ILCYM- Insect life cycle modeling: software for developing temperature-based insect phenology models with applications for regional and global pest risk assessments and mapping

Centro Internacional de la Papa, Integrated Crop Management Division, Apartado 1558, Lima 12, Peru; m.sporleder@cgiar.org

Abstract

Phenology models for insect pest species based on temperature are important analytical tools for predicting, evaluating, and understanding the dynamics of populations in ecosystems under a variety of environmental conditions, and more recently are also used in phytosanitary risk assessments. CIP had developed a temperature-driven, process-based full life-cycle phenology model for the potato tuber moth that predicted well life-table parameters for the pest in different agroecological zones. The model was further linked with geographic information systems (GIS) and atmospheric temperature databases, allowing simulation of pest risk indices on a worldwide scale. The approach used for developing and using a pest model can be principally adapted for other insect species. CIP further developed ILCYM, an open-source software package, which facilitates the development of phenology models on the basis of specific temperature-controlled experimental data by determining functions for temperature-driven processes in an insect species, i.e. development, mortality, reproduction, etc. compiling them into an overall pest model. A cohort up-dating algorithm and rate summation approach is used for simulating multidimensional age and stage structured populations. Tools for model validation are implemented. Further, ILCYM allows linking the models to a GIS environment; three generic risk indices (Establishment Index, Generation Index, and Activity Index) can be visualized. Life-table parameters can be forecasted over time for single locations. The paper presents the new up-dated version of ILCYM (version 2) and discusses how it might be used for supporting country-specific pest risk assessments, global climate change adaptation planning and improving pest management strategies.

Keywords: Insect Pests, Insect Pest Management, Temperature-dependency, Simulation Models, Computer-aided Decision Support, Potato Tuber Moth

Introduction

The distribution of insects and other poikilothermic animals is largely determined by climate. Insects cannot internally regulate their own temperature (exothermic organisms) and their development depends on the temperature to which they are exposed to in the environment. They require a certain amount of heat to develop from one developmental stage to another in their life-cycle (Uvarov 1931, Andrewartha and Birch 1955). Because of yearly variations in weather, calendar dates are faulty parameters for predicting pest population growth, and outbreaks and making management decisions. However, measuring the amount of heat accumulated over time provides a valuable physiological time scale that is biologically highly accurate. Phenology models are predicting the time of events in an organisms’ development. Such models for insect pests based on temperature are important analytical tools for evaluating, understanding and predicting the dynamics of pest populations in ecosystems under a variety of environmental conditions, and more recently are also used in phytosanitary risk assessments (Baker 1991, Jarvis and Baker 2001a, b).

CIP has developed a temperature-driven phenology model for the potato tuber moth, *Phthorimaea operculella* (Zeller) (Lepidoptera: Gelechidae) that predicted well life-table parameters for different agroecological zones (Sporleder et al. 2004); validated through field and laboratory data (Keller 2003). It was used to predict the establishment risk and potential pest activity in specific agroecologies according to temperature records (Kroschel and Sporleder 2006). Linked with geographic information systems (GIS) and atmospheric temperature the model allows simulating these risk indices on a worldwide scale or using it to predict future changes of these indices due to global warming (Sporleder et al. 2007, Sporleder et al. 2008). Possible applications of these models are manifold.
The approach used to develop and implement the *P. operculella* model can be principally used for other insect species. Therefore, CIP developed Insect Life Cycle Modeling software (ILCYM – version 2.0) to facilitate the development of further pest insect phenology models and provide analytical tools for studying pest population ecology. ILCYM software consists of two main modules: The first module (which can be called the “model builder”) facilitates researchers to develop insect pest phenology models based on temperature experiment data of a specific pest. This module also provides tools to analyze insect life-table and to validate developed models. The second module implements the phenology model in a GIS environment that allows for spatial simulation of pest activities (“pest risk mapping”).

**The purpose of ILCYM**

The objective of ILCYM software is to provide an open-source computer aided tool, especially for researchers in developing countries (p.e. entomologists, ecologists, researchers that are involved in integrated pest management (IPM) of agricultural and forest insect pests) that facilitate the development of generic phenology models based on temperature using advanced modeling techniques without having the mathematical knowledge, being experts in the field, or spending the time that is normally required to develop, implement and program a phenology model. We believe that the application of phenology models will lead to a better understanding of pests’ biology and ecology and in the long-term also support decision making in pest management programs. Such models might provide important information on quantitative pest population biology and “Pest risk mapping” provides important information in pest risk analysis (PRA). The present paper explains in briefly how the developed modeling approach works and how to handle and use the software package. Because ILCYM interactively leads the user through the steps of developing a PPM, or to conduct spatial simulations with a developed pest model, it is especially helpful for users who don’t want to learn programming languages before starting pest population modeling, and users who do not want to study mathematics before starting modeling.

**The model builder**

The model builder compiles full life cycle phenology models for insects species based on statistical analysis of experimental data, i.e. life-table or cohort-based development data, assessed over a range of constant temperatures. Different types of data might be analyzed; *a*) life-table data in which development time, survival, and reproduction are monitored for a group of individuals of the species of interest from egg to egg (i.e. fresh-laid eggs of the same age until the last egg of the offspring from these individuals is oviposited) - this data represent the full life history for a group of individuals of the population at a given temperature, or *b*) the data can be based on cohort studies. In the latter, a group of individuals of the same stage and age is monitored from the beginning to the end of a given life stage, i.e. fresh eggs to emerging larvae, or neonate larvae to pupation, etc. These experiments need to be conducted over a range of constant temperatures covering the temperature in which the insect species might develop; p.e. from 10°C to 30°C in 5°C intervals. If the female rate in the specie under study varies either with temperature to which the parent generation was exposed to or changes with female age, all offspring-eggs need to be reared to the adults’ stage for determination of sex.

If such data are generated, ILCYM provides a set of functions (predefined functions) that can be fitted to the data for describing variation in development times between the individuals of the population, temperature-dependency of median development times and survival in a given immature life stage, temperature-dependent adult senescence, and reproduction. The reproduction model is compiled assembling different functions (p.e. temperature-dependent total reproduction per female, age-dependent relative oviposition frequency) depending on how female rates are influenced by temperature and female age. The “model builder” interactively facilitates choosing the best-fitting functions for describing these temperature-driven processes in the insect specie’s development. These functions are model components, henceforth called “sub models” of the overall generic phenology model, or “modules” (the term “modules” is used because a sub-model can be replaced by another function that might describe better the specific process in the system) which are automatically implemented in the overall pest phenology model. The ILCYM “model builder” focuses strongly on “parameter determination” for models. The program produces outputs (statistics and graphs) that help to find the best parameters for a specific species on a statistical basis; parameters are stored in a database so that a model can be concluded part-by-part until its finalization.
Frequency distributions of insect development times are usually skewed toward the longer times and it is assumed that the intrinsic distributions of development times of an insect at different temperatures are of the same shape, i.e., the distributions at different temperatures will fall on top of each other when "normalized" by a selected value such as the mean or median of each distribution ("same shape property"); for further information on this theory see (Sharpe et al. 1977b, Curry et al. 1978a). In ILCYM three different probability density functions (PDF) of the general linear model type (GLM) are fitted to observed accumulated development frequencies in a parallel line approach using ln-development times as the explanatory variable. Fitted models (GLM) are the normal distribution, the logit model, and the complementary log-log (CLL) model having the following linearized forms:

Normal distribution model: 
\[ F(x) = \Phi(a_i + b \ln x) \]  
[1]

Logit model: 
\[ F(x) = \frac{1}{1 + \exp(-(a_i + b \ln x))} \]  
[2]

CLL model: 
\[ F(x) = 1 - \exp(-\exp(a_i + b \ln x)) \]  
[3]

where \( F(x) \) is the probability to complete development at time \( x \), and \( a_i \) and \( b \) are the parameters to be estimated; \( b \) is the common slope of the regression model representing the dispersion of development time in the life stage and \( a_i \)'s are the intercepts corresponding with temperatures \( i \). ILCYM selects the best-fitting model by examining well-known goodness of fit indicators, Akaike's Information Criterion (AIC) (Akaike 1973) and the Model Selection Criterion (MSC) (Scientist 1995). Median development times (\( T_{50} \)'s) for each constant temperature, \( i \), are calculated depending on the selected GLM using the following formulae:

If the normal distribution or logit model was selected:
\[ T_{50i} = \exp\left(-\frac{a_i}{b}\right) \]  
[4]

If the CLL model was selected:
\[ T_{50i} = \exp\left(-\ln(-\ln[0.5])a_i\right) \]  
[5]

Confidence intervals (95%) for the \( \ln(T_{50}) \) values are calculated by using the following formula:
\[ CI_{95i} = \ln T_{50i} \pm t_{95\%} \frac{1}{b} \sqrt{\text{parameter}} \]  
[6]

in which \( t_{95\%} \) is the value of the \( t \)-distribution corresponding to significance levels of 0.05 and the corresponding number of degrees of freedom, \( b \) is the parameter (slope) derived from the selected GLM (equations 1-3), and parameter is the number of parameters (\( a_i \)'s and \( b \)) estimated in the model.

Temperature-dependent development of insects from one stage to another does not follow a linear relationship and hence linear models are often not satisfactory to be used in phenology models, especially when the phenology model is likely to be applied for regions where the temperature fluctuates frequently below or above the temperature where development is in the linear range. Therefore, for describing the relationship between temperature and the development rate in a given insect life stage ILCYM provides a set of functions, including the modified version of the Sharpe & DeMichele model (Sharpe and DeMichele 1977b, Schoolfield et al. 1981) and its derivates (Ikemoto 2005b), Logan's models (Logan et al. 1976, Logan et al. 1979, Logan 1988), Stinner et al. model (Stinner et al. 1975), and other models (Taylor 1981, Lamb et al. 1984, Lamb 1992, Lactin et al. 1995, Brière et al. 1999) that have been employed successfully for many insect species for this purpose. Survivorship in immature life stages is calculated from the relative frequency of survivors. Different functions, including polynomial models, can be fitted by regression to describe the relation between mortality rate and temperature for each life stage. Users might choose the most appropriate model to be implemented in the overall model by evaluating goodness-of-fit criterions and residuals. Daily reproduction is modeled in ILCYM assembling two or more functions. Mean survival time of adults is generally determined for both sexes separately. As for immature
life stages, the inverse of the female life span (days–1) is plotted against temperature and a model similar to the functions used for describing temperature-dependent development employed to describe adult senescence. Senescence rate summation is then used to express the physiological age of females. Polynomial regression models are applied to find the relation between total oviposition per female and temperature. The cumulative proportion of eggs produced is plotted against female age expressed as normalized time, i.e. age in days divided by the median survival time, and different non-linear models can be fitted to the data to describe age-specific fecundity frequencies.

Once all functions required for describing the specie’s full life cycle are elected ILCYM compiles the overall model automatically. Therefore users do not need the knowledge for programming complex insect life cycle models. An example for an overall model, which was used for modeling potato tuber moth populations, is given in Figure 3. ILCYM uses a rate summation and cohort up-dating approach based on scheme proposed by Curry et al. (1978b) that was further described by Wagner et al. (1985) and Logan (1988). In published articles there is not so much discussion on including temperature-induced mortality in immature life stages and recruitment. Both are necessary for more realistic simulation and both are included in ILCYM (for further information on these merits see the chapters about “immature mortality” and “recruitment model”). The population is structured in different life stages, which are represented in a so called “box car train” (i.e. eggs [E], larvae [L], pupae [P], and adults [Af = adult females]) and into groups of individuals of the same age (i.e. cohorts), which are represented in a “box car” within each life stage (p.e. E0, E1, E2… , Ex; see Figure 3). Users can run simulations for one or multiple generations either in a deterministic or stochastic approach depending on the simulation purpose. Deterministic simulation is more useful for spatial simulations since it requires less computer time. For deterministic simulation, each “box car” contains three numbers, i.e. the age of the individuals in this group in days, the physiological age determined through development rate accumulation and the number of individuals in this age-group. When running the model each “box car” is up-dated in daily intervals, i.e. the value 1 is added to the age in days, the development rate calculated for the last 24-hour interval is added to the physiological age (rate summation) and the number of individuals in the cohort is minimized by the simulated daily mortality and the proportion which develops into the next life stage according to the temperature-dependent mortality function and age-dependent development distribution function, respectively.

For stochastic simulation, the model simulates development, mortality and reproduction of 100 individual. For each individual, simulation started from the egg stage (fresh egg). Development rates are accumulated as in deterministic modeling. Equation 1, 2 or 3, depending on the selected model, is used to calculate the probability that the individual develop from one stage to the next according to its physiological age. A random value between 0 and 1 is generated for each individual and used to determine the day when this particular insect develops to the next stage; i.e. the day when F(x) exceeded the random value. A second random value between 0 and 1 is generated to simulate survival of the individual; i.e. when the random value exceeded the value calculated by the mortality function included for describing temperature-dependent mortality the insects is considered to survive. This process is repeated for each immature life stage. The sex of an emerging adult is determined through generation of a further random value between 0 and 1; in case of a fixed female rate of 0.5 a random value <0.5 produced a female and a value >0.5 produced a male. The senescence rate of females is accumulated (=physiological age) and the relative oviposition frequency calculated for each day post emergence using the model determined to describe age-dependent oviposition frequencies. The resulting value is multiplied by the number of total oviposition resulting from the model selected describing temperature-depending total oviposition per female (no random variable is included as a stochastic component for variability in oviposition numbers per female). Adults died in the simulation when the physiological age exceeded a value that marks the end of the oviposition period.

For both, deterministic or stochastic simulation, a 15 min time step length is implemented in ILCYM. Model calculations are based on daily maximum and minimum air temperatures. Temperature in each 15 min time step is calculated by using a cosine function for half-day temperature predictions. The equation for the first half-day is:

\[
T_i = \frac{(Max - Min)}{2} \times \cos \left( \frac{\pi \times (i + 0.5)}{48} \right) + \frac{(Min + Max)}{2}
\]

in which \( T_i \) is the temperature (in °C) of time step \( i \) (\( i = 1, 2, 3, \ldots 48 \)), and \( Min \) and \( Max \) are daily minimum and maximum temperatures. The calculation is then repeated to obtain \( T_i \) for the second half-day employing the minimum temperature, \( Min \), of the following day in the equation.
Both deterministic and stochastic simulation over one generation produces a life-table for a given temperature (user input). Life-table parameters, i.e. net reproduction rate ($R_0$), mean generation time ($T$), intrinsic rate of natural increase ($r_m$), finite rate of increase ($\lambda$), and doubling time ($D_t$), can be calculated in ILCYM from modeling outputs as described in standard biological texts (Begon et al. 1990, Southwood and Henderson 2000). ILCYM determines the intrinsic rate of increase, $r_m$, employing the equation of Lotka (Birch 1948):

$$1 = \sum \exp(r_m \cdot x) \cdot l_x \cdot m_x$$

[8]

in which $x =$ age in days (including immature stages), $l_x =$ age-specific survival (including immature mortality) and $m_x =$ age-specific number of female offspring.

Figure 3. An schematic picture of an ILCYM temperature-based phenology/population model using a rate summation and cohort up-dating approach. See text for verbal description of the model.

Model validation

ILCYM provide several tools for validating developed phenology models. Repetitions of stochastic life table simulations for a given temperature allow determining confidence intervals for life table parameter means. Life tables of the same species established under fluctuating temperature can be used for comparing results (life table parameters) from real and simulated life tables. Individual functions (sub-models) used in the overall model can be examined and evaluated.

The environment for pest risk assessments

Once a full life-cycle pest model for a pest under study is completed the model can be used to simulate the species population growth potential spatially or for a given location over time using the GIS environment implemented in ILCYM. Three spatially referenced pest risk indices displaying the risk of establishment, numbers of generations per year, and an activity index can be computed spatially (see example in Figure 4). The simulation is based on daily maximum and minimum temperatures as inputs. For global or regional scale
simulations data based on WorldClim are preinstalled in ILCYM's pest risk modeling module. WorldClim is a set of global climate layers (grids) with a spatial resolution of 30 arc seconds (~1 km², downloadable at http://www.worldclim.org) and described in Hijmans et al. (2005). WordClim provides monthly aggregated climate variables. Because this aggregation might raise substantial problems of temporal scale, daily maximum and minimum temperatures are interpolated for each grid before simulation. For predicting the species' responses to climate change, similar maps can be generated for a scenarios of climate change using an atmospheric general circulation model (GCM) which is included in ILCYM. The latter data are described by Govindasamy et al. (2003) and forecast global climate for the year 2050. Specific spots (grids) in the map might be selected and the pest species population growth parameters can be simulated over time. For regional forecast at higher spatial resolution or higher data accuracy users are able load own temperature data, e.g. real data measured by meteorological stations for simulations over time or data sets from local meteorological Institutes for a given area for spatial simulations.

The approach for “risk mapping”, i.e simulating the risk of pest establishment and expansion, in ILCYM differs from the ‘match climate’ approach in which climate match functions seek out the potential exploitation of a non-indigenous invasive species to new areas by comparing the long-term meteorological data for each of a selected location where the species is absent with the location of origin or locations where the species prevails. In ILCYM the risk maps are simulated using the pests’ process-based phenology model that describes the basic physiological principals of insect species’ growth, i.e. its development, survival and reproduction.

![Figure 4. Survival index of *P. operculella* simulated globally using ILCYM. The index is 1 where all immature life-stages develop and survive throughout the year.](image)

**Conclusion**

ILCYM provides advanced insect modeling techniques and analysis tool that can be used efficiently by NARS scientists which are not experts in this field. The program interactively leads the user through the steps of developing a pest phenology/population model and aids conducting spatial simulations. Users do not need to learn programming languages or other specific mathematical knowledge to develop and design the model (abstraction); however, ILCYM restricts the modeler to certain predefined modeling designs and might not provide solutions for every problem.

**References**


