Estimating Frog Growth Curves when both Age and Sex are Unknown

Shirley Pledger and Ben Bell
School of Mathematics and Statistics
School of Biological Sciences
Victoria University of Wellington, N.Z.

Biometrics AR and SEEM
Bay of Islands

28 November 2023

Introducing ...

Ben Bell

Ecologist

Leiopelma hamiltoni

Hamilton's Frog (Maud Island)

The Studies

Long-term capture-recapture studies.

Maud Island - *Leiopelma hamiltoni* since 1976.

> Originally two 12m by 12m grids, then a third set up in 1984-5 by translocation of 100 frogs.

Coromandel - *Leiopelma archeyi* since 1982.

Annually on Maud Is. each grid visited for a few successive nights, but at Coromandel 10m x 10m grid day searches of rocks and logs during a single day.

The Frogs

Ground-dwelling, live under and between rocks.

Nocturnal, come to the surface on humid nights.

Ambush feeders, insectivores.

Catch food by mouth, not tongue.

Unusually long-lived for frogs, can reach 40+ years.

Females larger than males.

Reproduction

Males select a breeding site.

Attract females (chemically, not vocally).

Males clasp females around the groin (pelvic amplexus), which stimulates them to lay their eggs (about 10-15 for Hamilton's frog).

Males fertilise eggs externally.

Childcare

Males guard the eggs.

Babies hatch as froglets, not tadpoles.

They climb onto the father's back.

He keeps them there at the breeding site until they metamorphose into frogs and start living independently. This is in terrestrial species (*L. archeyi, L. hamiltoni*) - in *L. hochstetteri* there is less intensive parental care.

This happens around February.

Meanwhile, where are the females?

Off feeding, building up body reserves before breeding again.

It is known from captive populations that females can only breed every second year.

Males can breed every year.

Details of the study

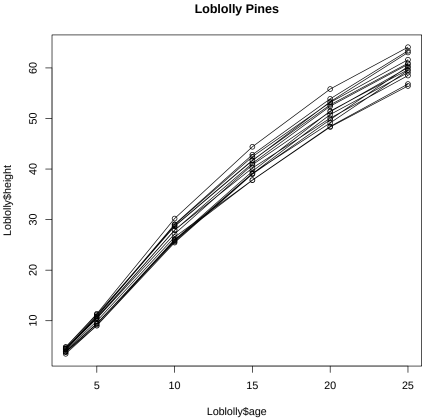
Frogs were marked at first capture, so if recaptured its identity is known.

At each capture the frog had its snout-vent length (SVL) and weight recorded.

Our objective: to fit two growth curves, female and male.

Why? General scientific interest, running and evaluating translocations, comparing populations, etc.

Growth Curves (i) A textbook example



Can find expected curve of height versus age, components of variance, individual variation (fast or slow growers) and leftover random variation.

Growth Curves (ii) The NZ frog world reality

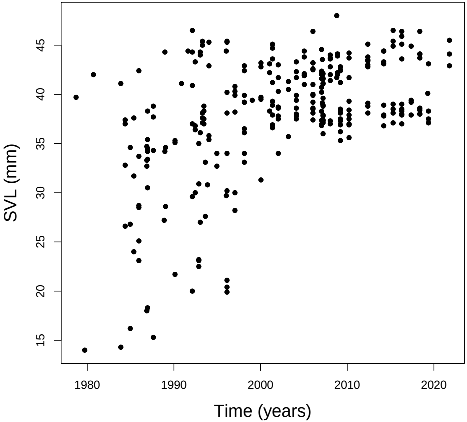
The Loblolly pines had:

- Repeated measures on the individuals
- Known age
- Regular synchronised measurement times
- Only one common curve.

For the frogs we have:

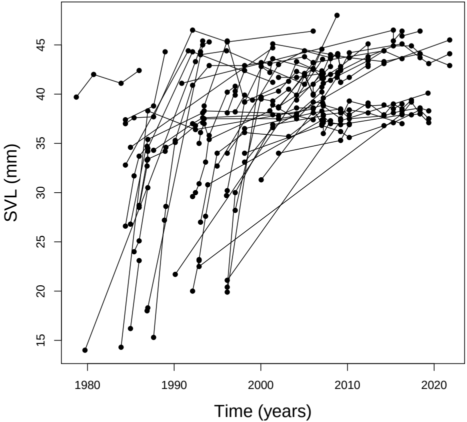
- Sparse, unsynchronised data of repeated measures on the individuals
- Unknown age - only the calendar dates of measurements
- Unknown sex - no visible sexual characteristics (although there is sexual dimorphism in adults, so two curves are needed).

Maud Island Grid 1 data, SVL by calendar time



Need to include repeated measure information.

Grid 1 SVL by time, repeated measures



Note fast growth if first caught young, which gives better information about date of “birth” (start of independent living).

Need age axis, align histories for zero age at DoB

Age: Assume "births" are in mid-February. Then can get fairly accurate estimate of DoB if caught in first three years.
Used six categories for age-group at first measurement:
0-1, 1-2, 2-3, 3-4, 4-5, 5+ years. High probability of correct allocation into first 3 categories (Ben's class A, B and C frogs).

Sex: Also need to allocate to Sex.
High probability of correct allocation if frog's SVL history includes steady high measurement (mature frogs, SVL hovering around asymptote).

Need two asymptotic growth curves, higher asymptote for females.

The von Bertalanffy curve (von B. 1960)

Using notation from R nlme package (Pinheiro and Bates 2000, repeated measures),

$$y = Asym - (Asym - R0)e^{-kx} \tag{1}$$

where y = expected length(SVL in mm), x = age (years), and the three parameters are $Asym$ = asymptote, $R0$ = response (expected SVL) at age 0 and k = instantaneous growth rate at age 0.

$$\frac{dy}{dx} = k(Asym - y) \tag{2}$$

so the rate of increase is proportional to the shortfall of y from $Asym$, implying that growth slows as the frog SVL nears the asymptote.

Features of the von Bertalanffy curve

Asym is the expected SVL being approached, so data points may fluctuate above or below the curve.

The variability of ultimate SVL may be partitioned into individual variation (e.g. a large frog with SVL consistently above the population average) and fluctuating unexplained residual error (e.g. measurement error).

Using a finite mixture model

We assume each frog belongs to one of 12 groups, with six ways of dealing with the missing AGFM (age group at first measurement), needed to estimate the DoB, crossed with two choices for which sex.

Each frog will be allocated to its most likely combination of AGFM and Sex.

Allocation to its AGFM enables us to estimate its DoB, while allocation to its Sex allows it to contribute to updating estimates of the curve parameters.

Using the GEM Algorithm

Generalized EM Algorithm, Dempster, Laird and Rubin 1977.

1. **Initialise:** Set up equations for two von B curves (six parameters) and twelve proportions π_1 to π_{12} for allocation of frogs to the 12 groups (only 11 independent pi's as they must add to 1, multinomial allocation).
2. **E step:** Use Bayes' theorem to calculate posterior probabilities for each frog to belong to each of the 12 groups. Assign it to the most likely group (i.e. use the most likely DoB to set the data on the Age axis, assign the frog to its most likely sex).
3. **M step:** Use the pseudo-data to refit the two von B curves, update the von B and π parameter estimates.

Cycle the E and M steps until the log likelihood and all the parameters have converged.

GEM: Exact allocation to most likely group at E step. If it converges it gives the correct parameter estimates for a local maximum (on a local optimum of the likelihood surface).

Visualising the EM algorithm at work

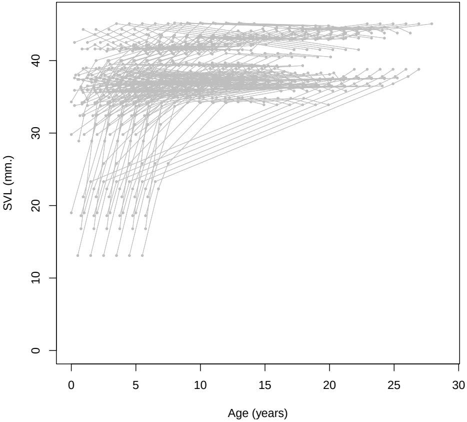
Suppose initial estimates of parameters for the two curves are assumed.

The data from each frog is a chain of SVL measurements with known time intervals between measurements. Assume there are six possible Age Group at First Measurement (AGFM) values, 0-1, 1-2, ...5+.

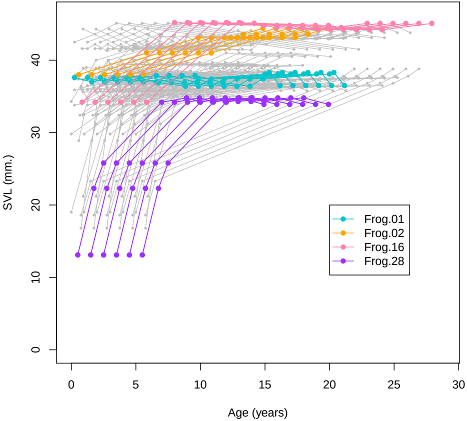
The next slide shows 26 of the Grid 1 frogs, transferred to an Age scale on the x axis, each with 6 possible placements.

Each frog has six pseudo-data chains showing these possible positions on the plot.

Six possible SVL chains for each of 26 frogs

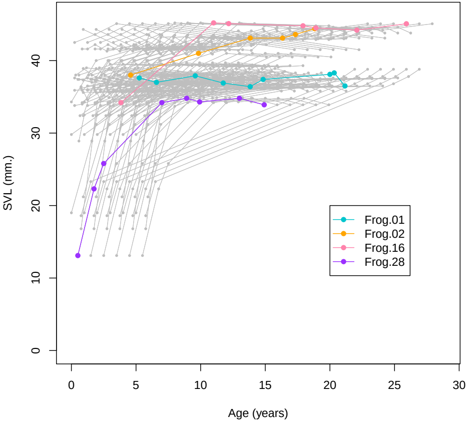


Pick out four frogs to follow

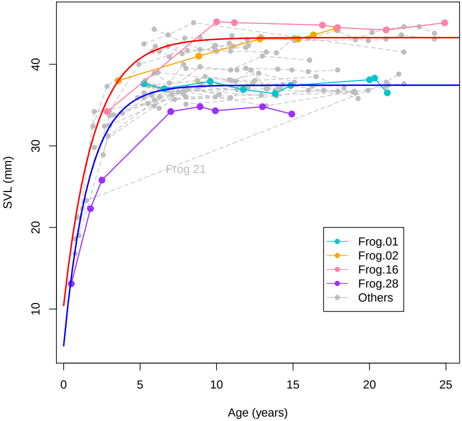


Using knowledge so far of the two curves, assign the four to their most likely age groups at first measurement (AGFMs) and Sex. Done in the E step.

Plot only the chosen pseudo-data for the chosen four

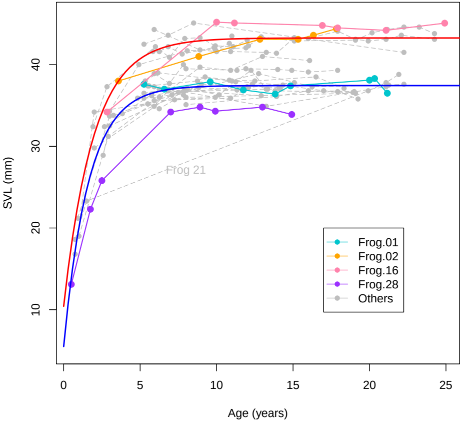


Do the same for the other frogs, re-fit the curves M step

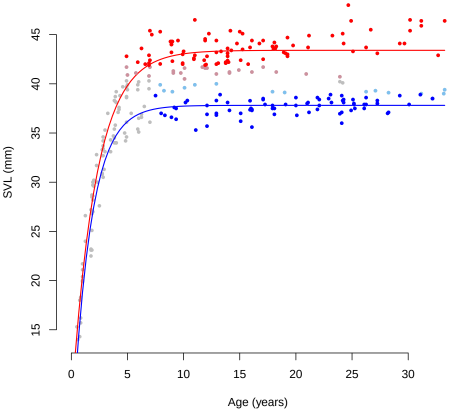


Continue to re-allocate to pseudo-data, update fitted curves, until converged.

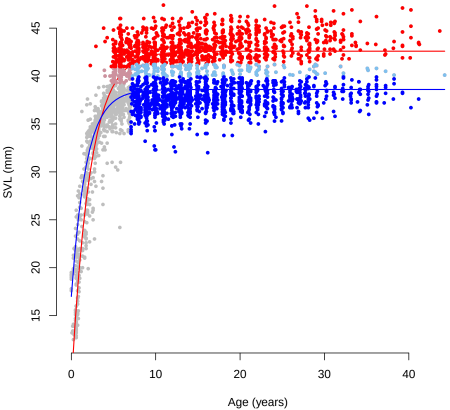
Results: Pseudo-data and final curves



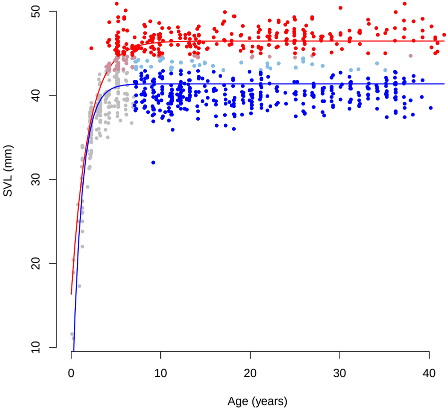
Results: Grid 1, Maud Island



Results: Grid 2, Maud Island



Results: Grid 3, Boat Bay, Maud Island



Results show these translocated frogs have grown larger than those left behind.

References

Bell, BD and Pledger SA (2023). Post-metamorphic body growth and remarkable longevity in Archey's frog and Hamilton's frog in New Zealand. New Zealand Journal of Ecology 47(2): online.

Dempster AP, Laird NM and Rubin DB (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society B 39: 1-38.

Pinheiro JC and Bates DM (2000). Mixed effects models in S ans S-PLUS. New York, Springer-Verlag, 544 pp.

von Bertalanffy L (1960). Principles and Theory of Growth. In: Nowinski WW ed. Fundamental aspects of normal and malignant growth. Amsterdam, Elsevier. Pp 137-259.