Research Article

Received: 9 May 2024 Revised: 3 October 2024 Published online in Wiley Online Library:

(wileyonlinelibrary.com) DOI 10.1002/ps.8585

Assessing pest control treatments from phenology models and field data

Diego F. Rincon,[a,b*] Javier Gutierrez-Illan[a] and David W. Crowder[a]

Abstract

BACKGROUND: Efficacy of insecticides is often determined from apparent yield loses due to a target pest. However, pests can affect yields even when controls work as expected. Further, most pest populations are monitored through adult counts without procedures to assess dynamics of immature stages. Here, we propose a framework to assess the efficacy of control treatments from adult counts in non-experimental setups based on the shifts in temporal patterns of adult emergence caused during the residual period of treatments applied to kill immatures. We use phenology models scaled to field counts to track the stage structure of pest populations across a season and produce reference population trajectories with and without the treatment. Field-collected trajectories are then classified as with or without an effective control through a time-sequential probability ratio test. The method was evaluated using pheromone trap captures of codling moth, *Cydia pomonella*, and four of the most widely implemented treatment programs in apple and pear orchards.

RESULTS: Simulations revealed that when field-collected trajectories are classified as treated with a control, there is 70% chance that the treatment program is $>$ 50% effective, or that the program is $<$ 66% effective when field-collected trajectories are classified as untreated, provided the trajectories are made of $\geq$ 15 pheromone traps.

CONCLUSION: This framework is a powerful, evidence-based tool to optimize the selection of inputs and application protocols for pest control and could be applied to virtually any pest that can be sampled regularly and whose phenology can be modeled as a function of degree-days.
© 2024 The Author(s). *Pest Management Science* published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

Supporting information may be found in the online version of this article.

Keywords: codling moth; *Cydia pomonella*; decision aid systems; insecticide effects models; insecticide control efficacy; mating disruption

1 INTRODUCTION

Pest managers often struggle to assess the efficacy of pest control treatments. Although pesticide registration by regulatory agencies requires efficacy data,[1] efficacy of most registered insecticides is hard to measure in fields, because conditions such as weather, application methods,[2] or pest behavioral adaptations[3] may affect pest control. Sound statistical frameworks have been developed to determine the control efficacy of treatments in experimental setups,[4] but there are few evidence-based procedures to aid growers to conduct such assessments with the monitoring data that they collect. While experimental setups may keep track of target immature life stages and include non-treated experimental controls,[5,6] growers only use count data on mobile pest stages which are not usually killed by insecticides and rely on their own experience and intuition to determine whether a strategy is effective. However, quick field evaluations of treatment programs are key to inform the selection of pest control inputs and application protocols and may favor growers' adaptability to increased pest pressure through experience-based learning.[7,8]

The phenology of many agricultural insect pests can be modeled as a function of heat accumulation during a season. Phenology models are used to predict the proportion of the pest population that has completed a given stage, and whether it is expected to increase or decrease based on temperature records and weather forecasts.[9,10] Phenology models are widely used to guide pest management decision-making about timing of insecticide treatments and when to start sampling monitoring networks,[11] and to track the effect of control treatments on the larval hatching pattern of pest populations throughout the season.[12] To increase their reach, phenology models have been implemented through web-based systems linked to weather station networks and forecast platforms to produce customized risk maps, and to inform pest management based on the cumulative proportion of individuals in a given stage.[13,14]

---

* Correspondence to: DF Rincon, Department of Entomology, Washington State University, Pullman, WA 99164, USA. E-mail: diego.rincon@wsu.edu

a Department of Entomology, Washington State University, Pullman, WA, USA

b Corporación Colombiana de Investigación Agropecuaria (AGROSAVIA), Centro de Investigación Tibaitatá, Mosquera, Colombia

Many tree-fruit pest management programs rely on phenology models to plan insecticide treatments.[13,15] The codling moth, *Cydia pomonella* (L.), is a major pest of tree fruit in part because the time window in which individuals are susceptible to insecticides is relatively narrow; and growers in turn benefit from models that inform the optimum timing for insecticide treatments.[12] However, once treatments are applied, it is difficult to assess whether they had the intended effect on the target population given that yield effects may not be immediately apparent. Furthermore, as most control tools are designed to kill a proportion of the target population and direct evidence of larvae killed by insecticides is difficult to collect, the assessment of recently applied treatments through measures of damage or adult captures in traps can be deceptive.

Here, we propose a method to assess treatment programs by combining probabilistic phenology models with field trap data. Prior work used codling moth heat-driven phenology models to track acute mortality events caused by insecticide treatments throughout subsequent pest generations.[12] We build on this framework to model emergence of codling moth adults that result from treatments applied to kill prior immature stages and produce reference population trajectories scaled to field counts with and without treatments. We use a time-sequential probability ratio test[16,17] to classify field-collected trajectories from pheromone trap captures into one of two categories: 'with control' or 'without control', based on the distinctive shape of codling moth trajectories produced due to the emergence pattern created after a successful treatment program. We tested our method with four of the most common control treatment programs for the codling moth.

2 MATERIALS AND METHODS

2.1 Phenology models and the effect of control treatments

The codling moth overwinters as last-instar larvae and pupates in spring. The emergence of the overwintering adults, egg laying, egg hatching, and the occurrence of subsequent generations throughout the season are driven by the accumulation of degree-days after 1 January, with lower and upper development thresholds of 10 °C and 31.1 °C, with horizontal cut-off.[11,18] The proportion of larval hatching or adult emergence, $P(\mathrm{DDs})$, for the two or three generations that occur across the season can be predicted as a function of cumulative degree-days with probabilistic phenology models, which are based on the Johnson SB distribution:

$$P(\mathrm{DDs})=\frac{\delta}{\lambda\times\sqrt[2]{2\pi}\times z(1-z)}\times\exp\left(-0.5\times\left(\gamma+\delta\times\log\left(\frac{z}{1-z}\right)\right)^{2}\right) \tag{1}$$

where DDs are cumulative degree-days since 1 January, $z=\frac{\mathrm{DDs}-\xi}{\lambda}$, γ and δ are shape parameters, and λ and ξ are scale and location parameters, respectively, and their coefficients for every stage and generation can be found elsewhere[12,18] (Supporting Information, Table S1).

Conventional control of codling moth is based on insecticides that target eggs or neonate larvae before they enter fruits. Treatments that target eggs are efficient because they can cover both eggs that are already laid and those that will be laid within the pesticide residual period. In contrast, treatments that target larvae are only expected to kill those that hatch within a residual period. Additional control for codling moth can be achieved through mating disruption, which reduces fecundity of adults by delaying mating and reducing female longevity.[12]

We used the framework of Jones,[12] where phenology models are used to track the effect of insecticide treatments throughout a season to assess spray timing. Briefly, insecticides typically cause acute mortality during a residual period, which is reflected in subsequent stages and generations. The proportion of the hatched larvae exposed to an insecticide treatment can be determined from Eqn (1) based on the degree-day accumulation when the treatment was made and the insecticide residual period, and then the effect can be tracked through the season. However, instead of tracking the mortality caused by insecticide treatments through subsequent larval generations as Jones[12] did, we tracked the effect on the adults of the same generation to assess the effect of the control using pheromone captures.

Following Jones,[12] the effect of an insecticide applied at DDs^* degree-days can be incorporated into a phenology model by:

$$P_L^*(\mathrm{DDs})=\begin{cases}P_L(\mathrm{DDs}),\mathrm{DDs}\notin\left(\mathrm{DDs}^*,\mathrm{DDs}^*+h\right)\\ P_L(\mathrm{DDs})\times(1-m),\mathrm{DDs}\in\left(\mathrm{DDs}^*,\mathrm{DDs}^*+h\right)\end{cases} \tag{2}$$

where $P_L^*(\mathrm{DDs})$ is the proportion of larvae that hatched and survived the insecticide treatment across cumulative degree-days, $P_L(\mathrm{DDs})$ is the proportion of larvae from the first summer generation that hatched from Eqn (1) (Table S1), and m is the mortality rate caused by the insecticide during the residual period h. It is important to note that Eqn (2) applies to situations when the larval population is only exposed upon hatching, but more advanced larval stages remain protected inside the fruits.

The proportion of the larval population affected by the insecticide treatment can be tracked through subsequent stages or generations by computing the cumulative proportions at DDs^* and DDs^*+h from Eqn (1), and then use them to determine the respective quantiles (i.e., degree-days) in phenology models of subsequent stages. Thus, the interval of degree-days at which the pattern of adult emergence is affected by the insecticide treatment can be denoted as $\left(\mathrm{DDs}'_1,\mathrm{DDs}'_2\right)$ and is given by:

$$\mathrm{DDs}'_1=Q_A\left[P'_L\left(\mathrm{DDs}^*\right)\right] \tag{3a}$$

and

$$\mathrm{DDs}'_2=Q_A\left[P'_L\left(\mathrm{DDs}^*+h\right)\right] \tag{3b}$$

where $Q_A(.)$ is the quantile function of Eqn (1) for the subsequent adult emergence (first summer adult generation) (Table S1), and $P'_L(\mathrm{DDs})$ is the cumulative density function of $P_L(\mathrm{DDs})$, so that $P'_L\left(\mathrm{DDs}^*\right)=P_L\left(\mathrm{DDs}^*\leq\mathrm{DDs}^*\right)$. The effect of the insecticide treatment on the adult emergence is then given by:

$$P_A^*(\mathrm{DDs})=\begin{cases}P_A(\mathrm{DDs}),\mathrm{DDs}\notin\left(\mathrm{DDs}'_1,\mathrm{DDs}'_2\right)\\ P_A(\mathrm{DDs})\times(1-m),\mathrm{DDs}\in\left(\mathrm{DDs}'_1,\mathrm{DDs}'_2\right)\end{cases} \tag{4}$$

where $P_A^*(\mathrm{DDs})$ is the proportion of adults that emerged considering an insecticide treatment that causes mortality rate m on the preceding larval stage across cumulative degree-days, and $P_A(\mathrm{DDs})$ is the proportion of adults that emerged according to Eqn (1).

The effect of mating disruption can be modeled in a similar way, except the effect occurs on the preceding adult generation and must include daily temperature data. Jones and Wiman[18] used a population model to estimate the effect of mating delay in degree-days due to mating disruption on codling moth population growth. We used a Gompertz equation to model the proportion of population growth not impacted by the effect of mating disruption on female survival and delayed mating as a function of degree-days, as proposed by Jones and Wiman[18]:

$$s_m = \exp(A(1-\exp(G\times \mathrm{DDs_D}))) \tag{5}$$

where s_m is the proportional of population growth as a function of degree-days built up during the mating delay in days, $\mathrm{DDs_D}$, and A and G are parameters which were set to 0.372 and 0.029, respectively from figure 2 in Jones and Wiman.[18]

If degree-day accumulation data, $\mathrm{DDs_C}$, are available for a growing season, one should be able to determine the proportion of adults that emerge daily considering the effect of a 3-day delay in mating caused by mating disruption[18,19] on the preceding generation by a five-step process: (i) compute the daily $\mathrm{DDs_D}$ for the 3-day delay in mating and the corresponding daily proportions of growth from Eqn (5), S_m, (ii) determine the distribution of proportions of adult emergence across calendar days for an untreated ('without control') population by $P'_A(\mathrm{DDs_C}) = P_A(\mathrm{DDs_C})/\sum P_A(\mathrm{DDs_C})$, so that proportions from the same generation sum up one, (iii) find the DDs at which adult emergence is impacted from $Q_A(P_L(\mathrm{DDs} \le \mathrm{DDs_C}))$ and the corresponding calendar days, (iv) assign each S_m computed in step (i) to the calendar days found in step (iii), (v) the resulting adult emergence per calendar day after the effect of mating disruption is given by:

$$P^*_A(\mathrm{DDs_C}) = P'_A(\mathrm{DDs_C})\times S_m \tag{6}$$

In cases where t_i are assigned with two or more S_m values, these are multiplied to get a single value, and those that are assigned no S_m values, adult emergence is approximated with a linear combination between the two closest known $P^*_A(\mathrm{DDs_C})$.

We modeled four treatment programs for codling moth[12] (Table 1). The first consists of two larvicide sprays that kill 90% of the exposed population with a 14-day residual period, with the first spray at 235 degree-days and the second 14 days after. The second program consists of an oil spray that kills 80% of the eggs, and then two larvicides, the first at 290 degree-days and the second 14 days after. As the first and second treatment programs were modeled in DDs, we assumed a daily accumulation of 8 DDs for the residual period (112 DDs total), and 9.5 DDs (133 DDs total) for the time between sprays. Although an ideal scenario involves an overlap between the first and second sprays, the residual period was shortened to account for rapid degradation.[20] Since the third and fourth programs involved mating disruption, they were modeled in calendar days using the temperature records collected from April and October 2023 from a weather station located in Pullman, WA, USA (46°41′24.0″, −117°09′00.0″) (https://weather.wsu.edu/). The third program consists of only mating disruption without any insecticides, and the fourth program included four virus sprays each causing 70% mortality on recently hatched larvae, 7 days apart. The modeled effect of an effective treatment for each treatment program on larval hatching and adult emergence is presented in Fig. 1.

2.2 Theoretical reference population trajectories

The effect of treatment programs on phenology models, and field data collected from pheromone traps throughout the season can be used as inputs to construct theoretical reference population trajectories with and without an effective control (Fig. 2). For treatment programs modeled in degree-days, the reference population trajectory without an effective control, $N_1(\mathrm{DDs}_i)$, is given by:

$$N_1(\mathrm{DDs}_i) = \int_{\mathrm{DDs}_i}^{\mathrm{DDs}_{i+1}} P_A(\mathrm{DDs})d\mathrm{DDs}\times\sum_{i=1}^{h} x_i \times\log\left(\frac{1}{\sum_{i=1}^{h}\int_{\mathrm{DDs}_i}^{\mathrm{DDs}_{i+1}} P^*_A(\mathrm{DDs})d\mathrm{DDs}}\right) \tag{7a}$$

and the reference population trajectory with an effective control, $N_0(\mathrm{DDs}_i)$, by:

$$N_0(\mathrm{DDs}_i) = \int_{\mathrm{DDs}_i}^{\mathrm{DDs}_{i+1}} P^*_A(\mathrm{DDs})d\mathrm{DDs}\times\sum_{i=1}^{h} x_i \times\log\left(\frac{1}{\sum_{i=1}^{h}\int_{\mathrm{DDs}_i}^{\mathrm{DDs}_{i+1}} P^*_A(\mathrm{DDs})d\mathrm{DDs}}\right) \tag{7b}$$

where a field-collected population trajectory, $\boldsymbol{X} = \{x_1,\ldots,x_h\}$, is made of mean captures per trap collected at corresponding ith

Table 1. Treatment programs and expected efficacies, as reported by Jones,[12] used to evaluate a procedure to assess the efficacy of control treatments for codling moth management

Treatment program	Percent survival per spray	Percent overall survival	Number of sprays	Timing (in degree-days)		
				First	Second	Subsequent
Conventional traditional	10, 10	35.5	Two larvicides	235	368 (14 days)	—
Delayed first cover	20, 10, 10	7.74	Oil and then two larvicides	210	290	423 (14 days)
Mating disruption[†]	—	75.49	—	—	—	—
Organic traditional[‡]	30, 30, 30, 30	40.9	Four virus sprays	235	+7 days	+7 days

[†] The overall survival of mating disruption programs depends on temperature.
[‡] The organic traditional treatment program is deployed in addition to mating disruption.

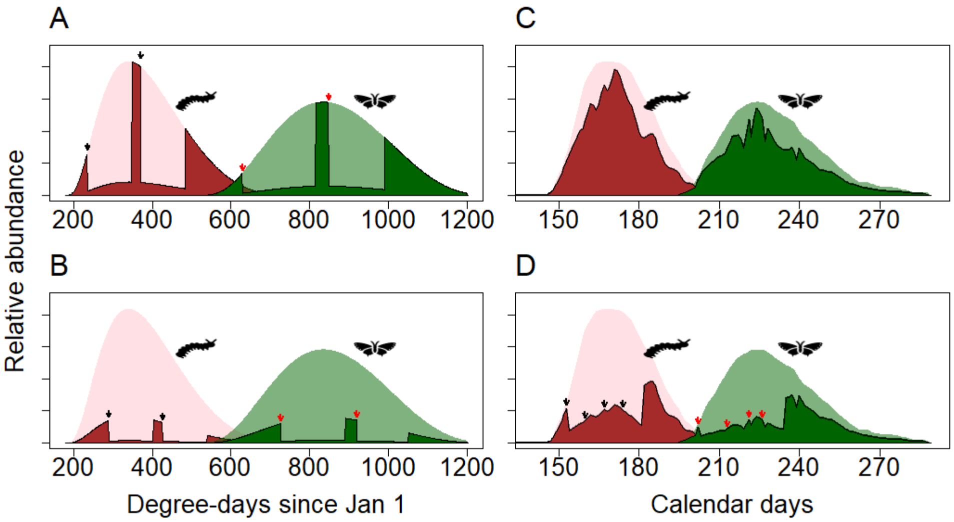

Figure 1. Theoretical effect of four treatment programs on the codling moth larva (brown for treated *verus* pink for untreated) and adult (dark green for treated *versus* light green for untreated) phenology for the first summer generation: conventional traditional (A), delayed first cover (oil application not shown) (B) (from Eqn (4)), mating disruption (C), and mating disruption plus organic traditional (D) (from Eqn (6)). Models that involve programs based solely on insecticide treatments (A and B) are showed in degree-days, while those that involve mating disruption are showed in calendar days with temperature records collected in Pullman (WA, USA) between April and October 2023 (C and D). Population size in *y*-axes is given in relative abundance as absolute quantities are context dependent. The black arrows indicate the timing of insecticide treatments and the red arrows the corresponding start of the effect on adult emergence.

DDs, and $N_1(\mathrm{DDs}_i)$ and $N_0(\mathrm{DDs}_i)$ are estimated at the same intervals. The first two terms in Eqns (7a) and (7b) produce a discretized version of $P_A(\mathrm{DDs})$ and $P^*_A(\mathrm{DDs})$ for the first summer adult generation (where $P^*_A(\mathrm{DDs})$ is obtained from $P^*_L(\mathrm{DDs})$ for the first summer larval generation) (Table S1) scaled to the total counts recorded within the season. The third term places reference population trajectories approximately equidistant to the field-collected trajectory (Fig. 2) and involves the ratio between the sum of proportions at sampling DDs intervals with and without an effective control. The numerator (without effective control) is set to 1 when the population trajectory includes the entire adult stage, and the denominator (with effective control) indicates the overall expected survival (Table 1). However, when the overall expected efficacy (i.e., 1 − survival) is lower than 63.2%, the third term in Eqns (7a) and (7b) becomes < 1 and produces unrealistic reference population trajectories, in which cases it is set to 1.

For treatment programs modeled in calendar days, reference population trajectories are produced using a similar approach, except that each element in the trajectory $\boldsymbol{X}$ is associated to a collection of corresponding *i*th day, t_i, each assigned with a number of cumulative degree-days, $\mathrm{DDs_C}(t_i)$. Thus, the reference population trajectory without effective control is given by:

$$N_1(t_i) = \int_{\mathrm{DDs_C}(t_i)}^{\mathrm{DDs_C}(t_{i+1})} P'_A(\mathrm{DDs_C})\,d\mathrm{DDs_C} \times \frac{\sum_{i=1}^{h} x_i}{\sum_{i=1}^{h}\int_{\mathrm{DDs}_i}^{\mathrm{DDs}_{i+1}} P'_A(\mathrm{DDs_C})\,d\mathrm{DDs_C}} \times \log\left(\frac{\sum_{i=1}^{h}\int_{\mathrm{DDs_C}(t_i)}^{\mathrm{DDs_C}(t_{i+1})} P'_A(\mathrm{DDs_C})\,d\mathrm{DDs_C}}{\sum_{i=1}^{h}\int_{\mathrm{DDs_C}(t_i)}^{\mathrm{DDs_C}(t_{i+1})} P^*_A(\mathrm{DDs_C})\,d\mathrm{DDs_C}}\right) \tag{8a}$$

and with effective control:

$$N_0(t_i) = \int_{\mathrm{DDs_C}(t_i)}^{\mathrm{DDs_C}(t_{i+1})} P^*_A(\mathrm{DDs_C})\,d\mathrm{DDs_C} \times \frac{\sum_{i=1}^{h} x_i}{\sum_{i=1}^{h}\int_{\mathrm{DDs}_i}^{\mathrm{DDs}_{i+1}} P'_A(\mathrm{DDs_C})\,d\mathrm{DDs_C}} \times \log\left(\frac{\sum_{i=1}^{h}\int_{\mathrm{DDs_C}(t_i)}^{\mathrm{DDs_C}(t_{i+1})} P'_A(\mathrm{DDs_C})\,d\mathrm{DDs_C}}{\sum_{i=1}^{h}\int_{\mathrm{DDs_C}(t_i)}^{\mathrm{DDs_C}(t_{i+1})} P^*_A(\mathrm{DDs_C})\,d\mathrm{DDs_C}}\right) \tag{8b}$$

Since the proportions of adult emergence across calendar days rarely sum up to 1, the sum of proportions at the sampling days t_i should be included in the denominator of the second term and the numerator of the third term. This sum of proportions could also be present in the second and third terms of Eqns (7a) and (7b), but they are assumed to equal 1 when proportions of emergence are modeled across DDs.

2.3 Assessment of field-collected trajectories

Pedigo and van Schaik[17] proposed modifying Wald's sequential probability ratio test[21] to work in time rather than in space, which is known as time-sequential probability ratio test. The motivation is to classify a pest population trajectory into either a problematic (outbreak) configuration, which is to be treated timely to prevent damage, or an undamaging (endemic) configuration, which does not require treatment. We use this approach to classify field-collected trajectories into one of two categories: consistent with a reference population trajectory with or without effective control.

The stop boundary curves were constructed assuming a negative binomial distribution of the sample counts and following

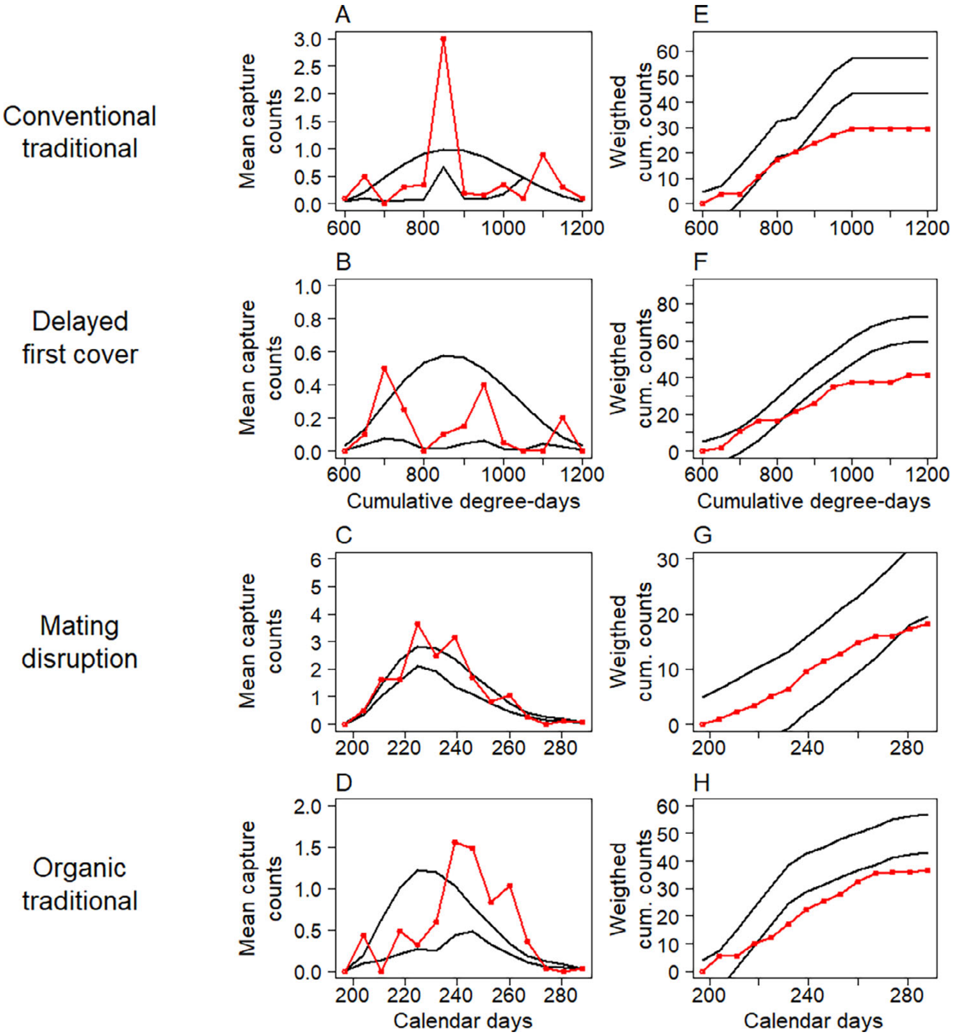

Figure 2. Reference population trajectories (black) and simulated codling moth pheromone trap counts (red) (A–D), compared to the corresponding stop lines and weighted cumulative trap counts (E–H) produced from the theoretical effect of four treatment programs on codling moth phenology. For illustration, all the simulated population trajectories assumed 90% control efficacy for treatment programs that involve insecticides (A, B, E and F), 70% control efficacy for viruses (D and H) and a 3-day-delay in mating for mating disruption (C, D, G and H).

Pedigo and van Schaik,[17] but considering the effect of sample size as proposed by Binns *et al*.,[16] such that the upper boundary for a population trajectory without an effective control is given by:

$$c_U = \log\left(\frac{1-\beta}{\alpha}\right) + b(r_i) \tag{9a}$$

and the lower boundary for a population trajectory with an effective control:

$$c_L = \log\left(\frac{\beta}{1-\alpha}\right) + b(r_i) \tag{9b}$$

where α is acceptable error rate when the control was effective, β is the error when the control was not effective, and they were set to 0.01 and 0.0001, respectively. Value for β were relatively conservative because mistaken classifications when control programs do not work are generally more costly than when they do. The second term in Eqns (9a) and (9b), $b(r_i)$, is a function of time in the broad sense, since r_i could be either days, t_i, or degree-days, DDs$_i$, and is expressed as:

$$b(r_i) = k\sum_{i=1}^{i} n_i \times \log\left\{\frac{k_0(r_i)[k_1(r_i)+N_1(r_i)]}{k_1(r_i)[k_0(r_i)+N_0(r_i)]}\right\} \tag{10}$$

where k is the common negative binomial dispersion parameter that describes codling moth captures in pheromone traps, $k_0(r_i)$ and $k_1(r_i)$ are the negative binomial dispersion parameters associated with $N_0(r_i)$ and $N_1(r_i)$, respectively, and n_i is the sample size at time r_i.

Before cumulative mean counts can be compared with boundaries, they must be weighed so that the dynamic change in separation between reference population trajectories is accounted for to place more emphasis on sampling times when the difference between trajectories is greatest. Thus, each value in any field-collected trajectory ***X*** must be weighted and multiplied by the corresponding sample size, before adding it to the previous sum and being compared with the corresponding boundaries:

$$x'_i=\sum_{i=1}^{i}x_i n_i \log\left\{\frac{k_0(r_i)k_1(r_i)N_1(r_i)[k_0(r_i)+N_0(r_i)]}{k_0(r_i)k_1(r_i)N_0(r_i)[k_1(r_i)+N_1(r_i)]}\right\} \tag{11}$$

where $\boldsymbol{X}'=\{x'_1,\ldots,x'_h\}$ is the weighed cumulative trajectory. The assessment is carried out by comparting the elements in $\boldsymbol{X}'$ with the corresponding elements in c_U and c_L. The field-collected trajectory is classified as 'without control' if there are more elements in $\boldsymbol{X}'$ that are greater than c_U compared to those that are less than c_L, or as 'with control' if more elements in $\boldsymbol{X}'$ are less than c_L compared to those that are greater than c_U. When there is the same number of elements in $\boldsymbol{X}'$ that are greater than c_U and less than c_L, the population trajectory is 'without control' (Fig. 2).

2.4 Estimation of negative binomial parameters for count data generation

Although model performance could ideally be assessed using experimental data from treated and untreated field codling moth populations, such data is not available. However, the use of computer-generated data offers the opportunity to test different pre-established control efficacies and sample sizes, which is hardly guaranteed or viable with experimental setups. To produce reliable pheromone trap capture counts, a 20-year codling moth pheromone capture dataset from the Okanagan-Kootenay Sterile Insect Release Program was used to model variance about codling moth captures per trap as a function of the mean. Data collection is described in detail by Rincon *et al.*[22] Briefly, the dataset used is made of 134 field-collected trajectories from apple orchards located around Kelowna, BC, Canada, between 1995 and 2015. Data were collected weekly for the 24-week period of codling moth flight from a variable pheromone trap density, which ranged between 0.01 and 0.13 traps per hectare (mean = 0.08 ± 0.04). Pheromone traps consisted of wing traps and Delta II traps baited with codlemone or a mixture of codlemone and pear ester (Trécé Inc., Adair, OK, USA).

Variation of k across mean trap captures was modeled using the Taylor's power law parameters, a and b, which were estimated with the linearized version, $\log(\mu)=\log(a)+b\log(\sigma^2)$ by ordinary least squares, with the means and variances of each of the sampling times from the 134 filed-collected trajectories. Since a large portion of the dataset includes zeros the variance–mean relationship showed the 'Poisson sampling effect',[23] Taylor's power law parameters were estimated through the split-domain technique, in which a separate set of parameters is estimated for each domain with a cut-off that is found through an iterative process.[22,24] Thus, the dispersion parameter of the negative binomial distribution, k, was estimated as[16]:

$$k=\begin{cases}\dfrac{\mu^2}{a_1\mu^{b_1}-\mu},\mu>\exp(c)\\[2ex]\dfrac{\mu^2}{a_2\mu^{b_2}-\mu},\mu\leq\exp(c)\end{cases} \tag{12}$$

where a_1 was estimated as 6.41 [standard error (SE) = 1.016], a_2 as 2.057 (SE = 1.053), b_1 as 1.61 (SE = 0.011), b_2 as 1.16 (SE = 0.014) and c as −2.54. The common k was set to 0.32, which is the k value obtained from Eqn (12) from the overall mean of the data.

2.5 Operational characteristics

Data were generated randomly from a negative binomial distribution with parameter k estimated from Eqn (13). The variability about Taylor's power law parameters was included as the root mean square error (RMSE) for the regression as described by Binns *et al.*,[16] such that:

$$k'=\begin{cases}\dfrac{\mu^2}{a_1\mu^{b_1}\exp(z(0,\sigma_1))-\mu},\mu>\exp(c)\\[2ex]\dfrac{\mu^2}{a_2\mu^{b_2}\exp(z(0,\sigma_2))-\mu},\mu\leq\exp(c)\end{cases} \tag{13}$$

where z is a normally distributed variable with mean 0 and standard deviation σ_1 or σ_2, which equal 0.293 and 0.590 and are the RMSE of the regression model for parameters a_1 and b_1, and a_2 and b_2, respectively. The total mean captures per population trajectory was set to 20, which was multiplied by the proportions derived from $P^*_A(\mathrm{DDs})$ or $P^*_A(\mathrm{DDs_C})$ (Eqns (4) and (6)), depending on whether the simulation was performed across DDs or $\mathrm{DDs_C}$, to set the mean and k' of the negative binomial at each sampling time of every generated population trajectory.

We tested model performance classifying sample population trajectories as 'with control' or 'without control' across different sample sizes (i.e., number of traps per sample): 5, 10, 15, 20, 25 and 30. The model was evaluated for the four treatments programs, using generated population trajectories set at the corresponding survival rates that result from 100% control efficacy of each treatment program (Table 1). Model performance was also examined across levels of control treatment efficacy, by introducing the overall efficacy rate as a factor of m in Eqn (4) to generate the sample population trajectories, which were set to a sample size of 25 traps per sampling time. In total, we tested ten different efficacy rates: 100%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10% and 0%. Each combination of treatment programs with sample sizes and control efficacies was simulated 1000 times, and the number of 'with control' and 'without control' classifications were recorded for each. To determine the efficacy of the model classifying sample population trajectories, the probability of obtaining a 'with control' or a 'without control' classification given different levels (ranges) of control efficacies were calculated from simulations. The proportions of classifications across efficacy rates were differentiated to obtain the independent probabilities for each, and then converted to cumulative probabilities across ranges of control efficacies.

All data analysis and simulations were completed in R version 4.2.3.[25] The Johnson SB function was implemented using the SuppDists package,[26] and data digitization from Jones and Wiman[18] for parameter estimation of Eqn (5) was performed using the digitize package.[27] All the simulations were run through parallel computing using the parallel package.[28]

3 RESULTS

Our model produced > 90% accurate classifications provided the sample size was 15 traps or more per sampling time for all the treatment programs when their efficacy was 100%, except for mating disruption alone, which reached a maximum of 84.7%. In general, the model was not reliable at classifying population trajectories treated with mating disruption alone even when sample size was increased to 30 traps (Fig. 3(A)). Simulations also revealed model performance across control efficacies varied with the treatment program being tested. When the model output was a 'with control' classification, 84.0% of the codling moth populations treated with mating disruption alone were correctly classified, but 41.2% of the untreated were classified incorrectly

(Fig. 3(B)). However, the model was accurate classifying population trajectories as 'with control' for the remaining programs. If we use an acceptable accuracy of 70%, simulations show that if a 'with control' classification is obtained, one should be able to consider that control efficacies are, in general, greater than 50% (Fig. 3(B)). The performance was best for the conventional traditional program, for which 'with control' classifications indicate (with 70% accuracy) that the treatment is ≥ 69.25% effective. For the delayed first cover and the traditional organic programs, when the model output is a 'with control' classification, simulations showed that treatments are ≥ 54.0%, and ≥ 51.03% effective, respectively, with a 70% accuracy (Fig. 3(B)).

When the model output was a 'without control' classification, 16.5% of the populations treated with mating disruption alone were incorrectly classified by the model, and only 60% of the untreated ones were correctly classified as such (Fig. 3(C)). For the remaining programs, simulations showed that when populations are classified as 'without control' by the model, there is a 70% chance that they are ≤ 66% effective (Fig. 3(C)). The most accurate results were obtained when assessing the conventional traditional program for which our model established with 70% accuracy that the treatment is ≤ 80.34% effective when a 'without control' classification is obtained. For the delayed first cover and the traditional organic programs, simulations showed that when a 'without control' classification is obtained, there is 70% chance that the treatment programs are ≤ 66.35% and ≤ 69.75% effective, respectively (Fig. 3(C)).

4 DISCUSSION

Our study is among the first to propose a method to assess treatment programs for pest control in non-experimental setups. Other approaches addressed field or laboratory evaluations of interventions with datasets that include non-treated experimental controls and/or direct observation of target stages, whose main purpose is applying for pesticide registration.[4,5] However, little attention has been paid to evaluation of pest control treatments carried out by growers to inform selection of inputs and adjustment of application methods. Our approach involves a four-step process, after a treatment program is applied: (i) produce probabilistic phenology models that include the theoretical expected effect of treatment programs for the sampled pest stage (Eqns (4) or (6)), (ii) collect data from at least 15 traps at weekly intervals during the emergence of the first adult generation, (iii) derive reference population trajectories from phenology models scaled to field counts (Eqns (7a) and (7b) or Eqns (8a) and (8b)), and (iv) classify the field-collected trajectory through a time-sequential probability ratio test (Eqns (9a) and (9b) and Eqn (11)). Since our model relies on scaled versions of heat-driven phenology models, the assessments have four key assumptions. First, the probabilistic phenology models are properly adapted to the region where evaluations are to be made, particularly to variations caused by latitude, as is the case for the codling moth.[29] Second, the sampling effort should be the same throughout the season. Third, any major deviations in the field-collected trajectories due to drivers other than heat accumulation are not considered; shifts in capture count patterns due to predation,[30] parasitism,[31] epizootics,[32] or migration[33] may produce incorrect classifications. Fourth, insecticides have a negligible effect on codling moth adult survival, flight, or attraction to pheromone traps, which is reasonable considering that most modern insecticides target larval stages or eggs and that adult intoxication by contact or ingestion is unlikely (but see Smith[34]).

The classification is based on the ratio between reference population trajectories at specific sampling times. When separation between reference trajectories is greatest, more weight is placed on the collected counts, and stop lines have greater slopes (Fig. 2). As a result, the model is relatively insensitive to changes in field-collected trajectories at times when reference population trajectories are closer to each other, and sensitive to changes when reference population trajectories are different. This condition makes the model efficient at classifying population trajectories based on the magnitude of counts and trajectory shape, but prone to mistakes due to delays of field-collected trajectories

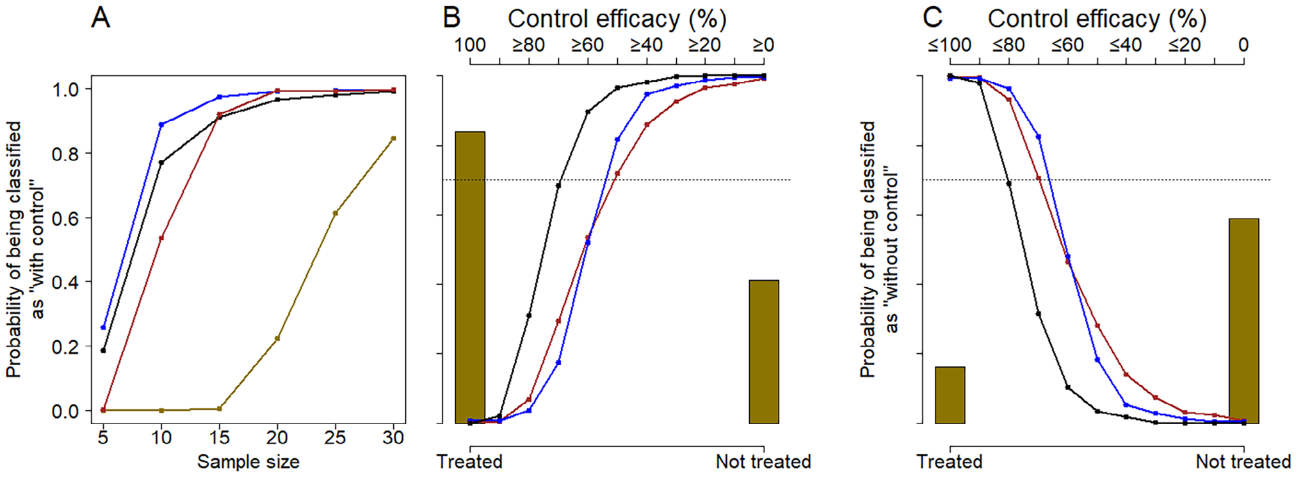

Figure 3. Operational characteristics of a procedure to assess pest control treatments from phenology models and pheromone trap counts for four treatment programs: conventional traditional (black), delayed first cover (blue), mating disruption (gold), and organic traditional (brown). In (A), the probability of a treated population trajectory being classified correctly as 'with control' is presented as a function of the number of traps per sampling time. In (B) and (C), the probability of a treated population trajectory being classified as 'with control' (B) or 'without control' (C) is presented as a function of ranges of control efficacy. The analysis in (A) assumes 100% control efficacy for all treatment programs, and the analyses in (B) and (C) were set to a sample size of 25 traps per sampling time. The horizontal dashed line in (B) and (C) denotes 0.7 probability (70% accuracy) as a reference. Every data point is based on 1000 simulations.

relative to heat-driven phenology. This is especially true when treatment programs produce clear peaks, such as the conventional traditional (Fig. 2(A)), in which a delay in field-collected trajectories may produce unrealistically high weighted counts and false 'without control' classifications. Caution is advised when classifications are carried out using trajectories collected from populations treated with mating disruption, which is known to delay between 2 to 3 days mate finding, and likely pheromone trap catching as well.[18,19]

One more effect of relying on the local separation between reference population trajectories is that the model becomes relatively insensitive when reference trajectories are similar in shape, even if they are different in the overall magnitude. For example, although the delayed first cover is the program with the lowest overall pest survival (Table 1 and Fig. 1), the resulting shape of the reference population trajectory is similar to that produced without an effective control (Fig. 2(C)). This condition makes the weighing factor of counts relatively constant across time, and the sole overall increased/reduced mean captures might not be enough for the model to classify a given field-collected population trajectory correctly. In fact, our model performed best classifying populations treated with the conventional traditional program, which produces reference population trajectories that differ in shape from those that result without an effective control, but with an overall survival that is similar to the organic traditional, which ranked third (Fig. 3(B)). One more instance of this limitation is the poor performance classifying populations treated with mating disruption alone. Beyond the relatively low efficacy (i.e., high overall survival) of mating disruption alone (Table 1), the similarity in shape of the produced reference population trajectories with those of produced without an effective control may be why our model failed classifying correctly populations treated with this program (Fig. 2(D)).

Beyond the limitations, these results show that our approach is robust classifying population trajectories according to the level of control achieved by pest control treatments, at least for some of the most commonly used treatment programs for codling moth management. In general, the procedure is sensitive to sample size regardless of the treatment program and provides accurate results when field-collected trajectories are obtained from monitoring networks made of ≥ 15 pheromone traps. Recent surveys indicate that the average size of an orchard in Washington State is about 40 ha (100 acres),[35] and most extension programs recommend trap densities that range from 0.1 to 10 per ha, depending on whether the aim is detection, monitoring or mass trapping.[36,37] Therefore, this procedure could be readily implemented in most orchards and produce useful control treatment assessments at least across Washington State, where nearly 68% of US apples are produced.[38]

Our approach showed to be better suited for treatment programs that involve insecticides that cause acute mortality events than for those that cause sustained, chronic control, such as those that involve mating disruption[39] and possibly other techniques like sterile insect release or biological control.[30,31] We established a 70% accuracy to interpret the model outcomes in terms of inferences about treatment programs efficacies, but more confidence could be added by increasing accuracy with a pay toll in the range of efficacies that can be inferred from the assessment. For instance, if we established a 90% accuracy instead, from a 'with control' outcome for populations treated with a conventional traditional or a delayed first covert programs one could infer that the treatment is ≥ 59.71% and ≥ 43.72% effective, respectively (Fig. 3(B)). Similarly, with a 90% accuracy, from a 'without control' outcome, one could infer that those treatments are ≤ 87.24% and ≤ 75.36% effective, respectively (Fig. 3(C)).

While biological and chemical control inputs are constantly developed for most major pests, strategies that assist growers' decision-making on the selection and application of available tools are still the main obstacle to the adoption of integrated pest management programs.[40,41] Several contributions have been produced to inform growers' decisions on the timing and need for interventions with chemical or biological inputs according to economic and biological criteria,[42] sampling protocols,[8] or weather patterns.[13] However, few studies have addressed the assessment of recent control treatments, which may inform the selection of inputs and application protocols. This study provides a framework that combines phenology modeling with sampling to guide the collection and evaluation of data and determine whether a recently applied treatment was effective or not with a predefined level of accuracy. This approach contributes to bridging the gaps for integrated pest management adoption and could be applied to other pests that can be sampled regularly and whose phenology can be modeled as a function of heat accumulation.

5 CONCLUSION

We present an approach that accurately classifies pest population trajectories from monitoring networks into one of two categories to assess recently applied control treatment programs. The approach relies on probabilistic phenology models to track the effect of acute mortality effects caused by treatment programs across the growing season, field-collected data to produce reference trajectories for treated and untreated populations, and a time-sequential probability ratio test to classify the trajectories accordingly. Simulations revealed that the method is robust across a reasonable range of sampling sizes, and control efficacies. End users should provide the timing of treatments either in calendar days or degree-days, the expected mortality rate, and a collection of counts from traps over time (i.e., field-collected trajectories). When field-collected trajectories are classified as 'with control', there is 70% chance that treatments are > 50% effective, and < 66% effective when they are classified as 'without control'. This framework is a powerful, evidence-based tool that can be used to optimize the selection of inputs and application protocols for pest management and could be applied to other pests that can be sampled regularly and whose phenology can be modeled as a function of heat accumulation.

ACKNOWLEDGEMENTS

The authors thank members of the Okanagan-Kootenay Sterile Insect Release Program for providing data, G. Dickel and R. Bruno for help with data analysis, and V. Jones for helpful discussions about the manuscript. The code produced for this study is available at https://github.com/rincondf/realtimeEval. The codling moth capture dataset used to fit variance–mean models for data-generation is available from the Okanagan-Kootenay Sterile Insect Release Board (https://www.oksir.org/) upon reasonable request to sirinfo@oksir.org. Research was supported by USDA-NIFA award 2022-67013-36422.

CONFLICT OF INTEREST STATEMENT

All authors agree there are no competing interests.

DATA AVAILABILITY STATEMENT

The code produced for this study is available at https://github.com/rincondf/realtimeEval. The codling moth capture dataset used to fit variance-mean models for data-generation is available from the Okanagan-Kootenay Sterile Insect Release Board (https://www.oksir.org/) upon reasonable request to sirinfo@oksir.org.

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Pelaez V, da Silva LR and Araújo EB, Regulation of pesticides: a comparative analysis. *Sci Public Policy* **40**:644–656 (2013).

2 Payne NJ, Factors influencing aerial insecticide application to forests. *Integr Pest Manage Rev* **5**:1–10 (2000).

3 Hoy CW, Head GP and Hall FR, Spatial heterogeneity and insect adaptation to toxins. *Annu Rev Entomol* **43**:571–594 (1998).

4 Sudo M, Yamanaka T and Miyai S, Quantifying pesticide efficacy from multiple field trials. *Popul Ecol* **61**:450–456 (2019).

5 EPPO, Design and analysis of efficacy evaluation trials. *EPPO Bull* **42**: 367–381 (2012).

6 FAO, Guidelines on efficacy evaluation for the registration of plant protection products (2006).

7 Shea K, Possingham HP, Murdoch WW and Roush R, Active adaptive management in insect pest and weed control: intervention with a plan for learning. *Ecol Appl* **12**:927–936 (2002).

8 Rincon DF, Rivera-Trujillo HF, Mojica-Ramos L and Borrero-Echeverry F, Sampling plans promoting farmers' memory provide decision support in *Tuta absoluta* management. *Agron Sustain Dev* **41**:33 (2021).

9 Crimmins TM, Gerst KL, Huerta DG, Marsh RL, Posthumus EE, Rosemartin AH *et al*., Short-term forecasts of insect phenology inform pest management. *Ann Entomol Soc Am* **113**:139–148 (2020).

10 Prasad YG and Prabhakar M, Pest monitoring and forecasting, in *Integrated Pest Management: Principles and Practice*, ed. by Abrol DP and Shankar U. CABI, Wallingford (2012).

11 Welch SM, Croft BA, Brunner JF and Michels MF, PETE: an extension phenology modeling system for management of multi-species pest complex. *Environ Entomol* **7**:487–494 (1978).

12 Jones VP, Using phenology models to estimate insecticide effects on population dynamics: examples from codling moth and obliquebanded leafroller. *Pest Manag Sci* **77**:1081–1093 (2021).

13 Jones VP, Brunner JF, Grove GG, Petit B, Tangren GV and Jones WE, A web-based decision support system to enhance IPM programs in Washington tree fruit. *Pest Manag Sci* **66**:587–595 (2010).

14 Damos P, Modular structure of web-based decision support systems for integrated pest management. A review. *Agron Sustainable Dev* **35**:1347–1372 (2015).

15 Samietz J, Graf B, Höhn H, Schaub L and Höpli HU, Phenology modelling of major insect pests in fruit orchards from biological basics to decision support: the forecasting tool SOPRA*. *EPPO Bull* **37**:255–260 (2007).

16 Binns MR, Nyrop JP and Wvd W, *Sampling and Monitoring in Crop Protection: The Theoretical Basis for Developing Practical Decision Guides*. CABI Pub, Wallingford, Oxon, UK; New York, N.Y (2000).

17 Pedigo LP and van Schaik JW, Time-sequential sampling: a new use of the sequential probability ratio test for Pest management decisions. *Bull Entomol Soc Am* **30**:32–36 (1984).

18 Jones VP and Wiman NG, Modeling the interaction of physiological time, seasonal weather patterns, and delayed mating on population dynamics of codling moth, *Cydia pomonella* (L.) (lepidoptera: Tortricidae). *Popul Ecol* **54**:421–429 (2012).

19 Jones VP, Wiman NG and Brunner JF, Comparison of delayed female mating on reproductive biology of codling moth and obliquebanded leafroller. *Environ Entomol* **37**:679–685 (2014).

20 Katagi T, Photodegradation of pesticides on plant and soil surfaces, in *Reviews of Environmental Contamination and Toxicology: Continuation of Residue Reviews*, ed. by Ware GW. Springer, New York: New York, NY, pp. 1–78 (2004).

21 Wald A, Sequential tests of statistical hypotheses. *Ann Math Stat* **16**: 117–186 (1945).

22 Rincon DF, Esch ED, Gutierrez-Illan J, Tesche M and Crowder DW, Predicting insect population dynamics by linking phenology models and monitoring data. *Ecol Model* **493**:110763 (2024).

23 Gaston KJ and McArdle BH, Measurement of variation in the size of populations in space and time: some points of clarification. *Oikos* **68**:357–360 (1993).

24 Perry JN and Woiwod IP, Fitting taylor power law. *Oikos* **65**:538–542 (1992).

25 R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2024).

26 Wheeler B and Pohlert T, SuppDists: Supplementary Distributions (2022).

27 Poisot T, Sachse R, Ashander J and Galili T, Use Data from Published Plots in R (2016).

28 R Core Team, Support for Parallel computation in R (2023).

29 Jones VP, Hilton R, Brunner JF, Bentley WJ, Alston DG, Barrett B *et al*., Predicting the emergence of the codling moth, *Cydia pomonella* (lepidoptera: Tortricidae), on a degree-day scale in north america. *Pest Manag Sci* **69**:1393–1398 (2013).

30 Boreau de Roincé C, Lavigne C, Ricard J-M, Franck P, Bouvier J-C, Garcin A *et al*., Predation by generalist predators on the codling moth versus a closely-related emerging pest the oriental fruit moth: a molecular analysis. *Agric For Entomol* **14**:260–269 (2012).

31 Mills N, Selecting effective parasitoids for biological control introductions: codling moth as a case study. *Biol Control* **34**:274–282 (2005).

32 Lacey LA, Thomson D, Vincent C and Arthurs SP, Codling moth granulovirus: a comprehensive review. *Biocontrol Sci Technol* **18**:639–663 (2008).

33 Margaritopoulos JT, Voudouris CC, Olivares J, Sauphanor B, Mamuris Z, Tsitsipis JA *et al*., Dispersal ability in codling moth: mark–release–recapture experiments and kinship analysis. *Agric For Entomol* **14**: 399–407 (2012).

34 Smith TJ, *Effects of Flight and Sublethal Pesticide Residues on Codling Moth, Cydia Pomonella (Linnaeus), Obliquebanded Leafroller, Choristoneura Rosaceana (Harris), and Convergent Ladybird Beetle, Hippodamia Convergens (Guérin-Méneville)*. Thesis of mater of science in entomology. Washington State University, Pullman, WA (2012).

35 Kershner J, Apple Farming in Washington. Washington State Historical Society https://www.historylink.org/File/21288 [accessed September 11 2023].

36 Riedl H, Croft BA and Howitt AJ, Forecasting codling moth phenology based on pheromone trap catches and physiological-time models. *Can Entomol* **108**:449–460 (1976).

37 Willson HR and Trammel K, Sex pheromone trapping for control of codling moth, oriental fruit moth, lesser appleworm, and three tortricid leafrollers in a new york apple orchard. *J Econ Entomol* **73**:291–295 (1980).

38 *National Agricultural Statistics Service, Press Release*. United States Department of Agriculture, Northwest Regional Field Office, Olympia, WA (2023). https://www.nass.usda.gov/Statistics_by_State/Washington/Publications/Fruit/2023/FR08_1.pdf [accessed Jan 2 2024].

39 Calkins CO and Faust RJ, Overview of areawide programs and the program for suppression of codling moth in the western USA directed by the United States Department of Agriculture—Agricultural Research Service. *Pest Manag Sci* **59**:601–604 (2003).

40 Parsa S, Morse S, Bonifacio A, Chancellor TCB, Condori B, Crespo-Pérez V *et al*., Obstacles to integrated pest management adoption in developing countries. *Proc Natl Acad Sci* **111**:3889–3894 (2014).

41 Peterson RKD, Higley LG and Pedigo LP, Whatever happened to IPM? *Am Entomol* **64**:146–150 (2018).

42 Rosenheim JA, Parsa S, Forbes AA, Krimmel WA, Law YH, Segoli M *et al*., Ecoinformatics for integrated pest management: expanding the applied insect ecologist's tool-kit. *J Econ Entomol* **104**:331–342 (2011).