

Pest Management

Degree-day-based model to predict egg hatching of *Philaenus spumarius* (Hemiptera: Aphrophoridae), the main vector of *Xylella fastidiosa* in Europe

Clara Lago,^{1,2,✉} Àlex Giménez-Romero,^{3,✉} Marina Morente,^{1,✉} Manuel A. Matías,^{3,✉} Aránzazu Moreno,^{1,✉} Alberto Fereres^{1,*}

¹Instituto de Ciencias Agrarias (ICA-CSIC), Serrano 115b, 28006, Madrid, Spain, ²Departamento de Producción Agraria, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas (ETSIAAB), Universidad Politécnica de Madrid (UPM), Avenida Puerta de Hierro, 2,4, 28040, Madrid, Spain, ³Instituto de Física Interdisciplinar y Sistemas Complejos (IFISC, CSIC-UIB), Campus UIB, 07122, Palma de Mallorca, Spain *Corresponding author, mail: afereres@ica.csic.es

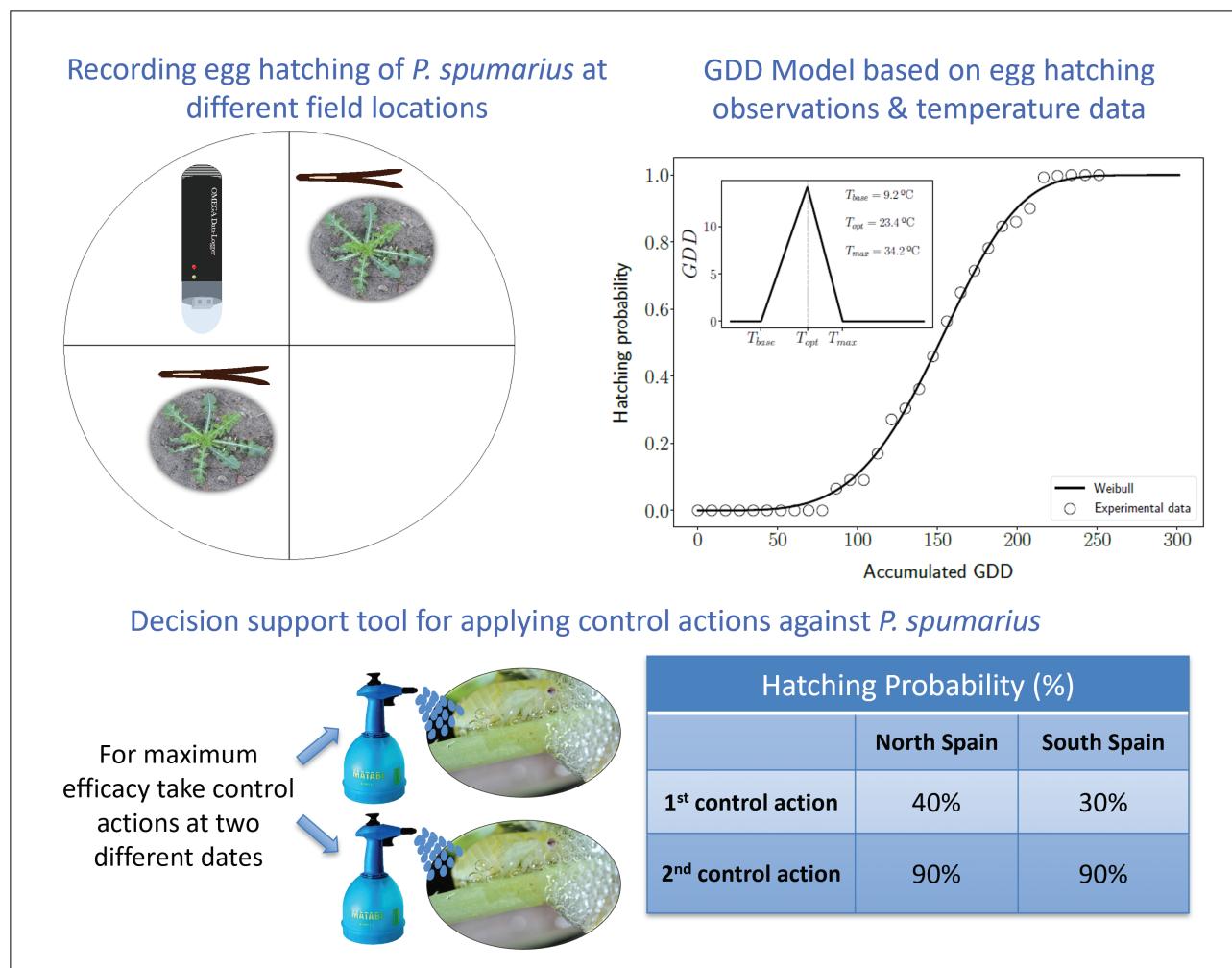
Subject Editor: Angel Acebes-Doria

Received on 28 June 2022; revised on 19 January 2023; accepted on 2 February 2023.

Philaenus spumarius L., the main vector of *Xylella fastidiosa* (Wells) in Europe, is a univoltine species that overwinters in the egg stage, and its nymphs emerge in late winter or spring. Predicting the time of egg hatching is essential for determining the precise times for deploying control strategies against insect pests. Here, we monitored *P. spumarius* eggs from oviposition to egg hatching together with the daily temperatures and relative humidities at four field locations that were located at different altitudes in central Spain. The collected data were used to build a growing degree day (GDD) model to forecast egg hatching in the Iberian Peninsula. Furthermore, the model was validated with field observations that were conducted in Spain. The model was then used as a decision-support tool to calculate the optimum timing for applying control actions against *P. spumarius*. Our results suggest that controlling nymphs at two different dates would target the highest percentages of nymphal populations present in the field. Our model represents a first step for predicting the emergence of nymphs and adopting timely control actions against *P. spumarius*. These actions could limit disease spread in areas where *X. fastidiosa* is present.

Key words: modelling, control, vector-borne disease, integrated pest management

Graphical Abstract



Introduction

Predicting epidemics can be challenging, especially after the emergence of a new pathogen in a novel ecosystem. The introduction of *Xylella fastidiosa* (Wells, 1987) (Xanthomonadales: Xanthomonadaceae) in Europe and the Mediterranean basin represents a major threat for agriculture (EPPO 2019, EFSA 2022). This xylem limited plant pathogenic bacterium is transmitted by xylem-sap feeders and *Philaenus spumarius* L. (1758) (Hemiptera: Aphrophoridae) is the only epidemiologically relevant vector of *X. fastidiosa* in Europe (EFSA 2015, Cornara et al. 2019). Therefore, a detailed understanding of its ecology and phenology is essential for developing an accurate forecasting model for controlling the vector. *Philaenus spumarius* is a univoltine species, with a single generation per year, which has an ovarian parapause and a winter diapause in the egg stage (Morente et al. 2018b, Antonatos et al. 2019, Bodino et al. 2020). Eggs overwinter until hatching occurs in early spring. After egg hatching, the pre-imago (nymph) passes through five instars and nymphal development takes about 5–6 wk until they become adults (Weaver and King 1954, Whittaker 1965, Ossiannilsson 1981, Fielding et al. 1999). In the fall mature females lay masses of eggs on plant debris on the soil until they die during winter (Morente et al. 2018b, Dongiovanni et al. 2019, Antonatos et al. 2019).

The study of the developmental times of ectotherms as a function of temperature, in particular insects, has a long history

(Rebaudo and Rabhi 2018). Insects need specific accumulations of heat units (HU) to reach certain development stages, which are commonly defined by the growing degree days (GDD) (also referred to as degree-days, heat units, or thermal units) (Herms 2004). Basically, the GDD is a measure of heat accumulation over time-based on insect development rates at temperatures between the lower and upper limits. Usually, the function that describes the temperature response exhibits a unimodal form, with minimum and maximum temperatures below or above which no development occurs.

Several authors have established correlations between *P. spumarius* phenology and temperature. Chmiel and Wilson (1979) estimated the upper and lower thresholds for nymphal development of *P. spumarius* in America by using the lowest coefficient of variation (CV) method described by Arnold (1959) and they calculated the HU by using the method described by Sevacherian et al. (1977). Similarly, Zajac et al. (1989) fitted linear regression functions to study the nymphal development of *P. spumarius*. One of the critical components of using GDD models is the determination of the starting points for degree-day accumulations (Kim et al. 2020). Ideally, it should be set up when insect development begins. The starting point of egg development of *P. spumarius* is assumed to happen after a winter diapause. In the GDD model introduced by Chmiel and Wilson (1979), they arbitrarily set the GDD accumulation starting

point at the 1st of January. Nevertheless, the precise date when egg development starts is totally unknown.

Forty years later, Bodino et al. (2019) calculated the days of development (DD) of the nymphal stages in the Apulia and Liguria regions of Italy, as a function of the number of hours in one year that was above a minimum temperature (8°C). Minimum temperature was calculated from assays performed at fixed temperatures (unpublished data). Thereafter, Beal et al. (2021) adapted the formula from Bodino et al. (2019) to study *P. spumarius* phenology in north coastal California, obtaining similar results, although slightly prolonged development for *P. spumarius* was observed in California compared to Italy. Despite both studies shed light for better understanding the phenology of nymphal development in relation with temperature, a forecasting tool to predict egg hatching and determine the precise moment for applying control measurements is still lacking. Furthermore, in many studies, a simple linear approximation was used to compute the GDD metric and used only the minimum temperature threshold (Johnson et al. 1998, Campbell and Hanula 2007, Bodino et al. 2019, Beal et al. 2021). Alternatively, complex nonlinear approximations have been used (Logan et al. 1976, Lactin et al. 1995, Briere et al. 1999). However, Quinn (2017) concluded in his extensive critical review that for most studies conducted thus far, more complex functions performed poorly relative to simpler ones. Moreover, degree-day models are usually based on laboratory assays that aim to experimentally determine the minimum, optimum, and maximum development temperatures, and such assays typically involve fixed temperature-controlled experiments (e.g., Diaz et al. 2007).

In the present study, we developed a degree-day model to forecast the egg hatching of *P. spumarius* by following the logical steps for model construction (Overton 1977): i) Data collection: We obtained field data from independent experiments conducted on egg hatching of *P. spumarius* in their natural environments and on-site temperature recordings at specific locations. ii) Model construction: Our model was based on a multilinear temperature response function with the minimum, optimal and maximum temperatures. iii) Model calibration: We calibrated the model and obtained the minimum, optimal and maximum temperatures that control egg development by using an optimization procedure to determine the parameters that provided the best fit to the available experimental data. iv) Model validation: Our model was validated with systematic and independent survey data of newborn nymphs in different regions in Spain. v) Model extension: The model was extended to predict egg hatching throughout the entire territory. Moreover, our GDD model was used as a decision support tool to determine the best timing for applying control measures to manage vector populations.

Materials and Methods

Insects and Plants

A total of 100 adult individuals of *P. spumarius* were collected at 'Pinilla del Valle' (Madrid, Centre Spain) (coordinates: 40.611108, -4.263008) during spring-summer, 2020, and were maintained inside bug-dorm cages ($0.5 \times 0.5 \times 0.5$ m) on *Sonchus oleraceus* L. plants (4–5 leaf stage) in a net house with no environmental controls (Temperature °C: MEAN + SE = 16 + 0.3, Max: 31.7, Min: 7.4; Humidity RH%: MEAN + SE = 63.5 + 1.3, Max: 99.6; Min: 20.0) at the ICA-CSIC facilities (Madrid, Spain). The proportion of sexes was one female per two males to ensure mating (Morente et al. 2018a, 2021). To obtain egg masses for the assays, which were conducted in October 2020, dry pine needles were placed on the substrate below the *S. oleraceus* plants to facilitate oviposition (Morente

et al. 2018a). The pine needles were checked once per week from 6-X-2020 to 4-XI-2020 to identify the egg masses (Fig. 1A, B). The dates when egg masses were observed on the pine needles were recorded. A total of 262 egg masses from five different oviposition dates were used in the field assay. The oviposition periods were the days preceding i) 8-X-2020, ii) 14-X-2020, iii) 22-X-2020, iv) 29-X-2020, and v) 4-XI-2020.

Data Collection: Monitoring Egg Hatching of *P. spumarius* Under Field Conditions

The egg masses from each oviposition date were divided into four equal parts, and each part was transferred to a different field location on the day after the eggs were collected. The four locations used for our field study consisted of natural environments located in central Spain. These sites were selected at different altitudes to provide a gradient in climate conditions: i) Alcalá de Henares (588 m) (coordinates: 40.521133, -3.290865); ii) Bustarviejo (1,222 m) (coordinates: 40.691827, -3.767162); iii) Mataelpino (1,086 m) (coordinates: 40.736439, -3.944309); and iv) Pedrezuela (880 m) (coordinates: 40.761003, -3.619030). At each field location, the egg masses were kept on an open 5.5-cm-diameter petri dish with small holes below to facilitate drainage. The petri dishes with egg masses were placed inside floorless mesh cages (50 cm height and 45 cm diameter) (Fyllen cloth basket 79 l, Inter IKEA Systems, Sweden). Two cages were placed at each field location: The cages were divided into four equal sectors by using cellulose acetate sheets. Since we had five oviposition dates, two and three plants were placed inside each of the two cages, respectively (Supp. Fig. S1). Egg masses from a given oviposition date were placed in a single plant and sector of the cage to avoid mixing nymphs after emergence. A few weeks before the first nymphs were expected to emerge in the field (e.g., 25-II-2021), one *S. oleraceus* plant was transplanted inside each cage division

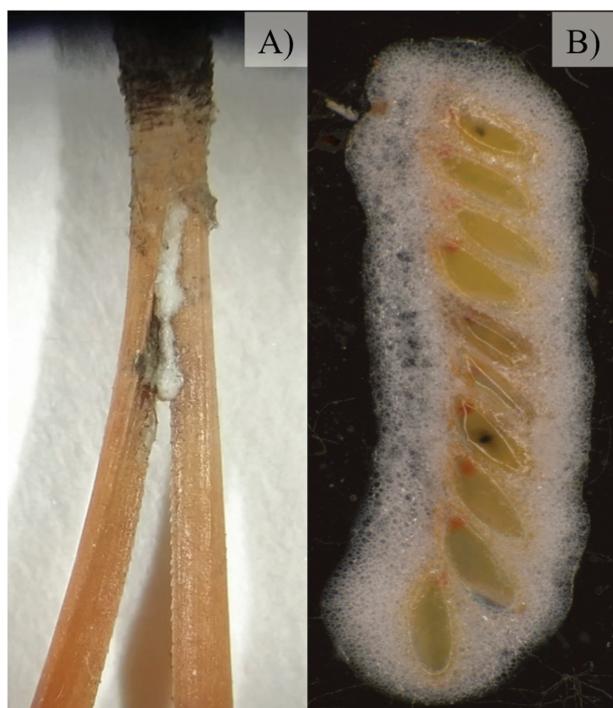


Fig. 1. A) *Philaenus spumarius* newly laid egg masses on a dry pine needle. B) *Philaenus spumarius* egg mass.

where the egg masses were placed to feed the newborn nymphs after emergence (Supp. Fig. S1). After transplanting, the egg masses were transferred from the petri dishes to the bare soil below the plants. In Mataelpino, the potted *S. oleraceus* plants inside the cages were watered weekly as opposed to the rainfed field locations. At this time, we checked each plant once a week to record egg hatching when newborn nymphs were observed. The newborn emerged nymphs were removed from the plants to facilitate further observations. All plants were inspected until no newborn nymphs were observed for two consecutive weeks.

Temperature and relative humidity (RH) were monitored hourly at each field location with temperature and humidity data loggers (OM-EL-USB-2, Omega Engineering, INC, Norwalk, Connecticut, USA) during the entire duration of the experiments. One data logger not exposed to direct sunlight was placed inside each cage at each field location to obtain the same temperature and RH data as that experienced by the egg masses. The data collected from the data loggers were downloaded every two weeks until the emergence of newborn nymphs. Thereafter, data were downloaded once a week. In the few cases where the data loggers failed to record weather data due to extreme weather conditions, the missing data were supplemented with external data collected from the nearest official meteorological station (Agencia Estatal de Meteorología de España, AEMET, Spain). After monitoring egg-hatching in each of our four experimental sites, we computed RH as mean values experienced by each egg during the experimental period per field site. Furthermore, we calculated the Wasserstein distance to measure the distance between probability distributions of GDD accumulation between each pair of field sites (two by two) (Panaretos and Zemel 2019).

Model Construction: GDD-based Model

In the current study, we used a multilinear temperature response function with minimum or base (T_{base}), optimal (T_{opt}), and maximum (T_{max}) temperatures to calculate the hourly contributions to the GDD. The use of a multilinear function is based on the principles of biochemical kinetics (i.e., on Arrhenius' Law) (Eq. 1) and is explained in detail in Supp. Document S2.

$$f(T) = \begin{cases} 0 & \text{if } T < T_{base} \\ T - T_{base} & \text{if } T_{base} \leq T < T_{opt} \\ m \cdot T + n & \text{if } T_{opt} \leq T < T_{max} \\ 0 & \text{if } T \geq T_{max} \end{cases} \quad (1)$$

In our model, the accumulated GDD values are directly related to the cumulative probabilities of egg hatching using the cumulative density function of the Weibull distribution where $k > 0$ and $\lambda > 0$ are the shape and scale parameters of the Weibull distribution, respectively.

$$P_{hatching}(GDD) = 1 - e^{(-(GDD/\lambda)^k)} \quad (2)$$

The starting date for GDD accumulation was the date when the winter diapause of *P. spumarius* ended. This date, in principle, should be determined by some metric of accumulation of cold temperatures (below a given threshold). However, the precise hours or days needed under a given temperature -cold requirement- and the precise length of the winter diapause of *P. spumarius* are unknown. Thus, we compared different arbitrarily fixed dates of diapause ending (January 1st, December 1st, and November 1st), and later validated the model with data obtained from field observations of newborn nymphs in the field from systematic surveys at different sites in Spain over the last six years (Supp. Table S3). The modeling was performed in Python programming language (Van Rossum and Drake 1994).

Model Calibration

The goal of our model was to provide estimates of the cumulative hatching probabilities of *P. spumarius* eggs by using the temperature data provided by the field experiments described above transformed into a GDD metric. All data were treated independently of location, and then pooled together. The experimental cumulative hatching probability at a given time was calculated as the number of nymphs that had already emerged at that given time over the total number of nymphs that had emerged by the end of the assay. Then, the cumulative GDD values between the end of diapause and the egg hatching dates had to be calculated so that the cumulative hatching probabilities could be expressed as functions of the accumulated GDD values. In this way, we could fit Eq. (2) to the experimental data and obtain an effective model for predicting egg hatching. However, the relationship between the cumulative hatching probabilities and accumulated GDD values depends on how the accumulated GDD is calculated; that is, it depends on the GDD (t) profile, namely on the cardinal temperatures, Eq. (1), and the starting date of accumulation. Thus, we first calibrated our model to obtain the GDD temperature profile and a starting date of GDD accumulation that provided the best-predicted hatching probabilities for our dataset.

We selected three plausible starting dates for GDD accumulation as described above: 1st of November, 1st of December, and 1st of January. Then we calculated the optimal GDD temperature profiles, by finding the values of T_{base} , T_{opt} , and T_{max} that gave the best fit of Eq. (2), which depends on Eq. (1), to the observed hatching probabilities of our dataset (the four field experimental sites pooled together). More precisely, we tested all the combinations with T_{base} from 4 to 12°C, T_{opt} from 20 to 28°C, and T_{max} from 30 to 42°C in steps of 0.2°C, predicting the hatching probabilities using Eq. (2) by using the accumulated GDDs computed with each profile and each starting date using Eq. (1). Finally, we selected the optimal profile (the optimal values of T_{base} , T_{opt} , and T_{max}) to define the profile that minimized the error between Eq. (2) and the experimental dataset for each starting date. To fit Eq. (2) to the experimental data we used the nonlinear least squares method implemented in the 'scipy.optimize.curve_fit' method in Python programming language.

Model Validation

To validate our model, we predicted the egg-hatching probabilities in the Iberian Peninsula by using the existing dataset from field observations recorded from 2016 to 2021 in Spain (Supp. Table S3) using the calibrated model considering three different dates of diapause termination date: 1st of November, 1st of December, and 1st of January. For each diapause termination date considered (1st of November, 1st of December, and 1st of January), an optimal GDD temperature profile can be calculated from our dataset (the 4 experimental sites). The hourly temperature data for these years were retrieved from the ERA5-Land dataset, which has a spatial resolution of $0.1^\circ \times 0.1^\circ$. Then, we computed the hatching probabilities with daily resolution and compared the predictions with the observational data from the field that were obtained from systematic surveys of newborn nymphs in different regions of Spain during six consecutive years (2016–2021) (Supp. Table S3). These data were obtained from systematic field surveys conducted by different researchers from public institutions in Spain (Andalucía, La Rioja, Madrid, Murcia, and Valencia) (see Acknowledgments). These data consisted of several records of GPS coordinates and dates where newborn nymphs were found in Spain, so we expected that our model would predict high probabilities of egg hatching (>50%) for each date and location where newborn nymphs were observed. Therefore, we conducted a

validation process to find out which diapause termination date fitted best to our six-year dataset.

Control Timing of *P. spumarius* Nymphs

After model validation, we used our calibrated and validated model to investigate the best timing to adopt control actions against *P. spumarius* nymphs based on field observations. To do so, we calculated the daily hatching probabilities of *P. spumarius* eggs in the Iberian Peninsula based on ERA5-Land temperature data (Muñoz Sabater 2019) and simulated applications of control actions at several probability levels. Thereafter, we calculated the efficacy of applying a control action at a given time, which is defined as the percentage of nymphs targeted in the field before they reach the adulthood if a control action was taken at that given time. Thus, the following algorithm was developed to determine the timing of control actions based on our model predictions. First, we selected a plausible range of diapause termination dates as the starting dates for our simulations. Although we calibrated and validated our model for a particular diapause termination date, we investigated if different dates with the same temperature profile were able to provide better results; then, we selected the ending date of each simulation as a plausible date on which no more egg hatching was expected. Afterward, we selected several probability levels to apply control actions (i.e., when any of these selected levels is crossed in any location, we simulated the application of a control action). For each diapause termination date in each simulation, the daily hatching probabilities were computed until the end of the simulation using the hourly temperature data from ERA5-Land. Then, we saved the dates at which every selected probability level (control action thresholds) was crossed at each location in which we have data for the presence of newborn nymphs in the field. (Supp. Table S3). Finally, the lags between the control action dates given by each probability level and observation dates were computed.

Each selected hatching probability (which matches an action date) has a given control action efficacy. It is known that nymphal development lasts for approximately 5–6 wk until they become adults, but the developmental rates decrease when temperatures are low and could extend to at least 100 days (Weaver and King 1954, Yurtsever 2000, Bodino et al. 2019). Thus, we followed a conservative approach to compute the control action efficacies and defined successful nymph control only if the action was applied after the newborn nymphs were observed and there was a lag of less than 30 days.

According to our results, there is high variability in the hatching dates (see the results section). Thus, to overcome this intrinsic variability in the timing of egg hatching of *P. spumarius*, we developed a strategy based on deploying control actions at two different dates to target the maximum number of nymphs as completely as possible. A list of combinations of two probability levels was selected (e.g., the first at 50% and the second at 90%) to apply control actions on two different dates. Moreover, the efficacies of applying control actions at a single date or at two different dates were compared.

In addition, we developed an R package script to calculate the daily cumulative hatching probabilities of *P. spumarius* eggs based on suitable temperature data. This package also calculates the optimum time to apply the first and second control actions, depending on the hatching probability achieved. This package and all relevant information regarding its use can be found on the GitHub repository: <https://github.com/agimenezromero/PSEggHatching>.

Results

Nymphal Emergence

A total of 435 *P. spumarius* nymphs emerged in the field assays at the four experimental sites. Newborn nymphs were detected feeding on *S. oleraceus* plants from 5-III-2021 to 1-V-2021. The first detection of newborn nymphs was recorded on 5-III-2021 at Alcalá de Henares (588 m), on 18-III-2021 at Bustarviejo (1,222 m), on 23-III-2021 at Pedrezuela (880 m) and 26-III-2021 at Mataelpino (1,086 m). Nymph emergences occurred at different moments and extended for approximately two months at all of the studied sites (Fig. 2). The results for the nymphal emergence dates per field location show that the intrinsic randomness of the hatching process is quite broad, which exhibits variations of up to one month for eggs subjected to the same environmental conditions. No correlation between the oviposition date and hatching date was observed after pooling data from all of the field sites ($\rho = -0.036$). Furthermore, there was no or poor correlation between the oviposition date and hatching date within specific field sites: Alcalá Henares ($\rho = -0.192$), Bustarviejo ($\rho = -0.289$), Mataelpino ($\rho = -0.016$), and Pedrezuela ($\rho = -0.316$).

Model Calibration

We calibrated the GDD temperature profile assuming different diapause termination dates: 1st of November, 1st of December, and 1st of January. We obtained an unrealistic temperature profile with high T_{opt} and T_{max} values when starting the accumulation of GDD on the 1st of January ($T_{base} = 9.2^\circ\text{C}$, $T_{opt} = 27.6^\circ\text{C}$, and $T_{max} = 41.8^\circ\text{C}$) and 1st of November ($T_{base} = 9^\circ\text{C}$, $T_{opt} = 27.4^\circ\text{C}$, and $T_{max} = 41.2^\circ\text{C}$). In contrast, more appropriate and realistic results were obtained when selecting the 1st of December as the diapause termination date ($T_{base} = 9.2^\circ\text{C}$, $T_{opt} = 23.4^\circ\text{C}$, and $T_{max} = 34.2^\circ\text{C}$).

As a visual example, Fig. 3 shows the final result of the model after the calibration procedure for the diapause termination date for 1st December (model calibration for diapause termination date for 1st of January and 1st of November are shown in Supp. Document S4). The dots represent the cumulative hatching probabilities that were retrieved from the dataset as a function of the accumulated GDD values. This accumulated GDD value was obtained using the optimal GDD profile that is shown in the inset ($T_{base} = 9.2^\circ\text{C}$, $T_{opt} = 23.4^\circ\text{C}$, and $T_{max} = 34.2^\circ\text{C}$). The black line is the best fit of Eq. (1) to the experimental data (dots) that was obtained with $k = 4.34$ and $\lambda = 164.86$, which had a relative error of 1%.

Model Validation and Predictions of Egg Hatching Dates in the Iberian Peninsula

It was found that our model predictions were only consistent with the field data when the diapause termination date was set to the 1st of December (Supp. Video S5 <https://zenodo.org/record/7465753>). When assuming the 1st of January, the model predicts egg hatching too late so that nymphs are observed about a month before our model predicts high hatching probabilities (Supp. Video S5 <https://zenodo.org/record/7465753>). On the other hand, when assuming November 1st, the model predicts a high probability of egg hatching too early, about 2 months before field observations (Supp. Video S5 <https://zenodo.org/record/7465753>). Nevertheless, considering the diapause termination as the 1st of December, some differences were also observed between model predictions and field observations from systematic surveys (Supp. Table S3) in southern Spain (latitudes below 40°) and northern Spain (latitudes above 40°). As shown in Supp. Video S5, in the South, 1st instar nymphs were detected in the field, when the model predicted high hatching probabilities (above 80%), while in the North 1st instar nymphs are detected when the model predicted low hatching probabilities (20–30%).

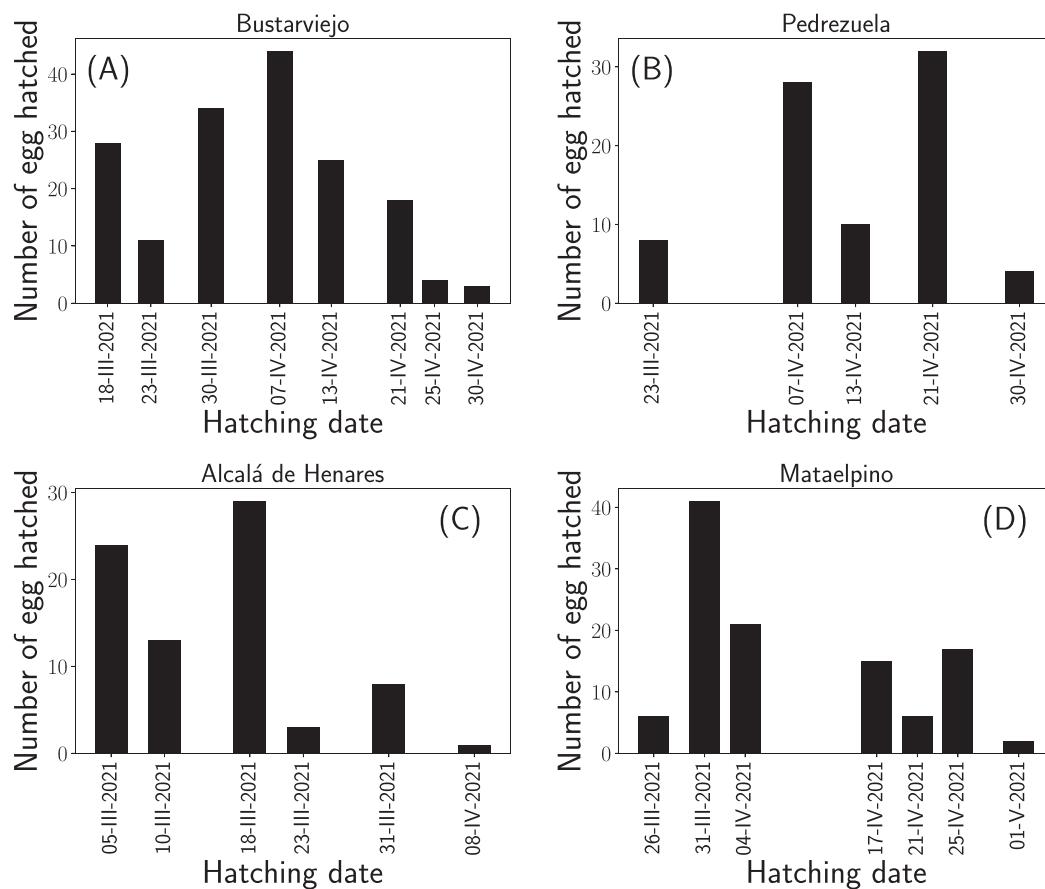


Fig. 2. Number of eggs hatched per date at each field site. The five oviposition dates were merged per each hatching date. Scouting weeks: 1 = 22-II-2021 to 28-II-2021; 2 = 01-III-2021 to 07-III-2021; 3 = 08-III-2021 to 14-III-2021; 4 = 15-III-2021 to 21-III-2021; 5 = 22-III-2021 to 28-III-2021; 6 = 29-IV-2021 to 04-IV-2021; 7 = 05-IV-2021 to 11-IV-2021; 8 = 12-IV-2021 to 18-IV-2021; 9 = 19-IV-2021 to 25-IV-2021; 26-IV-2021 to 02-V-2021. A) Bustarviejo. B) Pedrezaula. C) Alcalá Henares. D) Mataelpino.

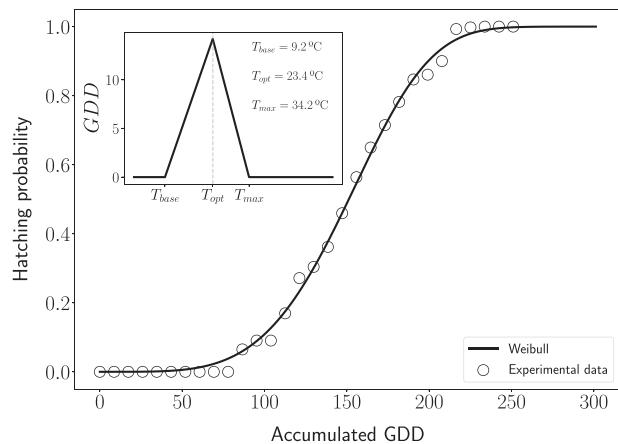


Fig. 3. Model calibration considering diapause breakage on December 1st. The main figure shows the hatching probability as a function of the accumulated GDD value. Dots represent the experimental data, and the black solid line shows the best fit using Eq. (1). The inset shows the GDD profile that is used to calculate the accumulated GDD values, which yielded the best fit between Eq. (1) and experimental data.

Impact of RH on GDD for Hatching and GDD Accumulation in the Field Sites

After model calibration and validation, considering diapause breakage on December 1st we explored the impact of RH on GDD needed for hatching. As shown in Fig. 4A, we didn't record humidity

values below 60% in Mataelpino and not below 75% in the rest of field points therefore, the relationship between RH and GDD at RHs below these values is unknown. Furthermore, there is no clear pattern or relation between RH and GDD in any of the field sites (Fig. 4A). As shown in Fig. 4B, the probability density function of egg hatching as a function of accumulated GDD is very similar in all field sites, except Bustarviejo and the boxplot of Fig. 4C again shows that the GDD needed for hatching is similar in all locations except Bustarviejo. More quantitatively, Fig. 4D shows the Wasserstein distance among the distributions, showing that Bustarviejo is the only one different from the others. Temperatures and RH registered in the four field locations are shown in Supp. Document S6.

Decision-Support Tool to Determine the Best Timing for controlling *P. spumarius* Nymphs

After the model validation, we calculated the precise timing for controlling *P. spumarius* nymphs most efficiently in different regions in Spain. The following initial conditions were fixed according to our results: i) We selected a plausible range of diapause termination dates: we selected the range 1st of November to 1st of January ii). The end date of each simulation was based on the field observations from the systematic surveys (2016–2021) (Supp. Table S3). Because the latest newborn nymphs were found on 27-V-2021 we assumed that the end date of each simulation was June 1st. With this algorithm in mind, the optimal dates for applying control actions (control timing) were selected to maximize the defined efficacy, which is the maximum percentage of targeted nymphs. In addition, we

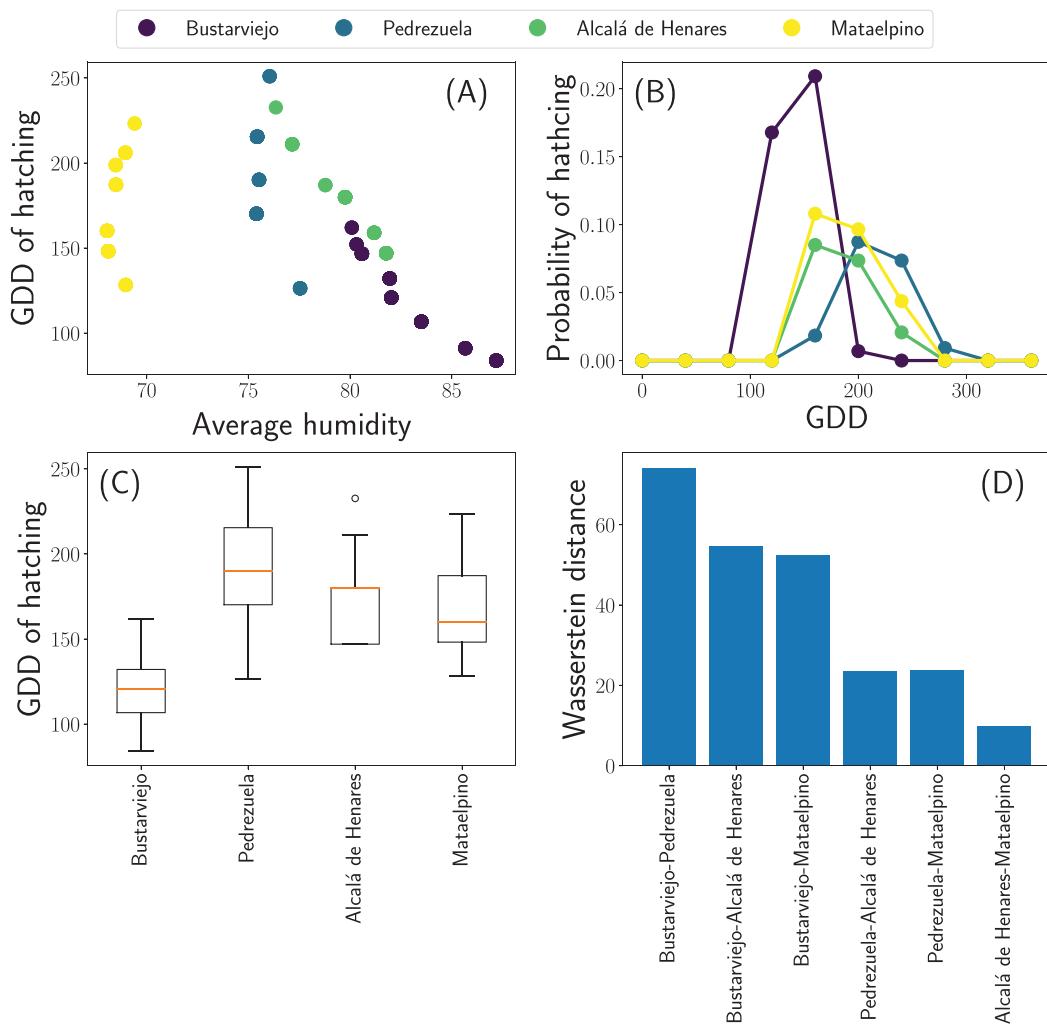


Fig. 4. A) GDD accumulation metric at the moment of egg hatching versus the RH values at the experimental locations. Each data point corresponds to one hatching event. B) Probability density function of egg hatching as a function of accumulated GDD C) Distribution of GDD needed for egg hatching on each field site. The horizontal orange lines represent median values. Boxes extend from the 25th to the 75th percentile of each group's distribution of values, and vertical extending lines indicate the range of values. D) Wasserstein Distance (WD) between probability distributions of GDD accumulation between each pair of field sites (two by two).

considered all possible date combinations for taking certain control actions and compared these simulations with all the hatching observations from both our assays and the systematic surveys (Supp. Table S3). This maximum efficacy changed depending on the diapause termination date (Fig. 5A–B). Consistent with the previous validation, it was found that the model efficacy is maximized if diapause termination is considered to occur in December. Interestingly, we found that assuming diapause termination in mid-December instead of 1st December is a more robust choice. In this way, if the diapause termination date slightly varies for a given particular year, we expect to maintain high efficacy (Fig. 5A–B). Furthermore, to optimize the control actions and to target the maximum number of nymphs as effectively as possible, a control strategy (control timing) applied at a single date was compared to a two-date application strategy. Our results clearly show that the best strategy consists of applying control actions at two different times to target the maximum nymphal population, achieving efficacies of 87 and 78% in northern and southern Spain, respectively (Fig. 5C–D, Supp. Tables S7–S8). In contrast, the maximum efficacy achieved by applying only one control action are 67% in northern Spain and 58% in southern Spain. Thus, by applying two control actions the

targeted population increases in a 30%. According to our results, for northern Spain (latitudes above 40°N) the first control actions should be taken when the accumulated egg-hatching probabilities reach 40% and the second when they reach 90%. On the other hand, for southern Spain (latitudes below 40°N), the first control actions should be taken when the accumulated egg-hatching probability reaches 30% and the second when it reaches 90% (Fig. 5, Supp. Tables S6–S7).

Discussion

In the present work, we built a GDD model obtaining the temperature thresholds (T_{base} , T_{max} , and T_{opt}) directly from experimental data obtained from field experiments under natural conditions. Our approach allows us to indirectly account for the effect of temperature fluctuations on the developmental rate, which is automatically incorporated into the model within the best-fit parameters. Temperature has significant nonlinear effects on insect development, which is slow when the temperatures approach either the upper or lower developmental thresholds (Murray 2020). For this reason, we provided a mathematical approach based on temperature effects on

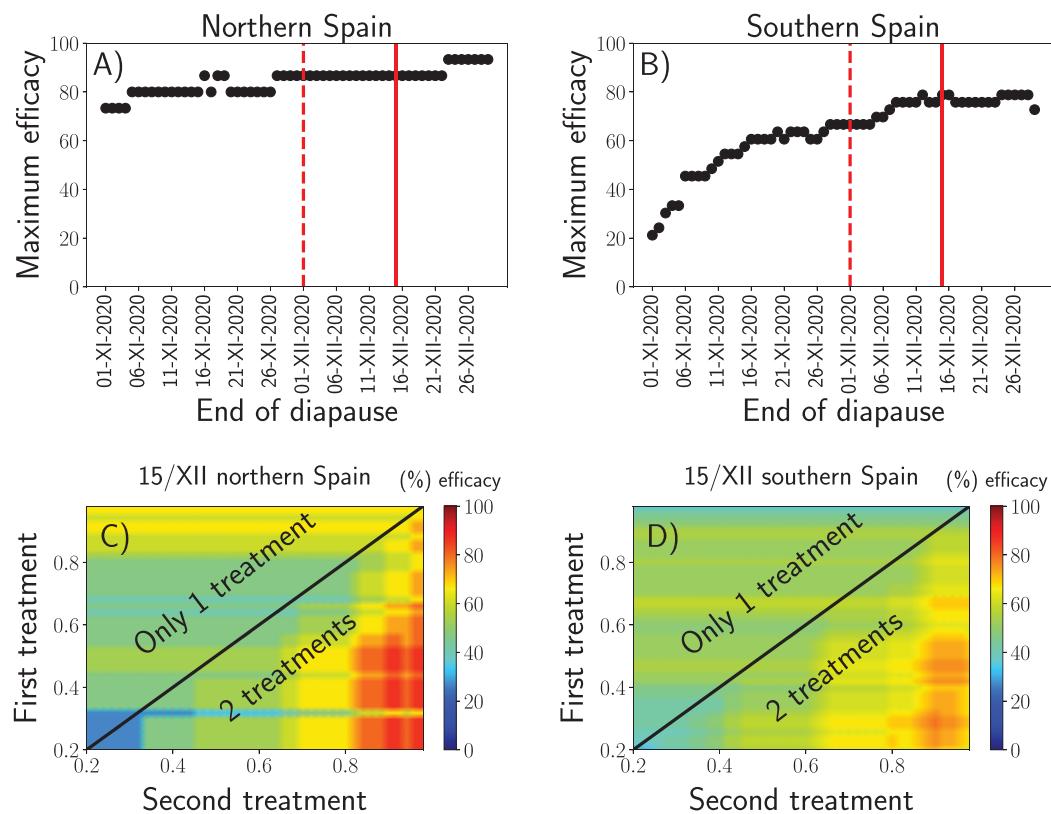


Fig. 5. (A and B) Maximum control action efficacies for each diapause breakage date from the selected range of dates to develop the algorithm (from November 1st to December 30th). The dashed line shows the diapause termination selected for model development (December 1st) and the solid line indicates the diapause termination date to achieve the maximum efficacy (December 10th in the north and December 12th in the south). A) Efficacy in northern Spain (latitude >40°N). B) Efficacies in southern Spain (latitude <40°N). (C and D) Examples of the efficacies of applying a control action depending on the egg hatching probability after applying one or two treatments (C) in northern Spain when considering diapause termination on December 10th and (D) in southern Spain when considering diapause termination on December 15th. This analysis was repeated for all diapause termination dates to obtain the maximum efficacies shown in (A and B).

the metabolic rates of ectotherms (Gillooly et al. 2001) to model the nonlinear development rates with a multilinear function. The model calibration and validation further support our approach of using a multilinear function to calculate the GDD metric and validate the direct use of field data to accurately determine the temperature thresholds.

Chmiel and Wilson (1979) modeled nymphal development of *P. spumarius* by using temperature data from meteorological stations, while in the present work, we installed data loggers inside the cages where the eggs were developing, similarly to the model of nymphal development constructed by Bodino et al. (2019). In this way, we obtained similar data as those experienced by the eggs (Bonhomme 2000). For model construction, it is essential to understand that *P. spumarius* overwinters at the egg stage experimenting a winter diapause (Witsack 1973). Nevertheless, the precise time when egg development starts is unknown, thus establish a starting point for GDD accumulation is challenging. In the GDD model introduced by (Chmiel and Wilson 1979), they arbitrarily set the GDD accumulation starting point on the 1st of January. For our model, we tested 1st of November, 1st of December, and 1st of January as possible ending diapause dates. When we tested December 1st as the starting date for GDD accumulation the temperature profile ($T_{min} = 9.2^\circ\text{C}$, $T_{opt} = 23.4$, $T_{max} = 34.2$) was more realistic than on 1st of January ($T_{min} = 9.2$, $T_{opt} = 27.6$, $T_{max} = 41.8^\circ\text{C}$) or November 1st ($T_{base} = 9^\circ\text{C}$, $T_{opt} = 27.4^\circ\text{C}$, and $T_{max} = 41.2^\circ\text{C}$) since maximum temperature yielded were well above the biological limit for survivorship of the insect, according to what we know about the life cycle parameters

and thermal requirements of *P. spumarius* (Weaver and King 1954, Halkka et al. 2006). Furthermore, when we tested December 1st as the starting date for GDD accumulation the temperature profile ($T_{min} = 9.2^\circ\text{C}$, $T_{opt} = 23.4$, $T_{max} = 34.2$) was more realistic and the model predictions matched well the predictions of 1st instar nymph emergence in Spain giving much more convincing results (see Supp. Video S5 <https://zenodo.org/record/7465753>).

The main reason for the discrepancy between the dates when the diapause ends in the study conducted by Chmiel and Wilson (1979) and the earlier date that we assumed in our study is presumably related to the striking differences in the winter climate between the two regions. While the former study was conducted in West Lafayette, Indiana, USA where winter temperatures are extremely low, our study was conducted in the Iberian Peninsula where temperatures are milder. Thus, *P. spumarius* populations from different regions may have different temperature developmental thresholds or different diapause termination requirements. The induction and termination of winter diapause are crucial components of the life cycle of insect species subjected to diapause (Posledovich et al. 2015, Tougeron 2019) and a geographic variation in diapause duration depending on temperature has been previously reported by several authors (e.g. Schmidt et al. 2005, Chuche and Thiéry 2009, Salman et al. 2019). Therefore, we assumed that *P. spumarius* populations that occur in Spain are adapted to milder winters and thus, they require fewer hours of cold for diapause termination.

Before applying any control actions, the best time of application, control timing, should be defined. For controlling *P.*

spumarius population is essential to apply actions at nymphal stage, before they reach adulthood, since nymphs have limited mobility while adults actively displace (Lago et al. 2021a, 2021b) and only adults contribute to the spread of *X. fastidiosa* to woody hosts (Cornara et al. 2016, 2018). There is a broad time window for nymphal emergence, about two months according to our results, and similar to those previously reported by several authors (Morente et al. 2018a, Bodino et al. 2019). In addition, nymphal development takes 5–6 wk until reaches adulthood (Weaver and King 1954, Yurtsever 2000, Bodino et al. 2019). Therefore, controlling nymphs at two different dates would target the highest percentages of nymphal populations present in the field, before they reach the adulthood. Moreover, the first control actions in the north should be taken when the accumulated egg-hatching probability reaches 40% and the second when it reaches 90%. For southern Spain, the first control actions should be taken when the accumulated egg-hatching probability reaches 30% and the second when it reaches 90%.

In the present study, we provided an R package script for practical use to compute GDDs in each location based on the local temperatures, and in turn, the model can determine the hatching probability at a given site and date (<https://github.com/agimenezromero/PSEggHatching>). Given an input dataset of hourly temperatures in a precise location from the starting date of GDD accumulation to a given date where the model is run, the package provides the current probability level for egg hatching at that given date and precise location. So, further explained, first, the user should somehow obtain hourly temperature data starting on the 1st of December. Then, each day after the 1st of December, the user could introduce this CSV (comma-separated values) data file containing up-to-date hourly temperature into the R program and the output would be the probability level that has been achieved up to that given date. If for that current date the probability reaches the first or second threshold the program warns the user that a treatment should be performed.

The developed model could be used as a decision-making tool to make recommendations regarding the best timing to adopt certain control actions against the main vector of *X. fastidiosa* and to reduce the risk of disease spread. Furthermore, these fitted models can provide estimates of *P. spumarius* biology, and potentially give valuable knowledge in future studies to predict biological aspects of the life cycle of other insect pests.

Acknowledgments

We would like to acknowledge our colleagues María Plaza and Marcos Ramírez for their help in the field experiments and maintaining the insect colonies. In addition, we acknowledge Martin Godefroid for his contribution to the initial experimental design. We are also thankful to Antonio Montserrat (IMIDA, Murcia, Spain), José Luis Ramos Sáez de Ojer (Consejería de Agricultura y Ganadería, La Rioja, Spain), Jose Manuel Durán and Manuel Ruiz Torres (Junta de Andalucía, Spain) for their help with the systematic survey observations used for model validation.

Funding

The work was funded by the Ministerio de Ciencia e Innovación under grant AGL2017-89604-R and fellowship PRE2018-083307 (C. L., M. M., A.M., and A. F.) and Grants RTI2018-095441-B-C22 (SuMaEco) and PID2021-123723OB-C22 (CYCLE) funded by MCIN/AEI/10.13039/501100011033 and by ERDF A way

of making Europe' (A. G. R. and M. A. M.) and Grant MDM-2017-0711 (María de Maeztu Excellence Unit) funded by MCIN/AEI/10.13039/501100011033 (A. G. R. and M. A. M.).

Author Contributions

CL: Data curation-EQUAL, Formal analysis-SUPPORTING, Investigation-Lead, Methodology-EQUAL, Writing – original draft-Lead, Writing – review & editing-Lead. AG-R: Formal analysis-Lead, Investigation-SUPPORTING, Methodology-EQUAL, Software-Lead, Validation-Lead, Writing – original draft-Equal, Writing – review & editing-Equal. MM: Data curation-EQUAL, Investigation-SUPPORTING, Methodology-SUPPORTING, Writing – original draft-Supporting. MM: Formal analysis-SUPPORTING, Investigation-SUPPORTING, Methodology-SUPPORTING, Software-Equal, Supervision-Equal, Validation-Equal, Writing – original draft-Supporting, Writing – review & editing-Supporting. AML: Conceptualization-EQUAL, Data curation-EQUAL, Funding acquisition-EQUAL, Investigation-EQUAL, Methodology-EQUAL, Project administration-Lead, Supervision-Equal, Writing – original draft-Equal, Writing – review & editing-Equal. AF: Conceptualization-Lead, Data curation-EQUAL, Funding acquisition-Lead, Investigation-EQUAL, Methodology-EQUAL, Project administration-EQUAL, Resources-Lead, Supervision-Lead, Writing – original draft-Equal, Writing – review & editing-Equal.

Supplementary Material

Supplementary data are available at *Environmental Entomology* online.

References

- Antonatos S, Papachristos DP, Kapantaikaki DE, Lytra IC, Varikou K, Evangelou VI, Milonas P. Presence of Cicadomorpha in olive orchards of Greece with special reference to *Xylella fastidiosa* vectors. *J Appl Entomol.* 2019;144(1–2):1–11. <https://doi.org/10.1111/jen.12695>
- Arnold C. The determination and significance of the base temperature in a linear heat unit system. *Proc Am Soc Hortic Sci.* 1959;74:430–445.
- Beal DJ, Cooper M, Daugherty MP, Purcell AH, Almeida RPP. Seasonal abundance and infectivity of *Philaenus spumarius* (Hemiptera: Aphrophoridae), a vector of *Xylella fastidiosa* in California Vineyards. *Environ. Entomol.* 2021;50(2):467–476. <https://doi.org/10.1093/ee/nvaa178>
- Bodino N, Cavalieri V, Dongiovanni C, Plazio E, Saladini MA, Volani S, Simonetto A, Fumarola G, Di Carolo M, Porcelli F, et al. Phenology, seasonal abundance and stage-structure of spittlebug (Hemiptera: Aphrophoridae) populations in olive groves in Italy. *Sci Rep.* 2019;9:1–17. <https://doi.org/10.1038/s41598-019-54279-8>
- Bodino N, Cavalieri V, Dongiovanni C, Saladini MA, Simonetto A, Volani S, Plazio E, Altamura G, Tauro D, Gilioli G, et al. Spittlebugs of mediterranean olive groves: Host-plant exploitation throughout the year. *Insects.* 2020;11(2):130. <https://doi.org/10.3390/insects11020130>
- Bonhomme R. Bases and limits to using ‘degree.day’ units. *Eur. J. Agron.* 2000;13(1):1–10. [https://doi.org/10.1016/S1161-0301\(00\)00058-7](https://doi.org/10.1016/S1161-0301(00)00058-7)
- Briere JF, Pracros P, Le Roux AY, Pierre JS. A novel rate model of temperature-dependent development for arthropods. *Environ. Entomol.* 1999;28(1):22–29. <https://doi.org/10.1093/ee/28.1.22>
- Campbell JW, Hanula JL. Efficiency of Malaise traps and colored pan traps for collecting flower visiting insects from three forested ecosystems. *J Insect Conserv.* 2007;11:399–408. <https://doi.org/10.1007/s10841-006-9055-4>
- Chmiel SM, Wilson CM. Estimation of the lower and upper developmental threshold temperatures and duration of the nymphal stages of the meadow spittlebug, *Philaenus spumarius*. *Environ. Entomol.* 1979;8(4):682–685. <https://doi.org/10.1093/ee/8.4.682>
- Chuche J, Thiéry D. Cold winter temperatures condition the egg-hatching dynamics of a grape disease vector. *Naturwissenschaften.* 2009;96:827–834. <https://doi.org/10.1007/s00114-009-0541-x>

- Cornara D, Sicard A, Zeilinger AR, Porcelli F, Purcell AH, Almeida RPP. Transmission of *Xylella fastidiosa* to grapevine by the meadow spittlebug. *Phytopathology*. 2016;106(11):1285–1290. <https://doi.org/10.1094/PHYTO-05-16-0202-R>
- Cornara D, Bosco D, Fereres A. *Philaenus spumarius*: when an old acquaintance becomes a new threat to European agriculture. *J Pest Sci* (2004). 2018;91:957–972. <https://doi.org/10.1007/s10340-018-0966-0>
- Cornara D, Morente M, Markheiser A, Bodino N, Tsai CW, Fereres A, Redak RA, Perring TM, Spotti Lopes JR. An overview on the worldwide vectors of *Xylella fastidiosa*. *Entomol Gen*. 2019;39:157–181. <https://doi.org/10.1127/entomologia/2019/0811>
- Diaz BM, Muñiz M, Barrios L, Fereres A. Temperature thresholds and thermal requirements for development of *Nasonovia ribisnigri* (Hemiptera: Aphididae). *Environ Entomol*. 2007;36(4):681–688. <https://doi.org/10.1093/ee/36.4.681>
- Dongiovanni C, Cavalieri V, Bodino N, Tauro D, Di Carolo M, Fumarola G, Altamura G, Lasorella C, Bosco D. Plant Selection and population trend of spittlebug immatures (Hemiptera: Aphrophoridae) in olive groves of the Apulia region of Italy. *J Econ Entomol*. 2019;112(1):67–74. <https://doi.org/10.1093/jee/toy289>
- EFSA. Scientific opinion on the risks to plant health posed by *Xylella fastidiosa* in the EU territory, with the identification and evaluation of risk reduction options. *EFSA J*. 2015;13(1):1–262. <https://doi.org/10.2903/j.efsa.2015.3989>
- EFSA. *Xylella fastidiosa*; 2022. (<https://www.efsa.europa.eu/en/topics/topic/xylella-fastidiosa>).
- EPPO. First report of *Xylella fastidiosa* in Israel; 2019. Available. EPPO Global Database. <https://gd.eppo.int/reporting/article-6551>
- Fielding CA, Whittaker JB, Butterfield JEL, Coulson JC. Predicting responses to climate change: the effect of altitude and latitude on the phenology of the Spittlebug *Neophilaenus lineatus*. *Funct Ecol*. 1999;13(1):65–73. <https://doi.org/10.1046/j.1365-2435.1999.00009.x>
- Gillooly JF, Brown JH, West GB, Savage VM, Charnov EL. Effects of size and temperature on metabolic rate. *Science* (1979). 2001;293(5538):2248–2251. <https://doi.org/10.1126/science.1061967>
- Halkka A, Halkka L, Halkka O, Roukka K, Pokki J. Lagged effects of North Atlantic oscillation on spittlebug *Philaenus spumarius* (Homoptera) abundance and survival. *Glob Change Biol*. 2006;12:2250–2262. <https://doi.org/10.1111/j.1365-2486.2006.01266.x>
- Herms D. Using degree-days and plant phenology to predict pest activity. In: Krischik V, Davidson J, editors, *IPM (integrated pest management) of mid-west landscapes*. 2004. p. 49–59.
- Johnson D, Bessin R, Townsend L. Predicting insect development using degree days. *Univ Kentucky Coll Agric Entfact*. 1998;123:1–2. <https://entomology.ca.uky.edu/files/efpdf1/ef123.pdf>
- Kim M-J, Baek S, Lee J-H. Egg hatching and first instar falling models of *Metcalfa pruinosa* (Hemiptera: Flatidae). *Insects*. 2020;11(6):345. <https://doi.org/10.3390/insects11060345>
- Lactin DJ, Holliday NJ, Johnson DL, Craigen R. Improved rate model of temperature-dependent development by arthropods. *Environ Entomol*. 1995;24(1):68–75. <https://doi.org/10.1093/ee/24.1.68>
- Lago C, Garzo E, Moreno A, Barrios L, Martí-Campoy A, Rodríguez-Ballester F, Fereres A. Flight performance and the factors affecting the flight behaviour of *Philaenus spumarius* the main vector of *Xylella fastidiosa* in Europe. *Sci Rep*. 2021a; 17608. <https://doi.org/10.1038/s41598-021-96904-5>
- Lago C, Morente M, de las Heras-Bravo D, Martí-Campoy A, Rodríguez-Ballester F, Plaza M, Moreno A, Fereres A. Dispersal of *Neophilaenus campestris*, a vector of *Xylella fastidiosa*, from olive groves to over-summering hosts. *J Appl Entomol*. 2021b;145(7):1–12. <https://doi.org/10.1111/jen.12888>
- Logan JA, Wollkind DJ, Hoyt SC, Tanigoshi LK. An analytic model for description of temperature dependent rate phenomena in arthropods. *Environ Entomol*. 1976;5(6):1133–1140. <https://doi.org/10.1093/ee/5.6.1133>
- Morente M, Cornara D, Moreno A, Fereres A. Continuous indoor rearing of *Philaenus spumarius*, the main European vector of *Xylella fastidiosa*. *J Appl Entomol*. 2018a;142(9):901–904. <https://doi.org/10.1111/jen.12553>
- Morente M, Cornara D, Plaza M, Durán JM, Capiscol C, Trillo R, Manuel R, Carmen R, Susana S, Pereira JA, et al. Distribution and relative abundance of insect vectors of *Xylella fastidiosa* in olive groves of the Iberian peninsula. *Insects*. 2018b;9:175. <https://doi.org/10.3390/insects9040175>
- Morente M, Cornara D, Moreno A, Fereres A. Parapause breakage as a key step for the continuous indoor rearing of *Philaenus spumarius*. *J Appl Entomol*. 2021;145(10):1062–1067. <https://doi.org/10.1111/jen.12931>
- Muñoz Sabater J. ERA5-Land hourly data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS); 2019. <https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.e2161bac?tab=overview>
- Murray, M. 2020. *Using degree days to time treatments for insect pests*. All Current Publications. 978. <https://climate.usu.edu/includes/pestFactSheets/degree-days08.pdf>.
- Ossiannilsson F. The Auchenorrhyncha (Homoptera) of Fennoscandia and Denmark. part 2: the families Cicadidae, Cercopidae, Membracidae, and Cicadellidae (excl. Deltcephalinae). *Fauna Entomol Scand*. 1981;7:223–593.
- Overton WS. A strategy of model construction. *Ecosyst Model Theor Pract*. 1977;49–73.
- Panaretos VM, Zemel Y. Statistical aspects of Wasserstein distances. *Annu Rev Stat Appl*. 2019;6:405–431. <https://doi.org/10.1146/annurev-statistics-030718-104938>
- Posledovich D, Toftegaard T, Wiklund C, Ehrlén J, Gotthard K. Latitudinal variation in diapause duration and post-winter development in two pierid butterflies in relation to phenological specialization. *Oecologia*. 2015;177(1):181–190. <https://doi.org/10.1007/s00442-014-3125-1>
- Quinn BK. A critical review of the use and performance of different function types for modeling temperature-dependent development of arthropod larvae. *J Therm Biol*. 2017;63:65–77. <https://doi.org/10.1016/j.jtherbio.2016.11.013>
- Rebaudo F, Rabhi VB. Modeling temperature-dependent development rate and phenology in insects: review of major developments, challenges, and future directions. *Entomol Exp Appl*. 2018;166(8):607–617. <https://doi.org/10.1111/eea.12693>
- Salman MHR, Bonsignore CP, el Fels AE.A, Giomi F, Hodar JA, Laparie M, Marini L, Merel C, Zalucki MP, Zamoum M, et al. Winter temperature predicts prolonged diapause in pine processionary moth species across their geographic range. *PeerJ*. 2019;7:e6530. <https://doi.org/10.7717/peerj.6530>
- Schmidt PS, Matzkin L, Ippolito M., Eanes WF. Geographic variation in diapause incidence, life-history traits, and climatic adaptation in *Drosophila melanogaster*. *Evolution (N. Y.)*. 2005;59(8):1721–1732. <https://doi.org/10.1111/j.0014-3820.2005.tb01821.x>
- Sevacherian V, Stern VM., Mueller AJ. Heat accumulation for timing Lygus control measures in a safflower-cotton complex13. *J. Econ Entomol*. 1977;70(4):399–402. <https://doi.org/10.1093/jee/70.4.399>
- Tougeron K. Diapause research in insects: historical review and recent work perspectives. *Entomol Exp Appl*. 2019;167(1):27–36. <https://doi.org/10.1111/eea.12753>
- Van Rossum G, Drake FL. *Python 3 reference manual*. Scotts Valley (CA): CreateSpace; 1994. <https://dl.acm.org/doi/abs/10.5555/1593511>
- Weaver CR, King DR. Meadow spittlebug, *Philaenus leucophthalmus* (L.). *Ohio Agric Exp Stat Res Bull*. 1954;74:1–99.
- Whittaker JB. The biology of *Neophilaenus lineatus* (L.) and *N. exclamations* (Thunberg) (Homoptera: Cercopidae) on Pennine moorland. *Proc R Entomol Soc Lond Ser A Gen Entomol* 1965;40(4–6):51–60. <https://doi.org/10.1111/j.1365-3032.1965.tb00313.x>
- Witsack W. Experimental and ecological investigations on forms of dormancy in homoptera-cicadina (Auchenorrhyncha). 2. On ovarian parapause and obligatory embryonic diapause in *Philaenus spumarius* (L.)(Aphrophoridae). *Zool Jahrb Abt Syst Okol Geogr Tiere*. 1973;100:517–562.
- Yurtsever S. On the polymorphic meadow spittlebug, *Philaenus spumarius* (L.) (Homoptera: Cercopidae). *Turk J Zool*. 2000;24(4):447–459. <https://journals.tubitak.gov.tr/zoology/vol24/iss4/13>
- Zajac MA, Hall FR, Wilson MC. Heat Unit model for the development of meadow spittlebug (Homoptera: Cercopidae) on strawberry. *Environ Entomol*. 1989;18(3):347–350. <https://doi.org/10.1093/ee/18.3.347>