

## DYNAMICAL EFFECTS OF PLANT QUALITY AND PARASITISM ON POPULATION CYCLES OF LARCH BUDMOTH

PETER TURCHIN,<sup>1</sup> SIMON N. WOOD,<sup>2</sup> STEPHEN P. ELLNER,<sup>3</sup> BRUCE E. KENDALL,<sup>4</sup> WILLIAM W. MURDOCH,<sup>5</sup>  
ANDREAS FISCHLIN,<sup>6</sup> JÉRÔME CASAS,<sup>7</sup> EDWARD MCCAULEY,<sup>8</sup> AND CHERYL J. BRIGGS<sup>9</sup>

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Connecticut 06269 USA

<sup>2</sup>Mathematical Institute, North Haugh, St. Andrews Fife, KY16 9SS, UK

<sup>3</sup>Department of Ecology and Evolutionary Biology, Corson Hall, Cornell University, Ithaca, New York 14853 USA

<sup>4</sup>School for Environmental Science and Management, University of California, Santa Barbara, California 93106 USA

<sup>5</sup>Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, California 93106 USA

<sup>6</sup>Institute of Terrestrial Ecology, Swiss Federal Institute of Technology Zurich ETHZ,  
CH-8952 Schlieren/Zurich, Switzerland

<sup>7</sup>University of Tours, IRBI-CNRS UPRESA 6035, F-37200 Tours, France

<sup>8</sup>Ecology Division, Department of Biological Sciences, University of Calgary, Calgary, Canada T2N 1N4

<sup>9</sup>Department of Integrative Biology, University of California, Berkeley, California 94720 USA

**Abstract.** Population cycles have been remarkably resistant to explanation, in part because crucial experiments are rarely possible on appropriate spatial and temporal scales. Here we show how new approaches to nonlinear time-series analysis can distinguish between competing hypotheses for population cycles of larch budmoth in the Swiss Alps: delayed effects of budmoth density on food quality, and budmoth–parasitoid interactions. We re-examined data on budmoth density, plant quality, and parasitism rates. Our results suggest that the effect of plant quality on budmoth density is weak. By contrast, a simple model of budmoth–parasitoid interaction accounts for 90% of the variance in budmoth population growth rates. Thus, contrary to previous studies, we find that parasitoid–budmoth interaction appears to be the dominant factor driving the budmoth cycle.

**Key words:** forest insects; larch budmoth; nonlinear time-series analysis; parasitism; plant quality; population cycles; Switzerland; trajectory matching; *Zeiraphera diniana*.

### INTRODUCTION

Population dynamics of forest insects have been a focus of numerous controversies in ecology. One area of contention is whether outbreak cycles are driven by density dependent processes or by exogenous forces (Royama 1981, Turchin 1990, Hunter and Price 1998). Within the density-dependence camp, arguments rage between proponents of intrinsic factors, e.g., maternal effects (Ginzburg and Taneyhill 1994) vs. trophic hypotheses (Berryman 1995).

Because dynamics occur on large spatial and temporal scales, in many cases it is not feasible, or at least it is very difficult and expensive, to test hypotheses about dynamic mechanisms experimentally. This is probably why there are so few experimental studies that successfully tested mechanistic hypotheses of population cycles, and only one, as far as we know, in a forest insect (snowshoe hares, Krebs et al. 1995; Red Grouse, Hudson et al. 1998; voles, Korpimäki and Norrdahl 1998; southern pine beetles, Turchin et al. 1999). If the only methodologically valid approach for elucidating the causes of population cycles is manipulative experiments in the field, then we should expect very slow progress in this area of ecology. On the other

hand, Kendall et al. (1999) showed, for a laboratory population, how a combination of time-series analysis and mechanistic modeling could serve as an equally valid approach for distinguishing between alternative hypotheses to explain population cycles. In this paper we show that this approach can be applied successfully to populations in the field, to explain the cause of cycles in a forest insect pest, the larch budmoth (LBM).

Larch budmoth (*Zeiraphera diniana*) populations in the Swiss Alps are among the most remarkable recorded cycles, oscillating with a period of 8–9 yr and ~100 000-fold change between peak and trough density (Fig. 1). Ecological theory offers several potential explanations for such regular population cycles in forest insects: parasitoid–host interaction (Hassell 1978), delayed effects of plant quality (Fischlin and Baltensweiler 1979, Edelstein-Keshet and Rausher 1989), pathogen–host interaction (Anderson and May 1980), and maternal effects (Ginzburg and Taneyhill 1994). A large empirical research effort has been devoted to studying the LBM cycle and trying to identify the causal mechanisms (see review in Baltensweiler and Fischlin 1988). As a result of this effort, the last two hypotheses can be rejected in this system. There is no empirical support for maternal effects, and although the granulosis virus infection was observed to cause substantial mortality during the first two intensively studied outbreaks (50% in the peak year of 1955 and

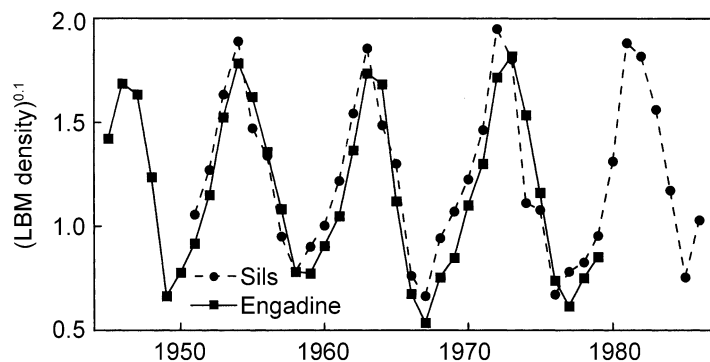


FIG. 1. Population fluctuations of larch budmoth density in the Upper Engadine Valley (for explanation of "Engadine" vs. "Sils" data, see *Methods: Sources of data*).

25% in the 1964 peak), subsequent outbreaks collapsed without being accompanied by any viral mortality (Baltensweiler and Fischlin 1988).

Plant quality (e.g., raw fiber and protein content) and parasitism have the necessary delayed effects to induce cycles. It takes two or more years for foliage quality to recover after heavy defoliation, which accompanies a budmoth population peak. Furthermore, field and laboratory bioassays show that poor food quality has a strong effect on budmoth survival and reproduction (Benz 1974, Omlin 1977, Fischlin 1982). Thus, the current explanation of budmoth cycles is based on their interaction with foliage quality (Baltensweiler and Fischlin 1988, Berryman 1999).

Measured parasitism rates vary between lows of 1–5% and highs of 80–90% (Delucchi 1982), and maximum parasitism rates are achieved ~2 yr after budmoth peaks. Previous studies concluded that mortality by parasitoid wasps does not cause the cycle, but merely tracks the budmoth population (Delucchi 1982, Baltensweiler and Fischlin 1988). However, as our analysis will show, this rejection of the parasitoid hypothesis was inappropriate.

The goal of our paper is to empirically distinguish between the two rival hypotheses for which there is empirical evidence in the larch budmoth (LBM) system: plant quality vs. parasitism. To do this, we formulated dynamic models of the LBM system that embody the rival hypotheses. We then use these mechanism-based models as statistical models for the purpose of testing hypotheses about which mechanism is best supported by the available data. Conceptually, we are simply using the standard approach to statistical hypothesis testing, but the highly nonlinear nature of the models complicates the implementation of the approach, as described below in *Methods: Models*.

## METHODS

### *Sources of data*

Systematic population census of larch budmoth in the Upper Engadine Valley (Switzerland) started in 1949 and with minor modifications continued until 1977 (Auer 1977, Baltensweiler and Fischlin 1988) (lesser quality data are also available for 1945–1948).

Data were collected at multiple sites throughout the valley separately (Auer 1977; these data are tabulated in Fischlin 1982), but because all sites oscillated in close synchrony, we can average them into one time series that we call "Engadine." The density of budmoth larvae (third instar) is expressed in number per unit of 1 kg of branches with foliage. For the period of 1952–1976 (with one year, 1968, missing) we have data on the percentage of larvae parasitized, also averaged over multiple sites (Delucchi 1982) omitting several data points that were interpolated by the author. Although numerous parasitoids are associated with larch budmoth, parasitism is dominated by the ichneumonid *Phytodietus griseanae* and three eulophid species, whose fluctuations are correlated ( $R^2 = 0.94$  between log *Phytodietus* abundance and log eulophid abundance,  $P < 0.001$ ). After 1977, sampling in the Upper Engadine valley continued on a reduced scale. At one site, Sils, data were collected in an uninterrupted sequence from 1951 to 1992: we refer to these as the "Sils" data (Baltensweiler 1993; A. Fischlin, unpublished data). During the period of 1961–1992, needle lengths of trees at Sils were also measured. Needle length is a good index of plant quality because it is well correlated with raw fiber and protein content of larch needles (Omlin 1977, Fischlin 1982). Furthermore, bioassay data of Benz (1974) indicated that needle length has a strong effect on larval survival and pupal mass (and pupal mass is closely connected to adult fecundity). Turchin (2003) calculated that the needle length of foliage with which LBM larvae were fed in the Benz (1974) bioassays explained 86% variance in a measure of LBM fitness (the product of larval survival and adult fecundity).

### *Models*

To decrease the chance that an inappropriate modeling choice would bias our results against the plant quality hypothesis, we modeled the effect of plant quality on budmoth dynamics with two alternative functional forms. The first (the "nonlinear" version) is a modified Ricker model in which the discrete rate of budmoth population increase is a hyperbolic (saturating) function of plant quality (for equations, see Table

TABLE 1. Dynamic models and model equations for larch budmoth populations in Switzerland.

Model and parameters	Equation
Plant quality I (the “nonlinear” version)	
$N_t$ : budmoth density	$N_{t+1} = \lambda_0 N_t \frac{Q_t}{\delta + Q_t} \exp[-gN_t]$
$Q_t$ : plant quality index	$Q_{t+1} = (1 - \alpha) \left( 1 - \frac{N_t}{\gamma + N_t} \right) + \alpha Q_t$
Plant quality II (alternative $N_t$ equation, the “linear” version)	
$N_t$ : budmoth density	$N_{t+1} = N_t \exp[u + vL_t - gN_t]$
$L_t$ : needle length	
Parasitoid–host	
$N_t$ : budmoth density	$N_{t+1} = \lambda_0 N_t \exp \left[ -gN_t - \frac{aP_t}{1 + ahN_t + awP_t} \right]$
$P_t$ : parasitoid density	$P_{t+1} = \phi N_t \left\{ 1 - \exp \left[ -\frac{aP_t}{1 + ahN_t + awP_t} \right] \right\}$
Tritrophic I (full model)	
$N_t$ : budmoth density	$N_{t+1} = \lambda_0 N_t \frac{Q_t}{\delta + Q_t} \exp \left[ -gN_t - \frac{aP_t}{1 + ahN_t + awP_t} \right]$
$P_t$ : parasitoid density	$P_{t+1} = \phi N_t \left\{ 1 - \exp \left[ -\frac{aP_t}{1 + ahN_t + awP_t} \right] \right\}$
$Q_t$ : plant quality index	$Q_{t+1} = (1 - \alpha) \left( 1 - \frac{N_t}{\gamma + N_t} \right) + \alpha Q_t$
Tritrophic II (simplified “linear” version)	
$N_t$ : budmoth density	$N_{t+1} = N_t \exp \left[ u + vL_t - \frac{aP_t}{1 + awP_t} \right]$
$P_t$ : parasitoid density	
$L_t$ : needle length	

*Notes:* All state variables and parameters are bounded below by 0. In addition,  $Q_t$ ,  $\alpha$ ,  $\phi$ ,  $\gamma$ ,  $h$ , and  $w$  are bounded above by 1. The proportion of larvae parasitized was computed as  $S_t = [1 - \exp\{-aP_t/(1 + ahP_t + awN_t)\}]$ . We assumed that the field data underestimate (some instances of parasitism are missed because field-collected larvae must be reared in the lab) so the expected observed parasitism rate was assumed to be  $\gamma S_t$  where  $\gamma \leq 1$  is a parameter to be fitted. The measure of plant quality in our data is needle length, and we assumed a linear relationship between needle length and actual plant quality,  $Q_t = b + cL_t$ , where  $b$  and  $c$  are parameters to be fitted.

1). The parameter  $g$  within the exponent quantifies direct density dependence. The second (the “linear” version) assumes that the realized per capita rate of budmoth population change ( $r_t = \log(N_t/N_{t-1})$ ) is a linear function of the plant quality index and LBM density. In the plant quality equation, quality next year is decreased by current-year herbivory in a hyperbolic fashion. Once herbivore damage becomes negligible (as a result of collapse of the budmoth population) quality increases gradually to the maximum, and the speed of recovery is regulated by the parameter  $\alpha$ .

The host–parasitoid model is based on the standard Nicholson–Bailey framework, but with self-regulation in the host and a general form of functional response (Beddington 1975) that combines effects of prey saturation and parasitoid mutual interference.

The tritrophic model simply combines the nonlinear version of the plant quality model and the parasitoid–host model. In addition, we investigated a simplified version of the tritrophic model, which combined the

“linear” version of the plant-quality model with the simplified form of parasitoid functional response (all equations in Table 1).

#### General philosophy of analysis

Our approach is based on the standard logic of statistical hypothesis testing. However, because our hypotheses are rich in mechanism, the mathematical models embodying these hypotheses are substantially more nonlinear and detailed than is usual in statistical hypothesis testing. As noted in the *Introduction*, we are focusing on two “monocausal” hypotheses (food quality and parasitism). In addition, it is possible that a composite explanation is correct (both food and parasitism jointly explain the LBM cycle). Once we have translated all hypotheses into models, we ask whether the observed data could plausibly have been generated by either of the simpler models or whether we require the more complicated model to explain them. The process involves fitting each model to the observed data

(estimating its parameters), so that model comparisons can be made with each model performing as well as its basic structure will allow (of course, both fitting and comparison must be performed using the same measure of fit). This is precisely the approach that underlies almost all statistical hypothesis testing.

The fact that standard statistical methods follow the framework that we use here can be seen by considering the theory underlying a familiar one-way ANOVA. In this case two models are compared, a null model that the data are all observations of random variables having the same mean, and an alternative model that they are observations of random variables whose mean depends on the level of some factor. Both models are fit to the data by least squares, and their fit compared by the difference in their residual sum of squares. The  $P$  value of the ANOVA is obtained by considering the distribution of difference in sum of squares that would be expected if the null model were true: we have to abandon the simpler model only if the improvement in model fit of the alternative is larger than could plausibly have occurred if the null model is true.

The models that we want to compare are highly nonlinear and this causes some difficulty in formulation and fitting. In particular, we must deal with two types of stochasticity: process error and measurement error. Process error is variability due to factors that are not explicitly modeled ("noise"). It affects the true values of population density. Measurement error is variability due to imprecisely measuring population density. Simultaneous treatment of both types of variability with highly nonlinear models is a relatively new development, requiring computationally intensive methods that have not yet received much testing (Ellner et al. 2002). We therefore chose to work with two complementary measures of fit, each addressing only one source of error. The first measure of model fit (one-step-ahead fitting) considers the average ability of the model to predict the population change over a year, given measurements of the system state at the beginning of the year: this amounts to neglecting measurement error. The second measure of model fit (global fitting, also known as "trajectory matching") measures the ability of the model to predict the entire series of data when iterated forward deterministically from starting conditions, which are also estimated by fitting: this amounts to neglecting process error, but yields accurate estimates if the period of oscillations is nonetheless constant or nearly so, as in LBM (Kendall et al. 1999, Wood 2001b). If conclusions based on both measures concur then we can have some confidence that neither of their key approximations is critical (Hilborn and Mangel 1997).

A key practical issue (especially for the trajectory-matching approaches) is that model fitting and inference are not straightforward for highly nonlinear, parameter-rich models when using relatively small data sets. In particular, the model fitting objectives (which

measure model fit) can show a good deal of structure, including local minima, at a variety of scales. Much of this variability is of no statistical significance, in the sense that it is of a scale below that of the sampling variability of the fitting objective itself. Nonetheless, such fine-scale structure can make model fitting difficult and tends to spoil the performance of the large-sample approximations such as Likelihood Ratio tests that would usually allow for statistical inferences about the fitted models. We deal with this issue by using bootstrapping methods, both as a means of escaping from small-scale local minima during fitting (see Wood 2001a) and for statistical comparison of models.

Formally, then, we treat each model comparison as testing a null hypothesis that a simple model is correct against an alternative that a more complicated model is needed. The sequence of steps is as follows:

- 1) Both models are fitted to data, and their difference in fit,  $\Delta$ , is noted.
- 2) We create replicate simulated data sets under the null hypothesis that the simpler model is true by bootstrap resampling the residuals from the simpler model and adding this resample to the fitted values from the simpler model.
- 3) Both models are fitted to each bootstrap-simulated data set and the difference in their fit  $\Delta^*$  is noted for each replicate.
- 4) The  $\Delta$  value is compared to the distribution of  $\Delta^*$  values generated by simulation—if it lies well within the distribution then the null hypothesis that the simple model is correct can be accepted, otherwise it will be rejected in favor of the alternative model. A  $P$  value can be obtained by assessing the proportion of  $\Delta^*$  values which are more favorable to the alternative model than the observed  $\Delta$ .

#### *One-step-ahead fitting*

The response variable was defined as the realized per capita rate of population change, defined as  $r_t = \ln(N_t) - \ln(N_{t-1})$  (Turchin 1990, Berryman 1991, Royama 1992). Each model used (see Table 1) could be recast to predict this quantity given measurements of some or all of  $N_{t-1}$ ,  $L_{t-1}$ , and  $P_{t-1}$  (LBM density, needle length [an index of food quality], and parasitism rate, respectively), depending on the model in question. Model parameters were estimated by minimizing the mean square difference between model predictions of  $r_t$  and observed  $r_t$  using the nonlinear regression routines in STATISTICA (Statsoft 1999).

As explained in *Methods: General philosophy of analysis*, bootstrap resampling of the model residuals was performed by resampling with replacement from the model residuals. We selected the subset of years (1962–1978) for which we had both needle length data and parasitism data. We used Sils budmoth data as the response variable in the regression analysis (thus, this procedure biases our results in favor of the plant quality hypothesis). Furthermore, we used the linear model  $r_t$



$= u + vL_{t-1}$ , because our analysis showed that it fitted the data better than the alternative model in Table 1 (given the one-step-ahead fitting objective). We also fitted the data with the parasitism model  $r_t = r_0 - aP_{t-1}/(1 + awP_{t-1})$ , and the tritrophic model  $r_t = u + vL_{t-1} - aP_{t-1}/(1 + awP_{t-1})$  (this is Tritrophic Model II in Table 1). Note that the test we constructed is conservative with respect to finding for the parasitism hypothesis.

#### Global fitting

Global fitting was performed using the methods described in Wood (2001b). Each model was fitted to all available data predicted by that model, using a least squares objective. In this case the model is not making predictions of the next year's population change given observations of state variables at the start of the year and the parameters, but is predicting each of the entire series of data given the model parameters and some starting conditions (which like the parameters, must be estimated). To stabilize variances, the LBM density was raised to the power 0.1 for fitting, and a linear transformation was applied to needle length in order that needle length residuals were comparable to transformed LBM and parasitism rate residuals. Parasitism rates were fitted on the usual arcsine square-root transformed scale. Bootstrapping of the residuals was performed by resampling with replacement on a series by series basis. Local minima were avoided during fitting by bootstrap restarting (Wood 2001a).

Again, the computer intensive methods described in the section *Methods: General philosophy of analysis* were used to test reduced models against the tritrophic model, using difference in residual sum of squares between the alternative models as  $\Delta$ . First, all four time series (Engadine, Sils, needle length, and parasitism rate) were fitted with the trajectory-matching method (1) by the tritrophic model and (2) by a reduced model in which food quality has no influence on the host but hosts have an impact on food quality (i.e.,  $\delta = 0$  in the tritrophic model of Table 1). The effect of plant quality was modeled using the first alternative in Table 1, because we found that it fitted data better than the linear model. Again, this procedure biases results in favor of the plant quality model. Secondly, the three series other than the parasitism rate were fitted (1) by the full model, and (2) by the plant quality model of Table 1. (The plant quality model cannot be used to predict parasitism rate: when we remove from the model the effects of parasitism on LBM, we also remove the effects of LBM on the parasites because both are functions of the same parasitism rate).

## RESULTS

### Exploratory regression analysis

The regression analysis indicated that the plant quality index has a surprisingly weak effect on budmoth

TABLE 2. Results of analyses of time-series data.

a) Nonlinear regression			
Model		Dependent variable	$R^2$
Plant quality		larch budmoth ( $r_t$ )	0.312
Plant quality		plant quality ( $Q_t$ )	0.615
Parasitoid–host		larch budmoth ( $r_t$ )	0.880
Parasitoid–host		parasitoid ( $P_t$ )	0.712
b) Results from trajectory matching			
Model		$R^2$	$P$ value
Full tritrophic		0.839	...
No feedback from plant quality to host		0.828	0.025
Plant quality only		0.443	<0.003

Notes:  $R^2$  is the coefficient of determination, and  $r_t = \log N_{t+1}/N_t$ . Other variables and equations are the same as in Table 1.

dynamics. The linear model fits the data better than the nonlinear alternative, but still explains only 31% of the variance in the realized per capita rate of change of budmoth,  $r_t = \ln(N_t) - \ln(N_{t-1})$  (Table 2a; the nonlinear model explains 15% of the variance). Interestingly, the last documented budmoth peak (in 1989, not shown in our graphs) was curtailed due to several years of unfavorable weather in winter and spring (Baltensweiler 1993). As a result, budmoth did not achieve the usual peak density and there was no widespread defoliation. The plant quality index did not collapse, but budmoth density did, suggesting that poor plant quality (as measured by needle length) may not be necessary for driving the budmoth population to a low level.

By contrast, parasitism rates explain almost 90% of variance in the per capita rate of change of budmoth numbers (Table 2a). In fact, a simple model  $r_t = r - aP_{t-1}/(1 + awP_{t-1})$  with only three fitted parameters explains the bulk of the effect of parasitism on budmoth rate of population change ( $R^2 = 86.5\%$ ). Using a different metric,  $\log N_t$  instead of  $r_t$ , and the same model, we find that 94.1% of the variance in  $\log N_t$  is explained by parasitism rates (see Fig. 2). We are not aware of any other field data sets in population ecology in which such a simple model can explain a comparable amount of variation. Interestingly, the effects of budmoth density on plant quality and on parasitism rates are almost equally strong, with >70% of variance explained in each case (Table 2a).

The analysis of joint effects of quality and parasitism in the reduced data set showed that the plant quality model explained 48.0% of the variance in budmoth rate of population change (the higher proportion of variance explained than in the full data set is primarily due to the reduced data set not including the last oscillation). The parasitism and tritrophic models explained 65.6% and 66.1% of variance, respectively.

The trajectory matching (global fitting) analysis yielded very similar results. The full tritrophic model provided an excellent fit to the data (Fig. 3), marginally

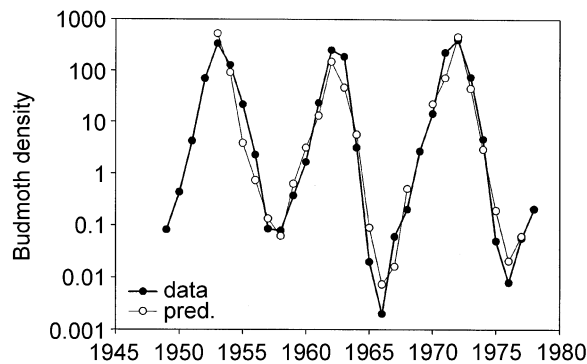


FIG. 2. Comparing predictions of the fitted parasitoid model to data. Predictions are obtained according to the following formula:  $\hat{N}_t = N_{t-1} \exp[r - aS_{t-1}N_{t-1}/(1 + awS_{t-1}N_{t-1})]$  where  $N_t$  and  $S_t$  are observed budmoth density and proportion parasitized, and  $r$ ,  $a$ , and  $w$  are fitted parameters.

better than the model omitting the effect of plant quality on budmoth (Table 2b). By contrast, the food quality model fitted poorly (Table 2b).

To summarize, exploratory one-step-ahead regression analysis indicated that parasitoids have a strong effect on budmoth dynamics, but it did not detect any effect of plant quality. Global fits suggested similar results albeit hinting at the possibility of a slight effect of plant quality. These results are consistent with the hypothesis that budmoth population cycles are driven primarily (or solely) by the moth's interactions with parasitoids, with plant quality fluctuating in response to changes of budmoth numbers.

#### Formal hypothesis testing

Using one-step-ahead fitting the parametric bootstrap test, described in *Methods*, indicated that the tritrophic model fitted data significantly better than plant quality model ( $P = 0.029$ ). By contrast, the parasitism model fitted the data as well as the tritrophic model ( $P = 0.64$ ). In other words, with this measure of fit, the analysis yields no evidence for an effect of plant quality on the rate of larch budmoth population growth. This result is particularly striking because our statistical test was designed to be biased in favor of plant quality hypothesis (see *Methods: Models*).

Using global fitting and testing the full tritrophic model against the null hypothesis of a reduced "parasitoid only" model (in which the feedback from plant quality to budmoth is eliminated) yielded an estimated  $P$  value of 0.025, indicating that the small effect of plant quality on host dynamics seen in the one-step-ahead regression analysis is nonetheless real and plays a role in shaping budmoth cycles. Testing the full model against the plant-quality-only model, yielded an estimates of  $P < 0.003$  with this measure of fit. Thus, there is strong evidence that plant quality alone cannot account for the observed host cycles, which lends firm support to both the exploratory results and the formal testing with the one-step-ahead approach.

## DISCUSSION

### Summary of LBM results

Our re-examination of data on budmoth density, plant quality, and parasitism rates suggests that the effect of plant quality on budmoth density is weak. By contrast, a simple model of budmoth–parasitoid interaction accounts for 90% of the variance in budmoth population growth rates. Thus, contrary to previous studies, we find that parasitoid–budmoth interaction appears to be the dominant factor driving the budmoth cycle. The fact that parasitism does not peak until budmoth is already in decline was previously taken as evidence against the parasitism hypothesis (Delucchi 1982). However, our model shows that this phase lag is consistent with parasitism driving the cycles (Fig. 3).

Our results, however, should not be interpreted as suggesting that plant quality plays no role in larch budmoth cycles. Although the regression approach could not detect a statistically significant effect of plant quality on larch budmoth population rate of change, the trajectory matching results indicate that the tritrophic model fits data slightly (and statistically significantly,  $P = 0.025$ ) better than the parasitism-only model.

The fact that the regression (one-step-ahead fitting) and trajectory matching (global fitting) approaches are not in perfect agreement about the effect of plant quality reflects the different model fitting criteria, which have different strengths and weaknesses. The regression analysis deals well with process noise by concentrating on short time scales, but consequently has low power to detect weaker interactions whose effects on the dynamics only become apparent over longer time

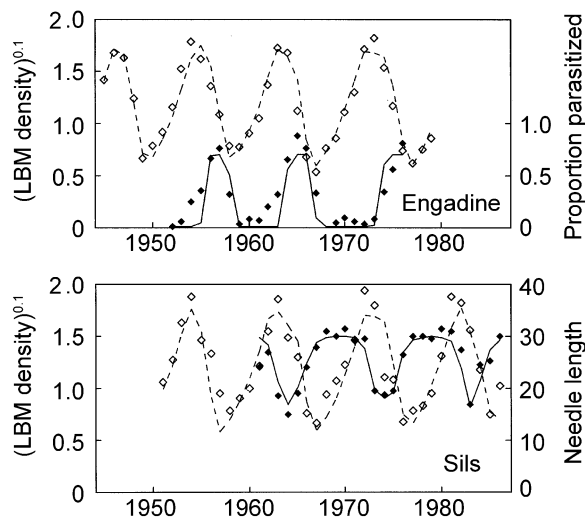


FIG. 3. Results of trajectory matching. Upper panel: Engadine data set (hollow symbols indicate budmoth density; filled symbols indicate proportion parasitized). Lower panel: Sils data set (hollow symbols indicate budmoth density; filled symbols indicate needle length).

scales. It is also sensitive to measurement errors, which are usually nontrivial in any ecological field study. In contrast, trajectory matching deals well with measurement error noise, and is sensitive to long time-scale patterns in the data, but makes an unrealistic assumption that process noise is absent. The observation that the two approaches agree on the role of parasitism in larch budmoth dynamics greatly strengthens this result.

Another feature of our approach that strengthens the main conclusion is that we did everything we could to avoid biasing the analysis against the food quality hypothesis. In fact, we consistently made choices that could be regarded as biasing the analysis in favor of the food-quality hypothesis, and nonetheless it failed to perform better than the parasitism hypothesis.

#### General implications

Models based on very different ecological mechanisms can produce dynamics that are qualitatively similar (Hunter and Dwyer 1998). For example, pathogen, plant quality, and parasitism models can all generate second-order oscillations (that is, cycles characterized by delayed density dependence) with periods of 7–10 yr typical of forest insect populations. Traditional approaches using qualitative comparisons between model predictions and data are unable to distinguish between dynamically plausible alternative hypotheses. Furthermore, since experimental manipulations at appropriate spatiotemporal scales are often impossible or prohibitively expensive, it would seem that we are unable to make further progress in elucidating the causes of forest insect cycles.

In a previous paper (Kendall et al. 1999), we argued that a way out of this apparent impasse is to develop approaches that combine mechanistic modeling with rigorous time-series statistics. Kendall et al. (1999) showed that in a laboratory setting such a synthetic approach can distinguish among rival mechanisms that predict qualitatively similar dynamics. In this paper, we show that the same approach can be successfully used in a field situation. The clear message from our analysis is that the dynamic data best support the hypothesis that LBM cycles are driven by parasitism, but modified slightly by an interaction with food quality. The reason we can obtain such a strong result is because we utilize the data more fully (performing detailed quantitative comparison of models with data) and because we employ not only data on LBM dynamics, but also on the dynamics of potential interactors (food quality and parasitoids).

Although technically quite involved, our approach has the same logical structure as standard hypothesis tests such as ANOVA: two models (e.g., effect present vs. effect absent) are compared based on their ability to fit the data, and the simpler is rejected in favor of the more complex if the improvement in goodness of fit is large enough. Few ecologists would dispute that an ANOVA table with  $P < 0.003$  (this is the  $P$  value

associated with rejecting the food quality hypothesis, see Table 2) should be regarded as a decisive result. The same force of evidence emerging from a dynamical model should be equally decisive. In fact, we would argue that it should count for more; our result is based on biologically plausible and empirically grounded dynamical models, while the assumptions of ANOVA are rarely subject to scrutiny.

#### ACKNOWLEDGMENTS

This paper is a result of a collaborative effort under the auspices of the Complex Population Dynamics working group at the National Center for Ecological Synthesis and Analysis, Santa Barbara, CA, USA. NCEAS is funded by NSF (grant number DEB-94-21535), the University of California-Santa Barbara, and the State of California. We thank Werner Baltensweiler for help with data sources and valuable criticism. Alan Berryman provided valuable critique.

#### LITERATURE CITED

- Anderson, R. M., and R. M. May. 1980. Infectious diseases and population cycles of forest insects. *Science* **210**:658–661.
- Auer, C. 1977. Dynamik von Lärchenwicklerpopulationen längs des Alpenbogens. *Mitteilungen der Eidgenössischen Anstalt für forstliches Versuchswesen* **53**:71–105.
- Baltensweiler, W. 1993. Why the larch bud-moth cycle collapsed in the subalpine larch-cembra pine forests in the year 1990 for the first time since 1850. *Oecologia* **94**:62–66.
- Baltensweiler, W., and A. Fischlin. 1988. The larch budmoth in the Alps. Pages 331–351 in A. A. Berryman, editor. *Dynamics of forest insect populations: patterns, causes, implications*. Plenum Press, New York, New York, USA.
- Beddington, J. R. 1975. Mutual interference between parasites or predators and its effect on searching efficiency. *Journal of Animal Ecology* **44**:331–340.
- Beniz, G. 1974. Negative Rückkopplung durch Raum- und Nahrungskonkurrenz sowie zyklische Veränderung der Nahrungsgrundlage als Regelsprinzip in der Populationsdynamik des Grauen Lärchenwicklers, *Zeiraphera diniana* (Guenee) (Lepidoptera, Tortricidae). *Zeitschrift fuer angewandte Entomologie* **76**:196–228.
- Berryman, A. A. 1991. Population theory: an essential ingredient in pest prediction, management and policy making. *American Entomologist* **37**:138–142.
- Berryman, A. A. 1995. Population cycles: a critique of maternal effects and allometric hypotheses. *Journal of Animal Ecology* **64**:290–293.
- Berryman, A. A. 1999. *Principles of population dynamics and their application*. Stanley Thornes, Cheltenham, UK.
- Delucchi, V. 1982. Parasitoids and hyperparasitoids of *Zeiraphera diniana* (Lepidoptera Tortricidae) and their role in population control in outbreak areas. *Entomophaga* **27**:77–92.
- Edelstein-Keshet, L., and M. D. Rausher. 1989. The effects of inducible plant defenses on herbivore populations. I. Mobile herbivores in continuous time. *American Naturalist* **133**:787–810.
- Ellner, S. P., Y. Seifu, and R. H. Smith. 2002. Fitting population dynamics models to time-series data by gradient matching. *Ecology* **83**:2256–2270.
- Fischlin, A. 1982. Analyse eines wald-insekten-systems: der subalpine Lärchen-arvenwald und der graue Lärchenwickler *Zeiraphera diniana* Gn (Lepidoptera, Tortricidae). Dissertation Number 6977. ETH, Zurich, Switzerland.
- Fischlin, A., and W. Baltensweiler. 1979. Systems analysis of the larch bud moth system. Part 1: the larch-larch bud

- moth relationship. *Mitteilungen der Schweizerischen Entomologischen Gesellschaft* **52**:273–289.
- Ginzburg, L. R., and D. E. Taneyhill. 1994. Population cycles of forest Lepidoptera: a maternal effect hypothesis. *Journal of Animal Ecology* **63**:79–92.
- Hassell, M. P. 1978. The dynamics of arthropod predator–prey systems. Princeton University Press, Princeton, New Jersey, USA.
- Hilborn, R., and M. Mangel. 1997. The ecological detective. Princeton University Press, Princeton, New Jersey, USA.
- Hudson, P. J., A. P. Dobson, and D. Newborn. 1998. Prevention of population cycles by parasite removal. *Science* **282**:2256–2258.
- Hunter, A. F., and G. Dwyer. 1998. Outbreaks and interactive factors: insect population explosions synthesized and dissected. *Integrative Biology* **1**:166–177.
- Hunter, M. D., and P. W. Price. 1998. Cycles in insect populations: delayed density dependence or exogenous driving variables? *Ecological Entomology* **23**:216–222.
- Kendall, B. E., C. J. Briggs, W. W. Murdoch, P. Turchin, S. P. Ellner, E. McCauley, R. M. Nisbet, and S. N. Wood. 1999. Why do populations cycle? A synthesis of statistical and mechanistic modeling approaches. *Ecology* **80**:1789–1805.
- Korpimäki, E., and K. Norrdahl. 1998. Experimental reduction of predators reverses the crash phase of small-rodent cycles. *Ecology* **79**:2448–2455.
- Krebs, C. J., S. Boutin, R. Boonstra, A. R. E. Sinclair, J. N. M. Smith, M. R. T. Dale, K. Martin, and R. Turkington. 1995. Impact of food and predation on the snowshoe hare cycle. *Science* **269**:1112–1115.
- Omlin, F. X. 1977. Zur populationsdynamischen Wirkung der durch Raupenfrass und Dungung veränderten Nahrungsbasis auf den Grauen Larchenwickler *Zeiraphera diniana* Gn. (Lep: Tortricidae). Dissertation Nr. 6064. ETH Zurich, Switzerland.
- Royama, T. 1981. Fundamental concepts and methodology for the analysis of animal population dynamics, with particular reference to univoltine species. *Ecological Monographs* **51**:473–493.
- Royama, T. 1992. Analytical population dynamics. Chapman and Hall, London, UK.
- StatSoft. 1999. STATISTICA for Windows. StatSoft, Tulsa, Oklahoma, USA.
- Turchin, P. 1990. Rarity of density dependence or population regulation with lags? *Nature* **344**:660–663.
- Turchin, P. 2003. Complex population dynamics: a theoretical/empirical synthesis. Princeton University Press, Princeton, New Jersey, USA.
- Turchin, P., A. D. Taylor, and J. D. Reeve. 1999. Dynamical role of predators in population cycles of a forest insect: an experimental test. *Science* **285**:1068–1071.
- Wood, S. N. 2001a. Minimising model fitting objectives that contain spurious local minima by bootstrap restarting. *Biometrics* **57**:240–244.
- Wood, S. N. 2001b. Partially specified ecological models. *Ecological Monographs* **71**:1–25.



## Supplementary Material

Peter Turchin et al. (2001). Dynamical effects of plant quality and parasitism on population cycles of larch budmoth.

### Trajectory Matching

Trajectory matching (also known as model calibration) estimates the parameters of a deterministic dynamic model such as a system of difference or differential equations, by fitting model output to a time series of observations. In order to give the fitting criterion used here we need to define some notation. Let  $X(t; \theta, x_0)$  denote the state vector of model solutions at time  $t$ , given the values of model parameters  $\theta$  and initial conditions  $x_0$  (i.e.  $X(0, \theta, x_0) = x_0$ ). The data are assumed to be measurements of a set of  $m$  scalar "readout variables"  $\mu_i = g_i(X)$ , at a series of times  $\{t_{i,j}\}_{j=1}^{N_i}$ , subject to measurement error with standard deviation  $\sigma_i$  (here we assume constancy of  $\sigma_i$  over time but this assumption can be relaxed). Note that the number of observations and the observation times may be different for each of the readout variables. In our model the state variables are budmoth density, parasite density, and food quality, and the readout variables are (budmoth density)<sup>0.1</sup>, the fraction of parasitized budmoth larvae, and needle length. The tenth-root transformation is very similar to the logarithmic transformation often used to stabilize variances in population time series data, but does not require special handling of zeros.

Let  $y_{ij}$  denote the empirical observation of readout variable  $i$  at the  $j$ th observation time (i.e. the observation of  $\mu_{ij}$ ). The fitting criterion is then to choose the values of  $\theta$  and  $x_0$  which minimize the objective function

$$F(x_0, \theta) = \sum_{i=1}^m \sum_{j=1}^{N_i} \frac{\left[ y_{ij} - g_i(X(t_{ij}; \theta, x_0)) \right]^2}{2\sigma^2}. \quad (S1)$$

(S1) was minimized using a constrained Quasi-Newton method as described in Gill et al. (1974, 1981). Initial parameter estimates were obtained by treating the first few terms of the autocorrelation function of the data as further data to be matched in (S1). Once rough estimates had been obtained these terms were dropped. Since non-linear discrete time models are prone to dynamic behaviour tending to introduce small scale local minima in objectives such as (S1), the fitting procedure was supplemented by "bootstrap

restarting” – a method for escaping local minima by perturbing the objective function through bootstrap resampling of the data  $y_{ij}$  (see Wood 2001 for full details). 40 bootstrap restarts were used per fit.

Initially the models were fitted to all available data (from 1945 until 1992), but after examination of fits and residual plots, it was decided to delete the final 6 sample times from the time series. The final peak-to-peak interval in LBM density is slightly different from the others in the data, and this phase drift is not explained by our model (possibly because we are using a deterministic model, or because some change in conditions occurred towards the end of the data set). Note that this deletion changes none of the conclusions presented in this paper – although it does somewhat increase the validity of the distributional assumptions underlying the generalized likelihood ratio tests reported below.

The best fits of the full model and the model in which budmoth dynamics are not affected by food quality produce residual plots with some problems. There is serial correlation of residuals, particularly in the parasitism and needle length series. However, normal probability plots of the residuals are reasonable and formal tests of the normality assumption fail to reject the normality assumption ( $p > 0.2$  for each series and for the combined residuals). We therefore believe that examination of the model residuals does not undermine the conclusions of this paper, but does caution against treating  $p$ -values as anything but approximate.

### **Significance tests**

Given that the objective function has small-scale irregularities and that any distributional assumptions we may make about the measurement errors are open to question, it seems sensible to approach formal hypothesis testing by computer intensive methods. In particular we follow Wood (2001) based on Davison and Hinkley (1987) and use semi-parametric bootstrapping to perform hypothesis testing.

For example, consider testing the null hypothesis that the full model with  $\delta = 0$  produced the data against the alternative that the full model without restrictions produced it. The restricted model was fitted to the data, and replicate datasets under the null were produced by resampling with replacement from the resulting residuals and adding these resampled residuals to the fitted values from the restricted fit (residuals were resampled on a series by series basis). Full and restricted models were then fitted to each

replicate, so that an approximate distribution of differences in fit could be obtained under the assumption that the restricted model was correct. By comparing this bootstrap distribution to the observed difference in fit a p-value for the difference between model fits is obtained. The obvious measure of fit to use for each model is that given by (S1): so let  $\hat{F}_{\delta=0}$  be the value of (S1) evaluated at the best fit of the restricted model and  $\hat{F}_{\delta>0}$  be the equivalent for the full model. The obvious statistic to use to measure difference in fit is then:  $\hat{F}_{\delta=0} - \hat{F}_{\delta>0}$ , but we expect better results from using an approximately pivotal statistic such as  $(\hat{F}_{\delta=0} - \hat{F}_{\delta>0}) / \hat{F}_{\delta>0}$  (see e.g. Cox and Hinkley, 1974), so this was the statistic used.

When comparing the tri-trophic model to the model with  $\delta = 0$ , both models were fitted to all 4 available time series. The  $\delta = 0$  model corresponds to the situation in which hosts reduce food quality, while food quality has no influence on hosts. However, when considering parasitoids, the model structure does not permit us to turn off the influence of parasitoids on hosts while maintaining the influence of hosts on parasitoids, without producing a model in which a proportion of hosts are simultaneously alive and dead. We therefore require a slightly more involved approach to testing the significance of parasitoids in explaining the host dynamics. Specifically, we fitted both the tri-trophic and the food quality model to the 3 series excluding parasitism rate and bootstrapped under the assumption that the food quality model is correct, exactly as before (but with the test statistic for model comparison involving only the 3 fitted series). The only problem with this approach is that it leaves open the possibility that the tri-trophic model only fits the real data better by virtue of allowing unrealistic parasitoid dynamics (since these are now unconstrained by the parasitism data). Fortunately this problem is very simple to address, provided that some overestimation of p-values is tolerable. For the calculation of the test statistic for the real data, we fitted the tri-trophic model to all data including the parasitism series, while calculating the observed test statistic using only the three series that both models can predict. The distribution under the null hypothesis was still produced using only the 3 series excluding parasitism (this makes the distribution under the null wider than if the tri-trophic model was fitted to all 4 series for each bootstrap replicate, but this is a price worth paying in order to ensure that there is no possibility of biasing estimated p-values downwards by this method). This procedure eliminates the possibility that the better fit of the tritrophic model is the result of allowing unrealistic parasitoid dynamics, but is strictly conservative in that the

estimated  $p$ -value can only be increased by the approach taken.

Alternatively, under the assumptions that the measurement errors are independent and Gaussian distributed, and that all departures observations and model output are due to the measurement errors (S1) is the negative log-likelihood of  $(\theta, x_0)$  given the data, so it is tempting to use a generalized likelihood ratio (GLR) test (e.g. Silvey, 1975). The given assumptions can only be strictly true if the dynamics of the real system are purely deterministic, which is unlikely to be true for a natural population. In practical terms neglecting stochasticity other than measurement error will mean that we have underestimated the error terms appropriate for describing the deviation of model from data and unless this problem is dealt with the  $p$ -values from such a test will be too high. In this circumstance of over-dispersion the standard approach (e.g. McCullagh and Nelder, 1989) is to use an F ratio statistic like  $(\hat{F}_{\delta=0} - \hat{F}_{\delta>0}) / \hat{F}_{\delta>0}$  and obtain  $p$ -values by comparison to the appropriate F distribution (although the theoretical justification for this is not strong). This procedure yields estimated  $p$ -values of 0.01 for the comparison of full and  $\delta = 0$  models, and an essentially zero  $p$ -value for the comparison of full and food quality models. Hence this approach agrees with the computer intensive tests that there is much stronger evidence for a parasitoid effect than for a food quality effect, but the estimated  $p$ -values are lower. The bootstrap analysis allows us to directly check the distributional assumptions underpinning the GLR test, and suggests that for the comparison of full and  $\delta = 0$  models the approximation will underestimate the  $p$ -values a little while for the other comparison the assumptions are rather poor and will substantially underestimate the  $p$ -values. Hence while it is re-assuring that the GLR approach leads to broadly the same conclusions as the bootstrap analysis, we believe that the computer intensive  $p$ -values are the more reliable.

### Measurement error variance

We used measurement error variances  $\sigma = 0.1, 1.0$ , and  $0.1$  for (budmoth density)<sup>0.1</sup>, needle length, and percent parasitism of budmoth larvae, respectively. The largest component of measurement error for budmoth density is not the sampling error, but the discrepancy between the density at the sampling location and the regional average density that the model is tracking. A rough estimate of this discrepancy was obtained by comparing our series (Engadine) with the series from Sils. The two series

(both tenth-root transformed) showed a linear relationship with no obvious trends in the variance of residuals from regression of Sils on Engadine or vice-versa. The regressions (in either direction) had root mean square (RMS) residual  $\sigma \approx 0.13$ . The discrepancy between two sites overestimates the discrepancy between any given site and the regional average. The amount of overestimation was estimated using data from Herren (1976), which gives short series from four locations. We computed the average  $\sigma$  (RMS residual) for each series predicting each other series, and the same for each series predicting the "regional mean" (average of all other series). The latter was smaller by a factor of 0.75. Applying this correction to the RMS residuals for Engadine-Sils predictions, we estimated  $\sigma \approx 0.1$  for the "measurement error" due to using Engadine data in place of the actual regional average. The value of  $\sigma$  for needle length (in mm) was based on the fact that measurements were rounded to the nearest 0.5mm, and the value for percent parasitism is based on our estimate that percent parasitized is measured with accuracy of about 5%.

### **Additional References**

Cox, D.R. and D.V. Hinkley 1974 Theoretical Statistics. Chapman and Hall, London

Davison, A.C. and D.V. Hinkley 1997 Bootstrap Methods and Their Application. Cambridge University Press, Cambridge.

Gill, P.E. G.H. Golub, W. Murray and M.A Saunders 1974. Methods for modifying matrix factorizations. Mathematics and Computation 28:505-535

Gill, P.E., W. Murray and M.H. Wright 1981 Practical Optimization. Academic Press, London.

Herren, H. R. 1976. Manipulation d'une population de la tordeuse grise du meleze en vue d'augmenter l'efficacite de ses parasitoids. Mitteilungen der Schweizerischen Entomologischen Gesellschaft 49:307-308.

McCullagh, P. and J. Nelder. 1989. Generalized Linear Models. 2nd Edition. Chapman and Hall, London.

Silvey, S.D. 1975 Statistical Inference. Chapman and Hall, London.

Turchin, P. and S.P. Ellner. Living on the edge of chaos: population dynamics of Fennoscandian voles. Ecology 81: 3099–3116.