*Ecological Archives E090-161-A1*

**Leo Polansky, Perry de Valpine, James 0. Lloyd-Smith, and Wayne M. Getz. 2009. Likelihood ridges and multimodality in population growth rate models. Ecology 90:2313-2320.**

Appendix A. The appendix includes more discussion on the theta-Ricker and generalized Beverton-Holt density-dependent models, a complete description of likelihood equations and methods of parameter estimation, elaboration on the simulation experiment, additional analyses of the *A. nisus* data set, TRPN best-fit and ALT models for 25 additional time series from the GPDD, and a presentation of the $r$-$\theta$ joint profile likelihood surface of a *C. oenas* data set for comparison with Bayesian posterior densities.

The appendix has six parts: A1) discussion of plausible *pgr* models defined by the theta-Ricker and generalized Beverton-Holt equations; A2) likelihood equations for all the stochastic models included in the analysis; A3) elaboration on the simulation experiment using a stochastic theta-Ricker with process noise model to obtain the frequency of best-fit $r$-$\theta$ estimates lying in the third quadrant of the $r$-$\theta$ plane and actual vs. nominal coverage plots to evaluate the accuracy of the likelihood ratio test for a limited sample size; A4) additional analyses of the focal *Accipiter nisus* (Eurasian sparrowhawk) population presented in the text; A5) theta-Ricker concave best-fit and supported convex ALT *pgr* models for 25 additional species from the Global Population Dynamics Database (GPDD) (NERC 1999); A6) Maximum likelihood based analysis of a *Columba oenas* (stock pigeon) density time series from the GPDD to compare with a Bayesian treatment of this model and data set in Ward (2006) and a least-squares point estimate by Sibly et al. (2005).

We denote the continuous population density at time $t$ by $N_t \geq 0$, a set of data by $\boldsymbol{N}^T = \{N_1,\ldots, N_T\}$, and define the associated *pgr* data by $\log_e(N_{t+1}/N_t)$. Using the natural logarithm of population density $N_t$ has useful statistical and heuristic advantages when conducting statistical inference with the models considered here (Dennis and Taper 1994, de Valpine and Hastings 2002, Turchin 2003). We denote the true state variable (the potentially unobserved true population density) by $x_t = \log_e(N_t)$, constituting the true population density trajectory $\boldsymbol{X}^T = \{x_1,\ldots, x_T\}$, and the observed density (data) denoted by $y_t$ which makes up the complete set of observations $\boldsymbol{Y}^T = \{y_1,\ldots, y_T\}$.

## SECTION A1. GEOMETRICAL DESCRIPTIONS AND CALCULUS OF TWO DETERMINISTIC PGR MODELS

We first review the qualitative features of the deterministic *pgr* curves for both the limiting behavior as $N_t \to 0$ or $N_t \to \infty$ (for fixed $t$) and the slope at the point equilibrium $N^*$ where the *pgr* is zero for the theta-Ricker and $\gamma$-BH models. We comment that the theta-Ricker with $\theta$ fixed at 1 (also known as the Ricker model) and $\gamma$-BH model with $\gamma$ fixed at 1 are special cases of a general model presented and discussed by Schnute (1985).

Theta-Ricker density dependence (Thomas et al. 1980) models the *pgr* curve as $g(N_t) = r(1 - (N_t/K)^\theta$, where $r$ is the maximum *pgr* which is approached monotonically as population density declines to 0, $K>0$ defines the density at which the *pgr* is 0, and $\theta$ describes the form of the density dependence (Fig. 1). The slope of the theta-Ricker model at the point equilibrium $N^* = K$ for all $t$ (the carrying capacity) is $-r\theta/K$. Thus, for $r$ and $\theta$ both negative or both positive, the *pgr* curve is decreasing at the point equilibrium. If either $r$ or $\theta$ equal 0, the

*pgr* curve is 0 for all $N_t$. If they are of opposite signs, the *pgr* curve is increasing over all $N_t$. Thus, the axes and the second and fourth quadrants of the $r$-$\theta$ plane are biologically implausible hypotheses for density-dependent regulated population growth. When $r$ and $\theta$ are both positive, the shape of the *pgr* curve is well known (Turchin 2003) and ranges between concave for $0<\theta<1$, linear at $\theta = 1$, and convex when $\theta>1$, with limiting behavior $pgr \to r$ as $N_t \to 0$ and decreasing towards $-\infty$ as $N_t \to \infty$ (Fig. 1). For both $r$ and $\theta$ less than 0, the $pgr \to \infty$ as $N_t \to 0$ and decreases towards $r$ as $N_t \to \infty$. Although describing locally similar *pgr* curves around $N^*$ (Fig. 1) as parameters from the first quadrant of the $r$-$\theta$ plane, third quadrant parameters are nevertheless implausible because of the unrealistic *pgr* values for low values of $N_t$.

The generalized Beverton-Holt ($\gamma$-BH) sets $g(N_t) = r - \ln(1 + (N_t / K)^{\gamma})$ (Maynard Smith and Slatkin 1973), where $\gamma$ controls the shape of density dependence and $K$ is the density for which $dg(N_t)/dN_t$ is maximized. Assuming $r>0$, positive values of $\gamma$ can produce concave ($0<\gamma\leq 1$) or sigmoidal ($\gamma>1$) declines in the *pgr* curve, with the $pgr \to r$ as $N_t \to 0$ and $pgr \to -\infty$ as $N_t \to \infty$. Although less tractable statistically in comparison to the theta-Ricker model (i.e. there is no convenient transformation allowing linear or power regression to be used for parameter estimation), the $\gamma$-BH model has been useful in estimating general shapes of *pgr* curves and is perhaps more biologically appropriate (Bellows 1981, Getz 1996). Additionally, the $\gamma$-BH model is an alternative model that may offer some stability to parameter estimation because of the following. The equilibrium $N^* = K(e^r - 1)^{1/\gamma}$ of the $\gamma$-BH model and the slope of the *pgr* curve at this point given by $-\gamma(e^r - 1)^{1-1/\gamma} / Ke^r$ imply that the parameter subspaces $r\leq 0$ and $\gamma\leq 0$ are biologically nonsensical regions for populations regulated by density dependence because $\gamma = 0$ creates a singularity in the model, $r\leq 0$ yields $N^* \leq 0$, and $r>0$ with $\gamma<0$ result in a *pgr* curve with positive slope at $N^*$. Thus, for *pgr* data that are decreasing with density, parameter estimates

would be expected only in the first quadrant of the $r$-$\gamma$ plane, rather than two quadrants as in the theta-Ricker model.

## SECTION A2. MODELS AND LIKELIHOOD EQUATIONS

Time series models and their likelihood functions are standard tools in population ecology, and we refer readers to Dennis and Taper (1994), Hilborn and Mangel (1997), or de Valpine and Hastings (2002) for clear introductions and examples of their construction in this field of science. Introductions with more general statistical and theoretical material can be found in Tong (1990) or, for example, Shumway and Stoffer (2000).

We include process noise, observation error, or both, as stochastic elements of our models to explain differences between the deterministic predictions and data. Throughout we assume both process noise and observation error are additive, independent and normally distributed, in logarithmic population space. The stochastic process model can be written as $x_{t+1} = F(x_t, \upsilon_t) = x_t + g(x_t) + \upsilon_t$, where $F : R \to R$ is the deterministic part of the model and $\upsilon_t$ is normally distributed with zero mean and standard deviation $\sigma_p$. The stochastic observation model $G$ we use sets $y_t = G(x_t, \varepsilon_t) = x_t + \varepsilon_t$, where $\varepsilon_t$ is normally distributed with zero mean and standard deviation $\sigma_o$. The set of model parameters is denoted by $\Theta$.

For density-dependent models with process noise only, observations $y_t$ are assumed to be exactly the values of $x_t$, and the likelihood of the parameters is the product of the conditional probability density function values at $y_{t+1}$, given $y_t$ and the parameters $\Theta$ of the model, for $t$=1,…,$T$-1. For the specific case of the theta-Ricker and $\gamma$-BH models with process noise, the log-likelihood equations are given by

$$\ell\left(r,K,\theta,\sigma_p\middle|\boldsymbol{Y}^T\right)=(T-1)\log\left(\frac{1}{\sigma_p\sqrt{2\pi}}\right)-\frac{1}{2\sigma_p^2}\sum_{t=1}^{T-1}\left[y_{t+1}-\left(y_t+r(1-(\exp(y_t)/K)^{\theta}\right)\right]^2 \quad \text{Eq. A.1}$$

and

$$\ell\left(r,K,\gamma,\sigma_p\middle|\boldsymbol{Y}^T\right)=(T-1)\log\left(\frac{1}{\sigma_p\sqrt{2\pi}}\right)-\frac{1}{2\sigma_p^2}\sum_{t=1}^{T-1}\left[y_{t+1}-\left(y_t+r-\log(1+(\exp(y_t)/K)^{\gamma})\right)\right]^2, \quad \text{Eq. A.2}$$

respectively. For these models, the least squares and maximum likelihood (ML) estimates of the structural parameters (i.e. all but the variances) are equivalent (Shumway and Stoffer 2000).

Two ways in which the process noise only parametric models described above may be incorrect/inadequate for real data are because of process misspecification or by omission of observation error. Process misspecification is beyond the scope of this research. However, we can include observation error. The simplest way is to model variation in data as a combination of the phenomenological model and observation error, but no process noise. For this model, the value of the true population density at $x_1$ must also be estimated and is included in the vector $\boldsymbol{\Theta}$. Given $x_1$, the theta-Ricker deterministic model, and a set of structural parameters, the deterministic trajectory is given by

$$\boldsymbol{X}^T=\left\{x_t:\ x_1 \text{ is given and } x_t=x_{t-1}+r(1-(\exp(x_{t-1})/K)^{\theta} \text{ for } t=2,...,T\right\} \quad \text{Eq. A.3a}$$

and the associated log-likelihood equation is

$$\ell\left(r,K,\gamma,\sigma_o,x_1\middle|\boldsymbol{Y}^T\right)=T\log\left(\frac{1}{\sigma_o\sqrt{2\pi}}\right)-\frac{1}{2\sigma_o^2}\sum_{t=1}^{T}\left[y_t-x_t\right]^2. \quad \text{Eq. A.3b}$$

The most realistic model we consider combines both process noise and observation error with a state-space approach. For a theta-Ricker model, the likelihood of $\boldsymbol{\Theta}=(r, K, \theta, \sigma_p, \sigma_o)$ is

$$L\left(\boldsymbol{\Theta}\mid\boldsymbol{Y}^T\right)=\int P\left(\boldsymbol{Y}^T\mid\boldsymbol{X}^T,\boldsymbol{\Theta}\right)P(\boldsymbol{X}^T\mid\boldsymbol{\Theta})\mathrm{d}\boldsymbol{X}^T \quad \text{Eq. A.4}$$

where $P\left(\boldsymbol{Y}^T \mid \boldsymbol{X}^T, \Theta\right)$ is the conditional probability density of the data, given the states and the error model *G*, and $P(\boldsymbol{X}^T \mid \Theta)$ is the conditional probability density of the states given the process model *F*. A general description of the numerical integration procedure we follow for calculating the likelihood equation Eq. A.4 is given in Kitagawa (1987) and de Valpine and Hastings (2002). We refer the reader to de Valpine and Hastings (2002) for a population ecology based example and the exact recipe we follow (with the exception of the initial state distribution, which is discussed below). This method is based on numerical integration of the likelihood and thus approximates the true likelihood, allowing us to use ML methods to obtain best-fit point estimates, construct joint profile likelihood surfaces, conduct likelihood ratio tests of alternative parameters (models), and use information-theoretic methods of model selection.

## SECTION A3. SIMULATION EXPERIMENT TO TEST FOR 3RD QUADRANT ML ESTIMATES

We performed a simulation experiment to evaluate the frequency with which point estimates of *r* and $\theta$ in the best-fit theta-Ricker model will be found in the third quadrant as discussed in section A1, when the true generating parameters are in the first quadrant and describe biologically plausible sigmoidal *pgr* curves. To maintain emphasis on likelihood surfaces in the absence of complicating factors such as observation error or model misspecification, we generated data with process noise but no observation error using $x_{t+1} = x_t + r(1 - (\exp(x_t)/K)^{\theta}) + \upsilon_t$ to simulate population density trajectories. For data time series $\boldsymbol{Y}^T$ that have information about *r* with points away from *K*, e.g. the population undergoes some perturbation or deterministic cycles induced by strong density-dependent effects, point estimates of the true generating values were very accurate and negative estimates of $\theta$ did not

occur using sample sizes $T$ = 10, 20, or 50 and for two levels of process noise $\sigma_p$ = 0.05 or $\sigma_p$ = 0.25. For stationary data time series without perturbations or cyclic deterministic attractors, negative estimates for $r$ and $\theta$ did occur. Figure 1B-C in the main text presents the distribution of $r$ and $\theta$ maximum likelihood estimates from Eq. A.1 fit to each of 300 time series from such a scenario where the generating parameters are $r = \log_e(1.2)$, $\theta = 2$, $K = 1$, and $\sigma_p = 0.05$.

There are a number of methods available for constructing confidence regions to evaluate the plausibility of alternative parameters (models) in the same model family. Some common approaches (e.g. the score and Wald statistics) are obtained from the expected or observed information matrices evaluated at the ML point and assume elliptical confidence regions (Severini 2000). The potential disjoint and hyperbolic regions of likelihood values similar to the ML solution in the parameter space defining the *pgr* curve of a theta-Ricker model (Figs. 1B-C) demonstrate that these methods could be highly inaccurate and not fully characterize confidence regions. (The shape of the regions might be made more elliptical by re-parameterization, which could improve the performance of these methods, but we use the standard parameterization to compare to previous work.)

We also used the simulated time series and associated $\ell(\hat{\boldsymbol{\Theta}})$ values of Eq. A.1 to evaluate the performance of the likelihood ratio test (LRT) with limited sample size. The alternative model in this case is the theta-Ricker with $\boldsymbol{\psi}_0$ fixed at the "true" values of ($r$, $K$, $\theta$, $\sigma_p$) used in the simulation, so we compare the sample quantiles defined by $\boldsymbol{\Lambda}_{\psi_0} := 2(\ell(\hat{\boldsymbol{\Theta}}) - \ell(\hat{\boldsymbol{\Theta}}_{\psi_0}))$ with the theoretical quantiles of a $\chi^2_4$ distribution. For this case of a perfectly correct model structure, we found that using the LRT to obtain quantiles when estimating confidence regions with unbounded ML parameters consistently exceeds (slightly overestimates) actual coverage using the nominal quantiles and is improved by bounding $r$ and $\theta$ estimates above 0, thereby forcing

the ML solution to correspond more closely to the consistent mode in the likelihood function (Fig. A1).

## SECTION A4. ADDITIONAL *ACCIPITER NISUS* MODEL ANALYSES AND DATA SETS

For each of the models described in section A2, we optimized the objective functions Eqs. A.1-A.4 using a Nelder-Mead algorithm given the *A. nisus* data time series from the GPDD (ID #6575) which has a GPDD assigned reliability rating of 3 on a scale of 1 (worst) to 5 (best). The results are shown in Table A1. We used multiple restarts with different initial parameter values to ensure the highest chances of convergence on global optima. For the density-dependent model including observation error but no process noise, Eq. A.3, the likelihood surface was quite complex due to estimating $x_1$, so we also included runs with a simulated annealing algorithm to obtain maximum likelihood parameter estimates. Table A2 summarizes these results.

**Table A1.** *A. nisus* population time series from two different sources.

| | Population Size | |
|---|---|---|
| Year | GPDD | Saether et al. (2002) Digitized |
| 1972 | 33 | |
| 1973 | 36 | |
| 1974 | 39 | |
| 1975 | 33 | 33 |
| 1976 | 34 | 34 |
| 1977 | 29 | 29 |
| 1978 | 36 | 36 |
| 1979 | 32 | 32 |
| 1980 | 38 | 38 |
| 1981 | 39 | 39 |
| 1982 | 32 | 32 |
| 1983 | 34 | 35 |
| 1984 | 34 | 34 |
| 1985 | 33 | 32 |
| 1986 | 33 | 32 |
| 1987 | 36 | 35 |
| 1988 | 35 | 34 |
| 1989 | 30 | 29 |

**Table A2.** Models and associated statistics given the *A. nisus* data set. Estimates of $\hat{N}^*$ rather than $\hat{K}$ are provided for comparison between model families. See main text for description of ALT models. Other model abbreviations are: RPN- Ricker ($\theta = 1$) with process noise; TRPN- theta-Ricker with process noise; TRSS- theta-Ricker with process noise and observation error; TROE- theta-Ricker with observation error; BHPN- Beverton-Holt ($\gamma = 1$) with process noise; $\gamma$-BHPN- generalized Beverton-Holt with process noise.

| Model and likelihood location | $r$ | $N^*$ | $\theta$, $\gamma$ | $\sigma_p$ | $\sigma_o$ | $\hat{N}_1$ | $\ell(\Theta)$ |
|---|---|---|---|---|---|---|---|
| RPN Eq. A.1 (with $\theta = 1$) best-fit | 1.13 | 34.30 | - | 0.08 | - | - | 18.61 |
| TRPN Eq. A.1 first quadrant best-fit | 6.35e2 | 34.20 | 1.81e-3 | 0.08 | - | - | 18.79 |
| TRPN Eq. A.1 third quadrant best-fit | -0.37 | 34.41 | -3.09 | 0.08 | - | - | 19.07 |
| TROE Eq. A.3 first quadrant best-fit | 0.17 | 34.72 | 17.06 | - | 0.07 | 35.94 | 21.48 |
| TROE Eq. A.3 third quadrant best-fit | -0.08 | 33.60 | -43.95 | - | 0.05 | 32.70 | 27.11 |
| TRSS Eq. A.4 interior first quadrant best-fit | 31.44 | 34.19 | 0.04 | 0.04 | 0.07 | - | 18.41 |
| TRSS Eq. A.4 interior third quadrant best-fit | -0.25 | 33.98 | -4.83 | 0.05 | 0.06 | - | 18.49 |
| TRSS Eq. A.4 ALT | 0.71 | 34.23 | 2.2 | 0.03 | 0.07 | - | 18.30 |
| BHPN Eq. A.2 (with $\gamma = 1$) best-fit | 9.15 | 34.18 | - | 0.08 | - | - | 18.63 |
| $\gamma$-BHPN Eq. A.2 best-fit | 5.17 | 34.21 | 1.15 | 0.08 | - | - | 18.79 |
| $\gamma$-BHPN Eq. A.2 ALT | 0.50 | 34.37 | 2.86 | 0.08 | - | - | 18.44 |

Several approximations are necessary when computing the integral in Eq. A.4 in practice, partly depending on the data used, which we mention here. First, we chose for the probability density function of $x_1$, prior to conditioning on any observations $y_t$, a normal density with mean equal to the average of $\boldsymbol{Y}^T$ and standard deviation 0.15. Other choices for this distribution (e.g. a normal density with mean set at the parameter $K$ and standard deviation 0.10 or $\sigma_p + \sigma_o$ ) did not qualitatively change the overall results, but do result in changes of the likelihood function value on an order of magnitude around 1. For example, preliminary analyses using a standard deviation of $\sigma_p + \sigma_o$ for the initial state resulted in global ML estimate at $(\hat{r},\hat{K},\hat{\theta},\hat{\sigma}_p,\hat{\sigma}_o)$ = (-0.58, 34.07, -2.12, 0.06, 0.05) with a log-likelihood value of 19.55. It is not appropriate to directly compare likelihood values for the state-space model Eq. A.4 with the other models due to treatment of the initial state and the resolution of the grid which influence numerical approximations of the likelihood; a fair comparison would involve approximating likelihood values based on a grid for all models A.1-A.4.

It is also necessary to put finite bounds on the lower and upper limits of possible values for $x_t$, for which we chose $x_{min} = \log_e(1)$ and $x_{max} = \log_e(1000)$ (see Table A1 for the values of this population). None of the probability distributions of Eq. A.4 were affected by these boundaries of the unobserved states. Finally, it is necessary to pick a resolution of the grid used to approximate the continuous space of the continuous random variables $x_t$ and $y_t$. We partitioned the interval $[x_{min}, x_{max}]$ into 4096 increments; finer partitioning did not appear to improve likelihood function values significantly.

The likelihood surface for the theta-Ricker state-space model Eq. A.4 requires some explanation and is relatively difficult to work with. Parameter estimation of even comparatively simple linear state-space models with Gaussian distributed noise can often result in maximum

likelihood values at the boundaries of the parameter space, and in particular at either $\sigma_p$ or $\sigma_o$ equal to 0 (Dennis et al. 2006). We found that for the theta-Ricker state-space model Eq. A.4 and the *A. nisus* data set, when fixed values of $r$ or $\theta$ are near zero, the optimization would converge on maximum at the boundary $\sigma_p = 0$, giving a model of essentially constant population size with data variation due to observation error. Allowing $\sigma_p \approx 0$ or $\sigma_o \approx 0$ leads to numerical difficulties for the state-space model likelihood, so we bounded these with minimum allowed values of 0.01. By initiating the optimization algorithm away from these boundaries at $r$ = +/-1, $K$ equal to the average of the population density, $\theta$ = +/-1, and $\sigma_p = \sigma_o = 0.05$, we were able to find maxima with $\sigma_p > 0.01$ and $\sigma_o > 0.01$ for the global and first-quadrant (of $r$-$\theta$) maximizations over all parameters (Table A2, Figure 1). We refer to these modes as the interior modes presented in Table A2; focusing on the interior modes increases both the biological plausibility of the model and chances of obtaining the consistent likelihood mode (Dennis et al. 2006).

When computing the joint profile likelihood surface, we first selected a grid of points in the $r$-$\theta$ plane to compute joint profile likelihoods (Fig. A3). Joint profile likelihoods were obtained by starting the optimization from several regions of the $\sigma_p$-$\sigma_o$ parameter subspace to obtain the best maximum value for fixed $\boldsymbol{\psi}_0 = (r_0, \theta_0)$. Likelihood values for several grid points of the $r$-$\theta$ joint profile likelihood surface in the third quadrant along the $r = 0$ axis were minutely higher (on the order of 1e-2) than the result of the global maximization, reflecting the convergence tolerance of the latter, even with restarts, and the inherent difficulty of optimization on a nearly flat, long curved ridge.

We considered several versions of the *A. nisus* data set to detect whether differences between our and Sibly et al.'s (2005) negative estimates and Saether et al.'s (2002) positive estimate was

an artifact of slight differences in the time series analyzed and not of the different fitting procedures. Although the particular references cited in Saether et al. (2002) and the GPDD differ, both time series describe the *A. nisus* population in Eskdale, Scotland through the 1970's and 1980's, as reported in Newton (1986), Newton and Rothery (1997), NERC (1999), and Saether et al. (2002). Because we do not have access to the reference cited for the data in Saether et al. (2002), we digitized the data from this figure with the results shown in Table A1. Note that the caption to Figure 2 of Saether et al. (2002) labels the symbols incorrectly. The empty squares show the *A. nisus* data, while the filled squares show data for the South Polar Skua *Catharacta maccormicki* (see the date ranges in their Supplementary Information to confirm). Close inspection of the data reveals a few minor differences from the GPDD data set. As shown in Table A1 the two time series clearly describe the same data set for the period 1975-1989, with a few points where the estimates differ by one. The power regression method (regression of the *pgr* data $\log_e(N_{t+1}/N_t)$ on $\{N_1^\theta,...,N_{T-1}^\theta\}$), over a grid of predetermined points used by Sibly et al. (2005) for the GPDD data set, our power regressions but without the grid for both the GPDD and digitized data sets, or maximum likelihood based estimates for both the GPDD and digitized data sets all yield negative estimates of $\theta$ (Table A3). Saether et al. (2002) also used a maximum-likelihood approach – although crucially they estimated *r* independently with a separate analysis – and obtained a positive estimate of $\theta$ larger than 1, contradicting the conclusion of concavity.

**Table A3**. Estimates of $\theta$ using maximum likelihood or power regression for three different versions of the *A. nisus* data set from Eskdale, Scotland. We include results of Sibly et al. (2005) and Saether et al. (2002) for ease of comparison.

| Method | Time range | $\hat{\theta}$ | CL lower | CL upper | Data source |
|---|---|---|---|---|---|
| Saether et al. (2002) | 1975-1989 | 2.57 | standard deviation 0.69 | | Saether et al. (2002) |
| power regression | 1975-1989 | -5.01 | -20 | 5.01 | digitized from Fig. 2 in Saether et al. (2002) |
| ML | 1975-1989 | -5.59 | | | digitized from Fig. 2 in Saether et al. (2002) |
| power regression | 1975-1989 | -3.98 | -15.8 | 6.31 | GPDD |
| ML | 1975-1989 | -4.27 | not provided | | GPDD |
| Sibly et al. (2005) | 1972-1989 | -3.16 | -12.6 | 10 | GPDD |
| ML | 1972-1989 | -3.09 | see text, figure 2A | | GPDD |

## SECTION A5. ADDITIONAL DATA SETS

Table A2.4 and Fig. A2.4 provide the specific details of the additional data sets referred to in the main text and supplement Fig. 3.

**Table A4.** Parameter estimates of the theta-Ricker with process noise model Eq. A.1 fit to an additional 25 data sets from the GPDD (ID number and reliability value are shown in parentheses). If the location of the mode for the maximum likelihood was in the third quadrant of the $r$-$\theta$ plane, we also found the mode for a local maximum in the first quadrant. LRT $p$-values are for rejecting the ALT model chosen along the likelihood ridge, given the global and local best-fit models, and follow a $\chi^2_2$ distribution. $T$- length of the time series. Model abbreviations: ML, gl- global maximum likelihood estimates; ML, lo- local maximum likelihood estimates; ALT- alt model with $r$ and $\theta$ chosen and Eq. A.1 maximized in $K$ and $\sigma_p$.

| Data set | Model | $r$ | $N^*$ | $\theta$ | $\sigma_p$ | LL | LRT $p$-value |
|---|---|---|---|---|---|---|---|
| *Abraxas grossulariata* | ML, gl | -0.07 | 3.94 | -10.84 | 0.03 | 31.77 | 0.50 |
| (6470, 2), Magpie moth, $T$ | ML, lo | 23.91 | 3.96 | 0.02 | 0.03 | 31.48 | 0.67 |
| = 17 | ALT | 0.06 | 3.97 | 6.95 | 0.03 | 31.08 | - |
| *Abraxas sylvata* (6467, 2), | ML, gl | -0.23 | 2.05 | -2.59 | 0.16 | 7.05 | 0.20 |
| Clouded magpie, $T$ = 17 | ML, lo | 465.71 | 2.14 | 1.37e-3 | 0.16 | 6.37 | 0.40 |
| | ALT | 0.30. | 2.20 | 1.86 | 0.17 | 5.44 | - |
| *Acrocephalus scirpaceus* | ML, gl | -0.69 | 4.60e3 | -0.26 | 0.30 | -3.60 | 0.30 |
| (9871, 2), Reed warbler, $T$ | ML, lo | 620.08 | 4.41e3 | 3.93e-4 | 0.30 | -3.66 | 0.32 |
| = 18 | ALT | 0.39 | 5.10e3 | 1.17 | 0.32 | -4.80 | - |

Table A4 continued

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Aegolius funereus* (9922, 3), Tengmalm's owl, *T* = 19 | ML, gl | -16.68 | 7.45 | -0.08 | 0.77 | -20.80 | 0.20 |
| | ML, lo | 604.94 | 7.64 | 2.15e-3 | 0.77 | -20.81 | 0.20 |
| | ALT | 1.18 | 10.56 | 1.10 | 0.84 | -22.41 | - |
| *Agrotis segetum* (6215, 2), Turnup moth, *T* = 17 | ML, gl | -0.63 | 3.11 | -1.30 | 0.10 | 14.65 | 0.40 |
| | ML, lo | 150.14 | 3.13 | 0.01 | 0.10 | 14.58 | 0.43 |
| | ALT | 0.19 | 3.19 | 3.97 | 0.10 | 13.73 | - |
| *Anas americana* (5051, -1), American wigeon, *T* = 17 | ML, gl | -1.22 | 9.42 | -0.75 | 0.35 | -5.70 | 0.45 |
| | ML, lo | 800.30 | 9.91 | 1.13e-3 | 0.35 | -6.49 | 0.51 |
| | ALT | 0.70 | 10.72 | 1.27 | 0.36 | -6.49 | - |
| *Caradrina morpheus* (6277, 2), Mottled rustic, *T* = 17 | ML, gl | -0.11 | 4.25 | -8.67 | 0.04 | 29.01 | 0.50 |
| | ML, lo | 71.55 | 4.28 | 0.01 | 0.04 | 28.69 | 0.69 |
| | ALT | 0.19 | 4.29 | 4.44 | 0.04 | 28.31 | - |
| *Columba oenas* (62, 4), Stock pigeon, *T* = 13 | ML, gl | 1.91 | 62.86 | 0.20 | 0.25 | -0.63 | 0.15 |
| | ALT | 0.77 | 58.79 | 1.31 | 0.30 | -2.52 | - |
| *Conophthorus resinosae* (9228, 2), Red pine cone beetle, *T* = 10 | ML, gl | 4.26 | 1.92e3 | 0.30 | 0.69 | -9.46 | 0.20 |
| | ALT | 0.56 | 2.19e3 | 1.36 | 0.83 | -11.07 | - |
| *Ennomoa fuscantaria* (6483, 2), Dusky thorn, *T* = 17 | ML, gl | -0.33 | 3.21 | -2.99 | 0.06 | 21.28 | 0.45 |
| | ML, lo | 298.91 | 3.23 | 3.40e-3 | 0.06 | 21.11 | 0.53 |
| | ALT | 0.24 | 3.26 | 4.08 | 0.07 | 20.48 | - |

Table A4 continued

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Falcaria lacertinaria* (6180, 2), Scalloped hook tip, $T = 17$ | ML, gl | -0.20 | 3.27 | -5.37 | 0.05 | 26.41 | 0.50 |
| | ML, lo | 142.92 | 3.29 | 0.01 | 0.05 | 26.17 | 0.64 |
| | ALT | 0.26 | 3.31 | 4.10 | 0.05 | 25.72 | - |
| *Felidae* (9693, 1), Cat, $T = 25$ | ML, gl | 7.26 | 1.36e4 | 0.09 | 0.68 | -24.68 | - |
| | ALT | 0.26 | 1.51e4 | 1.48 | 0.71 | -25.89 | 0.30 |
| *Gadus macrocephalus* (1810, 2), Pacific cod, $T = 13$ | ML, gl | 0.69 | 793.59 | 0.94 | 0.06 | 15.87 | - |
| | ALT | 0.52 | 794.66 | 1.73 | 0.07 | 14.66 | 0.30 |
| *Hirundo rustica* (9856, 2), swallow, $T = 18$ | ML, gl | 0.92 | 2.10e3 | 0.73 | 0.22 | 1.87 | - |
| | ALT | 0.17 | 2.28e3 | 3.25 | 0.23 | 0.95 | 0.40 |
| *Meles meles* (9692, 1), Eurasian badger, $T = 25$ | ML, gl | -5.01 | 1.58e3 | -0.10 | 0.42 | -13.38 | 0.50 |
| | ML, lo | 233.81 | 1.60e3 | 2.30e-3 | 0.42 | -13.39 | 0.50 |
| | ALT | 0.45 | 1.87e3 | 1.57 | 0.44 | -14.07 | - |
| *Ochlodes venata* (3621, 4), Large skipper, $T = 10$ | ML, gl | -1.00 | 41.50 | -1.58 | 0.24 | 0.11 | 0.12 |
| | ML, lo | 5.40e3 | 44.43 | 2.74e-4 | 0.26 | -0.76 | 0.28 |
| | ALT | 1.19 | 46.83 | 1.16 | 0.30 | -2.01 | - |
| *Peromyscus maniculatus* (6692, 2), Deer mouse, $T = 12$ | ML, gl | -10.76 | 5.26 | -0.10 | 0.57 | -9.50 | 0.15 |
| | ML, lo | 671.10 | 5.36 | 1.56e-3 | 0.57 | -9.51 | 0.15 |
| | ALT | 0.62 | 7.39 | 1.87 | 0.68 | -11.40 | - |

Table A4 continued

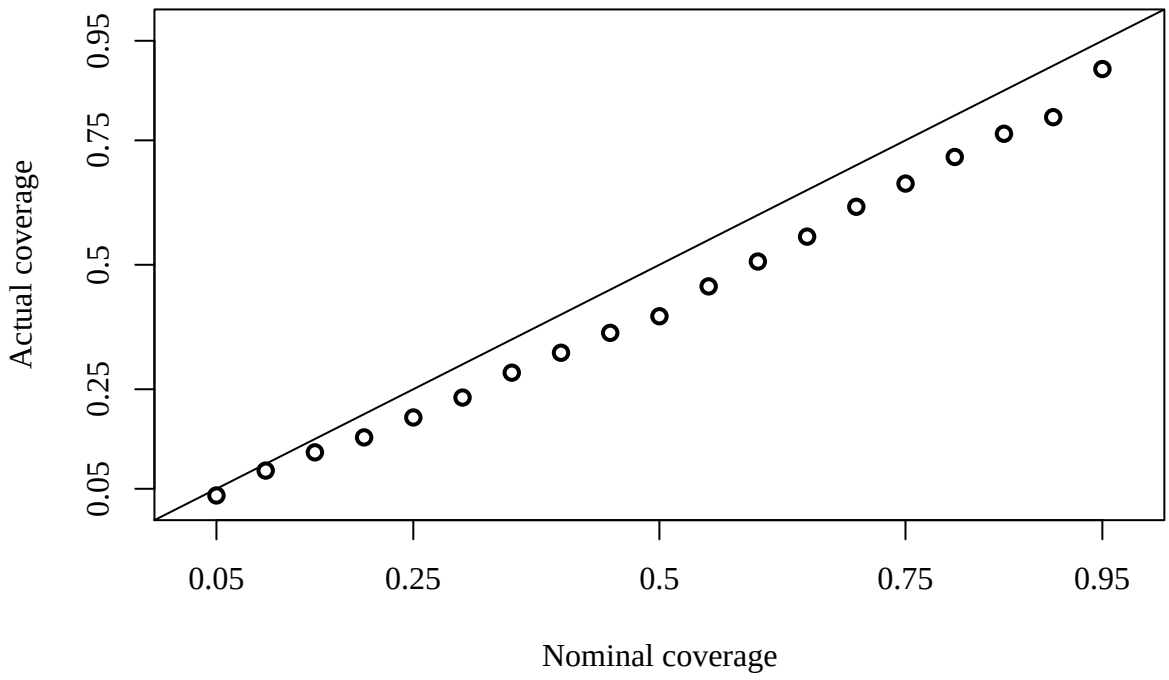
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Reinhart hippoglossoides* | ML, gl | -1.13 | 8.40e4 | -0.73 | 0.10 | 15.52 | 0.36 |
| (1715, -1), Greenland | ML, lo | 222.72 | 8.50e4 | 1.46e-3 | 0.10 | 15.50 | 0.37 |
| halibut, $T = 19$ | ALT | 0.39 | 8.70e4 | 1.35 | 0.11 | 14.31 | - |
| *Rhizedra lutosa* (6268, 2), | ML, gl | -0.06 | 2.66 | -10.83 | 0.08 | 17.63 | 0.30 |
| Large wainscot, $T = 17$ | ML, lo | 1.31e3 | 2.80 | 5.96e-4 | 0.08 | 16.76 | 0.71 |
| | ALT | 0.37 | 2.82 | 2.19 | 0.09 | 16.42 | - |
| *Salmo salar* (1626, -1), | ML, gl | 11.80 | 426.38 | 0.06 | 0.84 | -18.70 | 0.15 |
| Atlantic salmon, $T = 16$ | ALT | 0.10 | 480.98 | 1.43 | 0.96 | -20.60 | - |
| *Salmo trutta* (7023, -1), | ML, gl | -1.12 | 0.69 | -0.73 | 0.21 | 2.37 | 0.20 |
| Brown trout, $T = 17$ | ML, lo | 1.07e3 | 0.71 | 8.20e-4 | 0.21 | 2.25 | 0.23 |
| | ALT | 0.59 | 14.60 | 1.27 | 0.51 | 0.76 | - |
| *Selasphorus platycerus* | ML, gl | -1.33 | 12.43 | -0.91 | 0.43 | -5.11 | 0.20 |
| (9380, 4), Broad-tailed | ML, lo | 3.17e3 | 13.29 | 3.40e-4 | 0.45 | -5.52 | 0.30 |
| hummingbird, $T = 10$ | ALT | 0.59 | 14.60 | 1.27 | 0.51 | -6.72 | - |
| *Spiza americana* (9442, 2), | ML, gl | 1.30 | 60.88 | 0.70 | 0.23 | 1.39 | 0.40 |
| Dickcissel, $T = 28$ | ALT, l | 0.28 | 64.52 | 3.63 | 0.24 | 0.47 | - |
| *Thera obeliscata* (6428, 2), | ML, gl | -0.52 | 4.11 | -1.70 | 0.05 | 25.52 | 0.40 |
| Grey pine carpet, $T = 17$ | ML, lo | 0.06 | 4.19 | 14.59 | 0.05 | 25.51 | 0.41 |
| | ALT | 0.06 | 4.19 | 14.59 | 0.05 | 24.61 | |
| *Zeiraphera diniana* (6669, | ML, gl | -0.24 | 0.66 | -19.54 | 0.35 | -3.41 | 0.06 |
| -1), Larch budmoth, $T = 10$ | ML, lo | 849.79 | 1.48 | 9.04e-4 | 0.45 | -5.53 | 0.50 |
| | ALT | 0.65 | 1.61 | 1.05 | 0.48 | -6.22 | - |

## SECTION A6. ANALYSIS OF *COLUMBA OENAS* TIME SERIES

We analyzed a *C. oenas* population density time series in the GPDD (ID #62, reliability rating of 4), which has two previously published analyses by the theta-Ricker model: Sibly et al. (2005) reported a least-squares $\hat{\theta}$ estimate of 0.20 (-0.60, 2), and Ward (2006) used the same process model with observation error, applying a Bayesian treatment for parameter estimation to estimate a posterior mode for $\theta$ at 0.38. Similar to the analysis of the *A. nisus* data set, we used multiple restarts with different initial parameter values to ensure the highest chances of convergence on global optima and for the TROE model Eq. A.4 we also included runs with the simulated annealing optimization algorithm. We found the ML estimate for the TRPN model Eq. A.1 at $(\hat{r},\hat{K},\hat{\theta},\hat{\sigma}_p)$ = (1.91, 62.86, 0.20, 0.25); for the TROE model Eq. A.3 the MLE parameter estimate was $(\hat{r},\hat{K},\hat{\theta},\hat{\sigma}_0,N_1)$ = (1.91, 62.86, 0.20, 0.18, 0.92). Figure panels A5A and A5C show the $r$-$\theta$ joint profile likelihood for the TRPN and TROE models. The $r$-$\theta$ joint profile likelihood surface for both models show that the $r$-$\theta$ 0.95 confidence region based on the ML methods is qualitatively similar to the 95% central posterior density estimates for these parameters in Figure 1 of Ward (2006). An alternative TRPN model is given by ($r$, $K$, $\theta$, $\sigma_p$) = (0.77, 58.79, 1.31, 0.30), supported at a LRT $p$-value of 0.15 (Table A4, Fig. A5B).

Figure A1

a.) Unbounded parameter space




b.) Bounded parameter space

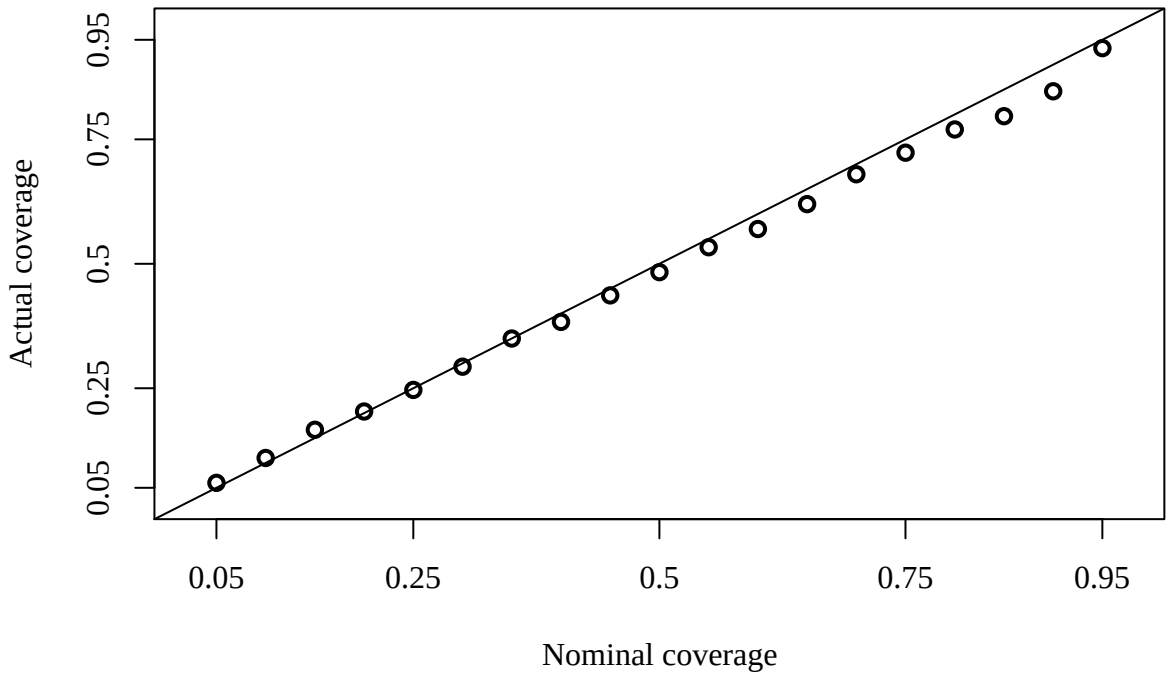


**FIG. A1**. Comparison of actual vs. nominal coverage using theoretical quantiles of the $\chi^2_4$ distribution for the distribution of $\Lambda_{\psi_0} = 2(\ell(\hat{\Theta}) - \ell(\hat{\Theta}_{\psi_0}))$ from the simulation experiment. Unbounded ML refers to $r$-$\theta$ third quadrant ML estimated parameters, bounded ML refers to first quadrant ML $r$-$\theta$ estimated parameters.

Figure A2

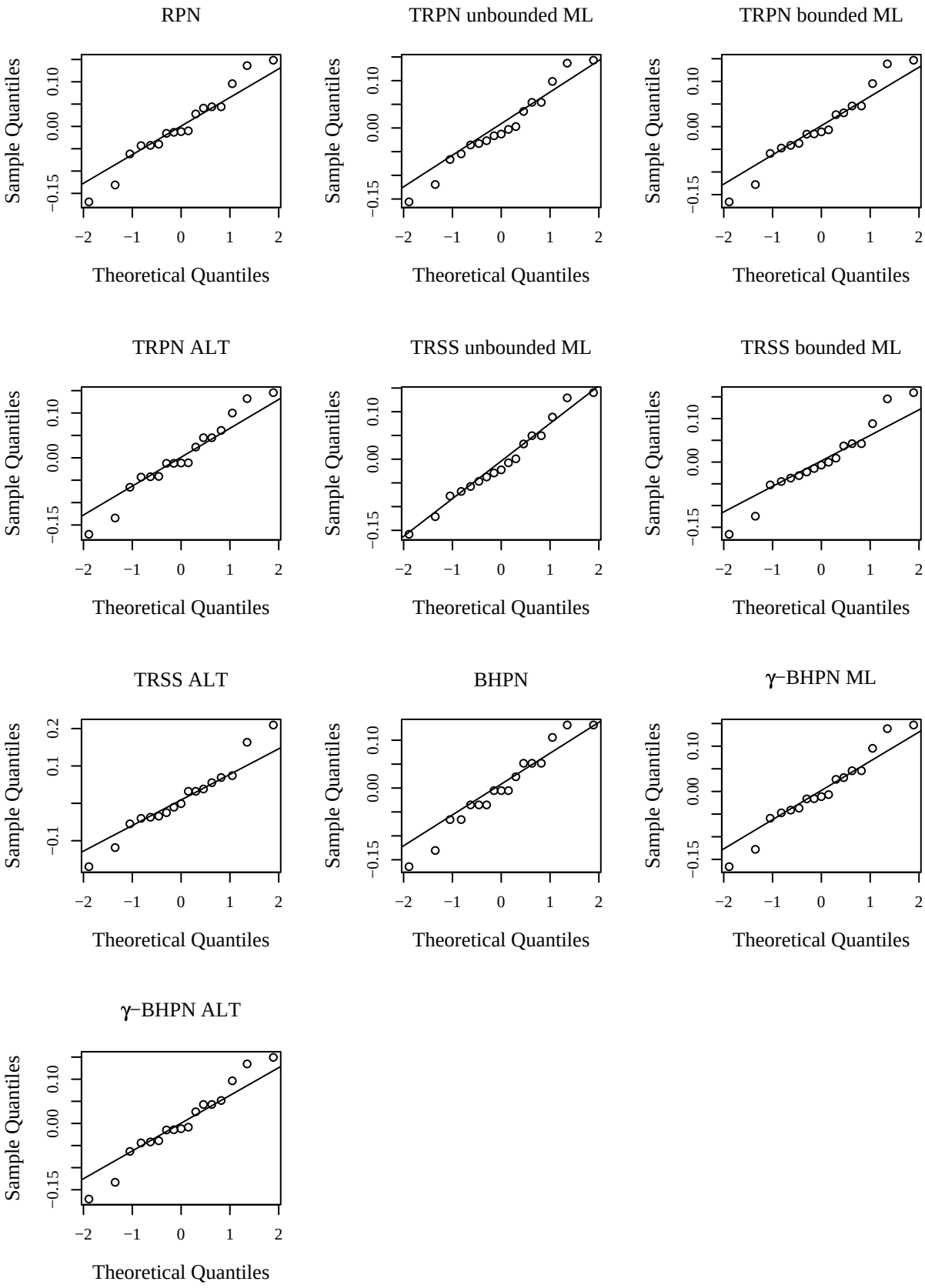

RPN
TRPN unbounded ML
TRPN bounded ML
TRPN ALT
TRSS unbounded ML
TRSS bounded ML
TRSS ALT
BHPN
γ−BHPN ML
γ−BHPN ALT
Sample Quantiles
Theoretical Quantiles

**FIG. A2**. Quantile-quantile plots of residuals from the models and data of Table 1 in the main text. Unbounded ML refers to $r$-$\theta$ third quadrant ML estimated parameters, bounded ML refers to first quadrant ML $r$-$\theta$ estimated parameters. Model abbreviations are: RPN- Ricker with process noise; TRPN unbounded ML- theta-Ricker with process noise; TRSS- theta-Ricker with process noise and observation error; BHPN- Beverton-Holt with process noise; $\gamma$-BHPN- generalized Beverton-Holt with process noise.

Figure A3

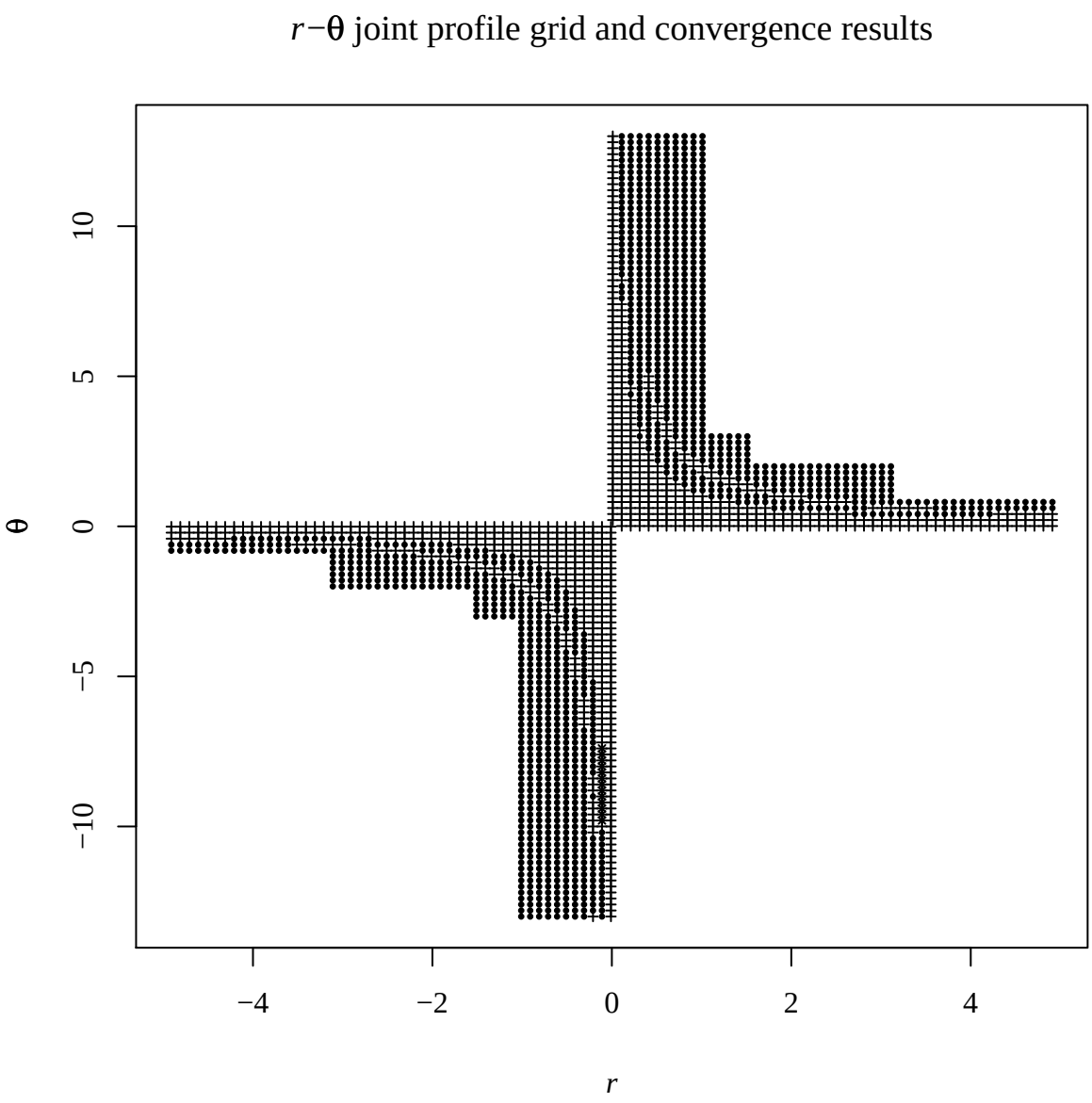

r−θ joint profile grid and convergence results
θ
10
5
0
−5
−10
−4
−2
0
2
4
r

**FIG. A3**. Grid of values $\psi_0 = (r_0, \theta_0)$ over which joint profile likelihood values were calculated by maximizing the likelihood function Eq. A.3 in the remaining parameters dimensions ($K$, $\sigma_p$, $\sigma_o$). Grid point symbols represent the following: solid points represent convergence away from the boundary $\sigma_p = \sigma_o = 0.01$, + represents convergence at $\sigma_p < 0.01$, and x shows values of $\psi_0$ for which joint likelihood values are slightly higher than the global interior maximum likelihood value.

Figure A4 (part 1 of 4)

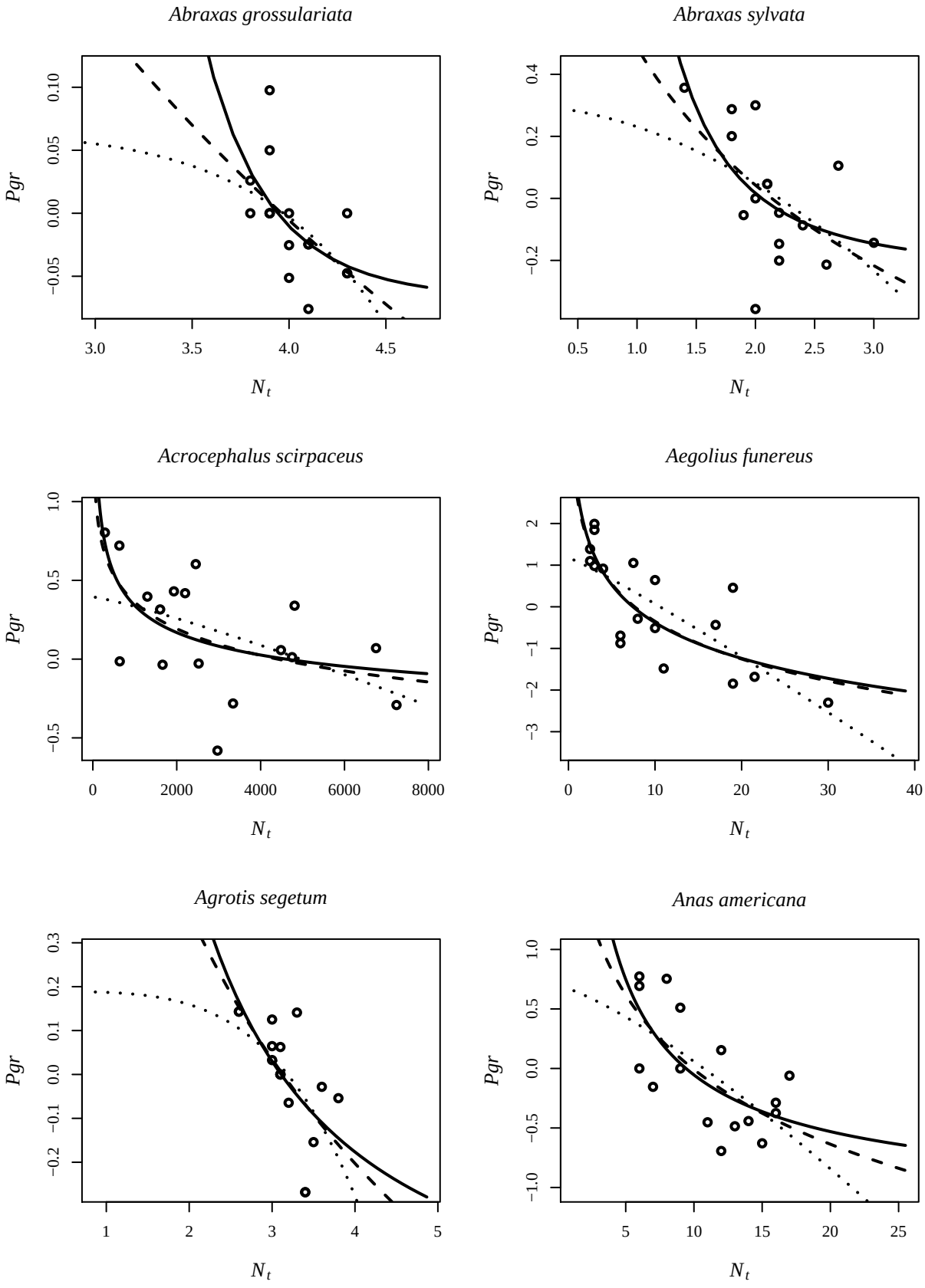


Figure A4 (part 2 of 4)

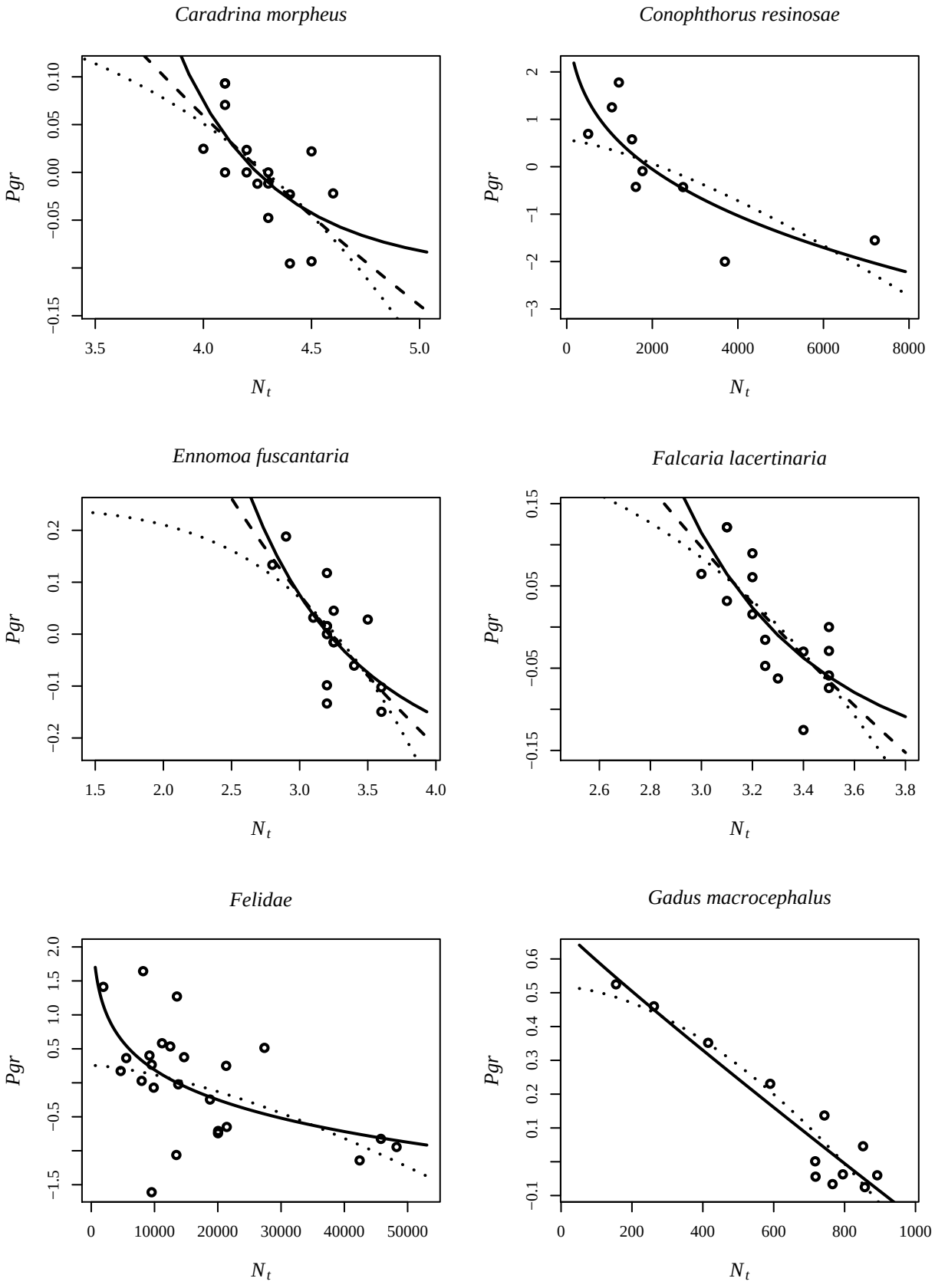


Figure A4 (part 3 of 4)

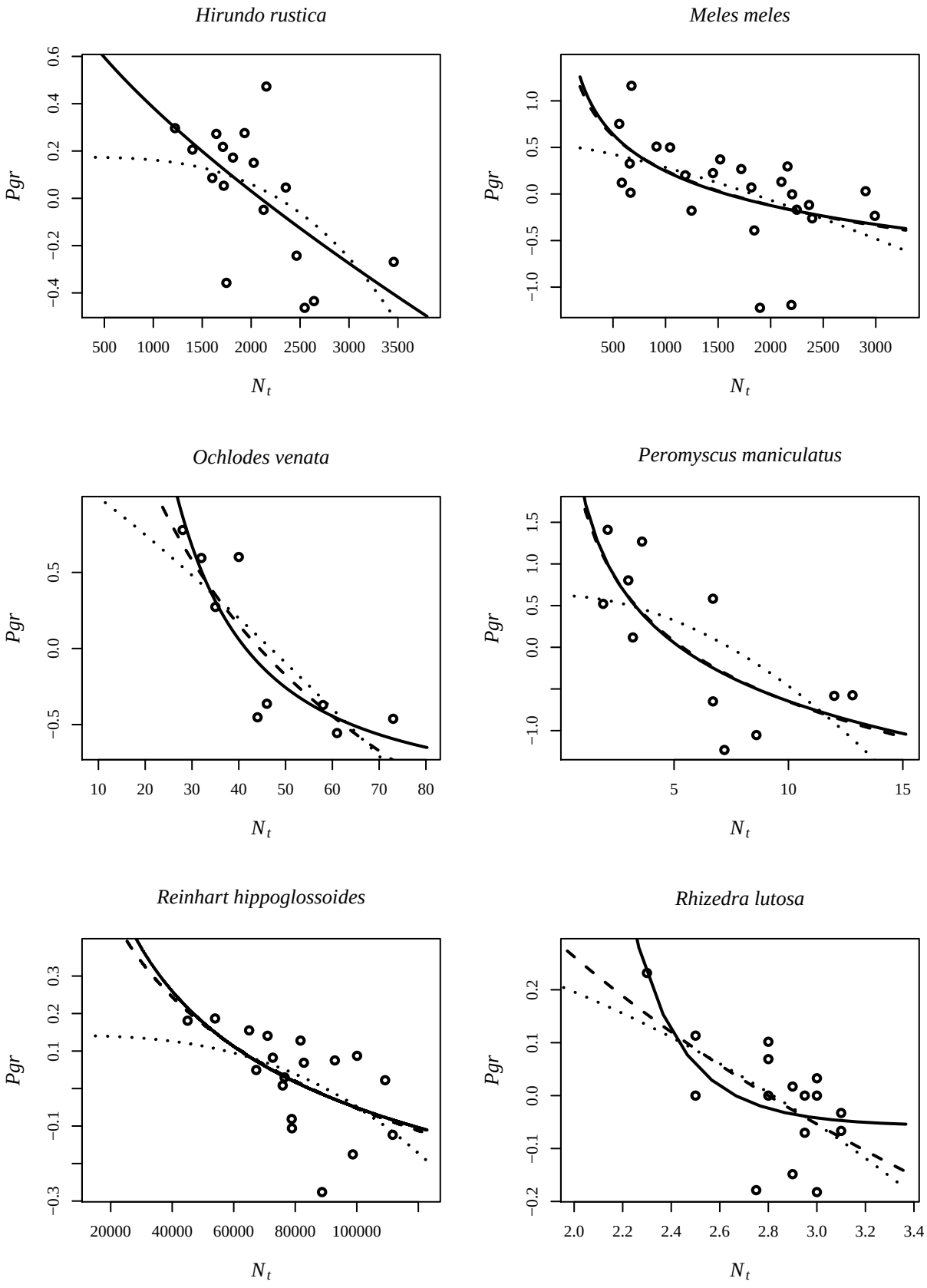


Figure A4 (part 4 of 4)

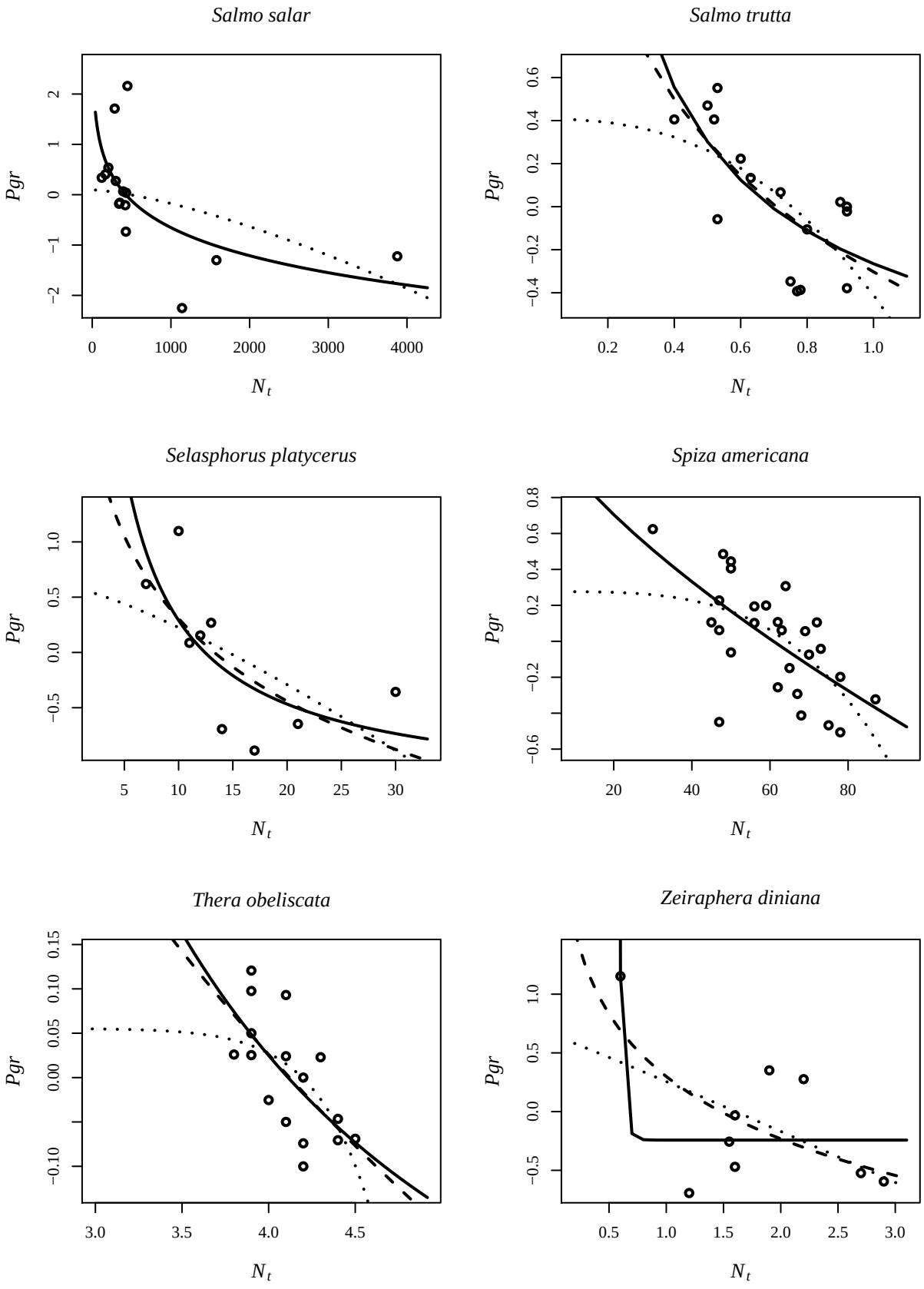
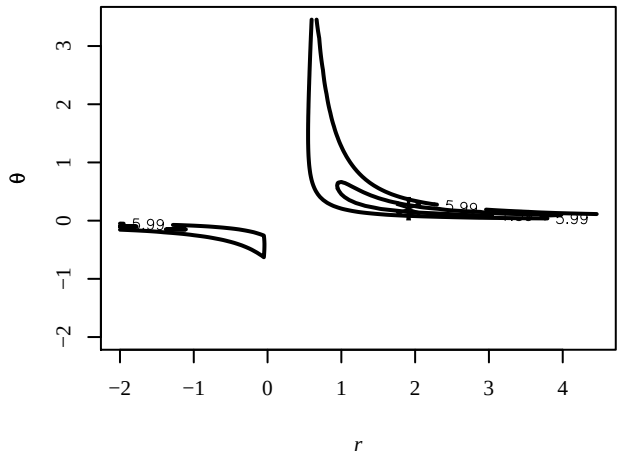


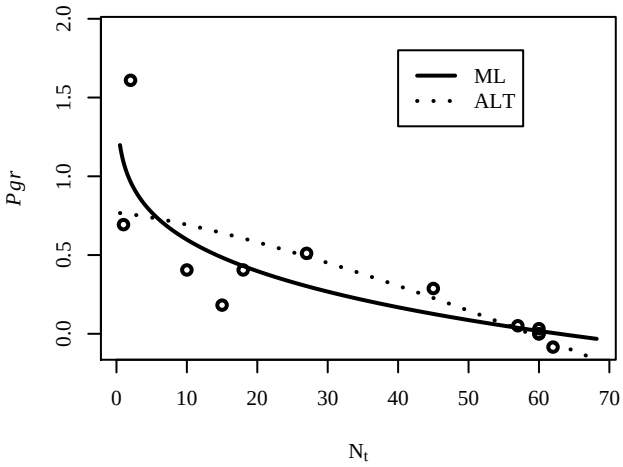
**FIG. A4**. *Pgr* data (circles) and predictions using models with global ML parameter estimates (solid line), local ML parameter estimates (dashed line, if different from the global ML ones; see Table A4), and ALT parameter estimates (dotted line).

# Figure A5

A) *C. oenas* TRPN $r-\theta$ joint profile likelihood region




B) *C. oenas* TRPN ML and ALT *pgr* models




C) *C. oenas* TROE $r-\theta$ joint profile likelihood region

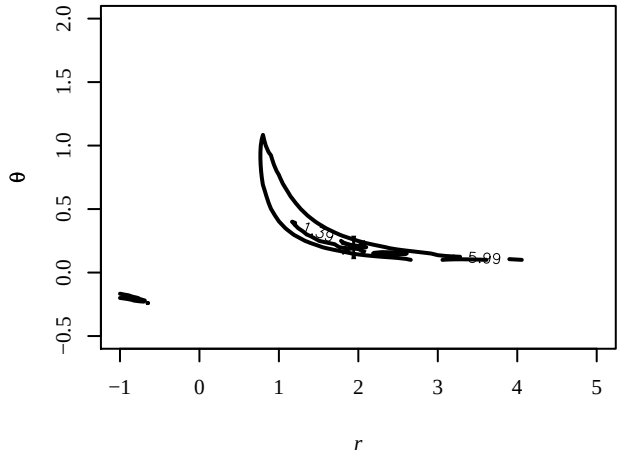


**FIG. A5**. For the *C. oenas* data set (A) shows the $r$-$\theta$ joint profile likelihood surface for a TRPN model with the ML estimated values of $r$ and $\theta$ located at *. (B) shows the *pgr* data (circles) and *pgr* predictions using the ML (solid line) and ALT (dashed line) parameters from Table A4. (C) shows the $r$-$\theta$ joint profile likelihood surface for a TROE model with the ML estimated values of $r$ and $\theta$ located at *. For panels A and C, the contour lines are drawn at the 0.5 and 0.95 quantiles of a chi-squared distribution with 2 degrees of freedom.

## LITERATURE CITED


Bellows, T. S. 1981. The Descriptive Properties of Some Models for Density Dependence. Journal of Animal Ecology 50:139–156.

Burnham, K. P., and D. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York, New York, USA.

de Valpine, P., and A. Hastings. 2002. Fitting population models incorporating process noise and observation error. Ecological Monographs 72:57–76.

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.

Dennis, B., and M. L. Taper. 1994. Density-Dependence in Time-Series Observations of Natural-Populations - Estimation and Testing. Ecological Monographs 64:205–224.

Getz, W. M. 1996. A hypothesis regarding the abruptness of density dependence and the growth rate of populations. Ecology 77:2014–2026.

Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, New Jersey, USA.

Kitagawa, G. 1987. Non-Gaussian state-space modeling of nonstationary time-series. Journal of the American Statistical Association 82:1032–1041.

Maynard Smith, J., and M. Slatkin. 1973. Stability of Predator-Prey Systems. Ecology 54:384:391.

National Environment Research Council (NERC) Centre for Population Biology, I. C. 1999. The Global Population Dynamics Database.

Newton, I., and M. Marquiss. 1986. Population regulation in sparrowhawks. Journal of Animal Ecology 55:463–480.

Newton, I., and P. Rothery. 1997. Senescence and reproductive value in Sparrowhawks. Ecology 78:1000–1008.

R Development Core Team (2007). R: A Language and Environment for Statistical Computing. In: Vienna, Austria.

Saether, B. E., S. Engen, and E. Matthysen. 2002. Demographic characteristics and population dynamical patterns of solitary birds. Science 295:2070–2073.

Schnute, J. 1985. A general theory for analysis of catch and effort data. Canadian Journal of Fisheries and Aquatic Sciences 42:414:429.

Severini, T. A. 2000. Likelihood Methods in Statistics. Oxford University Press, Oxford, UK.

Shumway, R. H., and D. S. Stoffer. 2000. Time Series Analysis and Its Applications. First edition. Springer-Verlag, Harrisonburg, Virginia, USA.

Sibly, R. M., D. Barker, M. C. Denham, J. Hone, and M. Pagel. 2005. On the regulation of populations of mammals, birds, fish, and insects. Science 309:607–610.

Thomas, W. R., M. J. Pomerantz, and M. E. Gilpin. 1980. Chaos, asymmetric growth and group selection for dynamic stability. Ecology 61:1312–1320.

Tong, H. 1990. Non-linear Time Series: a dynamical systems approach. Oxford University Press, New York, New York, USA.

Turchin, P. 2003. Complex population dynamics : a theoretical/empirical synthesis. Princeton University Press, Princeton, New Jersey, USA.

Ward, E. J. 2006. A new BEAST: Bayesian software tools for ecological trend analysis. Wildlife Society Bulletin 34:1420–1424.