

# **Feature matching versus likelihood for dynamic systems: Nicholson's blowflies as a case study**

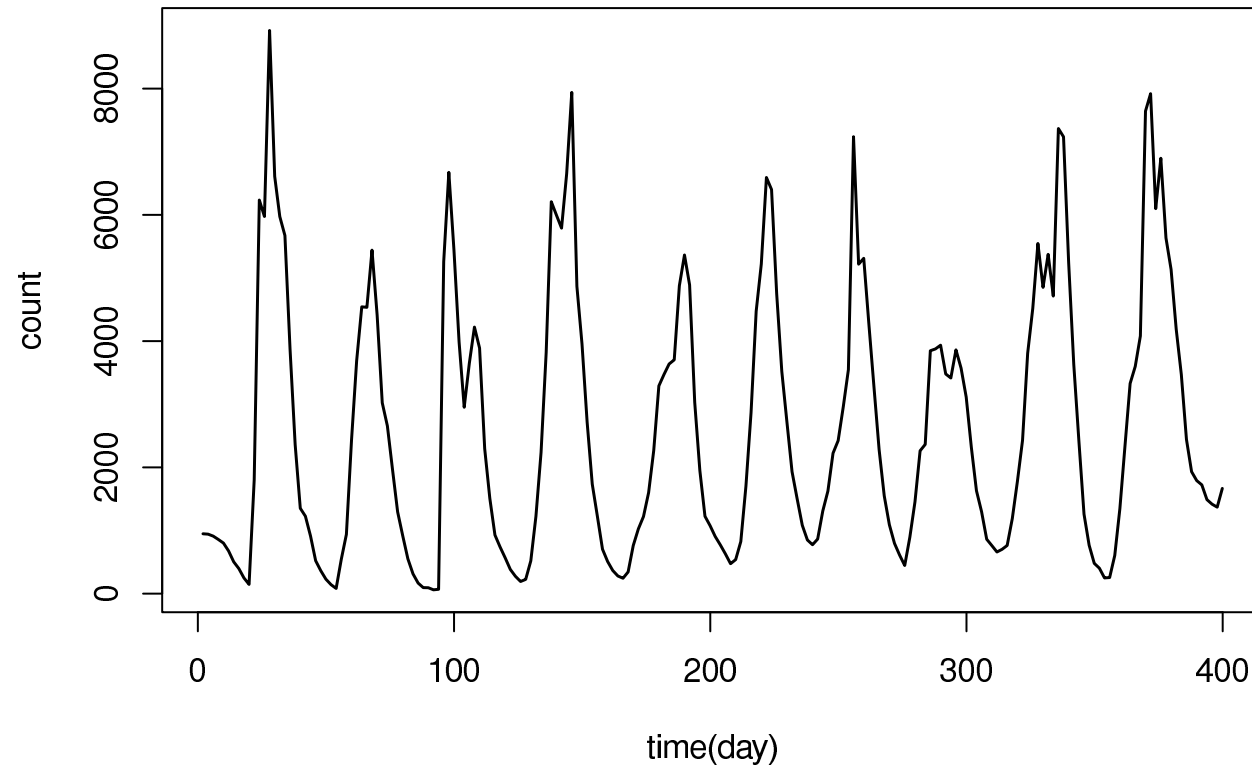
**Ed Ionides**

**Michigan Statistical Learning seminar, 26 April 2011**

- Nicholson and his blowflies.
- Likelihood and feature matching for dynamic systems.
- Two analyses of one of Nicholson's experiments.

- **A. J. Nicholson** (1895-1969) does not have a Wikipedia page. He was born in Ireland, studied in England, and moved to Australia. Primarily an entomologist, he became interested in the regulation of populations. Initial theoretical investigations (Nicholson and Bailey, 1935) led to extensive lab experiments with blowfly populations (*Lucilia cuprina*).
- One experiment: Adults in a cage were given excess carbohydrate but competed for limited protein required for egg production. Eggs were removed and larvae were provided with excess food. Every second day, newly hatched adults were counted and returned to the cage. At these times, dead adults were removed and counted. This gives an indirect count of live adults.
- Nicholson (1954, 1957) argued for the importance of density dependent factors based on a fairly qualitative analysis of his data.

## Cycles of adult blowfly population



- Generation time is  $\approx 16$  days.
- Cycles last  $\approx 40$  days. A double peak is often evident.

- Nicholson's data were modeled by May, Maynard Smith and Oster (who all have Wikipedia pages) and others in the 1970s.
- **Gurney et al** (*Nature*, 1980) found a good, simple model:

$$dN(t)/dt = P N(t - \tau) \exp\{-N(t - \tau)/N_0\} - \delta N(t)$$

$P$  is per-capita offspring without competition;  $N_0$  is the population size maximizing total offspring.

- **Wood** (*Nature*, 2010) wrote  $N(t) = R_t + S_t$ , the sum of recruits and survivors over the time interval  $[t, t + \Delta]$ , and set

$$R_t \sim \text{Poisson}\left[P N(t - \tau) \exp\left\{-N(t - \tau)/N_0\right\} \Delta e_t\right],$$

$$S_t \sim \text{Binomial}\left[N(t), \exp\left\{-\delta \Delta \varepsilon_t\right\}\right].$$

Here,  $e_t$  and  $\varepsilon_t$  are Gamma noise on the rates, with mean 1 and variances  $\sigma_p^2 \Delta^{-1}$  and  $\sigma_d^2 \Delta^{-1}$ .

## Nicholson's blowflies as a motivation for feature matching

- Wood (2010) used a synthetic likelihood made by matching features of the data (e.g., the sample covariance function) to features of simulations.

**Caution:** This work is motivated by a false claim that one cannot work with the actual likelihood.

- Xia and Tong (2011) used a model based on Gurney et al (1980) to argue that matching features can produce a better fit (to some specific feature of interest) than likelihood when the model is mis-specified.

**Caution:** Quantitative fitting of statistically mis-specified models is delicate. The model of Gurney et al (1980) is qualitative not quantitative in a statistical sense.

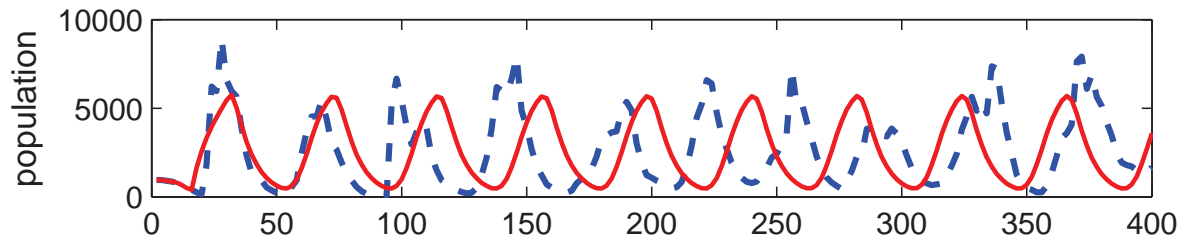
## Likelihood in one slide

- Data  $y_1, \dots, y_n$  are modeled via a density  $f(y_1, \dots, y_n | \theta)$  depending on an unknown parameter vector  $\theta$ .
- The **maximum likelihood estimator** (MLE) is the value  $\hat{\theta}$  maximizing  $f(y_1, \dots, y_n | \theta)$ .
- **Amazingly**,  $\hat{\theta}$  is (asymptotically) “optimal” in usual situations.
- **Even better**, values of the maximized log likelihood allow comparison between different models.
- For dynamic models,  $f$  may be hard to compute. Nevertheless, the model defines  $f$  implicitly, and Monte Carlo methods often allow one to work with it.

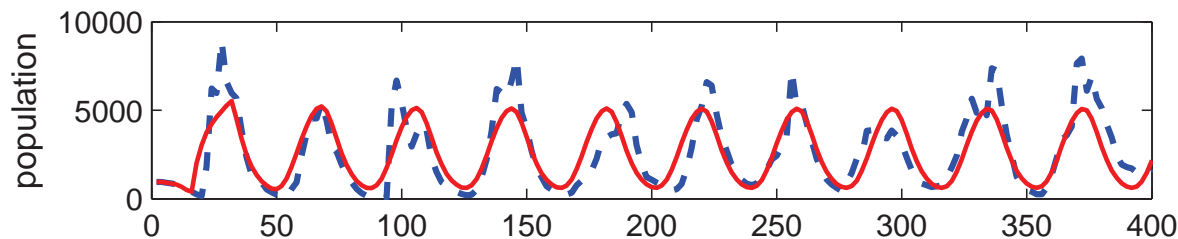
**Xia and Tong** fitted a discrete time model based on Gurney et al, using:

(i) least-square 1-step average prediction error, called APE(1). This is the MLE for additive Gaussian process noise with no measurement error.

(ii) least-square prediction errors from  $k$ -step skeleton forecasts for  $1 \leq k \leq T$ , called APE(T).



Skeleton for  
APE(1) estimate



Skeleton for  
APE(T) estimate

## Reanalysis (Ionides, 2011)

- the log likelihoods of the APE(1) and APE(T) estimates, under Xia and Tong's implicit model, are **-1568.5** and **-1569.5** respectively.
- a chi-square approximation with 5 d.f. says differences of  $\leq 5.5$  log units are statistically small (all in the same 95% confidence interval).
- simulations to check this asymptotic approximation are fairly irrelevant when the result is unambiguous.
- do either of these models fit well? A non-mechanistic log-ARMA(2,2) model had a maximized log likelihood of **-1524.3** estimating 6 parameters.
- a model based on Wood (2010) gave a maximized log likelihood of **-1465.4** which comfortably beats the log-ARMA(2,2) benchmark.



- The MLE “optimally” combines information at all time scales if the model is correct: this is not obvious from the usual calculation of the likelihood via 1-step forecasts.
- A mis-specified model can perhaps be fitted to match some features but not all features simultaneously. The  $APE(T)$  criterion is constructed to match long-term features. What does it miss?
- Emphasizing a fit to long-term features necessarily de-emphasizes the short-term.
- ‘Mechanism’ usually describes increments or derivatives of a system (short-term behavior). Perhaps the  $APE(T)$  parameter estimates match less well with their biological interpretation?

- $X(t) = cX^\alpha(t - \tau) \exp\{-X(t - \tau)/N_0\} + \nu X(t - 2)$

	APE(1)	APE(T)	interpretation
$c$	20.1	592	maximum offspring per-capita per 2 day
$\alpha$	0.85	0.26	exponent to stabilize time discretization
$\alpha N_0$	499	344	population maximizing total offspring
$2/(1 - \nu)$	8.33	5.67	mean life expectancy

- Mechanistically,  $c \approx 20$  and  $\alpha$  should be a little less than 1.
- The APE(1) estimate is more biologically interpretable than APE(T).

- The skeleton of the MLE for the Wood (2010) model matches the periodicity of the data, and the model fits well by likelihood criteria.
- Parameter estimates for this model are biologically interpretable.
- The fit is good, compared to a simple benchmark.
- Likelihood is a reasonable way to fit reasonable models. It is also a good criterion for comparing models and diagnosing mis-specification.

## Technical details of the model based on Wood (2010)

$N(t + \Delta) = R_t + S_t$  with

$$R_t \sim \text{Poisson}[P N(t - \tau) \exp \{ - N(t - \tau)/N_0 \} \Delta e_t],$$

$$S_t \sim \text{Binomial}[N(t), \exp \{ - \delta \Delta \varepsilon_t \}].$$

- $e_t$  and  $\varepsilon_t$  are Gamma with mean 1 and variances  $\sigma_p^2 \Delta^{-1}$  and  $\sigma_d^2 \Delta^{-1}$ .
- $\Delta = 1$  day is exactly the dynamic model of Wood (2010).
- If  $\Delta$  divides  $\tau$ , we get a discrete-time Markov process  

$$X(t) = (N(t), N(t - \Delta), N(t - 2\Delta), \dots, N(t - \tau)).$$
- To complete a POMP model, one needs to specify initial conditions and a measurement process.
- Estimates  $\hat{P} = 3.28$ ,  $\hat{N}_0 = 680$ ,  $\hat{\delta} = 0.161$ ,  $\hat{\sigma}_p = 1.4$ ,  $\hat{\sigma}_d = 0.75$  for  $\Delta = 1$  day. Preferred to  $\Delta = 2$  via logliks of **-1465.4** versus **-1471.4**.

## Initial values

- Write Nicholson's recorded data as  $y_1, \dots, y_T$  where  $y_k$  gives the adult blowflies counted at time  $t_k = 2k$  day, and  $T = 200$ .
- For comparison with Xia and Tong, I fixed  $\tau = 14$  day and required that the model should provide a likelihood for  $y_9, y_{10}, \dots, y_T$ .
- **Caution:** to compare likelihoods, the models should correspond to densities for exactly the same data, measured on the same scale.
- The initial state  $X(t_8)$  can be constructed using  $y_1, \dots, y_8$ .
- With  $\Delta = 2$  day, I set  $N(t_k) = y_k$  for  $k \in \{1, \dots, 8\}$  rather than treating the initial conditions as unknown parameters.
- For general  $\Delta$ , I specified  $X(t_8)$  using a cubic spline interpolation of  $y_1, \dots, y_8$ .

**The measurement model:**  $y_k \sim \text{Negbinom}(N(t_k), \sigma_y^{-2})$ .

- A negative binomial distribution conditional on  $N(t_k)$  with mean  $N(t_k)$  and variance  $N(t_k) + [\sigma_y N(t_k)]^2$ .
- Nicholson's blowfly counts were known to have some error (Brillinger et al., 1980).
- Here,  $\hat{\sigma}_y = 0.027$  is small, but to justify the logarithmic scaling (environmental noise) we need a more subtle interpretation:
  - ◇ The fertility of adults varies according to their age and potentially for other unmodeled biological reasons.
  - ◇  $N(t)$  describes fertility in units of ideal, standardized adults rather than simply measuring the actual number of adults.
  - ◇ Fluctuation in the calibration between adults present and their reproductive potential is described by measurement error

## Evaluation and maximization of the likelihood

- Can be done using sequential Monte Carlo (Ionides et al., 2006; King et al., 2008; Bretó et al., 2009; He et al., 2010; Laneri et al., 2010; Bhadra et al., 2011; Ionides et al., 2011).
- The synthetic likelihood of Wood (2010) shares some advantages of likelihood, except:
  - ◇ It cannot be compared with log-ARMA likelihoods, etc.
  - ◇ For complex data and models, it is hard to find low-dimensional functions of the data which capture all key features of the data.
  - ◇ There may be subjectivity in choosing features to match.
  - ◇ For non-stationary models, synthetic likelihood may be computationally inefficient.

## References

- Bhadra, A., Ionides, E. L., Laneri, K., Pascual, M., Bouma, M., and Dhiman, R. C. (2011). Malaria in Northwest India: Data analysis via partially observed stochastic differential equation models driven by Lévy noise. *JASA*, to appear.
- Bretó, C., He, D., Ionides, E. L., and King, A. A. (2009). Time series analysis via mechanistic models. *Annals of Applied Statistics*, 3:319–348.
- Brillinger, D. R., Guckenheimer, J., Guttorp, P., and Oster, G. (1980). Empirical modelling of population time series: The case of age and density dependent rates. In Oster, G., editor, *Some Questions in Mathematical Biology*, pages 65–90. American Mathematical Society, Providence.
- Gurney, W. S. C., Blythe, S. P., and Nisbet, R. M. (1980). Nicholson's blowflies revisited. *Nature*, 287(5777):17–21.
- He, D., Ionides, E. L., and King, A. A. (2010). Plug-and-play inference for disease dynamics: Measles in large and small towns as a case study. *Journal of the Royal Society Interface*, 7:271–283.



- Ionides, E. L. (2011). Discussion on “Feature matching in time series modeling” by Y. Xia and H. Tong. *Statistical Science*, To appear.
- Ionides, E. L., Bhadra, A., Atchadé, Y., and King, A. A. (2011). Iterated filtering. *Annals of Statistics*, Pre-published online.
- Ionides, E. L., Bretó, C., and King, A. A. (2006). Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences of the USA*, 103:18438–18443.
- King, A. A., Ionides, E. L., Pascual, M., and Bouma, M. J. (2008). Inapparent infections and cholera dynamics. *Nature*, 454:877–880.
- Laneri, K., Bhadra, A., Ionides, E. L., Bouma, M., Yadav, R., Dhiman, R., and Pascual, M. (2010). Forcing versus feedback: Epidemic malaria and monsoon rains in NW India. *PLoS Computational Biology*, 6:e1000898.
- Nicholson, A. J. (1954). An outline of the dynamics of animal populations. *Australian Journal of Zoology*, 2:9–65.
- Nicholson, A. J. (1957). The self-adjustment of populations to change. *Cold Spring Harbor Symposia on Quantitative Biology*, 22:153–173.

- Nicholson, A. J. and Bailey, V. A. (1935). The balance of animal populations. Part I. *Proceedings of the Zoological Society of London*, 105(3):551–598.
- Wood, S. N. (2010). Statistical inference for noisy nonlinear ecological dynamic systems. *Nature*, 466:1102–1104.
- Xia, Y. and Tong, H. (2011). Feature matching in time series modelling. *Statistical Science*, To appear.