Hare/lynx dynamics – combining ecological modeling with data using particle filter methods.

A study in new methodology that enables continuous time statistical inference.
100 years of catchment data for the Hudson Bay company. The lynx data has some holes. Looks like there is some semi-regular cycles there with cycle time ≈ 10 years. Widely analyzed, but usually using standard time series analysis tools. Idea: Why not adapt the data to processes from theoretical ecology? Problem: How do you do that?
Differential equations have been used for modeling abundance dynamics since Malthus: \( \frac{dy(t)}{dt} = r y(t) \)

A second order term (density dependency) can give the system a carrying capacity (rather than shooting off into infinity):

\[ \frac{dy(t)}{dt} = r(y(t) - y(t)^2 / K) \]

Lotka-Volterra developed a formula for predator-prey relationships (H=hare, L=lynx):

\[ \frac{dH(t)}{dt} = a_H H(t) - c H(t) L(t) \]
\[ \frac{dL(t)}{dt} = -a_L L(t) + k c H(t) L(t) \]
Holling expanded the Lotka-Volterra system by limiting the impact of hare abundance on individual lynx multiplication rates to below a fixed level:
\[
dH(t)/dt = a_H H(t) - \frac{c H(t)}{d + H(t)} L(t)
\]
\[
dL(t)/dt = -a_L L(t) + \frac{k c H(t)}{d + H(t)} L(t)
\]
If lynx numbers should crash, the number of hares should still stay limited:
\[
dH(t)/dt = a_H H(t) - b_H H(t)^2 - \frac{c H(t)}{d + H(t)} L(t)
\]
\[
dL(t)/dt = -a_L L(t) + \frac{k c H(t)}{d + H(t)} L(t)
\]
Lynx might survive on a steady supply of other prey, when hares are scarce (with higher probability the less hares there are):
\[
dH(t)/dt = a_H H(t) - b_H H^2 - \frac{p(H) c H L}{d + H(t)}
\]
\[
dL(t)/dt = [p(H)(a_L L - b_L L^2) + (1 - p(H))(-a_L L + \frac{k c H}{d + H} L)]
\]
where \(p(H) = e^{-H/H_n}\)
Hares could be (are) in a predator-prey relationship to some plants:
\[
dR(t)/dt = a_P P(t) - b_P P(t)^2 - \frac{c_P P(t)}{d_P + P(t)} H(t)
\]
\[
dH(t)/dt = -a_H H(t) + \frac{k c_P P(t)}{d_P + P(t)} H(t) - \frac{c_H H(t)}{d_H + H(t)} L(t)
\]
\[
dL(t)/dt = -a_L L(t) + \frac{k c_H H(t)}{d_H + H(t)} L(t)
\]
Holling model with carrying capacity:

\[ \frac{dH(t)}{dt} = a_H H(t) - b_H H(t)^2 - \frac{cH(t)}{d+H(t)} L(t) \]

\[ \frac{dL(t)}{dt} = -a_L L(t) + \frac{k c H(t)}{d+H(t)} L(t) \]

While the terms have ecological meaning, it might be difficult to relate to the parameter values. (What is a big value for c?)

But we can find other ways to write the same system, by first finding the equilibrium state (where \( \frac{dH}{dt}=\frac{dL}{dt}=0 \)): \( L_0, H_0 \). Then:

\[ \frac{dh(t)}{dt} = \frac{\log \varphi}{t_H} (1 - r_0 e^{h(t)}) - \frac{t_L + t'_L}{t_L + t'_L e^{h(t)}} e^{l(t)} \]

\[ \frac{dl(t)}{dt} = \frac{\log \varphi (e^{h(t)} - 1)}{t_L + t'_L e^{h(t)}} \]

where \( h(t) = \log \left( \frac{H(t)}{H_0} \right) \), \( l(t) = \log \left( \frac{L(t)}{L_0} \right) \)

\( r_0 \) is the ratio between \( H_0 \) and the carrying capacity

\( t_H \) is the doubling time of hares in the absence of lynx

\( t_L \) is the half-life of lynx in the absence of hares

\( t'_L \) is the doubling time of lynx when hares are plenty
# Pros and cons of using differential equations for modeling ecology

## Pros:
- Allowing continuous time modeling.
- Terms can be ecologically meaningful.
- The properties of the dynamics can be used for saying something general about properties of ecological systems.

For instance, Kolmogorov showed that for some general requirements, two-component predator-prey systems either had stable fixed points or stable cycles. May argued that systems with more components ought to be less stable.

## Con:
- Nature just isn’t that neat. There’s all kind of relevant factors outside of our control which affects the systems => stochasticity.
- Abundances are count data.

PS: Stochastic systems can have properties that deterministic systems lack.

## Con of that con:
- Standard statistical time series models are discrete in time and linear (i.e. not on the form of our ecologically meaningful terms).
Stochastic processes – how they can differ from deterministic ones

- Even if you know everything up to now, the future state of the system will have an element of uncertainty.
- Finite probability that the system will crash. Means eventually it will happen (though it may take a ridiculous amount of time).
- Where you before got an equilibrium state, you now can have a stationary distribution.
- Can get cycles even when the deterministic version of the same system stabilizes.

The system might look like it is changing smoothly with time with a regular periodicity. None the less, the changes are driven by stochasticity. Remove those and the system looses all dynamics (red lines). Also, the periodicity is not 100% regular.
Stochastic models

- Classic statistical time series modeling (ARMA models, transfer functions).
  - Discrete in time.
  - Are not readily ecologically interpretable, like the differential equations from ecology.
  - Causal connections get unrealistic treatment (we do not expect that hare numbers react to lynx numbers from exactly one year ago, do we?)
  - It’s been done!

- Continuous stochastic time series – stochastic differential equations (SDEs).
  - Readily ecologically interpretable, since one can make them from the differential equations from ecology.
  - Causal connections are instantaneous (but takes time to build up).
  - Not many are doing inference using SDEs. (Though for good reasons, it turns out.)

- Birth-death models
  - Even more realistic.
  - More difficult to specify and might be more difficult to simulate from and do inference on. Less tradition for it.
  - The extra realism might be entirely unnecessary, as long as abundances > 1000.
Stochastic differential equations (SDEs)

- Looks a lot like ordinary differential equations, except they have an extra term that signifies small stochastic contributions:
  
  **DE:** \( dX(t) = g(X(t))dt \)
  
  **SDE:** \( dX(t) = g(X(t))dt + \sigma(X(t))dB(t) \)

- \( dB(t) \) can be seen as the changes in a random walk, \( B(t) \), (Wiener process) over an infinitesimal time.

- As such: \( dX(t) = \sigma dB(t) \) has the simple solution \( X(t) = X(0) + \sigma B(t) \)

- Ornstein-Uhlenbeck:
  \[
  dX(t) = -\frac{1}{\Delta t}(X(t) - \mu)dt + \sigma \sqrt{\frac{2}{\Delta t}} dB(t)
  \]

- Vectorial linear SDEs (may have several hidden components):
  \[
  dX(t) = -A(X(t) - \mu(t))dt + \Sigma(t)dB(t)
  \]
Questions:

- Can we estimate the parameters from hare/lynx catchment data?
- Will these estimates suggest deterministic cyclic behavior or stochastically upheld such?
- To what degree will we be able to detect cyclic behavior at all?
- Can we find better models, like independent lynx survival or plants as a third component?
- Can we plug climate (NAO) in as a regression coefficient in one of the parameters? I.e. have climatic variations mattered?
- How feasible is it to answer any of these questions when he have such nonlinear SDEs as the one above?

Stochastic Holling model
(re-parametrized)

\[
dk(t) = \left( \frac{\log(\mathcal{L})}{t_H} \right) \left( 1 - r_0 e^{h(t)} + \frac{t_L + t'_L}{t_L + t'_L e^{h(t)}} e^{l(t)} \right) dt + \sigma_H dB_H(t)
\]

\[
dk(t) = \left( \frac{\log(\mathcal{L})}{t_L + t'_L e^{h(t)}} \right) dt + \sigma_L dB_L(t)
\]
In addition to the stochasticity inherent in the ecological process (the actual abundance), there's also observational noise. The statistical model thus falls into the category of hidden time series modeling. Such models have two components:

1. $X_k$: State of the process (abundance) at time of observation $k$, $t_k$. Updates according to a system equation (SE) $f(X_{k+1}|X_k)$. This is sufficient to specify a Markov chain type of process (per def).

2. $Z_k$: Observation number $k$. Related to the state by an observational equation (OE): $f(Z_k|X_k)$.

Horizontal arrows represent the system equation in action, while vertical arrows represent the observational equation.
Procedure: Go stepwise through the observations from 1 to n.

For observation k, assume you have $f(X_k | Z_k, \ldots, Z_1)$.

1. The law of total probability (LOTP) on that and SE gives you $f(X_{k+1} | Z_k, \ldots, Z_1)$.
2. LOTP on 1 and OE (observational equation), gives you $f(Z_{k+1} | Z_k, \ldots, Z_1)$.
3. Use Bayes' formula to calculate $f(X_{k+1} | Z_{k+1}, \ldots, Z_1)$. (You will have to look at what you got at step 1 and 2 plus OE for the previous time point to do this).
4. Go back to step 1, now for $X_{k+2}$.

Likelihood: $L = f(Z_1, \ldots, Z_n | \theta) = f(Z_1) f(Z_2 | Z_1) \ldots f(Z_{k+1} | Z_k) \ldots f(Z_n | Z_{n-1})$.

If all densities are normal (as in linear SDEs), all this work can be done analytically and this is called the Kalman filter.
If we can’t do all these steps analytically, we can have a set of “particles” that represents the distribution of the process state at any given observation...

1. Assume the particles, $X_k^{(1)}, \ldots, X_k^{(J)}$ represent a sample from $f(X_k|Z_k, \ldots, Z_1)$.
2. Make a proposal distribution for how to evolve each particle into time $k+1$, $g(X_{k+1}|X_k)$. Now you have $X_{k+1}^{(1)}, \ldots, X_{k+1}^{(J)}$
3. Calculate the weight of each particle, $j$:
   $$w_j = \frac{f(X_k^{(j)} | X_k^{(j)}) f(Z_{k+1} | X_{k+1}^{(j)})}{g(X_k^{(j)} | X_k^{(j)}, Z_{k+1}^{(j)})}$$
4. Estimate the likelihood contribution:
   $$f(Z_{k+1} | Z_k, \ldots, Z_1) \approx \frac{1}{J} \sum_{j=1}^{J} w_j$$
5. Normalize the weights so they sum to one.
6. Resample all particles at time $k+1$ according to $w_j$. You now have a sample of $f(X_{k+1}^1|Z_{k+1}^1, \ldots, Z_1)$!
7. If you want to do process inference, keep track of which particles at time $k$ gave rise to which resampled particles at time $k+1$.
8. Go back to 1.

(A pilot study showed this to be very efficient when the proposal distribution was chosen wisely.)
Particle filtering when you have no system update equation, \( f(X_{k+1} | X_k) \)

A non-linear SDE only give the probability of infinitesimal updates. We do not have system equation \( f(X_{k+1} | X_k) \)!
(Actually, even the ordinary differential equation (ODE) itself is not analytically solvable.)

Just as for ODE, there are however numerical ways to simulate ourselves forward (stochastic versions of Runge-Kutta). It’s even so possible to simulate from the distribution we’re after.

Keep in mind that the weight is \( w_j = \frac{f(X^{(j)}_{k+1} | X^{(j)}_k) f(Z_{k+1} | X^{(j)}_{k+1})}{g(X^{(j)}_{k+1} | X^{(j)}_k, Z_{k+1})} \). This means that if we use \( g(X_{k+1} | X_k, Z_{k+1}) = f(X_{k+1} | X_k) \) as out proposal distribution (by simulation), it cancels with the unknown \( f(X_{k+1} | X_k) \) above.

Procedure: For each time point \( k \), simulate particles forward to time \( k+1 \), calculate mean observational probability and then re-sample the particles according to each observational probability.

In summary, we can estimate a likelihood just from the distribution of the observational noise plus the ability to simulate from the underlying system! (But it’s no necessarily all that efficient.)
Particle filters give a stochastic estimate of the likelihood:

- “Importance sampling” used for calculating Bayesian Model Likelihood (BML, probability of data given model). Used for model comparison.
- Used the Holling model as the zero-hypothesis for the system equation. Try to expand from there.

- Assumed expected catchments were exactly one hundredth of the mean abundance over the year (identification restriction).

- Used negative binomial distribution for the observation equation, with an expectancy from 2 and an unknown form parameter, r. (low r means high observation noise. High r means Poisson distribution.)

- Simulated 1000 years of data to test if the method could estimate the model parameters. The results were good.
Results

- Inference takes a looooon time! (1 month)
- Process inference looks good!
- Independent lynx survival better than basic Holling model.
- 3 component plant-hare-lynx model didn’t work out.
- Parameters are however very uncertain.
- Cyclic behavior remains largely unexplained. The stochasticity washes out the cyclicity inherent in the deterministic system.
- Residual trends suggest a better model might be found.
- Since BMLs are stochastically estimated, NAO regression on each parameter turned out to be difficult, as the BML difference in models was quite low. ANOVA suggested there were differences though, and NAO dependency in the hares doubling time, $t_H$, was the best model.
Used linear SDEs, with the possibility of layered hidden processes affecting the two that produce the observations. Went through a lot of such models.

Results: One hidden layer for both hares and lynx. Best model that was easily interpretable looked like this (causal diagram).

Interpretation: Abundances found in the lowest layer, which has the right characteristics of a predator-prey relationship. Processes on top react quickly to changes in abundances. Hunter activity?

Cyclicity caught quite well! Estimated cycle time, 10 years!
Linear models caught the semi-periodicity rather well, while ecological modeling (so far) didn’t. The reason might have been that these models so far lacked the extra layer representing how human hunting activity is affected by the abundances. The extra stochasticity in these layers would be projected into the only layer of the ecological models instead, and thus the cyclicity in the abundances are washed out.

Human hunting activity can also be affected by economic consideration. Perhaps pelt prices should be examined? Keep in mind, catchments ≠ abundance.

The best ecologically motivated model described the marginal distribution quite a bit better than the linear model, though.

Searching for the right model tend to be an iterative approach, but re-programming and 1 month (or more) execution time does not encourage this.

When looking at multiple models (like the NAO study), you really want the BMLs to differ a lot, or you end up having to redo the analysis multiple times.

All in all, the PMCMC methodology is rather taxing on computer resources (and programming resources).

Histogram of log-lynx catchments.

Blue – linear model
Red – ecological model (from simulations)
References