

Chapter 6

Flour Beetle Dynamics: A Case Study

Small beetles of the family Tenebrionidae, commonly known as flour beetles (genus *Tribolium*), can be found occasionally in their native forest habitats. But for thousands of years their preferred habitat has been milled flour and other stored grain products. These insects are hardy and prolific animals, traits which make them important agricultural pests and quite resistant to eradication. These same traits, however, make them easy to culture and manipulate in the laboratory. Flour beetles have been used extensively as animal models in genetics and population ecology. The details of their life cycle are well known.

Flour beetles undergo *complete metamorphosis* which encompasses four life cycle stages: egg, larva, pupa, and adult. The tiny eggs—as many as 500 from a single female—hatch in approximately four days under laboratory conditions and barely can be seen by the unaided eye. The larvae feed voraciously and eventually grow to the size of rice grains. After about 14 days, each larva encloses itself in a pupal case which consists of a thin outer covering, the pupal cuticle. During the pupal stage, the larva completely reorganizes itself to become an adult. The new adult emerges another 14 days later, and sheds its pupal case as little pieces of dried tissue called frass. See Fig. 6.1.

One might expect that populations of flour beetles, when cultured under constant environmental conditions, would grow exponentially at first, then equilibrate due to resource limitations. But they do not. Instead, their numbers typically exhibit sustained oscillations. The factor that regulates population size and causes these oscillations is cannibalism. In general, the mobile stages (larvae and adults) cannibalize the immobile stages (eggs and pupae). These cannibalistic encounters are crucial to our understanding of flour beetle population dynamics.

Cannibalism is not a significant source of nutrition for these beetles and is not caused by limitations of food or space. Instead, the incidence of this behavior is under genetic control. Varying degrees of cannibalism can be

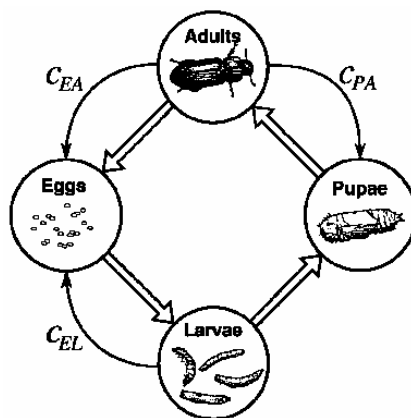


Figure 6.1: Life-cycle diagram for the flour beetle *T. castaneum*

artificially selected for in the laboratory to create different genetic strains. As far as the modeling process is concerned, *Tribolium* cannibalism is assumed to occur at random, when the mobile larvae and adults encounter immobile eggs and pupae. Apparently, mobile cannibalistic animals tunnel through the flour and eat whatever they encounter, be it flour or immobile forms of their own species. Hence, the rate of cannibalism is inversely proportional to habitat volume.

6.1 Data

The time series data in Table 1 are from laboratory populations of the corn oil sensitive strain of *Tribolium castaneum* Herbst, cultured by Desharnais and Costantino (1980) in half pint milk bottles containing 20g of corn oil media (90% wheat flour, 5% brewer's yeast, and 5% liquid corn oil). The cultures were kept in an unlighted incubator at $33 \pm 1^\circ$ C and a relative humidity of $56 \pm 11\%$. Every two weeks, all life cycle stages (except eggs) were censused, and the animals (including eggs) were returned to the incubator in fresh media. Small (feeding) larvae were counted as L-stage, large (nonfeeding) larvae, pupae, and callow (immature) adults were counted as P-stage, and sexually mature adults were counted as A-stage.

6.2 The LPA model

6.2.1 Simplifying assumptions

We must now simplify the system by making assumptions. Ideally, the assumptions should address two kinds of questions. First, what are the main deterministic trends in the system, or better, what deterministic mechanisms are most important in driving the system? Second, what kind of process error is in the system, and what kind of measurement error is in the data?

Deterministic

- A1) The egg stage can be ignored.
- A2) In the absence of cannibalism, the per capita larval recruitment rate is a constant $b > 0$ larvae per adult per unit time.
- A3) In the absence of cannibalism, the per capita death rate for larvae is μ_l per unit time, where $0 < \mu_l < 1$.
- A4) In the absence of cannibalism, the per capita death rate for pupae is 0.
- A5) In the absence of cannibalism, the per capita death rate for adults is μ_a per unit time, where $0 < \mu_a < 1$.
- A6) In the presence of cannibalism, the larval recruitment rate is decreased due to cannibalism of eggs by larvae. It is decreased by a Ricker factor of the form e^{-cL} .
- A7) In the presence of cannibalism, the larval recruitment rate is decreased due to cannibalism of eggs by adults. It is decreased by a Ricker factor of the form e^{-cA} .
- A8) In the presence of cannibalism, the pupal per capita survival rate has the form e^{-cA} due to the cannibalism of pupae by adults.

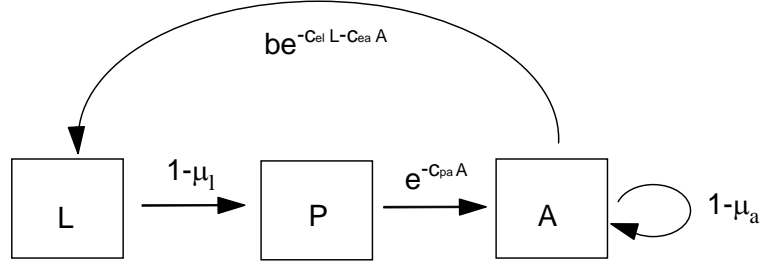
Stochastic

Measurement error

- A9) There is no measurement error.

Process error

- A10a) Environmental stochasticity is the main source of noise in the system.
- A10b) (*Alternative hypothesis*) Demographic stochasticity is the main source of noise in the system.

Figure 6.2: Leslie diagram for *T. castaneum*

- A11) Stochastic perturbations are uncorrelated across life cycle stages, and are also uncorrelated in time.

The above premises are called modeling assumptions because they are translated into model equations. However, it is important to remember that modeling assumptions are always *biological hypotheses*, which are tested when the model is evaluated against data. In fact, assumptions A10a and A10b are alternative hypotheses. These two alternatives will give rise to two alternative models.

6.2.2 The deterministic LPA model

Let

$$\begin{aligned} L_t &= \text{number of L-stage animals at time } t \\ P_t &= \text{number of P-stage animals at time } t \\ A_t &= \text{number of A-stage animals at time } t \end{aligned}$$

Then assumptions (A1)-(A8) can be conceptualized in the following Leslie Diagram:

The corresponding nonlinear Leslie model is known as the ‘‘LPA model’’:

$$\begin{aligned} L_{t+1} &= bA_t e^{-c_{el}L_t - c_{ea}A_t} \\ P_{t+1} &= (1 - \mu_l) L_t \\ A_{t+1} &= P_t e^{-c_{pa}A_t} + (1 - \mu_a) A_t. \end{aligned} \tag{6.1}$$

6.2.3 NLAR models

Given a discrete-time model of the form

$$x_{t+1} = f(x_t), \quad (6.2)$$

where x_t is a real number or a vector of real numbers, how does one include stochasticity? If noise is additive (as it is in the bone model in Chapter 2), then the stochastic version of the model is

$$X_{t+1} = f(X_t) + \sigma E_t, \quad (6.3)$$

where X_t is a random variable and E_t is a random standard normal variable. When the standard deviation σ is set to zero, the deterministic skeleton (6.2) is recovered. Stochastic models of the form (6.3) are called *nonlinear autoregressive* (NLAR) models, and have nice statistical properties. In matter of fact, however, process noise, whether environmental or demographic, is not in general additive.

Two results from statistics make it possible to approximate environmental and demographic noise with NLAR models. First, environmental noise is approximately additive on the log scale. Second, demographic noise is approximately additive on the square root scale. That is, the NLAR model

$$\ln X_{t+1} = \ln(f(X_t)) + \sigma E_t$$

does a good job if the noise in the system is mostly environmental, while

$$\sqrt{X_{t+1}} = \sqrt{f(X_t)} + \sigma E_t$$

does a good job if the noise is mostly demographic.

In general, if a stochastic system with deterministic skeleton (6.2) can be approximated by an NLAR model of the form

$$g(X_{t+1}) = g(f(X_t)) + \sigma E_t$$

for some function g , then g is called a *variance stabilizing transformation* for that system. So, the log and the square root functions are the variance stabilizing transformations for environmental and demographic noise, respectively.

6.2.4 The alternative stochastic LPA models

Under assumption A10a, environmental stochasticity is the main source of variability in the system. The appropriate NLAR model is therefore

$$\begin{aligned} \ln(L_{t+1}) &= \ln(bA_t e^{-c_{el}L_t - c_{ea}A_t}) + \sigma_1 E_{1t} \\ \ln(P_{t+1}) &= \ln((1 - \mu_l)L_t) + \sigma_2 E_{2t} \\ \ln(A_{t+1}) &= \ln(P_t e^{-c_{pa}A_t} + (1 - \mu_a)A_t) + \sigma_3 E_{3t}, \end{aligned}$$

where the E_t are standard normal random variables with mean zero and standard deviation one. This model can be rewritten as

$$\begin{aligned} L_{t+1} &= bA_t e^{-c_{el}L_t - c_{ea}A_t} e^{\sigma_1 E_{1t}} \\ P_{t+1} &= (1 - \mu_l) L_t e^{\sigma_2 E_{2t}} \\ A_{t+1} &= (P_t e^{-c_{pa}A_t} + (1 - \mu_a) A_t) e^{\sigma_3 E_{3t}}. \end{aligned} \quad (6.4)$$

This is called the “environmental noise LPA model”.

Under assumption A10b, demographic stochasticity is the main source of variability in the system. The appropriate NLAR model is therefore

$$\begin{aligned} \sqrt{L_{t+1}} &= \sqrt{bA_t e^{-c_{el}L_t - c_{ea}A_t} + \sigma_1 E_{1t}} \\ \sqrt{P_{t+1}} &= \sqrt{(1 - \mu_l) L_t + \sigma_2 E_{2t}} \\ \sqrt{A_{t+1}} &= \sqrt{P_t e^{-c_{pa}A_t} + (1 - \mu_a) A_t + \sigma_3 E_{3t}}, \end{aligned}$$

where again the E_t are standard normal random variables with mean zero and standard deviation one. This model can be rewritten as

$$\begin{aligned} L_{t+1} &= \left(\sqrt{bA_t e^{-c_{el}L_t - c_{ea}A_t} + \sigma_1 E_{1t}} \right)^2 \\ P_{t+1} &= \left(\sqrt{(1 - \mu_l) L_t + \sigma_2 E_{2t}} \right)^2 \\ A_{t+1} &= \left(\sqrt{P_t e^{-c_{pa}A_t} + (1 - \mu_a) A_t + \sigma_3 E_{3t}} \right)^2. \end{aligned} \quad (6.5)$$

This is called the “demographic noise LPA model”. Note that if the random variables happen to be negative enough, the quantities inside the squares on the right hand side of equation (6.5) can be negative. In this case they should be set to zero.

Let (l_t, p_t, a_t) and $(l_{t+1}, p_{t+1}, a_{t+1})$ be the actual observed data triples at time t and $t + 1$. Since we are thinking of the models (6.4) and (6.5) as a surrogates for the biological system, we can think of these data triples as realizations of the random variable triple (L, P, A) . The conditional *one step residual errors* at time $t + 1$ are the actual measurements $(l_{t+1}, p_{t+1}, a_{t+1})$ at time $t + 1$ minus the values predicted by the deterministic skeleton, given the data point (l_t, p_t, a_t) at time t . For the environmental LPA model, we must consider the *log-residuals*

$$\begin{aligned} \varepsilon_{1t} &= \ln(l_{t+1}) - \ln(ba_t e^{-c_{el}l_t - c_{ea}a_t}) \\ \varepsilon_{2t} &= \ln(p_{t+1}) - \ln((1 - \mu_l) l_t) \\ \varepsilon_{3t} &= \ln(a_{t+1}) - \ln(p_t e^{-c_{pa}a_t} + (1 - \mu_a) a_t) \end{aligned}$$

For the demographic LPA model, we must consider the *square root-residuals*

$$\begin{aligned} \varepsilon_{1t} &= \sqrt{l_{t+1}} - \sqrt{ba_t e^{-c_{el}l_t - c_{ea}a_t}} \\ \varepsilon_{2t} &= \sqrt{p_{t+1}} - \sqrt{(1 - \mu_l) l_t} \\ \varepsilon_{3t} &= \sqrt{a_{t+1}} - \sqrt{p_t e^{-c_{pa}a_t} + (1 - \mu_a) a_t} \end{aligned}$$

Note that the residuals ε_{it} themselves are realizations of the random variables $\sigma_i E_{it}$ in the stochastic models. Therefore, the alternative hypotheses A10a and A10b boil down to the hypotheses that either a) *the log-residuals are normally distributed with mean zero*; or b) *the square root-residuals are normally distributed with mean zero*.

6.3 Model parameterization

There are six parameters in the LPA model: $b, \mu_l, \mu_a, c_{el}, c_{ea}, c_{pa}$. We will use the historical beetle data in Table 1 to estimate the values of these parameters by the method of nonlinear least squares. Ideally, we should randomly divide the data set into two parts. One part would be used for model calibration (parameter estimation), and the other part would be reserved for independent model evaluation (validation). In this case, however, we are estimating six parameters, so we need a lot of data to calibrate the model. We will therefore use all of the historical data for model calibration. In order to validate the model, we will use it to make predictions which we will then test against a new data set.

We use the method of conditioned least squares (CLS) to parameterize the model.

6.3.1 Conditioned least squares parameter estimation (CLS)

Let $\theta = (b, \mu_l, \mu_a, c_{el}, c_{ea}, c_{pa})$ be the vector of parameter values, and suppose the data time series is of duration $t = 0, 1, 2, \dots, q$. The residual sum of squares is

$$RSS(\theta) = \sum_{t=0}^{q-1} \varepsilon_{1t}^2 + \sum_{t=0}^{q-1} \varepsilon_{2t}^2 + \sum_{t=0}^{q-1} \varepsilon_{3t}^2$$

where the ε_{it} are the one-step residual errors. Thus, for the environmental noise LPA model, we have

$$\begin{aligned} RSS(\theta) = & \sum_{t=0}^{q-1} [\ln(l_{t+1}) - \ln(ba_t e^{-c_{el}l_t - c_{ea}a_t})]^2 + \\ & \sum_{t=0}^{q-1} [\ln(p_{t+1}) - \ln((1 - \mu_l)l_t)]^2 + \\ & \sum_{t=0}^{q-1} [\ln(a_{t+1}) - \ln(p_t e^{-c_{pa}a_t} + (1 - \mu_a)a_t)]^2. \end{aligned}$$

For the demographic noise LPA model, we have

$$\begin{aligned}
 RSS(\theta) = & \sum_{t=0}^{q-1} \left[\sqrt{l_{t+1}} - \sqrt{ba_t e^{-c_{el}l_t - c_{ea}a_t}} \right]^2 + \\
 & \sum_{t=0}^{q-1} \left[\sqrt{p_{t+1}} - \sqrt{(1 - \mu_l)l_t} \right]^2 + \\
 & \sum_{t=0}^{q-1} \left[\sqrt{a_{t+1}} - \sqrt{(p_t e^{-c_{pa}a_t} + (1 - \mu_a)a_t)} \right]^2.
 \end{aligned}$$

RSS must be minimized as a function of θ in order to get the best model fit. The minimizing parameters for RSS are called the *conditioned least squares* (CLS) parameters. The CLS parameters for the environmental noise model (6.4) are listed in Table 2. In exercise (1) you will complete Table 2 by parameterizing the demographic noise model.

model	\hat{b}	$\hat{\mu}_l$	$\hat{\mu}_a$	\hat{c}_{el}	\hat{c}_{ea}	\hat{c}_{pa}	RSS	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\sigma}_3$
enviro	10.57	0.5129	0.1065	0.009463	0.009649	0.01817	54.39	0.5256	0.6545	0.1054
demo										

Table 2

The case study is continued in the exercises.

6.4 Exercises

1. **Model selection** In this problem you will parameterize the two alternative stochastic models presented in this chapter.
 - (a) Reproduce all the values in Table 2 for the environmental noise LPA model.
 - (b) Fill in the second row of Table 2 for the demographic noise LPA model.
 - (c) The alternative hypotheses A10a and A10b boil down to the hypotheses that either a) the log-residuals are normally distributed with mean zero; or b) the square root-residuals are normally distributed with mean zero. Using Excel, construct two histograms: one for the log-residuals and one for the square root-residuals. Print these out. Which stochastic model appears to be better?
 - (d) Can you find a statistical test to make your answer in (c) statistically rigorous? State the name of the test. For extra credit, carry out the test.
2. **Model predictions** In this problem you will explore the dynamics of the LPA model and make testable predictions.

- (a) Write a program to simulate the deterministic LPA model. Use the parameter values estimated from the environmental noise model (Table 2). Produce a time series of length 20, given the initial condition $(0, 0, 30)$. Plot the time series for L , P , and A on the same graph. Attach (i) your program, (ii) the time series (as a matrix with 20 rows and 3 columns), and (iii) the graph.
 - (b) Write a program to simulate the environmental noise LPA model (6.4). You can include the stochastic model in your program written in (2a), or you can write a separate program. Produce a typical stochastic time series of length 20, given the initial condition $(0, 0, 30)$. Plot the time series for L , P , and A on the same graph. Attach (i) your program, (ii) the time series (as a matrix with 20 rows and 3 columns), and (iii) the graph.
 - (c) Use the computer to graph the bifurcation diagram for larval numbers, using μ_a as a bifurcation parameter with range $0 \leq \mu_a \leq 1$.
 - (d) The bifurcation diagram in (2c) makes predictions about the behavior of the flour beetle dynamics as a function of the adult death rate μ_a . Design a laboratory experiment to test these predictions. Show your experimental design to a biologist and ask for feedback. (Hint: Your design will involve a control, treatments, replicates, and experimental protocol. Note that in the protocol you will need to specify a method for manipulating the parameter μ_a in the laboratory.)
 - (e) Prepare time series graphs that show the deterministic model predictions for each of your experimental treatments.
 - (f) Prepare time series graphs that show typical stochastic model simulations for each of your experimental treatments.
3. **Model validation: test of predictions** In this problem you will test the predictions of the LPA model.
- (a) @
4. In this problem you will carry out a stability analysis for the extinction state of the LPA model.
- (a) Prove that the extinction state $(0, 0, 0)$ is an equilibrium for the LPA model.
 - (b) Find the linearization of the LPA model at the equilibrium $(0, 0, 0)$.
 - (c) Find the eigenvalues of the linearization. The eigenvalues will depend on the parameters. (If you do not know how to find the determinant of a 3×3 matrix, you can look this up in an introductory linear algebra text or calculus text.)
 - (d) Find conditions on the parameters that guarantee stability / instability of the extinction state $(0, 0, 0)$.

- (e) Define the *inherent net reproductive number* n by $n = b(1 - \mu_l) / \mu_a$. This is the expected number of adults arising from one adult in the absence of cannibalism. Rewrite your conditions from part (4d) in terms of n . Give a biological interpretation to these conditions.