

Neural Dynamic N-mixture Model

A deep learning framework for inferring demographic rates from count data

Speaker:

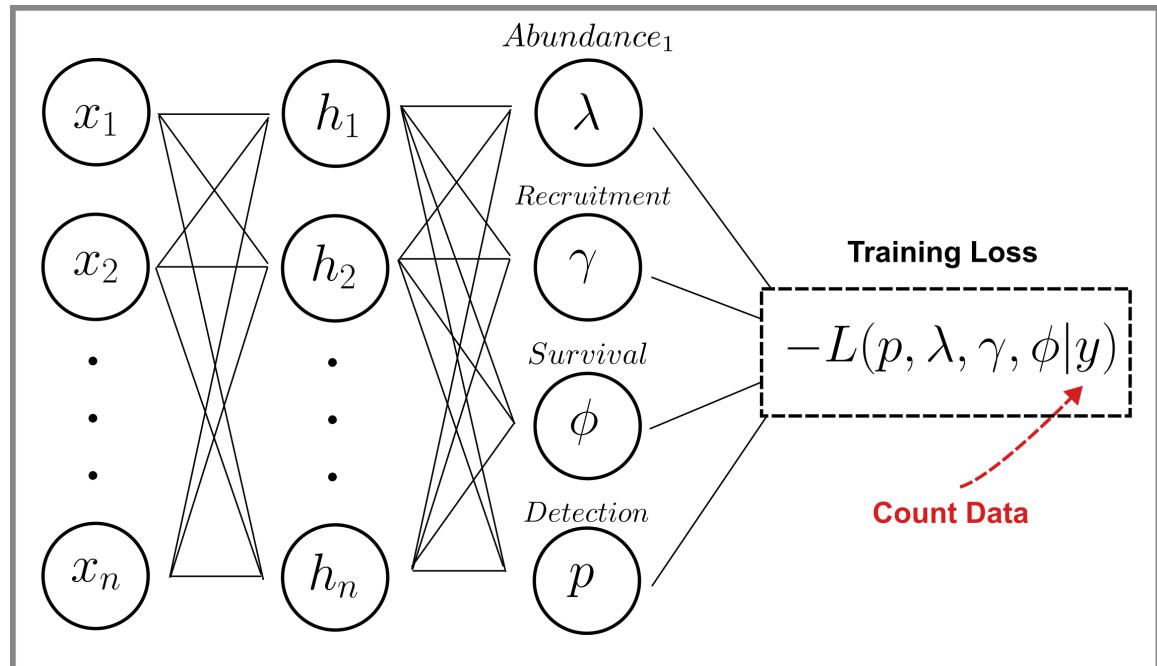
François Leroy (he/him)

Supervised by:

Marta Jarzyna



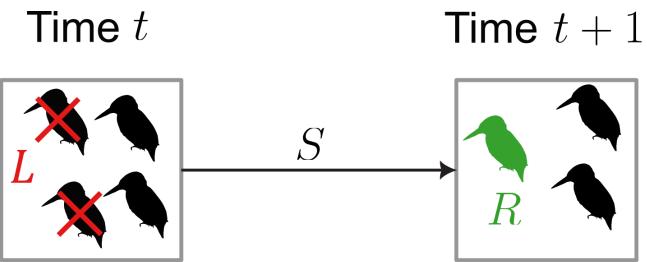
The Ohio State University
2025-10-06



Beyond abundance change

Demographic rates

- Changes in abundance are valuable, but demographic mechanisms offer deeper ecological insights
- Species' extinctions event are the result of increasing difference between recruitment and loss
- Changes in demographic rates can be early warnings of species extinction
- Monitoring survival and recruitment can help tailoring efficient conservation policies



$$N_t = L + S$$

$$N_{t+1} = S + R$$

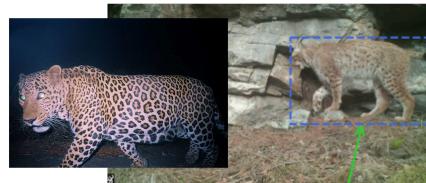
$$\Delta N = N_{t+1} - N_t = R - L$$

Beyond abundance change

Problem

- Data for demographic rates models are per individual (i.e. *individual encounter history data*)
- They are costly in resources and time
- Individual identification is not equal for all taxa
- Can be invasive/traumatic
- Limited spatial and temporal extent

Individual encounter history data



Less invasive
Lower effort
Broadly applicable
Resource efficient



More invasive
Higher effort
Taxon specific
Resource intensive

Beyond abundance change

Using abundance to infer demographic rates

- However, abundance data are available at large spatial and temporal scale
- The idea of inferring birth/immigration (μ) and death/emigration (λ) from experimental data is not new
- While overall change in abundance gives information about $\lambda - \mu$, the volatility of the time-series provide information about $\lambda + \mu$ making λ and μ identifiable

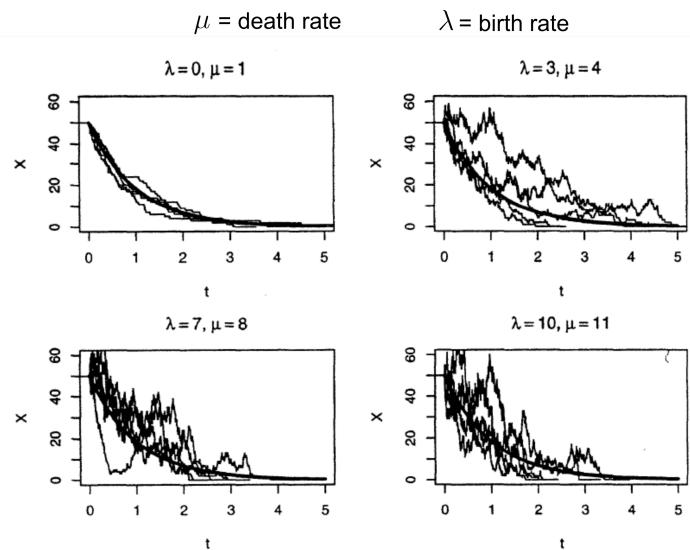


Figure 1.3 Five realisations of a stochastic linear birth-death process together with the continuous deterministic solution for four different (λ, μ) combinations, each with $\lambda - \mu = -1$ and $x_0 = 50$

Dynamic N-mixture model

Hidden Markov Model

Observation process

$$y_{i,j,t} \sim \text{Binomial}(N_{i,t}, p)$$

y = observed abundance

N_{it} = latent "true" abundance

p = detection probability

State process

$$N_{i,1} \sim \text{Poisson}(\lambda)$$

λ = abundance at $t = 1$

$$S_{i,t+1} \sim \text{Binomial}(N_{i,t}, \phi_{i,t})$$

ϕ = survival probability

$$R_{i,t+1} \sim \text{Poisson}(\gamma_{i,t})$$

γ = number of recruits

$$N_{i,t+1} = S_{i,t+1} + R_{i,t+1}$$

Limitations of Existing Frameworks

Limitations of hierarchical framework

- Linear or simple polynomial effects of covariates, even though true ecological responses are often unknown *a priori*
- MCMC algorithm scales poorly because inherently sequential, little possibility of parallelization
- Sensitive to priors and initial values

Limitations of Neural Networks

- Usually lack inferential power, making it less relevant for ecological insights
- Doesn't distinguish between ecological process and imperfect detection



METHODS

Neural hierarchical models of ecological populations

Maxwell B. Joseph*

Abstract

Neural networks are increasingly being used in science to infer hidden dynamics of natural systems from noisy observations, a task typically handled by hierarchical models in ecology. This

Neural hierarchical model

ECOLOGY LETTERS

Ecology Letters, (2020)

doi: 10.1111/ele.13462

METHODS

Neural hierarchical models of ecological populations

Abstract

Maxwell B. Joseph* 

Neural networks are increasingly being used in science to infer hidden dynamics of natural systems from noisy observations, a task typically handled by hierarchical models in ecology. This

- Combines flexibility and scalability of neural networks with the inferential power of hierarchical models
- Combines decades of development in hierarchical modelling for ecological data with the flexibility and predictive power of Neural Networks
- Loss function: tailored from the model's specific Likelihood (i.e. negative log-likelihood)
- Output activation function according to the parameter to infer

Neural Dynamic N-mixture model

Hidden Markov Model

State process

$$N_{i,1} \sim \text{Poisson}(\lambda)$$

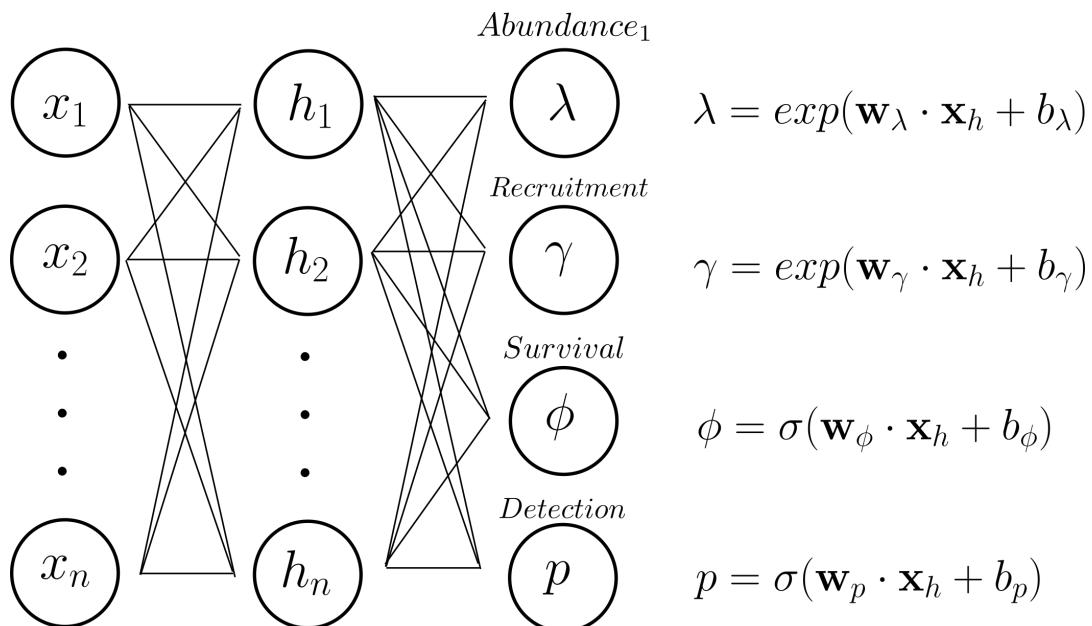
$$R_{i,t+1} \sim \text{Poisson}(\gamma_{i,t})$$

$$S_{i,t+1} \sim \text{Binomial}(N_{i,t}, \phi_{i,t})$$

$$N_{i,t+1} = S_{i,t+1} + R_{i,t+1}$$

Observation process

$$y_{i,j,t} \sim \text{Binomial}(N_{i,t}, p)$$



Neural Dynamic N-mixture model

Hidden Markov Model

State process

$$N_{i,1} \sim \text{Poisson}(\lambda)$$

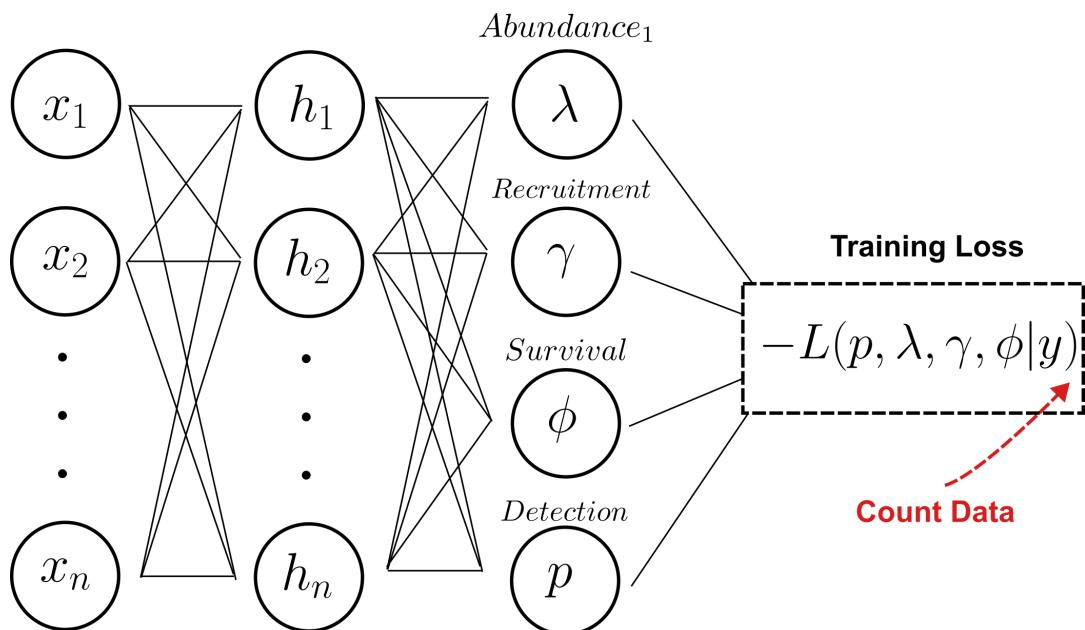
$$R_{i,t+1} \sim \text{Poisson}(\gamma_{i,t})$$

$$S_{i,t+1} \sim \text{Binomial}(N_{i,t}, \phi_{i,t})$$

$$N_{i,t+1} = S_{i,t+1} + R_{i,t+1}$$

Observation process

$$y_{i,j,t} \sim \text{Binomial}(N_{i,t}, p)$$



Example with the N-mixture model

Observation process

$$y_{i,j,t} \sim \text{Binomial}(N_{i,t}, p)$$

State process

$$N_i \sim \text{Poisson}(\lambda)$$

Integrated likelihood:

$$L(p, \lambda \mid y_{it}) = \prod_{i=1}^R \left(\sum_{N_i=0}^{\infty} \left(\prod_{t=1}^T \text{Bin}(y_{it}; N_i, p) \right) \text{Pois}(N_i; \lambda) \right)$$

Example with the N-mixture model

Observation process

$$y_{i,j,t} \sim \text{Binomial}(N_{i,t}, p)$$

State process

$$N_i \sim \text{Poisson}(\lambda)$$

Integrated likelihood:

$$L(p, \lambda \mid y_{it}) = \prod_{i=1}^R \left(\sum_{N_i=0}^K \left(\prod_{t=1}^T \text{Bin}(y_{it}; N_i, p) \right) \text{Pois}(N_i; \lambda) \right)$$

- Marginalize over $K >> y_{it}$

Dynamic N-mixture model

State process

$$\begin{aligned}N_{i,1} &\sim \text{Poisson}(\lambda) \\R_{i,t+1} &\sim \text{Poisson}(\gamma_{i,t}) \\S_{i,t+1} &\sim \text{Binomial}(N_{i,t}, \phi_{i,t}) \\N_{i,t+1} &= S_{i,t+1} + R_{i,t+1}\end{aligned}$$

Observation process

$$y_{i,j,t} \sim \text{Binomial}(N_{i,t}, p)$$

Integrated likelihood

$$\begin{aligned}\mathcal{L}(p, \lambda, \gamma, \phi \mid y_{it}) &= \prod_{i=1}^R \left[\sum_{N_{i1}=0}^{\infty} \cdots \sum_{N_{iT}=0}^{\infty} \left\{ \left(\prod_{t=1}^T \text{Bin}(y_{it}; N_{it}, p) \right) \right. \right. \\&\quad \times \text{Pois}(N_{i1}; \lambda) \cdot \prod_{t=2}^T P_{N_{it}, N_{it+1}} \Big\} \Big] \end{aligned}$$

Dynamic N-mixture model

State process

$$\begin{aligned}N_{i,1} &\sim \text{Poisson}(\lambda) \\R_{i,t+1} &\sim \text{Poisson}(\gamma_{i,t}) \\S_{i,t+1} &\sim \text{Binomial}(N_{i,t}, \phi_{i,t}) \\N_{i,t+1} &= S_{i,t+1} + R_{i,t+1}\end{aligned}$$

Observation process

$$y_{i,j,t} \sim \text{Binomial}(N_{i,t}, p)$$

Integrated likelihood

$$\begin{aligned}\mathcal{L}(p, \lambda, \gamma, \phi \mid y_{it}) &= \prod_{i=1}^R \left[\sum_{N_{i1}=0}^K \cdots \sum_{N_{iT}=0}^K \left\{ \left(\prod_{t=1}^T \text{Bin}(y_{it}; N_{it}, p) \right) \right. \right. \\&\quad \times \text{Pois}(N_{i1}; \lambda) \cdot \prod_{t=2}^T P_{N_{it}, N_{it+1}} \Big\} \Big] \end{aligned}$$

Transition matrix

- Key part of Hidden Markov Models
- Involves the transition between abundance at time t and $t + 1$ and the demographic parameters
- Each element of the transition matrix P represents the transition probability from state $N_{it} = j$ to state $N_{it+1} = k$ with the discrete convolution:

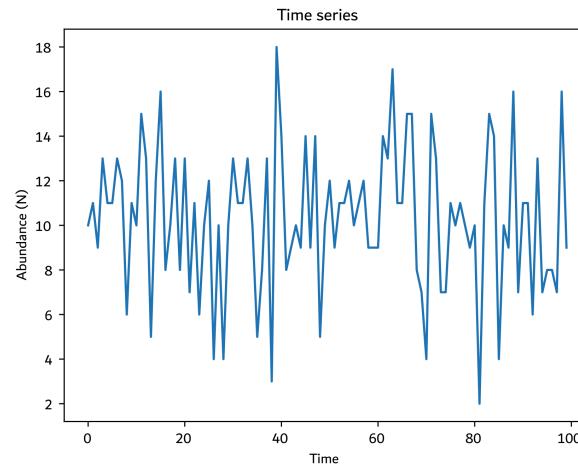
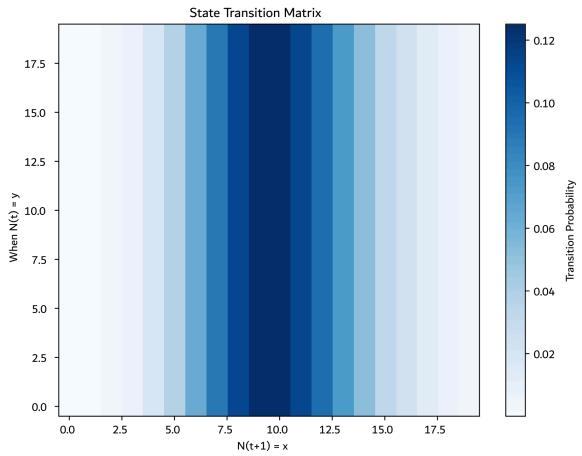
$$P_{jk} = \sum_{c=0}^{\min(j,k)} \text{Bin}(c; j, \phi) \cdot \text{Pois}(k - c; \gamma)$$

$$N_{i,t} \begin{bmatrix} P_{K0} & P_{K1} & P_{K2} & \cdots & P_{KK} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{20} & P_{21} & P_{22} & \cdots & P_{2K} \\ P_{10} & P_{11} & P_{12} & \cdots & P_{1K} \\ P_{00} & P_{01} & P_{02} & \cdots & P_{0K} \end{bmatrix}$$

Transition matrix

A

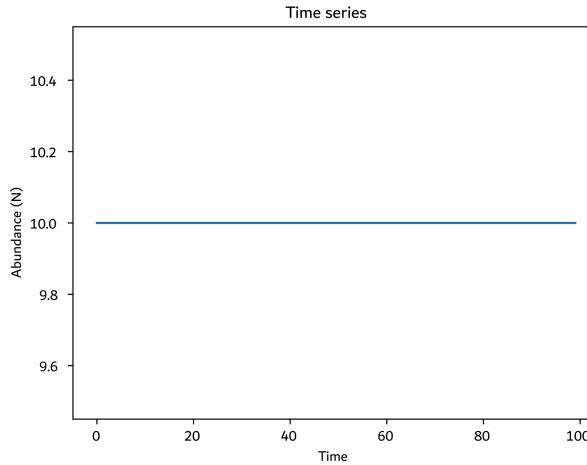
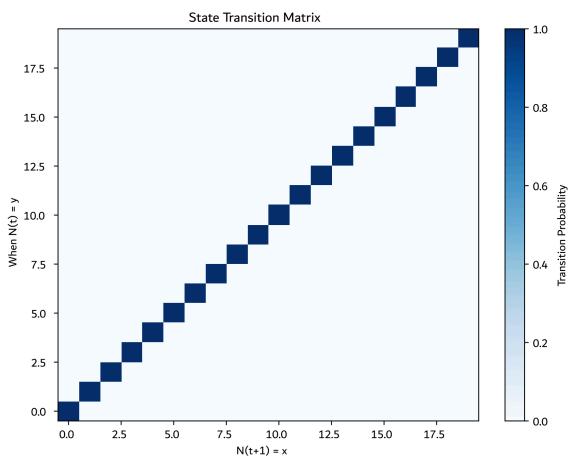
Survival probability = 0 / Recruitment = 10



When $N(t) = y$

B

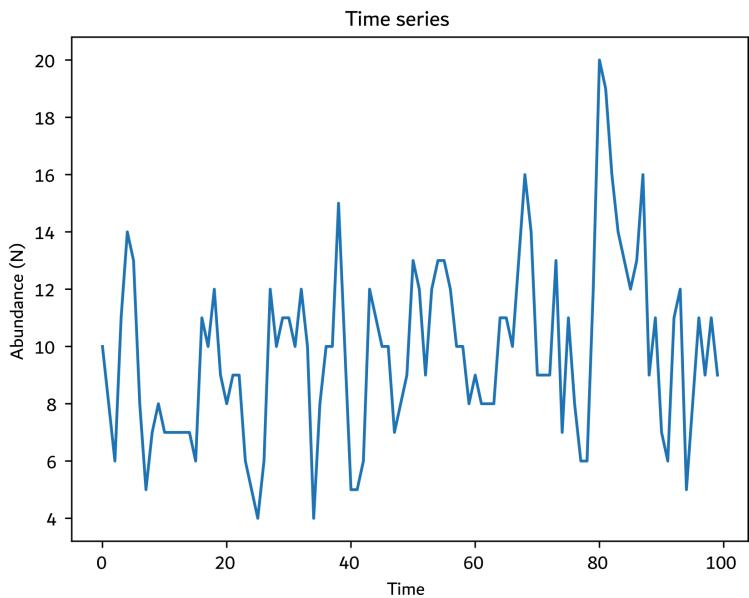
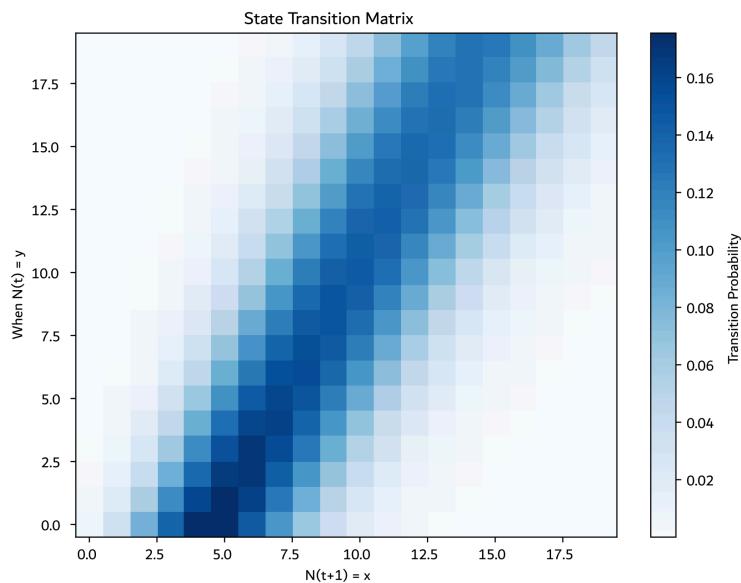
Survival probability = 1 / Recruitment = 0



Transition matrix

C

Survival probability = 0.5 / Recruitment = 5



Transition matrix - Optimization

- One transition matrix **per time step and per time-series** \Rightarrow Bottleneck

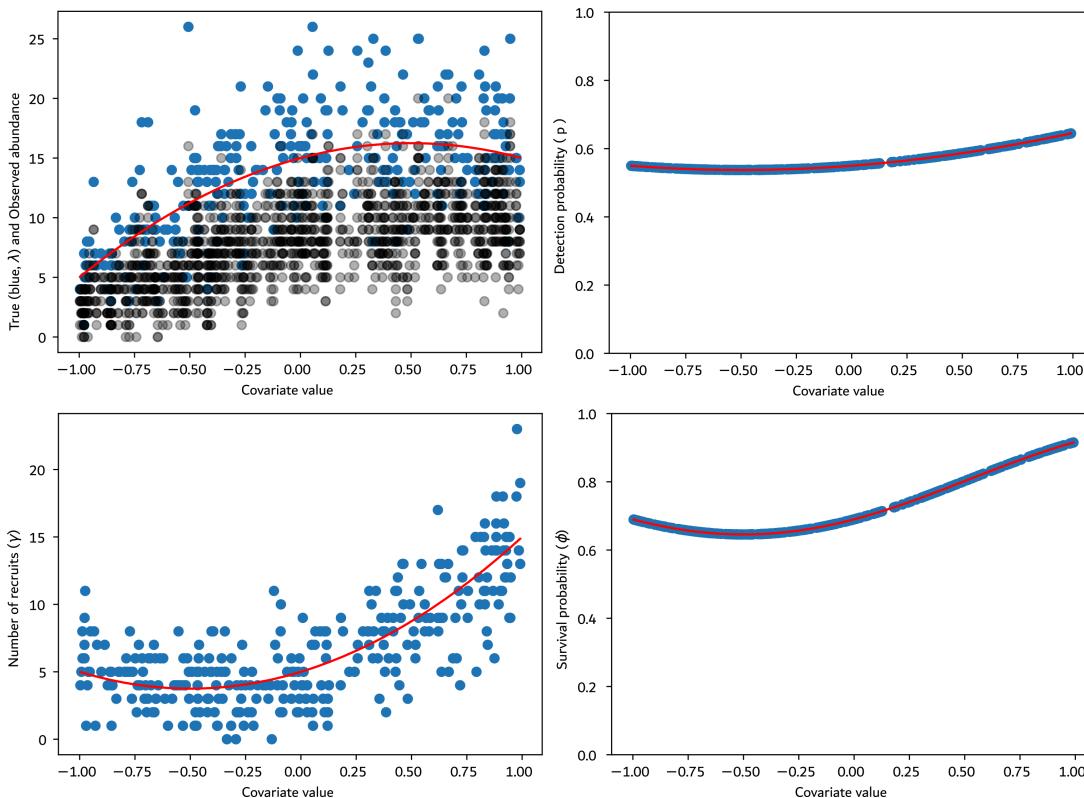
$$P_{jk} = \sum_{c=0}^{\min(j,k)} \text{Bin}(c; j, \phi) \cdot \text{Pois}(k - c; \gamma)$$

$$N_{i,t} \begin{bmatrix} P_{K0} & P_{K1} & P_{K2} & \cdots & P_{KK} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{20} & P_{21} & P_{22} & \cdots & P_{2K} \\ P_{10} & P_{11} & P_{12} & \cdots & P_{1K} \\ P_{00} & P_{01} & P_{02} & \cdots & P_{0K} \end{bmatrix}^{N_{i,t+1}}$$

- **Fast, heavy memory implementation:** vectorisation of P over every time steps and batch size
- **Time gain:** e.g. from 88 minutes in Bayesian framework (JAGS) to ca. 15 minutes with this implementation (CUDA, 8GB VRAM, 300 epochs)

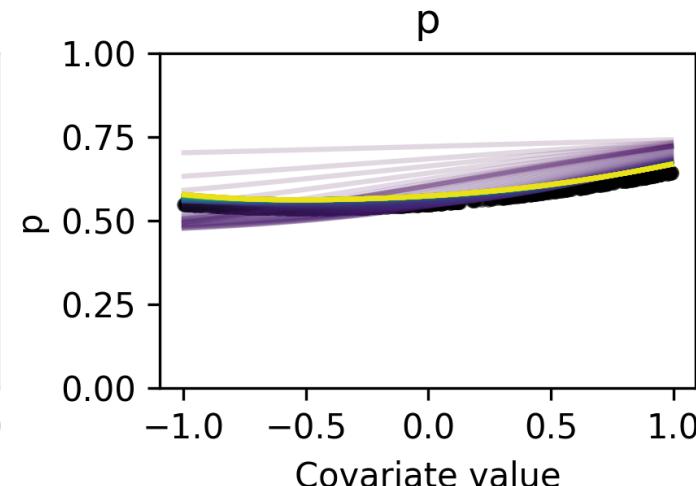
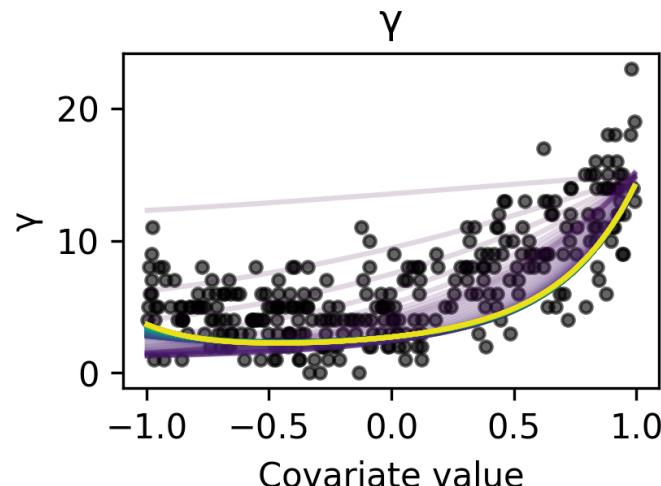
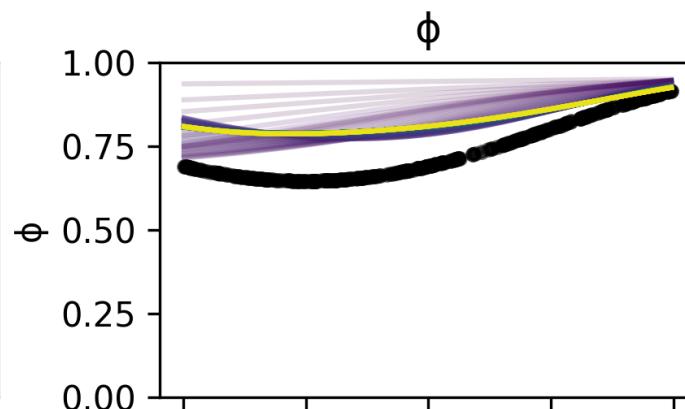
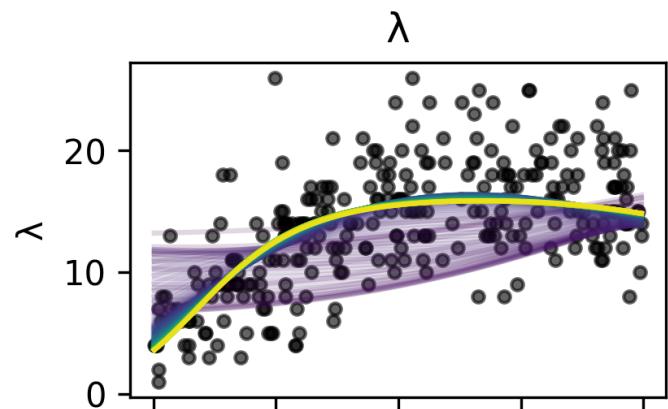
Data simulation

- Simulating quadratic relationship between a covariate x and the parameters p , λ , γ , and ϕ
- 300 sites, 15 years long time series, 5 repeated counts



Results of the simulation

- Simple MLP: 1 hidden layer, 64 hidden units, 300 epochs



Next step

- So far the architecture of the network is a Multilayer Perceptron
- Next step is to use a Convolutional Neural Network with this Loss Function
- Potential to use satellite images as input
- This will tell us which features of the landscape are enhancing survival and recruitment
- Could give real insights on landscape management and provide guidelines for conservation strategies
- Use this framework on large spatio-temporal data

Limitations

- In real life, parametric assumptions are rarely met: the field sampling won't often sample from Poisson and Binomial distribution
- Especially true for Recruitment (γ) and Survival (ϕ), where the input data doesn't contain clear information (unlike abundance N)
- Tends to overestimate survival rate over recruitment
- Marginalization over K : if too small, biased likelihood, if too big, slow computation/memory heavy

Summary

- Neural Dynamic N-mixture model
 - Merges hierarchical ecological modeling with deep learning flexibility
 - Enables fast, parallelizable inference on GPUs
- Interpretability retained
 - Demographic parameters (survival, recruitment, detection, abundance) are explicitly inferred from count data
- Performance advantage
 - Scales better than classical Bayesian MCMC approaches
 - Recovers demographic parameters reliably
- Future potential
 - Extends demographic rate estimation from count data to applications with computer vision

Acknowledgements



JARZYNA LAB



Marta Jarzyna



- Tanya Berger-Wolf
- David Rolnick
- Justin Kitzes
- Sara Beery
- Kaitlyn Gaynor
- Laura Pollock
- Graham Taylor

And many more...



Finally, Maxwell Joseph for helpful discussions and guidance