

FIELD-SCALE ROLES OF DENSITY, TEMPERATURE, NITROGEN, AND PREDATION ON APHID POPULATION DYNAMICS

PERRY DE VALPINE^{1,3} AND JAY A. ROSENHEIM²

¹Department of Environmental Science, Policy, and Management, University of California, 137 Mulford Hall #3114, Berkeley, California 94720-3114 USA

²Department of Entomology, University of California, One Shields Avenue, Davis, California 95616-8584 USA

Abstract. Robust analyses of noisy, stage-structured, irregularly spaced, field-scale data incorporating multiple sources of variability and nonlinear dynamics remain very limited, hindering understanding of how small-scale studies relate to large-scale population dynamics. We used a novel, complementary Bayesian and frequentist state–space model analysis to ask how density, temperature, plant nitrogen, and predators affect cotton aphid (*Aphis gossypii*) population dynamics in weekly data from 18 field-years and whether estimated effects are consistent with small-scale studies. We found clear roles of density and temperature but not of plant nitrogen or predators, for which Bayesian and frequentist evidence differed. However, overall predictability of field-scale dynamics remained low. This study demonstrates stage-structured state–space model analysis incorporating bottom-up, top-down, and density-dependent effects for within-season (nearly continuous time), nonlinear population dynamics. The analysis combines Bayesian posterior evidence with maximum-likelihood estimation and frequentist hypothesis testing using average one-step-ahead residuals.

Key words: agriculture, field studies; *Aphis gossypii*; Bayesian and frequentist evidence; computational statistics methodology; density dependence; Markov chain Monte Carlo; Monte Carlo kernel likelihood; population dynamics; San Joaquin Valley, California, USA; stage-structure; state–space model.

INTRODUCTION

Population dynamics of insects in agricultural fields may be affected by density dependence, complex trophic interactions (Rosenheim 1998), exogenous factors such as weather and plant conditions (Cisneros and Godfrey 2001), movement and landscape factors (Kareiva 1990, de Roos et al. 1991, Tschardt and Brandl 2004), phenotypic plasticity (Nylin and Gotthard 1998, Miner et al. 2005, Stillwell and Fox 2005), and other processes. Field-scale, within-season time series of herbivore and natural-enemy abundances and other factors have the potential to shed light on several related questions: How do results from small-scale field and laboratory studies relate (or “scale-up”) to field-scale processes (Carpenter 1996, Pascual and Levin 1999, Steele and Forrester 2005)? What is the relative importance of top-down vs. bottom-up factors on herbivore dynamics (Shurin et al. 2002)? What is the overall predictability of herbivore dynamics? However, analysis of such data has been very limited in practice due to serious statistical challenges of incorporating process and measurement variability for fitting structured population models to data, and thus robust analysis of the large-scale dynamics of such systems has been very limited. Here we use a novel synthesis of Bayesian and frequentist state–space models to evaluate the roles of predators, density dependence,

temperature, and plant nitrogen on aphid population dynamics at field scales in 18 cotton fields of California’s San Joaquin Valley (USA) and their relationship to extensive small-scale studies.

We chose the cotton aphid, *Aphis gossypii*, as a model organism whose ecology is well known from experiments in the laboratory, in small field enclosures, and in small, replicated experimental field plots, but is unstudied at the scale of agricultural fields. Laboratory and small-scale field studies suggest that cotton aphid populations may be influenced by many factors. First, *A. gossypii* development and reproduction rates are affected by temperature (Slosser et al. 1998, 2004), with maximum reproduction at a relatively low temperature for a hot-temperate herbivore, ~25°C (Akey and Butler 1989). Second, as with many phloem feeding insects, *A. gossypii* appears often to be nitrogen limited, and supplemental N fertilization of cotton plantings generally results in enhanced aphid densities (Slosser et al. 1997, 1998, 2004, Cisneros and Godfrey 2001, Nevo and Coll 2001). Third, intraspecific competition appears to be important, especially at higher aphid densities, contributing to subsequent population declines (Slosser et al. 1998, 2004). Finally, predators, and in particular *Chrysoperla* spp. lacewings, appear to have the potential to suppress cotton aphid populations when tested singly. However, this potential appears to be lost, because lacewing larvae are themselves consumed by other members of the diverse predator community (Rosenheim et al. 1993, 1999, Cisneros and Rosenheim 1997, Rosenheim 2001).

Manuscript received 30 November 2006; revised 16 May 2007; accepted 4 June 2007. Corresponding Editor: F. He.

³ E-mail: pdevalpine@nature.berkeley.edu

In summary, previous studies predict (1) that aphid dynamics will be affected by temperature, plant nitrogen, and density dependence, and (2) that lacewings have the strongest predation rate per individual but do not typically develop large enough populations to suppress *A. gossypii*. While some of the factors thought to shape cotton aphid dynamics could in principle be manipulated in full-field experiments (e.g., nitrogen supplementation), others could not (e.g., predator species composition). Whether or not insights obtained at small scales can be extended to the full scale of agricultural fields is unknown.

Our analysis synthesizes three advances in computational statistics methods, allowing complementary Bayesian and frequentist insights. First, advances in Markov chain Monte Carlo (MCMC) and other numerical-integration methods allow incorporation of process noise and measurement error in a nonlinear state-space model framework. Most state-space modeling has used univariate (population size) models with a fixed time step and no external driving variables either for linear (Dennis et al. 2006) or nonlinear models (de Valpine and Hastings 2002, Calder et al. 2003). Examples of more advanced applications include age-structured models (Meyer and Millar 1999, Besbeas et al. 2002, de Valpine 2003, Newman et al. 2006); work on plant disease dynamics (Gibson 1997); and work on pea aphid-parasitoid dynamics using a two-species, stage-structured model with daily time steps allowing time-varying demographic parameters (Gross et al. 2005). Here we use a two-stage model with density dependence, accommodate irregularly spaced data by using a one-day time step for the state dynamics, and allow effects of temperature, plant nitrogen, and predator populations in fairly general ways.

Second, we give maximum-likelihood estimates (MLEs) for a nonlinear, stage-structured state-space model, calculated using the Monte Carlo kernel likelihood method (de Valpine 2003, 2004, de Valpine and Hilborn 2005). Obtaining MLEs has been a challenging problem for these models (de Valpine 2004), so most analyses have been limited to Bayesian results and have not included MLEs or any estimate of the statistical significance of model components.

Third, even given MLEs, frequentist hypothesis testing poses challenges because likelihood ratio values are difficult to estimate (Meng and Wong 1996, Chib and Jeliazkov 2001, Mira and Nicholls 2004), and it is unclear whether asymptotic chi-squared results for them will be accurate. Residuals from state-space models—another possible basis for frequentist hypothesis testing—are nontrivial to define and calculate, and thus have not typically been included in analyses. Here we define “residuals” as the mean one-step-ahead prediction error given all previous data (for a given time series) given the MLE parameters. These one-step-ahead residuals are approximately independent and allow frequentist hypothesis testing of residuals vs. tempera-

ture, nitrogen, and predator data series, but calculating them requires a new MCMC sampler for each datum. In addition, use of one-step-ahead residuals allows evaluation of the differences in predictive power between models directly in terms of root-mean-squared-error (i.e., average mean residual).

Complementary Bayesian and frequentist results are given based on MCMC posterior distributions, maximum-likelihood estimates, and hypothesis testing using the one-step-ahead residuals given MLE parameters. In the results, Bayesian evidence about parameter(s) comprises pairwise marginal posterior distributions of parameters, highest posterior marginal density estimates, the ratio of highest posterior marginal density to density at zero (or some value of interest), and the fraction of a posterior with density lower than the density at zero. Frequentist evidence comprises maximum-likelihood estimates and regression tests of one-step-ahead predictions vs. possible explanatory variables. For example, the significance of temperature is tested by calculating one-step-ahead residuals from a model without temperature and regressing those against temperature. The combination of Bayesian and frequentist evidence represents a practical, best-of-both-worlds approach without getting bogged down in philosophical debate (Clark 2005). The types of evidence given here do not represent every type of Bayesian or frequentist evidence that could be considered.

APHID AND NATURAL-ENEMY DATA

Data were collected from 10 fields in 1993 and 8 fields in 1994, with no overlap of sites between years. We refer to these data as “field-scale” because they are sampled from 1–2 ha sections of agricultural fields. Approximately every week (range: 5–9 d), all nymph- and adult-stage aphids were collected and counted from the fifth-mainstem-node leaf of 50 plants (range: 44–55). At very high densities, a subsample of ~10% of the collected aphids was counted. Counts are reported and modeled as number of aphids per 50 leaves.

Five of the most common and potentially important predator taxa were counted from searches of 7 (median; range 4–10; one sample of 3) whole plants and 6 (median; range 3–10) sweep-net samples. Whole-plant counts included lacewing eggs and larvae (*Chrysoperla carnea*, *C. comanche*, *Chrysopa nigricornis* (see Plate 1), and *C. oculata*, in order of decreasing frequency); *Orius tristicolor* (Hemiptera) (and rarely *O. insidiosus*) nymphs; and nymphs of three *Geocoris* spp. (*G. pallens*, *G. punctipes*, and [rarely] *G. atricolor*). Sweep-net counts included the same *Chrysoperla*, *Chrysopa*, *Orius*, and *Geocoris* species as well as nymphs of the larger hemipteran predators *Nabis* spp. and *Zelus renardii*. Sweep-net samples were obtained in a standardized manner, and when both types of data were available (for *Chrysoperla*, *Chrysopa*, *Orius*, and *Geocoris*) the joint distribution of whole-plant counts and sweep-net counts was modeled as a mixture of gamma-distributed



PLATE 1. An adult lacewing, *Chrysopa nigricornis*, one of several predators of the cotton aphid, *Aphis gossypii*, in cotton. Photo credit: J. A. Rosenheim.

abundances with Poisson measurements by each method, with a constant ratio of expected counts from the two sampling methods (implemented numerically), and the adjusted sweep-net count given the whole-plant count was used in the model. Herbivorous spider mites (*Tetranychus* spp.) were counted from the same leaf samples as *Aphis gossypii*. Predator data were treated as fixed covariates for the aphid model rather than treating their dynamics in the state-space framework because their dynamics do not appear to be strongly coupled to the aphids and doing so would require additional parameters and assumptions. Predator estimates were linearly interpolated for days between measurements.

Plant densities were 5.7–26.6 plants/m, and aphid densities were rescaled by multiplying by plants per meter divided by 10.9 to correspond to the number of aphids per 50 leaves at a plant density per 10.9 m, the most common plant spacing. Percentage nitrogen was measured in weekly petiole samples, which was also interpolated between measurements to provide daily covariates. It is recognized that petiole percentage nitrogen does not directly measure phloem composition (Youssefi et al. 2000, Karley et al. 2002), but it is used here as a general measure of plant status that could be predicted to correlate with aphid dynamics (Slosser et al. 1997, 1998, 2004, Cisneros and Godfrey 2001, Nevo and Coll 2001). Temperature records were obtained from the nearest weather station for each site. Figures reporting all data and estimated aphid models are in the Appendix.

APHID POPULATION MODEL

Inferences were conducted in the framework of a standard stage-structured model extended to incorporate bottom-up (plant nitrogen) and abiotic (temperature) effects, endogenous feedbacks (density dependence), top-down effects (predators), and environmental stochasticity. Dynamics for aphid nymphs (A_1) and adults (A_2) with daily time steps are as follows:

$$\begin{bmatrix} A_1(t+1) \\ A_2(t+1) \end{bmatrix} = \begin{bmatrix} R_1(t)\exp[v_1(t)] & 0 \\ 0 & R_2(t)\exp[v_2(t)] \end{bmatrix} \cdot \begin{bmatrix} S_1(t)[1-M(t)] & F(t) \\ S_1(t)M(t) & S_2(t) \end{bmatrix} \begin{bmatrix} A_1(t) \\ A_2(t) \end{bmatrix}. \quad (1)$$

Factors $R_1(t)$ and $R_2(t)$ incorporate density dependence and predation, and $v_1(t)$ and $v_2(t)$ are environmental stochasticity random variables, explained more below; t is day and F is daily fecundity. Daily nymph and adult survival rates, $S_1(t)$ and $S_2(t)$, respectively, and nymph-to-adult maturation rate, $M(t)$, each are affected by nitrogen (centered at 1.4%) and temperature (in degrees Fahrenheit, centered at 32) on a logit scale:

$$\begin{aligned} X(t) = \text{logit}^{-1} \{ & \text{logit}(X) + \beta_{T,X}[T(t) - 32] \\ & + \beta_{N,X}[N(t) - 1.4] \} \end{aligned} \quad (2)$$

for $X = S_1, S_2$, or M . Here $T(t)$ and $N(t)$ are temperature and nitrogen, with coefficients $\beta_{T,X}$ and $\beta_{N,X}$, respectively; X is the value of $X(t)$ when $T(t) = 32$ and $N(t) = 1.4$; and $\text{logit}(z) = \log[z/(1-z)]$. (All statistical models were based on the Fahrenheit scale; for SI conversion: $^{\circ}\text{F} - 32 = ^{\circ}\text{C} \times 1.8$.) The fecundity model is

$$F(t) = \max\{F + \beta_{T,F}[T(t) - 32] + \beta_{N,F}[N(t) - 1.4], 0.01\} \quad (3)$$

where temperature and nitrogen have linear effects with coefficients $\beta_{T,F}$ and $\beta_{N,F}$, respectively, down to a negligible minimum of 0.01. In the hypotheses considered here, the temperature and nitrogen effects ($\beta_{T,X}$ and $\beta_{N,X}$) are estimated only for fecundity ($X = F$) and maturation rate ($X = M$). This choice avoids overparameterization yet allows an effect on input into each stage class.

Factors $R_1(t)$ and $R_2(t)$ incorporate Gompertz density dependence and predation with a saturated functional response, $R_i(t) = \exp[-\delta_i a(t) - \alpha_P P(t)/A(t)]$, for $i = 1$ or 2 . Here $A(t) = A_1(t) + A_2(t)$ is total aphid density; $a(t) = \log[A(t)]$; δ_i is the strength of density dependence on stage $i = 1$ or 2 ; $P(t)$ is the abundance of predators, which are tested singly to limit the number of parameters estimated in any one model; and α_P is predation strength for predator P . Predation strength is assumed to be the same on nymphs and adults because the predators are likely able to consume both stages, and preliminary fits with stage-specific predation rates did not improve results. Preliminary fits with a Type II functional response indicated that only a saturation level could be reasonably estimated from these data: with α_P aphids consumed per predator per day, daily prey survival is $1 - \alpha_P P(t)/A(t)$, which is approximately $\exp[-\alpha_P P(t)/A(t)]$. In some studies the simplest one-parameter functional response is considered to be the Holling Type I linear response, with no prey saturation of predators, but with the high aphid abundances common in these data that would seem unrealistic.

Environmental stochasticity was modeled as bivariate log-normal with the same variance for nymphs and adults, i.e., $(v_1(t), v_2(t))$ has mean $(0, 0)$, variances σ_v^2 , and correlation ρ_v . The data estimates of nymph and adult densities are $D_1(t)$ and $D_2(t)$, respectively. Measurement error was modeled as bivariate log-plus-one normal with the same variance for nymphs and adults, i.e., $\log(D_i(t) + 1)$ has mean $\log(A_i(t) + 1)$ and variance σ_D^2 , for $i = 1$ or 2 . This was based on the experience that adults and nymphs were equally detectable. Note that the survival and maturation rates can range from 0 to 1, but the environmental stochasticity is considered broadly to include immigration, emigration, variability in stage transitions, and other factors.

An important point about the model (Eq. 1) is that even without density dependence the linearity of the deterministic “skeleton” matrix model is combined with non-Gaussian (log-normal) noise. For state-space models with (multivariate) linear dynamics and additive Gaussian noise, the Kalman Filter can calculate integrated likelihoods exactly (Dennis et al. 2006). For the population matrix model (Eq. 1), however, the noise is Gaussian on a log scale, for which the model is nonlinear. Thus, even the simplest matrix model with realistic noise leads to a nonlinear and/or non-Gaussian state-space model, so the Kalman Filter does not apply.

MCMC and parameter prior distributions

A customized Markov chain Monte Carlo (MCMC) sampler was developed to sample the posterior distribution of all values of parameters and latent states (A_1 and A_2). Flat priors were used for all parameters, with the range limits: $F > 0.01$; $0.001 < (S_1, S_2, M) < 0.999$; $-0.95 < \rho < 0.95$; and $-2 < \beta_{N,X} < 2$. For F and other parameters the unbounded flat priors were improper, but the MCMC nevertheless mixed well (i.e., there was no infinite integral problem). The first three range constraints promoted MCMC mixing by keeping parameters off of boundary values, while the last constraint avoided poor mixing (possibly non-integrable) in a very low posterior density region.

Sequential model building

A sequential model-building approach was used, with models fit in the order: Gompertz density dependence with two parameters (GDD2, $\delta_1 \neq \delta_2$); GDD2 + temperature effects on F and M (GDD2temp2), GDD2 + nitrogen effects on F and M (GDD2nit2), GDD2 + temperature and nitrogen effects on F and M (GDD2temp2nit2); GDD2 + temperature effect on F (GDD2tempF); and GDD2tempF + single-predator effects (GDD2tempFpred1). In the final models that include predators (GDD2tempFpred1), only the effect of temperature on F was included because it appeared as the only statistically significant temperature or nitrogen effect, with concordant Bayesian evidence, from the GDD2temp2 and GDD2nit2 models. The predator

model (GDD2tempFpred1) was run separately for each predator data series.

Lacewings were of special interest because of their potential role as strong aphid predators that are suppressed by intraguild predation, but the lacewing larvae data were relatively noisy, so the lacewing egg data were also considered as a surrogate measure of overall lacewing abundance.

Initial results that all predators showed evidence for a positive predation parameter was biologically plausible but raised the question of some unknown analysis artifact giving spurious results. As one way to check whether the positive predation parameters were somehow an artifact of the estimation methods, the herbivorous spider mite data were used as a “control” predator sequence, i.e., the model was fit as if herbivorous mites might be predators. While mites might have indirect interactions with aphids, such as resource competition or apparent competition via shared predators, the interaction would not necessarily be expected to follow a predator model. Values of α_P for different predators correspond to either sweep-net or whole-plant count data, which are not directly comparable, but for species with both types of data (jointly incorporated), α_P is reported for both sweep-net and whole-plant scales.

Maximum-likelihood estimation and one-step-ahead residuals

The Monte Carlo kernel likelihood (MCKL) algorithm uses weighted kernel density estimates of the posterior parameter dimensions of the MCMC sample to estimate the likelihood (de Valpine 2003, 2004, de Valpine and Hilborn 2005). In this case the weights were all uniform and the kernel was truncated (and renormalized) at the parameter boundaries. With many parameter dimensions (e.g., nine for model GDD2: $F, S_1, M, S_2, \delta_1, \delta_2, \sigma_v, \rho, \sigma_D$), the kernel estimation requires a large bandwidth and may be inaccurate. To improve estimates, after initial estimation the algorithm was iterated with one or two parameter dimensions at a time (chosen so that highly correlated dimensions were paired) for 5–10 iterations. This was done by generating an MCMC sample of the latent states and one or two parameter dimensions with the others fixed, and using MCKL to estimate the likelihood surface for those parameters given the others, which can be quite accurate.

The one-step-ahead predictions for $A_1(t)$ and $A_2(t)$ were the mean latent states at time t from an MCMC sampler run only for states up to (and including) t given data prior to but *not* including t , with parameters fixed at the MLE (maximum-likelihood estimate); thus they represent the model prediction for time t conditioned on all the observations prior to t . Because the extent of data used for the MCMC run that estimates the one-step-ahead prediction for time t depends on t , a separate MCMC run is required for each t , but these are efficient

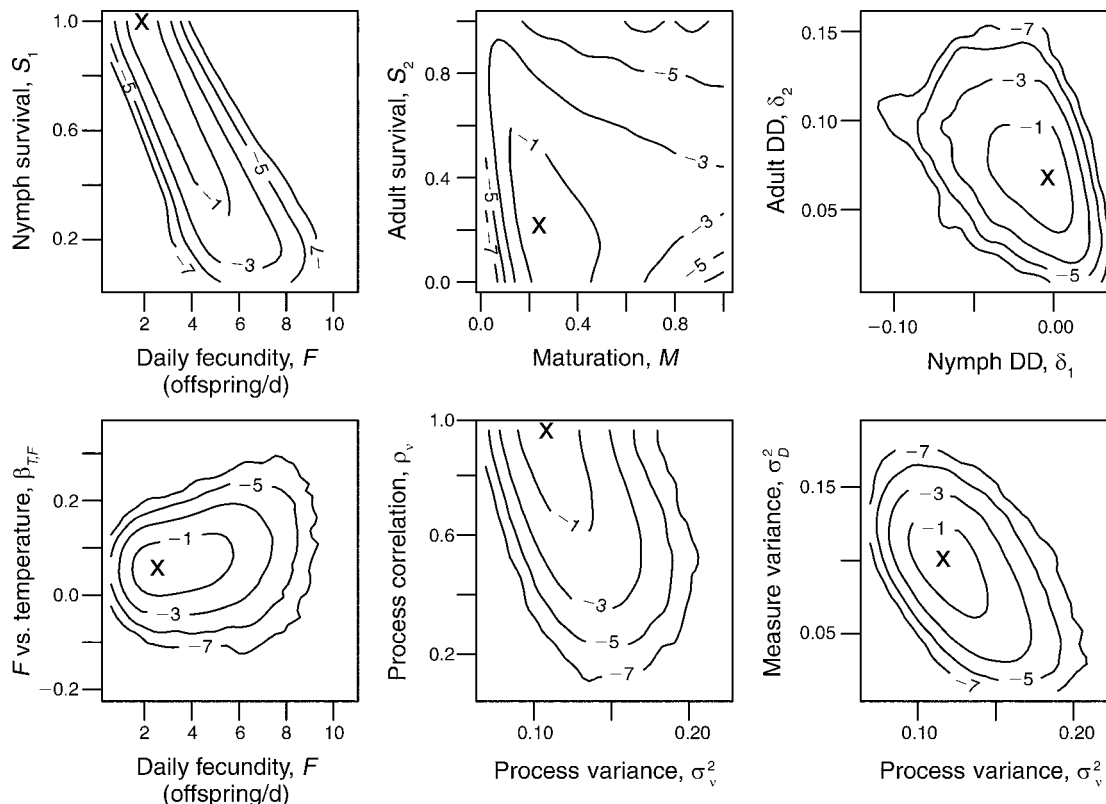


FIG. 1. Bayesian two-dimensional marginal posterior densities (from normal kernel density estimation) for a model with two density-dependence parameters, one temperature parameter, and no predators (i.e., GDD2tempF). Each subfigure shows contours and log-density differences of -1 , -3 , -5 , and -7 from the maximum ("X"), corresponding to posterior density ratios of 0.37, 0.050, 0.0067, and 0.00091, respectively. Parameter pairings were chosen to illustrate the clearest parameter correlations and biological interpretability. Roughness of contour lines at low densities reflects sparsity of sample points. Maximum-likelihood estimates were: fecundity, $F = 1.94$; nymph survival, $S_1 = 0.95$; maturation, $M = 0.20$; adult survival, $S_2 = 0.23$; density dependence (DD) on nymphs, $\delta_1 = -0.00071$; density dependence on adults, $\delta_2 = 0.079$; $\beta_{T,F} = 0.059$; process variance, $\sigma_v^2 = 0.10$; process correlation between nymphs and adults, $\rho = 0.95$; and measurement variance, $\sigma_D^2 = 0.12$.

runs because the parameters are fixed. One-step-ahead residuals (observed minus predicted) should be independent for a reasonably fitting model, and indeed no field's residuals showed significant autocorrelation for any models. Residuals were regressed against explanatory variable (temperature, nitrogen, predators) either to see whether those variables should be added to the model or, if already included, whether the model adequately captures their role. Significance levels of residual correlations were calculated by bootstrap because of apparent heteroscedasticity, and regressions of residuals vs. a possible predictor variable did not correspond exactly to the form of a relationship for building the predictor variable into the state-space model. Prior to these regressions, ANOVA was used to test the assumption of no between-field variation in residuals—i.e., whether the fields are really comparable—and Tukey's hsd method was used to identify different fields for removal from the correlation analyses against temperature, nitrogen, and predators.

RESULTS

Bayesian marginal posterior densities give insight into parameter estimability and correlations (Figs. 1 and 2, Table 1). For example, there is a ridge between higher F (daily fecundity) and lower S_1 (daily survival rate for aphid nymphs), so the data do not strongly distinguish between these parameters. A similar trade-off occurs between maturation rate (M) and adult survival rate (A_2), with apparently high uncertainty in these parameters. Nevertheless, the posterior distribution for the basic demographic parameters was generally biologically plausible. For example, Akey and Butler (1989), Xia et al. (1999), and Kersting et al. (1999) reared *Aphis gossypii* from different regions in labs across a range of temperatures and found nymph development times in the range of 4.5–8.5 days (corresponding roughly to $M = 0.12$ – 0.22), net nymph survival in the 70–95% range (corresponding to $S_1 = 0.94$ – 0.98 for a 6-d nymph stage), total adult-stage duration in the range of 7–15 d (corresponding roughly to $S_2 = 0.86$ – 0.93), and nymph

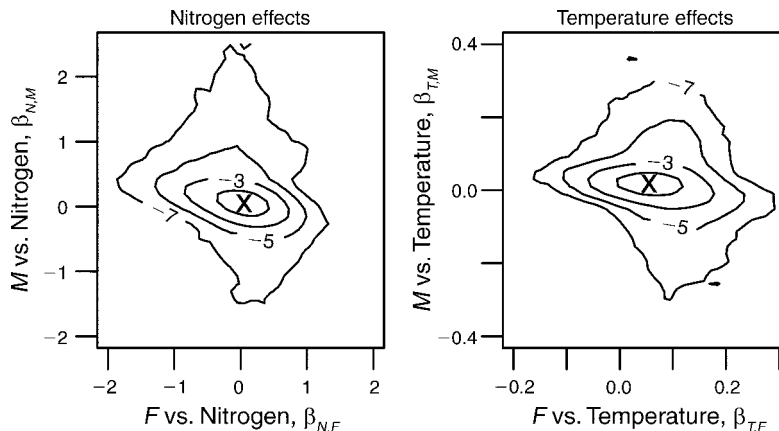


FIG. 2. Bayesian two-dimensional marginal posterior densities for models with either two nitrogen or two temperature parameters (GDD2nit2 and GDD2temp2, respectively). Each subfigure shows contours and log-density differences of $-1, -3, -5,$ and -7 from the maximum (“X”), corresponding to posterior density ratios of 0.37, 0.050, 0.0067, and 0.00091, respectively. The strongest evidence is for an effect of temperature on fecundity.

production per day (F) in the range of 1–3, all approximated from their data for relevant temperatures. With the possible exception of adult survival, these values are reasonably consistent with the posterior distributions (for GDD2tempF in Fig. 1). Density-dependence parameters were strongly nonzero for adults (δ_2), and appeared to be different for nymphs and adults (only 0.0006 of the posterior had $\delta_1 > \delta_2$).

Residual autocorrelation and differences between fields were tested for the baseline model GDD2 (Gompertz density dependence with two parameters). No field showed residual autocorrelation—supporting the approximate sequential independence of residuals—but there were significant differences between fields ($F_{17,168} = 2.55, P = 0.001$). Tukey’s hsd method identified fields 18 and 2 as the source of strong differences, and when they were removed, differences among fields were no longer significant ($F_{15,153} = 1.41, P$

$= 0.15$). All subsequent residual analysis omits those fields (with no meaningful effect on results).

Temperature but not nitrogen was significantly correlated with the one-step-ahead residuals from Model GDD2 (Table 2; the Appendix shows one-step-ahead predictions). When temperature effects on F and M (Model GDD2temp2) or just F (Model GDD2tempF) were included, the residual-vs.-temperature relation was random, indicating that one temperature component sufficiently captured its effect. When nitrogen effects on F and M were included (Model GDD2nit2), the residual-vs.-temperature relation was almost unchanged; temperature and nitrogen were significantly, but weakly, correlated ($R^2 = 0.03, P = 0.02$). These results accord qualitatively with the Bayesian posterior, which has

TABLE 1. Bayesian evidence for temperature and nitrogen effects on aphids.

Parameter	HPMD†	Evidence‡
Model GDD2temp		
F vs. Temp. ($\beta_{T,F}$)	0.062	[0.10, 23.1, 0.20]
M vs. Temp. ($\beta_{T,M}$)	0.018	[0.22, 6.42, 0.44]
Model GDD2nit		
F vs. Nit. ($\beta_{N,T}$)	0.056	[0.47, 8.8, 0.83]
M vs. Nit. ($\beta_{N,M}$)	0.058	[0.32, 4.8, 0.76]

Notes: Model GDD2temp2 has Gompertz density dependence with two parameters plus temperature effects (Temp.) on F (fecundity) and M (maturation rate). GDD2nit2 is a similar model for nitrogen (Nit.) effects.

† HPMD is the highest posterior marginal density estimate for the one-dimensional marginal distribution.

‡ Evidence is reported as [proportion of posterior density below zero (i.e., one-tailed evidence), ratio of highest posterior marginal density to density at 0, proportion of posterior density with lower density than 0 (i.e., two-tailed evidence)].

TABLE 2. Frequentist evidence and RMSE for temperature and nitrogen effects on aphid nymphs and adults.

Model	RMSE	Residuals vs. Temp.		Residuals vs. Nit.	
		R^2	P	R^2	P
GDD2					
Nymph	0.71	0.01	0.12	0.01	0.26
Adult	0.73	0.03	0.02	0.01	0.18
GDD2temp2					
Nymph	0.67	0.00	0.99	0.01	0.26
Adult	0.69	0.00	0.55	0.00	0.66
GDD2nit2					
Nymph	0.69	0.02	0.10	0.00	0.52
Adult	0.72	0.03	0.02	0.00	0.53
GDD2tempF					
Nymph	0.67	0.00	0.99	0.01	0.26
Adult	0.70	0.01	0.30	0.01	0.20

Notes: For each model, one-step-ahead residuals from the MLE (maximum-likelihood estimate) parameters were used to estimate RMSE (root mean squared error), and correlations with temperature and nitrogen. P values are bootstrapped; boldface indicates $P < 0.05$.

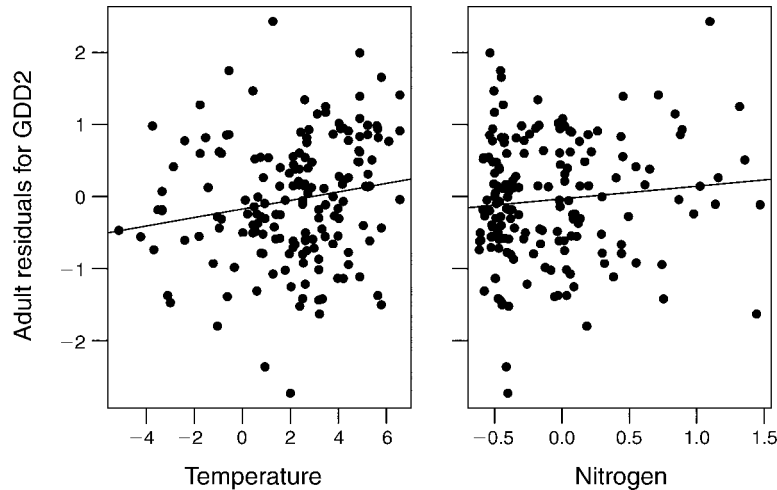


FIG. 3. Correlations of one-step-ahead adult residuals from Model GDD2 vs. temperature and nitrogen.

nitrogen parameters straddling zero and strongest evidence for the temperature effect on F (Table 1, Fig. 2). However, taken alone, the Bayesian evidence for the temperature effect might have been judged as equivocal, with 10% of the $\beta_{T,F}$ posterior below 0 and 20% with lower posterior density than 0. The temperature effect on fecundity was retained for testing of predator effects (Models GDD2tempFpred1).

All predators showed evidence for positive predation parameters (Fig. 3, Table 3). As a “control” analysis of a non-predator (put in the model as a predator), spider mites had a posterior centered near zero with 61% negative values. Regardless of possible mite–aphid interactions, this suggests that positive posteriors for the true predators are not an analysis artifact. Both Bayesian and frequentist analysis gave similar point estimates of individual predation rates, α_P , with α_P

decreasing in the order: *Zelus*, lacewing larvae, *Nabis*, *Orius*, *Geocoris* for sweep-net data; and lacewing larvae, *Geocoris*, lacewing eggs, *Orius* for whole-plant data. While the statistical strength of each ordering is unclear, the overall order accords well with predator-enclosure experiments, with lacewing larvae having clearly the highest predation rate and *Nabis* having a higher rate than the smaller species *Orius* and *Geocoris*. The clearest surprise is the high predation rate of *Zelus*, a large species but one that could potentially positively affect aphids by suppressing other predators but does not appear to do so at field scales.

Estimates of the net effect of each predator on aphid population growth were calculated using the median population growth for each predator and for aphids. The posterior distribution (Bayesian) of population growth rate (λ) was calculated, giving population effects in decreasing

TABLE 3. Bayesian estimates and evidence for predator effects.

Predator	HPMD†		Evidence‡	Median§	Aphid growth
	SN	WP			
None	N/A	N/A	N/A	432	[1.46, 1.67, 1.91]
<i>Orius</i>	0.0036	0.0018	[0.045, 25.8, 0.17]	6.7	[1.41, 1.61, 1.84]
<i>Geocoris</i>	0.0016	0.0077	[0.033, 28.7, 0.15]	19.3	[1.39, 1.60, 1.83]
Lacewing eggs	N/A	0.0022	[0.010, 65.1, 0.06]	23.3	[1.35, 1.55, 1.78]
Lacewing larvae	0.0075	0.011	[0.015, 59.3, 0.07]	4.7	[1.38, 1.59, 1.82]
<i>Zelus</i>	0.012	N/A	[0.05, 25.4, 0.18]	1.4	[1.42, 1.63, 1.86]
<i>Nabis</i>	0.0069	N/A	[0.09, 15.4, 0.31]	2.0	[1.44, 1.65, 1.87]
Mites	$1.2 \times 10^{-5}\¶$		[0.61, 7.8, 0.53]	214	[1.48, 1.72, 1.93]

Notes: Maximum-likelihood estimates are: *Orius*, 0.0032; *Geocoris*, 0.0015; lacewing eggs, 0.0019; lacewing larvae, 0.0072; *Zelus*, 0.011; and *Nabis*, 0.0066. N/A means not available or not applicable.

† HPMD is the highest posterior (one-dimensional) marginal density for α_P , on the scale of sweep-net (SN) and/or whole-plant (WP) counts.

‡ Evidence is reported as [proportion of posterior density below zero (i.e., one-tailed evidence), ratio of highest posterior marginal density to density at 0, proportion of posterior density with lower density than 0 (i.e., two-tailed evidence)].

§ The median abundance across all data, used to estimate effect on aphid population growth rate. For the “none” row, median aphids per 50 leaves is given.

|| The [5th, 50th, 95th] percentiles of posterior distribution of aphid population growth for median aphid and predator populations. For the “none” row, no predator is included.

¶ Herbivorous mites were counted from washed leaves, not sweep nets or whole-plant searches.

TABLE 4. Frequentist evidence for predator effects and root-mean-squared-error (RMSE) for predator models.

Predator	Nymphs			Adults		
	RMSE	Resid. R^2	Resid. P	RMSE	Resid. R^2	Resid. P
<i>Orius</i>	0.69	0.00	0.92	0.71	0.00	0.66
<i>Geocoris</i>	0.66	0.03	0.02	0.67	0.02	0.05
Lacewing eggs	0.69	0.00	0.68	0.70	0.01	0.18
Lacewing larvae	0.66	0.00	0.58	0.69	0.00	0.98
<i>Zelus</i>	0.68	0.00	0.72	0.70	0.01	0.29
<i>Nabis</i>	0.69	0.08	0.0001	0.71	0.08	0.0001
Mites	0.66	0.00	0.49	0.69	0.01	0.19

Notes: RMSE is reported for the GDD2tempFpred1 model with each predator. Residual R^2 and P values give correlations between residuals of the GDD2tempF model vs. predator time series. Boldface indicates $P < 0.05$. There were no significant correlations of residuals vs. predator divided by estimated total aphids, corresponding to a saturated functional response.

order of: lacewing eggs, lacewing larvae, *Geocoris*, *Orius*, *Zelus*, *Nabis*. Incorporating frequentist uncertainty in parameter estimates for such a calculation is not as straightforward, but simply using the MLE parameters gives an identical ranking of net effects. Again, while evidence for the relative order of each predator—even if each is judged as meaningfully different from zero—appears weak, the pattern of *Orius* and *Geocoris* (smaller, more abundant predators that are more likely to eat aphids) having larger population-level effects than *Nabis* and *Zelus* (larger, less abundant predators able to eat other predators as well as aphids) is consistent with our predictions, while the predicted role of lacewings was that it would have small population effect due to intraguild predation.

While the two methods accord in ranking individual and population-level effects, they differ in assessing the weight of evidence for each effect. Bayesian evidence essentially follows the ranking of individual effects, while frequentist analysis of GDD2tempF residuals vs. predators (Table 4) revealed statistically significant effects only for *Nabis* and *Geocoris* (both negative; Fig. 4) with significance decreasing in the order: *Nabis*, *Geocoris*, lacewing larvae, lacewing eggs, *Zelus*, *Orius*.

These correlations did not account for functional-response saturation such as in the GDD2tempFpred1 model, and correlations of residuals vs. predators/(Total aphids), where (Total aphids) is the mean estimated (latent) aphid abundance, revealed no significant correlations. Thus, frequentist evidence suggests that *Nabis* and *Geocoris* are significantly associated with below-expected *Aphis gossypii* densities, but this pattern is not associated with a functional-response model.

While the density, temperature, and predator effects, but not the lack of nitrogen effect, are plausibly consistent with small-scale studies, they fall short of explaining net aphid dynamics. All of the estimated effects contributed only negligible improvements to the residual root-mean-squared-error, which was in the range 0.67–0.71 and 0.69–0.73 for nymphs and adults, respectively. These values correspond to about two thirds of roughly one-week predictions falling within about a factor of 2 from the actual data, which cannot be considered strong prediction. Weak contributions to predictive power accord with the low R^2 values of the frequentist correlations for temperature and predator effects, and also may explain the discrepancies between strength of Bayesian and frequentist predator evidence.

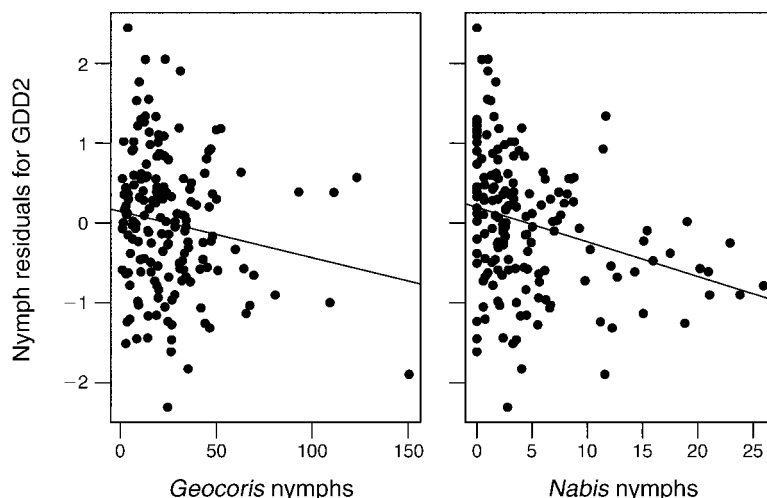


FIG. 4. Correlations of one-step-ahead nymph residuals from Model GDD2tempF vs. *Nabis* and *Geocoris*.

DISCUSSION

We conducted a complementary Bayesian and frequentist state-space model analysis that incorporates stage-structured population dynamics with short time steps and allows top-down, bottom-up, density-dependence, and abiotic effects. Estimates of basic aphid demographic rates were biologically reasonable, although with high statistical variability, with the possible exception of adult survival. Estimates of stage-specific density dependence suggest that unstructured models of density dependence may not capture field dynamics of some populations. Noisy effects of temperature and some predators were detected but contributed little explanatory power.

Thus, while the nonlinear state-space model approach and joint Bayesian and frequentist analysis has allowed insight into noisy statistical relationships, the overall predictability of field-scale aphid dynamics is only barely improved by including data on most of the major factors elucidated by small-scale experiments. This is a sobering result that raises the questions of when and how well field-scale predictability of insect dynamics can be achieved from small-scale experiments. It also suggests that combining maximum-likelihood estimation and residual analysis of overall predictive power may contribute to robust application of state-space population models in a manner complementary to Bayesian posterior analysis.

What processes might drive aphid dynamics that were beyond the scope of our data and model? Both herbivore and/or predator populations could be driven by landscape-scale processes, so that field scales are too small to estimate important driving variables; or population fluctuations could occur at smaller scales than we measured, so that our large-scale data obscured important processes. Unmeasured factors such as soil biota and watering regimes, neighboring crops, and even short-term evolution (Via and Shaw 1996) could have affected aphids. Phenotypic plasticity of morphology and life history is known to be large for *A. gossypii*. Had the stage-structured model achieved better predictability, one might have argued that it successfully reflected phenotypic plasticity (e.g., as a density or temperature effect), but it is possible that it was somehow structurally inadequate or that data were not sufficiently accurate to estimate all model effects. Most natural populations are spatially open, subject to a wide range of environmental conditions and disturbances, and are expected to undergo complicated dynamics. The questions of whether, when, and at what scale a relatively complete statistical accounting of population dynamics can be achieved is an important area for future research.

ACKNOWLEDGMENTS

J. A. Rosenheim was supported by the California State Support Board of Cotton Incorporated, the University of California Statewide IPM program, and the USDA NRICGP (grants 94-37302-0504, 96-35302-3816, and 01-02894). P.

de Valpine was supported by a USDA NRICGP post-doctoral grant during the initial stages of this work.

LITERATURE CITED

- Akey, D. H., and G. D. Butler. 1989. Developmental rates and fecundity of *Apterous Aphis gossypii* (Homoptera, Aphididae) on seedlings of *Gossypium hirsutum*. *Southwestern Entomologist* 14:295–299.
- Besbeas, P., S. N. Freeman, B. J. T. Morgan, and E. A. Catchpole. 2002. Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. *Biometrics* 58:540–547.
- Calder, C., M. Lavine, P. Muller, and J. S. Clark. 2003. Incorporating multiple sources of stochasticity into dynamic population models. *Ecology* 84:1395–1402.
- Carpenter, S. R. 1996. Microcosm experiments have limited relevance for community and ecosystem ecology. *Ecology* 77:677–680.
- Chib, S., and I. Jeliazkov. 2001. Marginal likelihood from the Metropolis-Hastings output. *Journal of the American Statistical Association* 96:270–281.
- Cisneros, J. J., and L. D. Godfrey. 2001. Midseason pest status of the cotton aphid (Homoptera: Aphididae) in California cotton: Is nitrogen a key factor? *Environmental Entomology* 30:501–510.
- Cisneros, J. J., and J. A. Rosenheim. 1997. Ontogenetic change of prey preference in the generalist predator *Zelus renardii* and its influence on predator-predator interactions. *Ecological Entomology* 22:399–407.
- Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. *Ecology Letters* 8:2–14.
- Dennis, B., J. M. Ponciano, S. R. Lele, M. L. Taper, and D. F. Staples. 2006. Estimating density dependence, process noise, and observation error. *Ecological Monographs* 76:323–341.
- de Roos, A. M., E. McCauley, and W. G. Wilson. 1991. Mobility versus density-limited predator-prey dynamics on different spatial scales. *Proceedings of the Royal Society of London B* 246:117–122.
- de Valpine, P. 2003. Better inferences from population-dynamics experiments using Monte Carlo state-space likelihood methods. *Ecology* 84:3064–3077.
- de Valpine, P. 2004. Monte Carlo state-space likelihoods by weighted posterior kernel density estimation. *Journal of the American Statistical Association* 99:523–536.
- de Valpine, P., and A. Hastings. 2002. Fitting population models incorporating process noise and observation error. *Ecological Monographs* 72:57–76.
- de Valpine, P., and R. Hilborn. 2005. State-space likelihoods for nonlinear fisheries time-series. *Canadian Journal of Fisheries and Aquatic Sciences* 62:1937–1952.
- Gibson, G. J. 1997. Investigating mechanisms of spatiotemporal epidemic spread using stochastic models. *Phytopathology* 87:139–146.
- Gross, K., A. R. Ives, and E. V. Norpheim. 2005. Estimating fluctuating vital rates from time-series data: a case study of aphid biocontrol. *Ecology* 86:740–752.
- Kareiva, P. 1990. Population-dynamics in spatially complex environments: theory and data. *Philosophical Transactions of the Royal Society of London B* 330:175–190.
- Karley, A. J., A. E. Douglas, and W. E. Parker. 2002. Amino acid composition and nutritional quality of potato leaf phloem sap for aphids. *Journal of Experimental Biology* 205:3009–3018.
- Kersting, U., S. Satar, and N. Uygun. 1999. Effect of temperature on development rate and fecundity of apterous *Aphis gossypii* Glover (Hom., Aphididae) reared on *Gossypium hirsutum* L. *Journal of Applied Entomology* 123:23–27.
- Meng, X. L., and W. H. Wong. 1996. Simulating ratios of normalizing constants via a simple identity: a theoretical exploration. *Statistica Sinica* 6:831–860.

- Meyer, R., and R. B. Millar. 1999. Bayesian stock assessment using a state-space implementation of the delay difference model. *Canadian Journal of Fisheries and Aquatic Sciences* 56:37–52.
- Miner, B. G., S. E. Sultan, S. G. Morgan, D. K. Padilla, and R. A. Relyea. 2005. Ecological consequences of phenotypic plasticity. *Trends in Ecology and Evolution* 20:685–692.
- Mira, A., and G. Nicholls. 2004. Bridge estimation of the probability density at a point. *Statistica Sinica* 14:603–612.
- Nevo, E., and M. Coll. 2001. Effect of nitrogen fertilization on *Aphis gossypii* (Homoptera: Aphididae): Variation in size, color, and reproduction. *Journal of Economic Entomology* 94:27–32.
- Newman, K. B., S. T. Buckland, S. T. Lindley, L. Thomas, and C. Fernandez. 2006. Hidden process models for animal population dynamics. *Ecological Applications* 16:74–86.
- Nylin, S., and K. Gotthard. 1998. Plasticity in life-history traits. *Annual Review of Entomology* 43:63–83.
- Pascual, M., and S. A. Levin. 1999. From individuals to population densities: searching for the intermediate scale of nontrivial determinism. *Ecology* 80:2225–2236.
- Rosenheim, J. A. 1998. Higher-order predators and the regulation of insect herbivore populations. *Annual Review of Entomology* 43:421–447.
- Rosenheim, J. A. 2001. Source-sink dynamics for a generalist insect predator in habitats with strong higher-order predation. *Ecological Monographs* 71:93–116.
- Rosenheim, J. A., D. D. Limburg, and R. G. Colfer. 1999. Impact of generalist predators on a biological control agent, *Chrysoperla carnea*: direct observations. *Ecological Applications* 9:409–417.
- Rosenheim, J. A., L. R. Wilhoit, and C. A. Armer. 1993. Influence of intraguild predation among generalist insect predators on the suppression of an herbivore population. *Oecologia* 96:439–449.
- Shurin, J. B., E. T. Borer, E. W. Seabloom, K. Anderson, C. A. Blanchette, B. Broitman, S. D. Cooper, and B. S. Halpern. 2002. A cross-ecosystem comparison of the strength of trophic cascades. *Ecology Letters* 5:785–791.
- Slosser, J. E., R. Montandon, W. E. Pinchak, and D. R. Rummel. 1997. Cotton aphid (Homoptera: Aphididae) response to nitrogen fertility in dryland cotton. *Southwestern Entomologist* 22:1–10.
- Slosser, J. E., M. N. Parajulee, D. L. Hendrix, T. J. Henneberry, and W. E. Pinchak. 2004. Cotton aphid (Homoptera : Aphididae) abundance in relation to cotton leaf sugars. *Environmental Entomology* 33:690–699.
- Slosser, J. E., W. E. Pinchak, and D. R. Rummel. 1998. Biotic and abiotic regulation of *Aphis gossypii* Glover in west Texas dryland cotton. *Southwestern Entomologist* 23:31–65.
- Steele, M. A., and G. E. Forrester. 2005. Small-scale field experiments accurately scale up to predict density dependence in reef fish populations at large-scales. *Proceedings of the National Academy of Sciences (USA)* 102:13513–13516.
- Stillwell, R. C., and C. W. Fox. 2005. Complex patterns of phenotypic plasticity: Interactive effects of temperature during rearing and oviposition. *Ecology* 86:924–934.
- Tscharntke, T., and R. Brandl. 2004. Plant-insect interactions in fragmented landscapes. *Annual Review of Entomology* 49: 405–430.
- Via, S., and A. J. Shaw. 1996. Short-term evolution in the size and shape of pea aphids. *Evolution* 50:163–173.
- Xia, J. Y., W. van der Werf, and R. Rabbinge. 1999. Influence of temperature on bionomics of cotton aphid, *Aphis gossypii*, on cotton. *Entomologia Experimentalis et Applicata* 90:25–35.
- Youssefi, F., P. H. Brown, and S. A. Weinbaum. 2000. Relationship between tree nitrogen status, xylem and phloem sap amino acid concentrations, and apparent soil nitrogen uptake by almond trees (*Prunus dulcis*). *Journal of Horticultural Science and Biotechnology* 75:62–68.

APPENDIX

Figures presenting the aphid, predator, temperature, and nitrogen data from each field and year, with estimated conditional aphid distribution given maximum-likelihood parameters (*Ecological Archives* E089-028-A1).