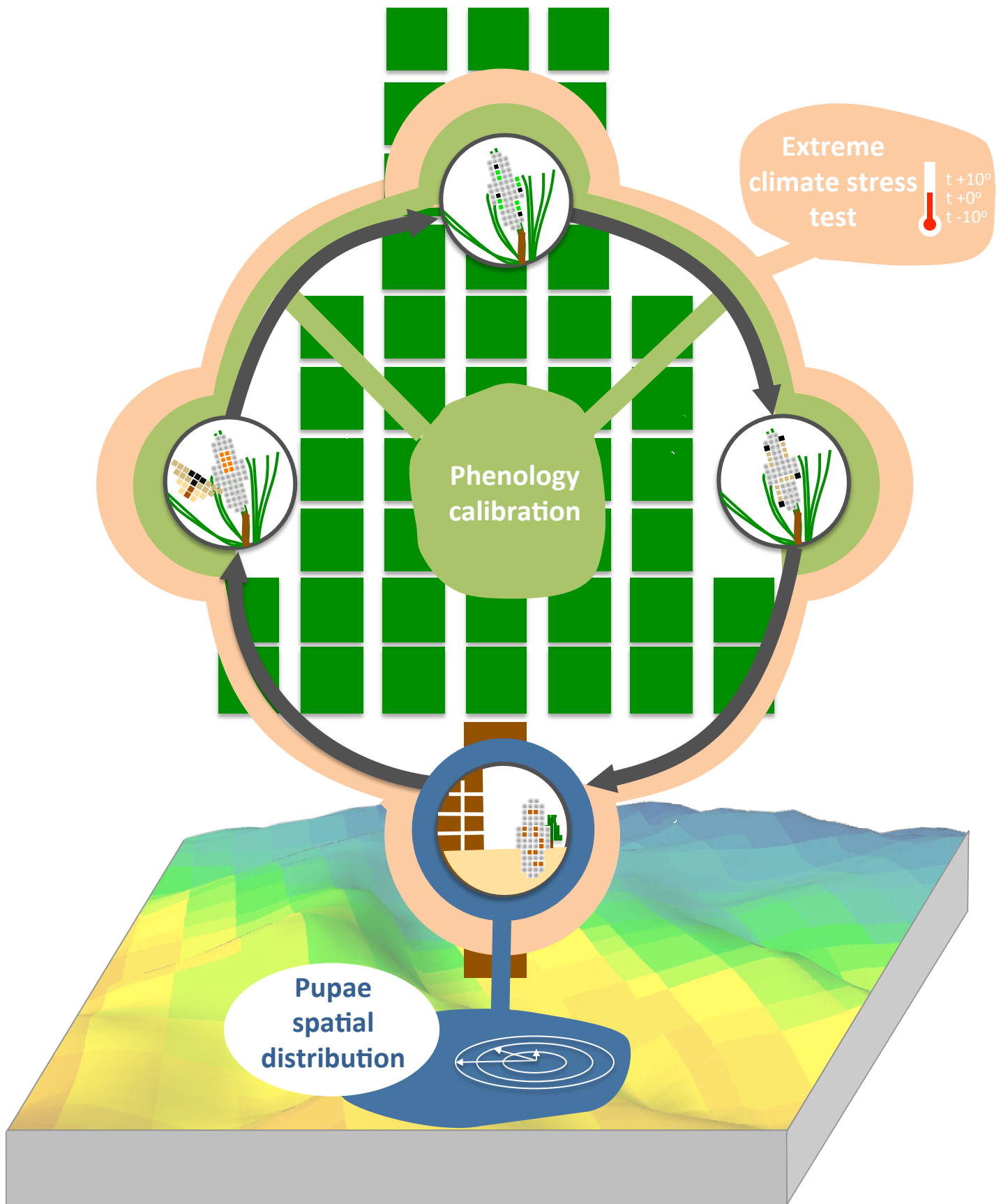


Figure3  
[Click here to download high resolution image](#)

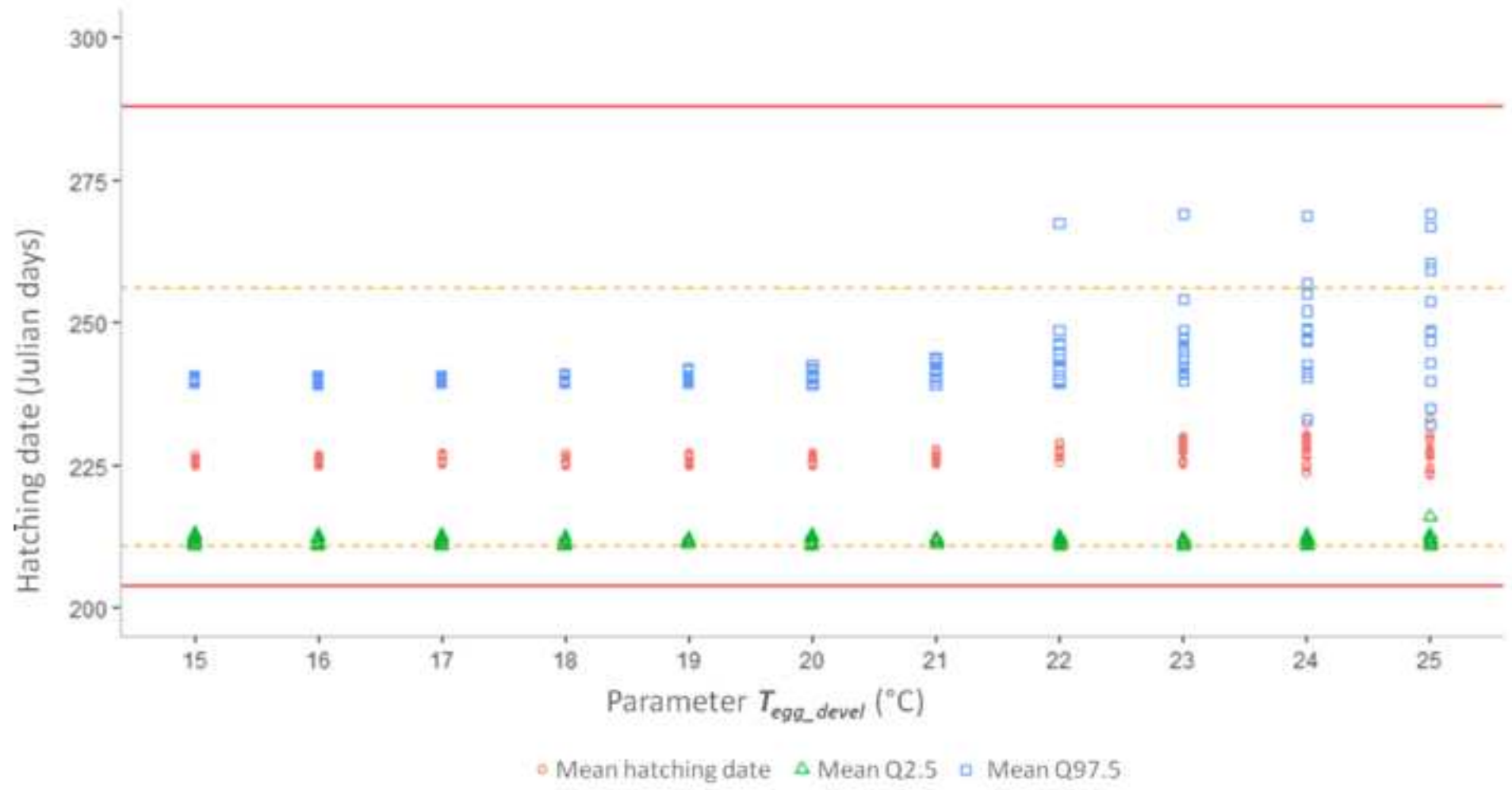


Figure4A  
[Click here to download high resolution image](#)

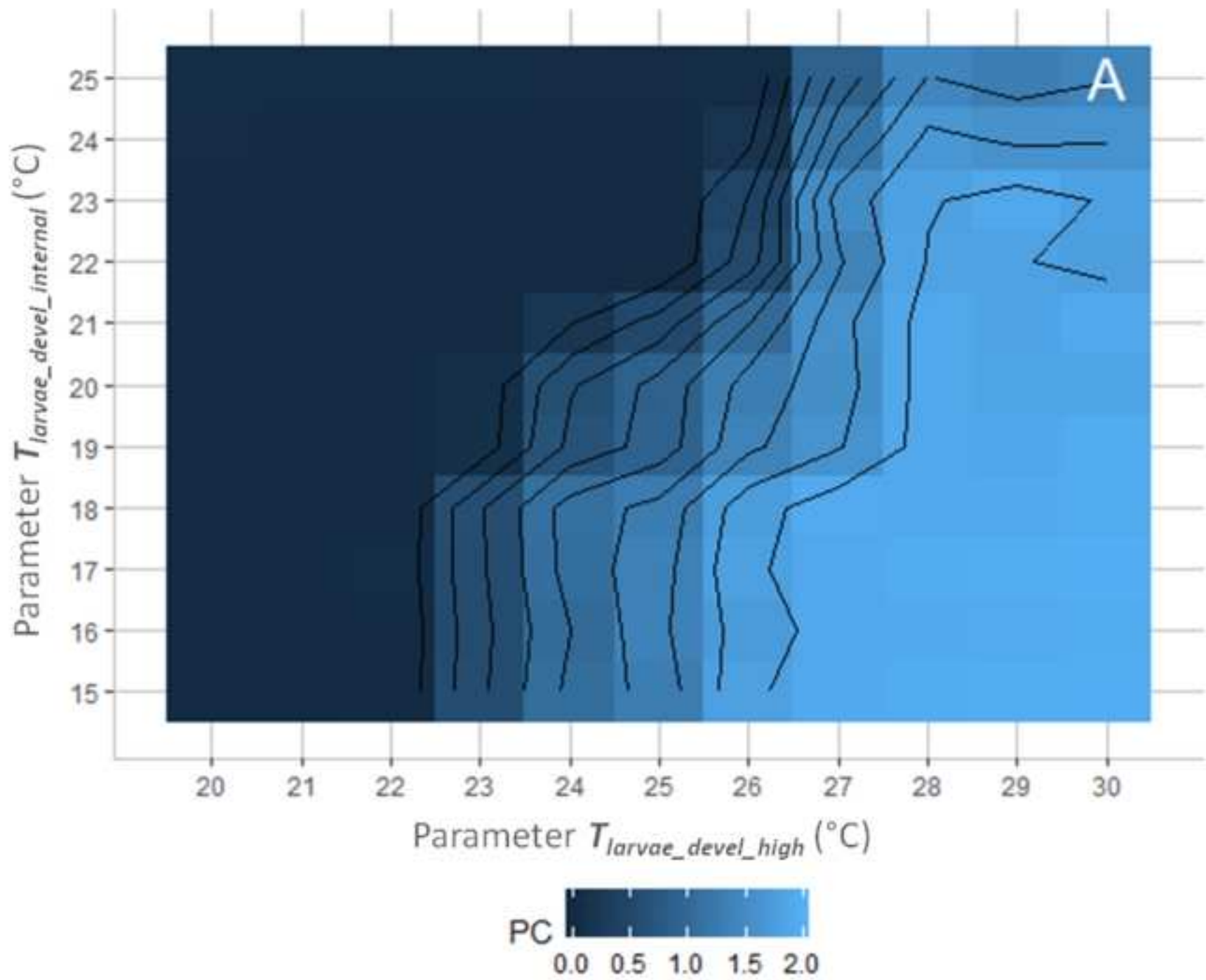


Figure4B  
[Click here to download high resolution image](#)

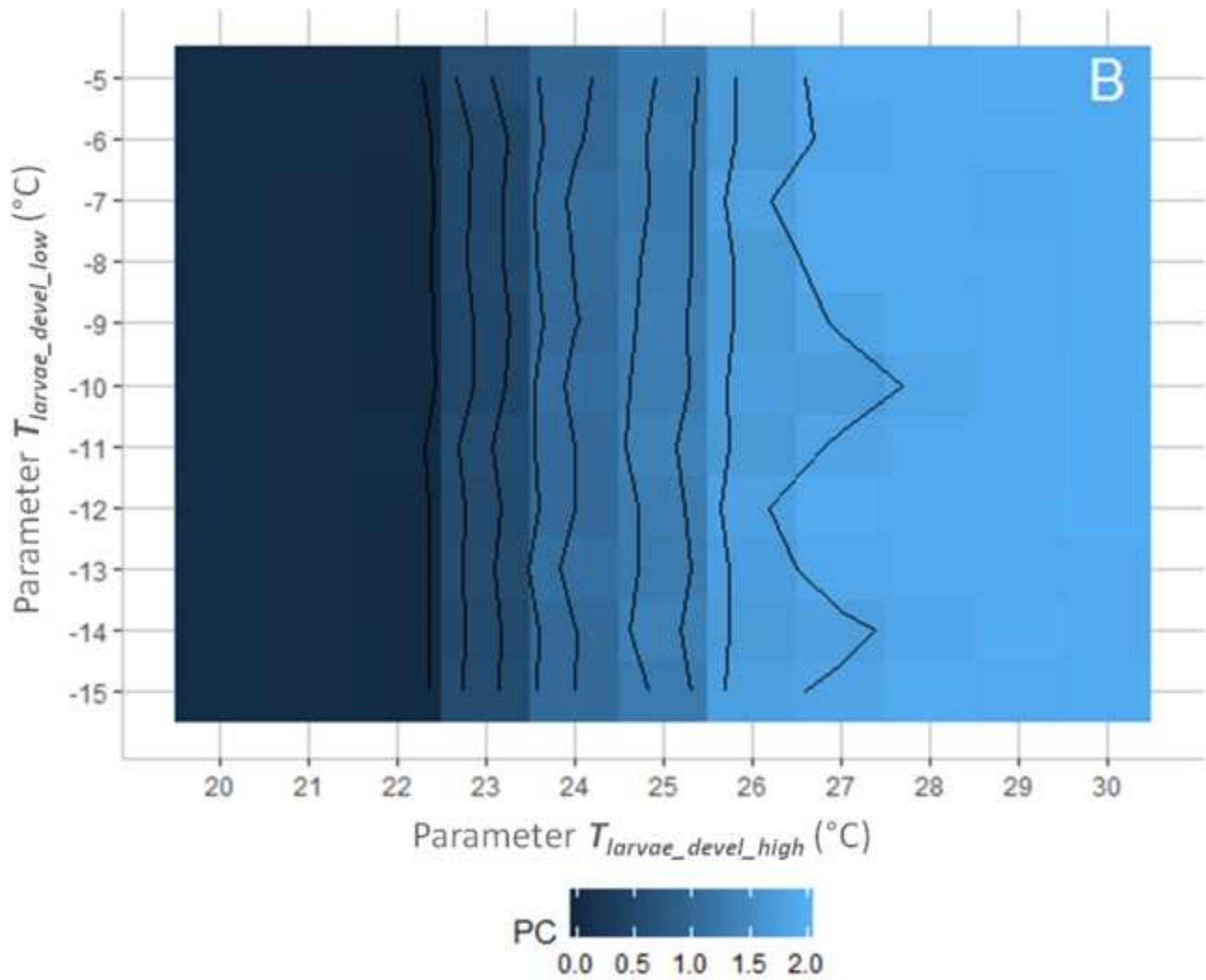


Figure4C

[Click here to download high resolution image](#)

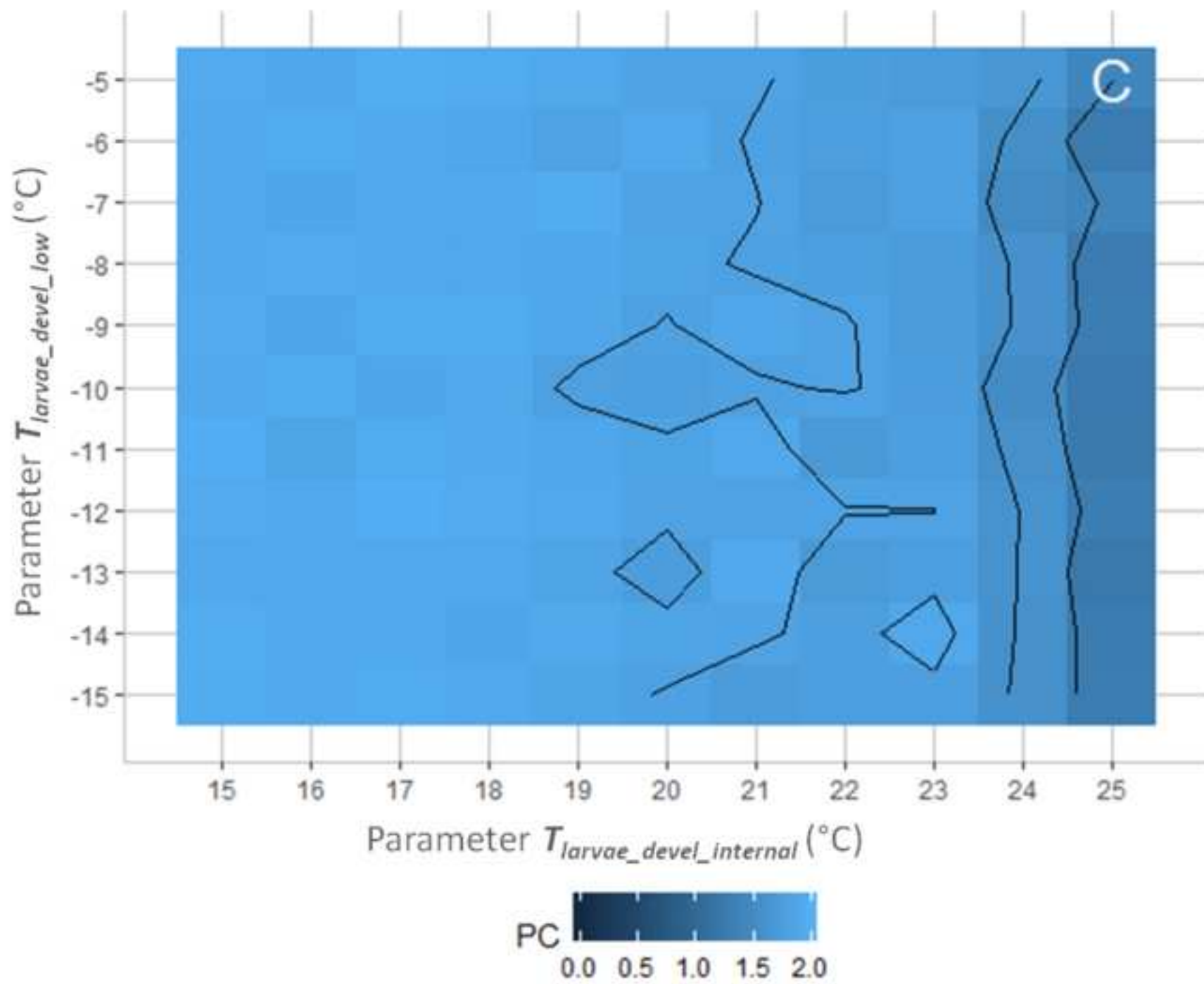


Figure5  
[Click here to download high resolution image](#)

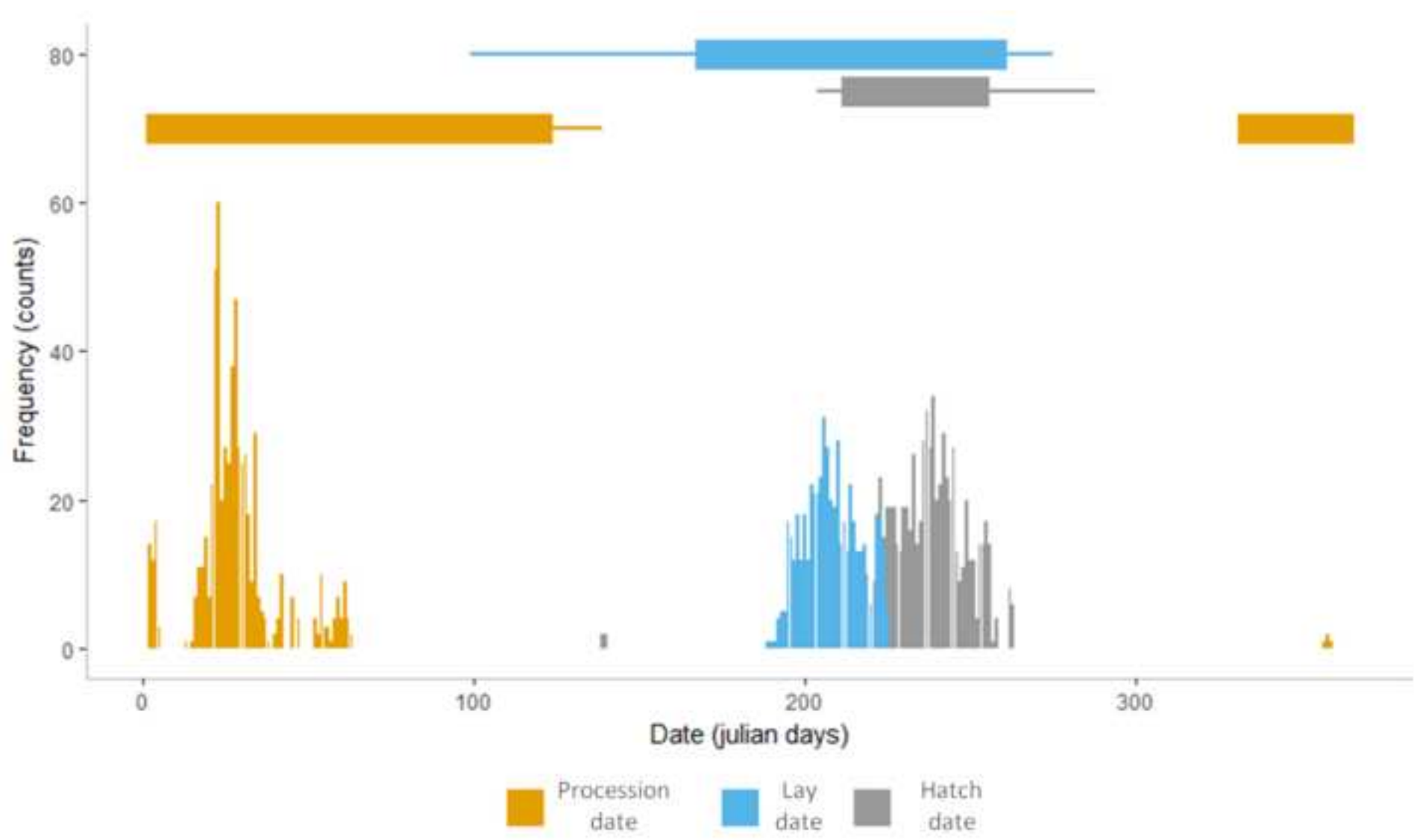


Figure6  
[Click here to download high resolution image](#)

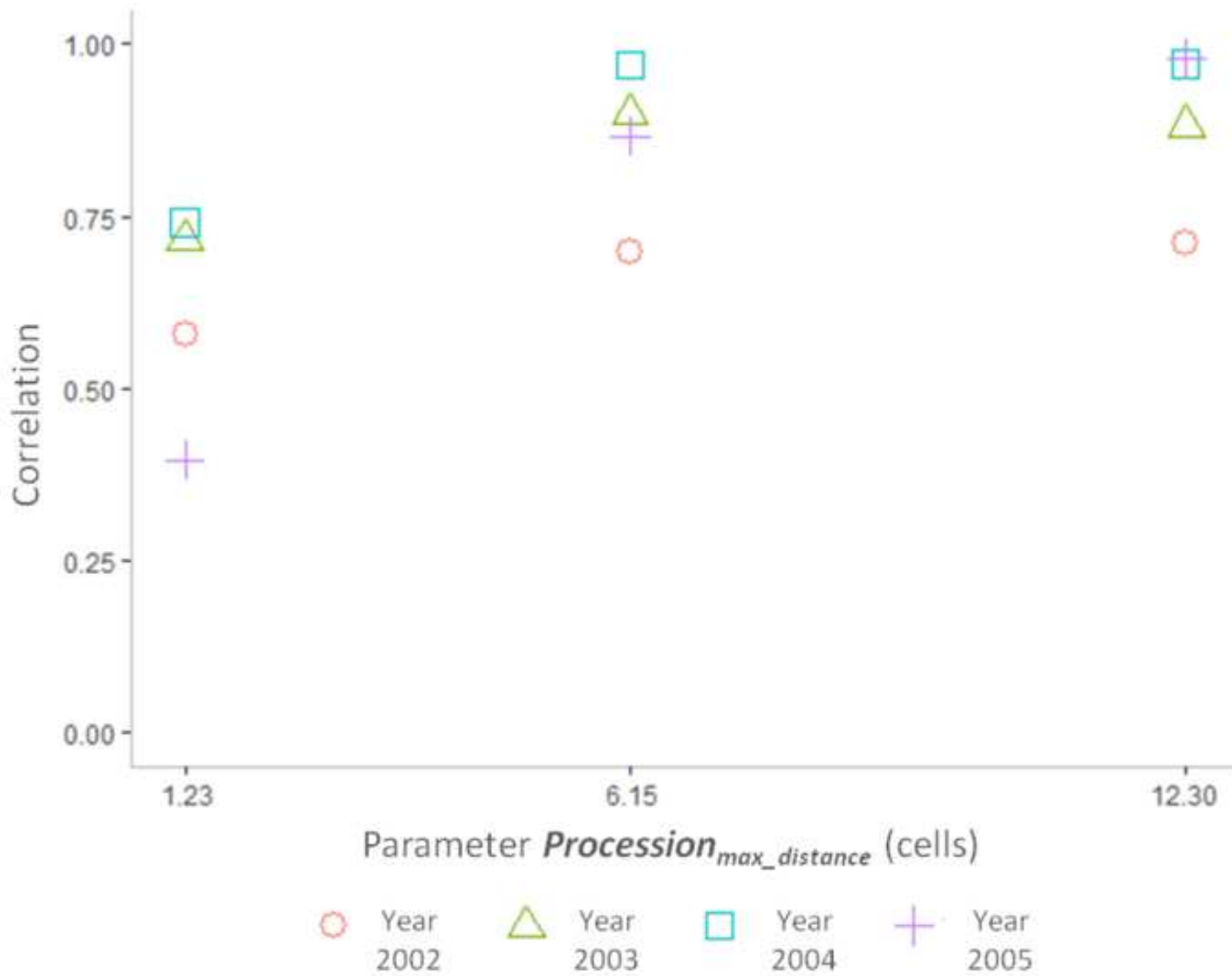




Figure7AB  
[Click here to download high resolution image](#)

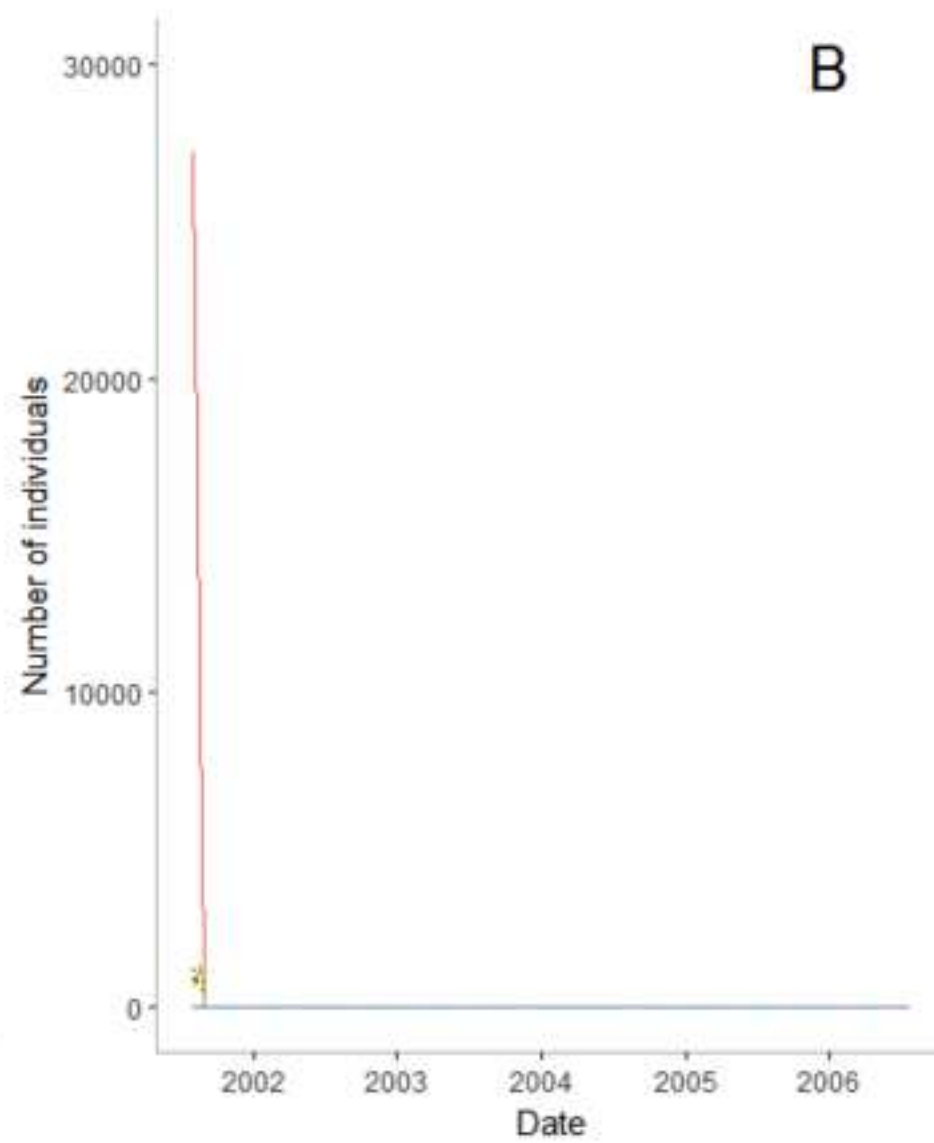
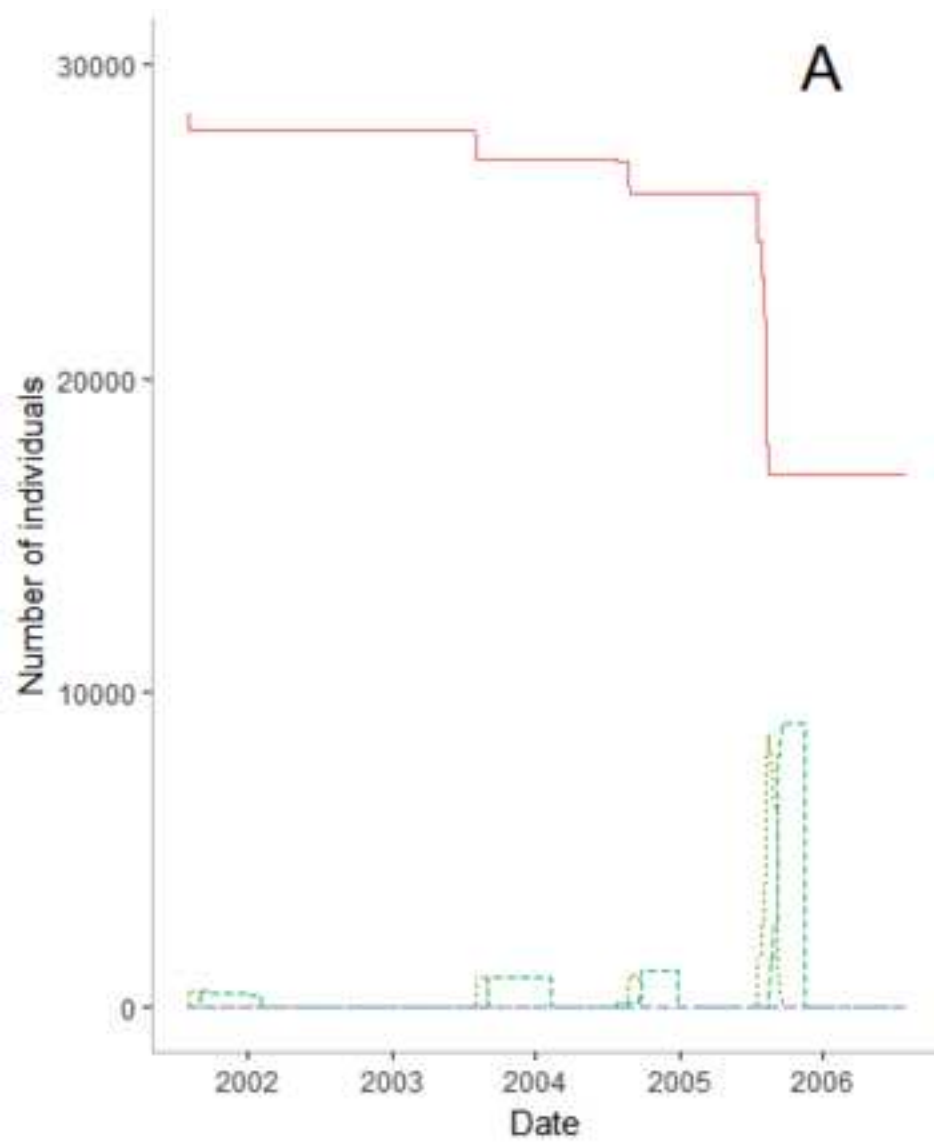


Figure7C  
[Click here to download high resolution image](#)

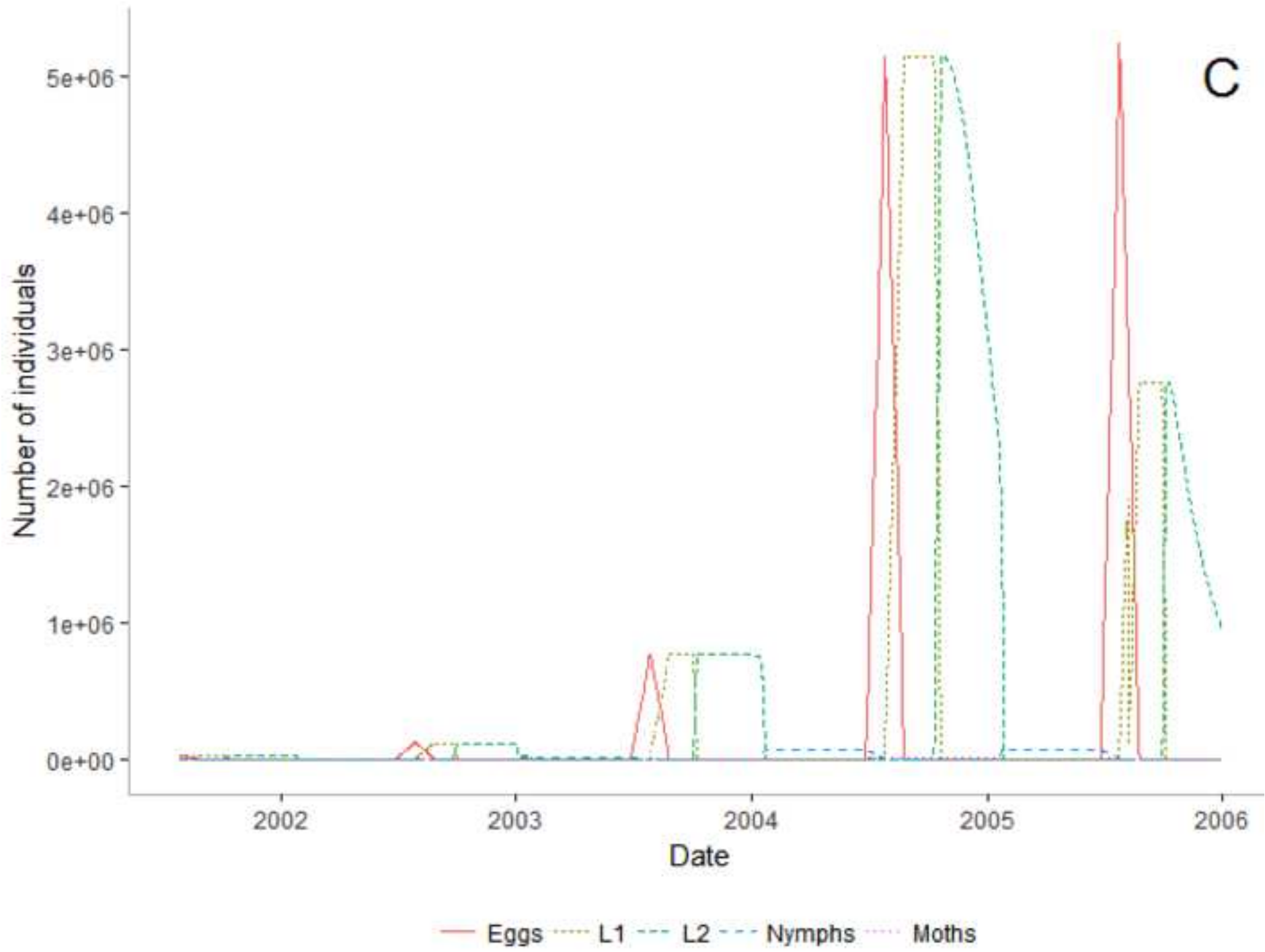


Figure8

[Click here to download high resolution image](#)

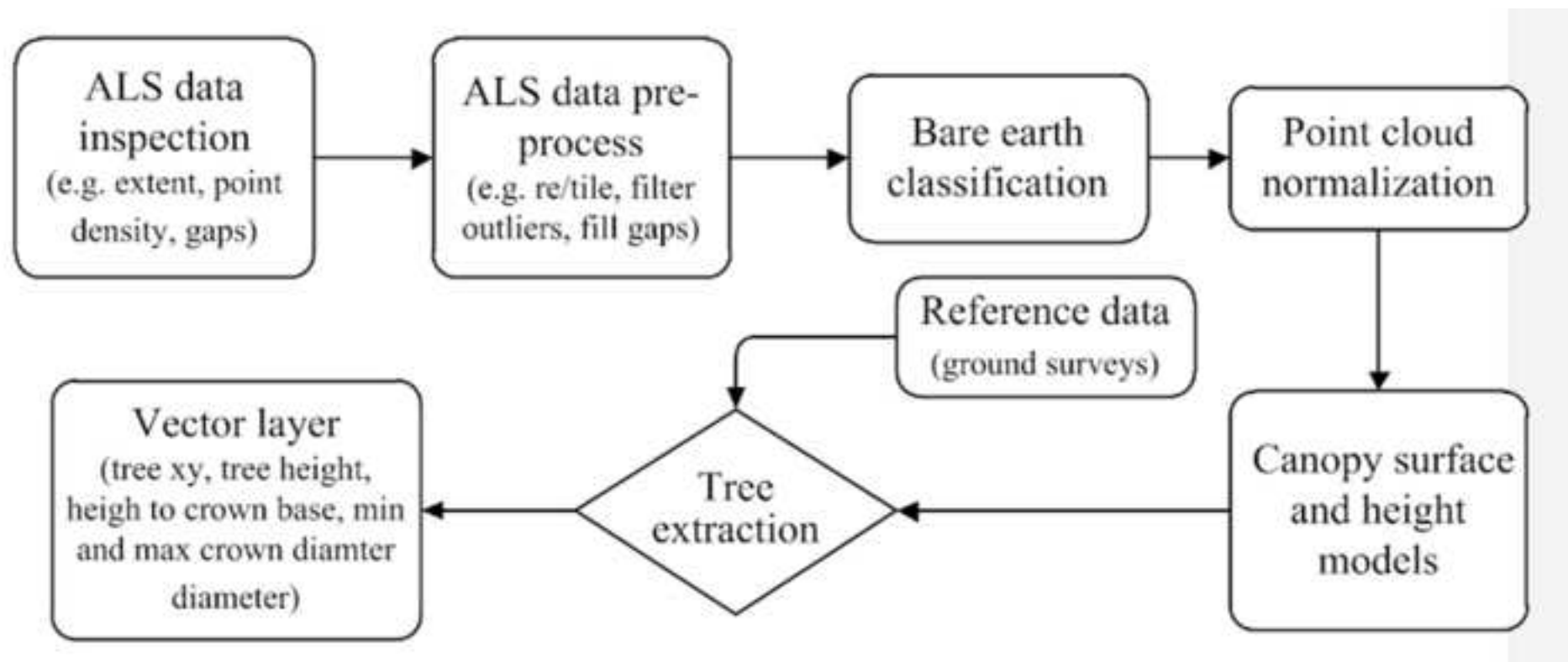
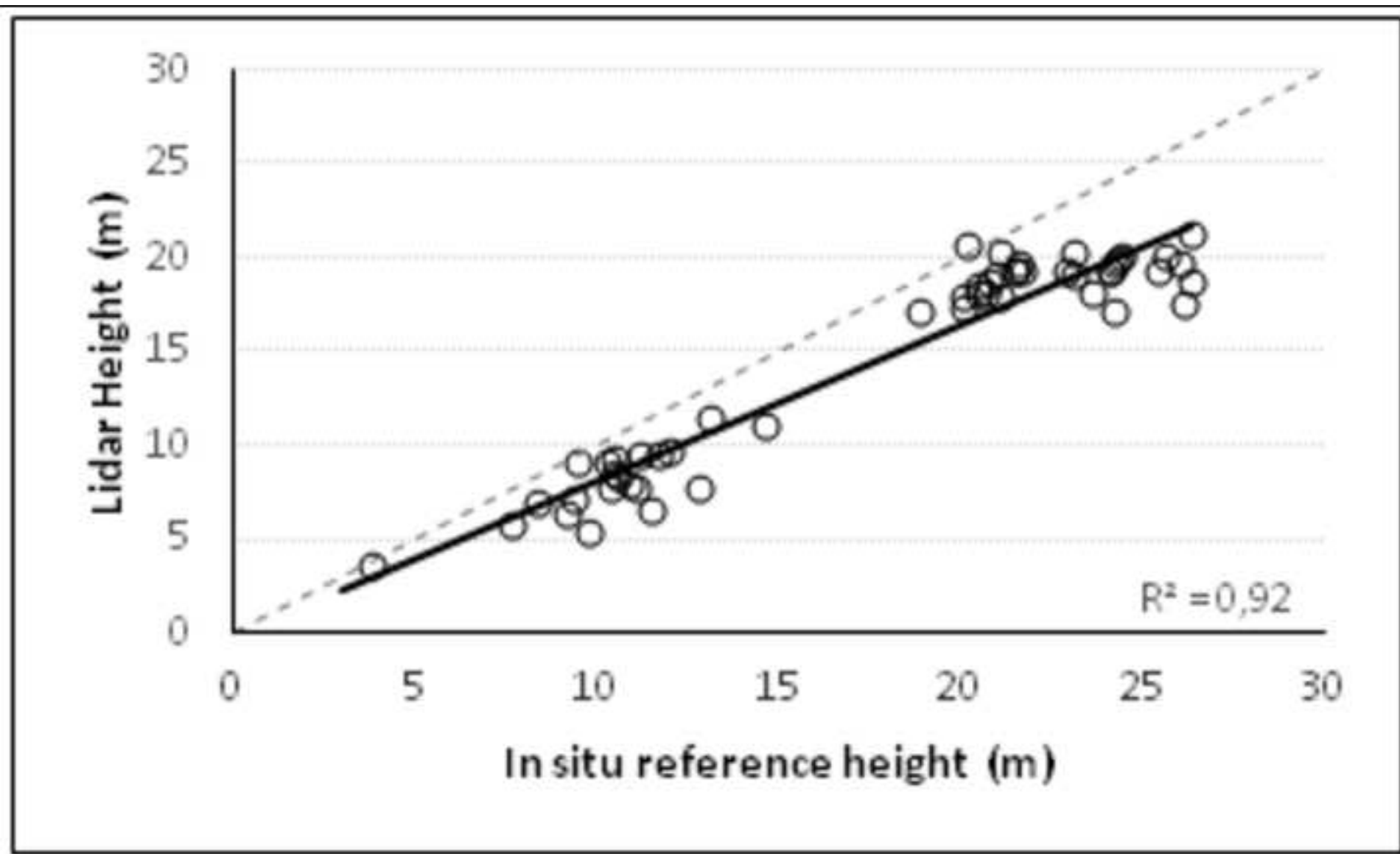
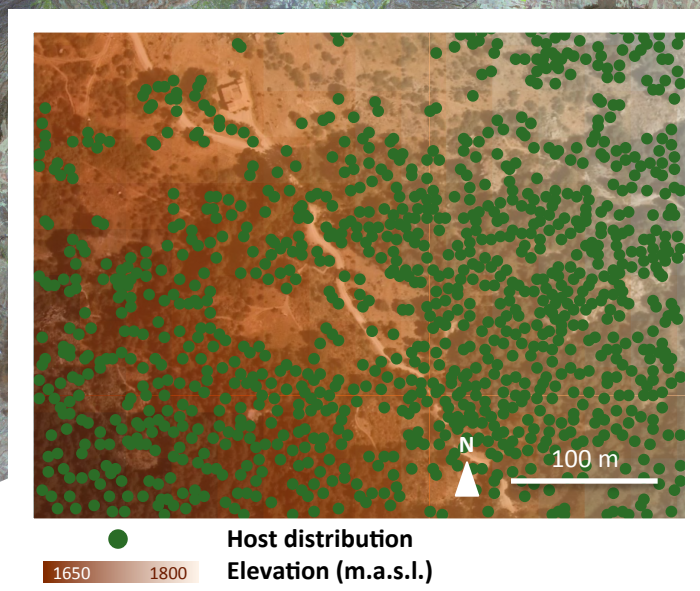
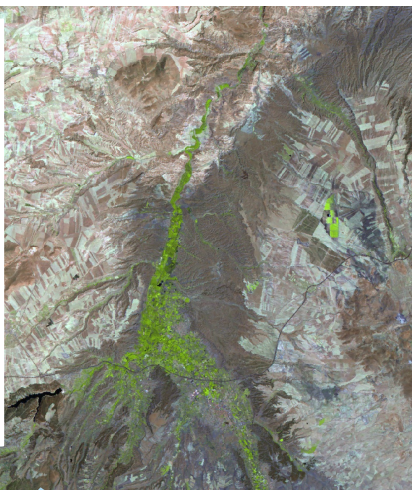


Figure9  
[Click here to download high resolution image](#)





Location of the virtual  
experimental area

Location of the data  
used to calibrate the  
carrying capacity of *P.  
halepensis* stands

Sierra Nevada  
protected area

N  
20 Km

appendices\_reviewed

[Click here to download Supplementary material for online publication only: ms\\_ODD\\_INSTAR\\_Appendices\\_reviewed.docx](#)