

Supplementary information

1. Continuous dynamics for the hybrid model

The continuous dynamics of the hybrid dynamical system employed for the multi-season model assumes Lotka-Volterra, predator-prey interactions for species on Bylot and logistic growth for species off Bylot. There are three "seasons" in the model: $q = 0$ is winter, $q = 1$ is a low summer, and $q = 2$ is a peak summer. Let \mathbf{B} be a vector of species biomasses where the component B_1 is the biomass of the brown lemming, B_2 is the biomass of the collared lemming, B_3 is the biomass of the snow goose, B_4 is the biomass of the arctic fox, and B_5 is the biomass of the snowy owl. Allowing the dot symbol (i.e. \dot{B}_i) to be the time derivative, the dynamics for winter are given by the interactions between the resident species (brown lemming, collared lemming, and arctic fox):

$$\begin{aligned}\dot{B}_1 &= b_1^{(0)} B_1 \left(1 - \frac{B_1}{K_1^{(0)}} \right) - \alpha_{14}^{(0)} B_1 B_4 \\ \dot{B}_2 &= b_2^{(0)} B_2 \left(1 - \frac{B_2}{K_2^{(0)}} \right) - \alpha_{24}^{(0)} B_2 B_4 \\ \dot{B}_3 &= b_3^{(0)} B_3 \left(1 - \frac{B_3}{K_3^{(0)}} \right) \\ \dot{B}_4 &= B_4 (-b_4^{(0)} + \alpha_{41}^{(0)} B_1 + \alpha_{42}^{(0)} B_2) - \eta_4^{(0)} B_4^2 \\ \dot{B}_5 &= -b_5^{(0)} B_5 - \eta_5^{(0)} B_5^2,\end{aligned}\tag{1.1}$$

For the low summer dynamics, the predator-prey equations involve the resident species as well as the snow goose:

$$\begin{aligned}\dot{B}_1 &= b_1^{(1)} B_1 \left(1 - \frac{B_1}{K_1^{(1)}} \right) - \alpha_{14}^{(1)} B_1 B_4 \\ \dot{B}_2 &= b_2^{(1)} B_2 \left(1 - \frac{B_2}{K_2^{(1)}} \right) - \alpha_{24}^{(1)} B_2 B_4 \\ \dot{B}_3 &= b_3^{(1)} B_2 \left(1 - \frac{B_3}{K_3^{(1)}} \right) - \alpha_{34}^{(1)} B_3 B_4 \\ \dot{B}_4 &= B_4 (-b_4^{(1)} + \alpha_{41}^{(1)} B_1 + \alpha_{42}^{(1)} B_2 + \alpha_{43}^{(1)} B_3) - \eta_4^{(1)} B_4^2 \\ \dot{B}_5 &= -b_5^{(1)} B_5 - \eta_5^{(1)} B_5^2,\end{aligned}\tag{1.2}$$

and for the peak summer all species, including the snowy owl, are involved in Lotka-Volterra dynamics:

$$\begin{aligned}\dot{B}_1 &= b_1^{(2)} B_1 \left(1 - \frac{B_1}{K_1^{(2)}} \right) - \alpha_{14}^{(2)} B_1 B_4 - \alpha_{15}^{(2)} B_1 B_5 \\ \dot{B}_2 &= b_2^{(2)} B_2 \left(1 - \frac{B_2}{K_2^{(2)}} \right) - \alpha_{24}^{(2)} B_2 B_4 - \alpha_{25}^{(2)} B_2 B_5 \\ \dot{B}_3 &= b_3^{(2)} B_2 \left(1 - \frac{B_3}{K_3^{(2)}} \right) - \alpha_{34}^{(2)} B_3 B_4 - \alpha_{35}^{(2)} B_3 B_5 \\ \dot{B}_4 &= B_4 (-b_4^{(2)} + \alpha_{41}^{(2)} B_1 + \alpha_{42}^{(2)} B_2 + \alpha_{43}^{(2)} B_3) - \eta_4^{(2)} B_4^2 \\ \dot{B}_5 &= B_5 (-b_5^{(2)} + \alpha_{51}^{(2)} B_1 + \alpha_{52}^{(2)} B_2 + \alpha_{53}^{(2)} B_3) - \eta_5^{(2)} B_5^2\end{aligned}\tag{1.3}$$

Data: i - starting data index , j - ending data index , k - number of data points added per iteration

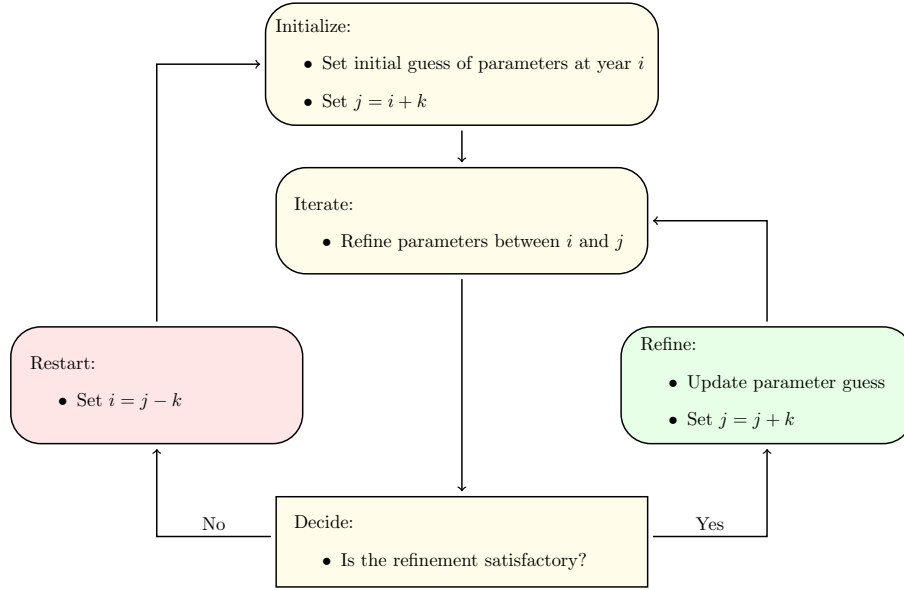


Figure 1. Flow chart of the refinement algorithm to set unknown parameter values for our hybrid system model using the yearly data available for summer biomasses. In our implementation of this algorithm, we assume that $k = 3$ or we are aggregating the time-series in three-year segments.

2. Algorithm for parameter estimation refinement

We develop an optimisation scheme to help refine parameter estimates for unknown parameters (i.e. intrinsic growth rates, intraspecific terms, and winter model parameters) by comparing the model output to the yearly (summer) biomass data. For dynamical systems which are sensitive to parameters, small differences in values can cause large changes in output. The algorithm we develop allows us to incrementally improve parameter estimates by slowly adding more time-points such that we always have a good guess to start the refinement procedure. Fig. 1 details the steps of the algorithm: initialise, iterate, decide, refine or restart.

Initialise: An initial guess is chosen, which should be the best guess we have at any point. To begin, we estimate the intrinsic growth rates using Legagneux et al. [1], which assumes that the equilibrium biomass is equal to the long-term average. We create a local distribution of initial guesses around the best guess and execute the next step for all guesses.

Iterate: MATLAB fmincon function is used to minimise a cost function, thus, refining and improving parameter estimates constrained by suitable upper and lower bounds. Data between years i and j is used. The cost function has two parts:

- (i) Weighted least squares: brown lemmings have the highest weight in an effort to capture multi-annual cycles. Foxes have the least weight to allow for temporal fluctuations.
- (ii) Equilibrium cost: We only look for estimates which have positive equilibria in each season.

Decide: After every iteration a decision must be made. As it is an incremental parameter refinement, a refinement is rejected as being further from the "optimal" parameters if: a) the cost is too large or the distance between subsequent parameter estimates is increasing (i.e. we are not

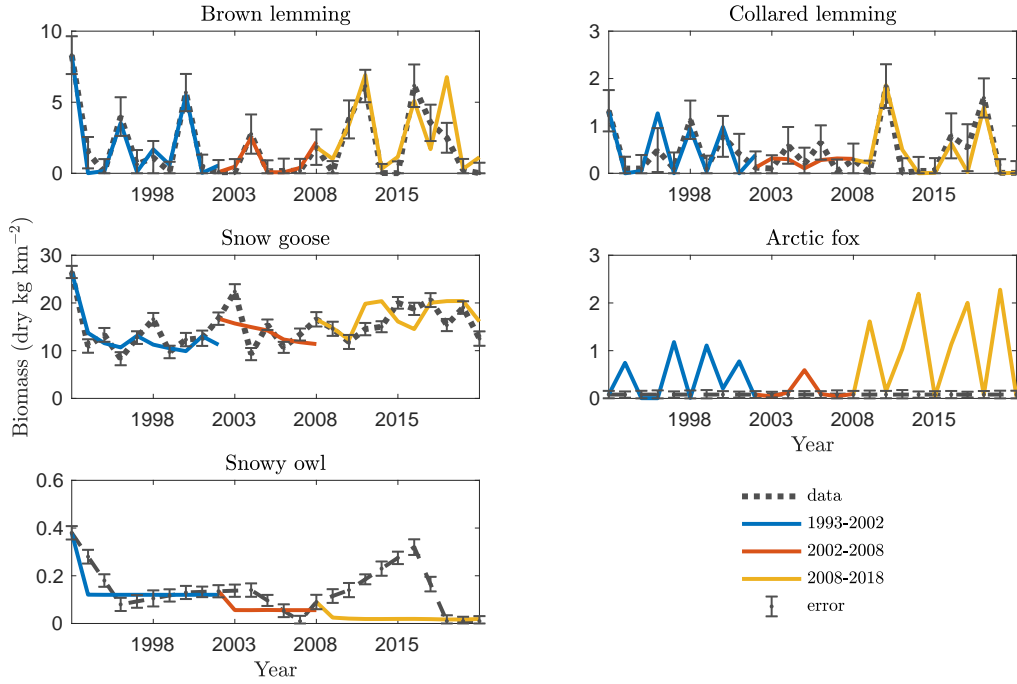


Figure 2. Model output using parameter estimation scheme. Three segments are inferred. The blue segment (1993-2002) is used for this study.

converging to a parameter set) and b) the qualitative dynamics are poor (i.e. the parameters are such that there is no variation in the lemming biomasses).

Refine: If a parameter refinement is accepted, more data is added in units of three years and another refinement is performed.

Restart: If a parameter refinement is rejected, a new set of parameters are generated and refined from the best guess that exists so far. In this case, the cost is no longer computed using data which was used for the previous set of parameters. This step helps to improve the parameter estimates overall.

This algorithm generates three segments from the biomass data available for the Bylot Island community. These are shown in Fig. 2.

It is possible that the non-stationarity of parameters in the model are driven by such environmental factors such as spring snow cover. The set of parameters used in this study are from the first segment (i.e. 1993-2002). The reason for this choice is threefold: 1) they seem to represent "historical" brown lemming cycling, 2) they have the lowest cost of the three segments, and 3) they capture important qualitative features desired in the model. For this segment, the intrinsic growth rates are given by

$$b^{(0)} = (3.7588, 8.7888, 0.1394, -0.8003, 6.8620) \quad (2.1)$$

$$b^{(1)} = (0.9068, 9.9698, 2.1270, -9.9993, 3.1959) \quad (2.2)$$

$$b^{(2)} = (0.4576, 9.9690, 16.1250, -9.9982, -3.5495). \quad (2.3)$$

The winter interaction coefficients are given by

$$\alpha^{(0)} = \begin{bmatrix} -0.1481 & 0 & 0 & -13.5012 & 0 \\ 0 & -6.8874 & 0 & -19.9908 & 0 \\ 0 & 0 & -0.0437 & 0 & 0 \\ 1.7848 & 0.2341 & 0 & -1.0430 & 0 \\ 0 & 0 & 0 & 0 & -56.6439 \end{bmatrix} \quad (2.4)$$

the low summer coefficients are

$$\alpha^{(1)} = \begin{bmatrix} -0.0006 & 0 & 0 & -29.8506 & 0 \\ 0 & -1.1080 & 0 & -83.7546 & 0 \\ 0 & 0 & -0.0413 & -2.6821 & 0 \\ 0.7939 & 2.2275 & 0.0713 & -278.5454 & 0 \\ 0 & 0 & 0 & 0 & -187.4275 \end{bmatrix} \quad (2.5)$$

, and for a peak summer

$$\alpha^{(2)} = \begin{bmatrix} -0.0001 & 0 & 0 & -7.3207 & -1.2696 \\ 0 & -694.9786 & 0 & -8.4118 & -8.1196 \\ 0 & 0 & -0.7034 & -0.9229 & -0.1563 \\ 0.1947 & 0.2237 & 0.0245 & -176.5767 & 0 \\ 0.1123 & 0.7183 & 0.0138 & 0 & -65.6308 \end{bmatrix} \quad (2.6)$$

The diagonal components are proportional to the intraspecific competition terms (i.e. carrying capacities). We do not expect these to correspond to independently measured quantities (i.e. in the field).

3. Convergent cross mapping (CCM)

(a) Interpreting CCM

To conduct a CCM, two manipulations need to be done to the time series [2]. First, the time-series for each species needs to be normalised to allow comparison between them. This is accomplished by subtracting from it the mean and dividing by the variance. Second, the time-series must exhibit stationarity; that is, the mean and variance are time independent. We check for stationarity by inspection, choosing only simulations which best satisfy this constraint.

rEDM (Empirical Dynamic Modelling) is an R package which computes cross map skill based on these time-series as Pearson's correlation coefficient, ρ [3]. The CCM function requires several inputs in addition to the time series of the target and the causal variables. One of these inputs is the optimal embedding dimension for each species is the number of lagged coordinates for that species needed to best reconstruct the state space attractor (it may not be the same as the dimensionality of the corresponding dynamical system). We use rEDM simplex projection to identify the embedding dimension for each species. The second input depends on a splitting of the time-series into two halves: a "library" set, which is used to create reconstructions, and a prediction set, which is used to make forecasts on. We chose to take time-data once per month as this is the most common method for collecting data. We compute the averaged cross map skill as the average ρ over 100 simulated runs (which differ from each other due to the inherent randomness of owls coupling to the summer dynamics) with invalid time-series removed (i.e. those which had cross maps with NA).

The cross map for the time-series generated from the multi-season model is shown in the main text. Below we show the results of the cross mapping for the static model in Fig. 3 and every 2 months of summer are taken from the time-series and concatenated in Fig. 4.

(b) Significance testing

Suppose that we want to determine whether the converged value of Pearson's correlation coefficient, ρ , associated with the causal link $x \rightarrow y$ is significantly different than zero. We use the method developed by Aftab et al. [4]. Since ρ is obtained from using a sample of points in the state space attractor of x , it has a distribution which we can transform via Fisher's z -transformation to a normally distributed variable or z -score:

$$z = \frac{1}{2} \ln \left(\frac{1 + \rho}{1 - \rho} \right). \quad (3.1)$$

We then standardize to the unit normal,

$$Z^* = \frac{z}{\sqrt{\frac{1}{N-3}}}. \quad (3.2)$$

For a $Z^* = 1.96$ indicating a p -value of 0.05 and a library size of $N = 600$ (multi-season and summer models) $\rho > 0.080$ for the causation to be significant. For a library size of $N = 3000$ (static model), $\rho > 0.036$ for the causation to be significant.

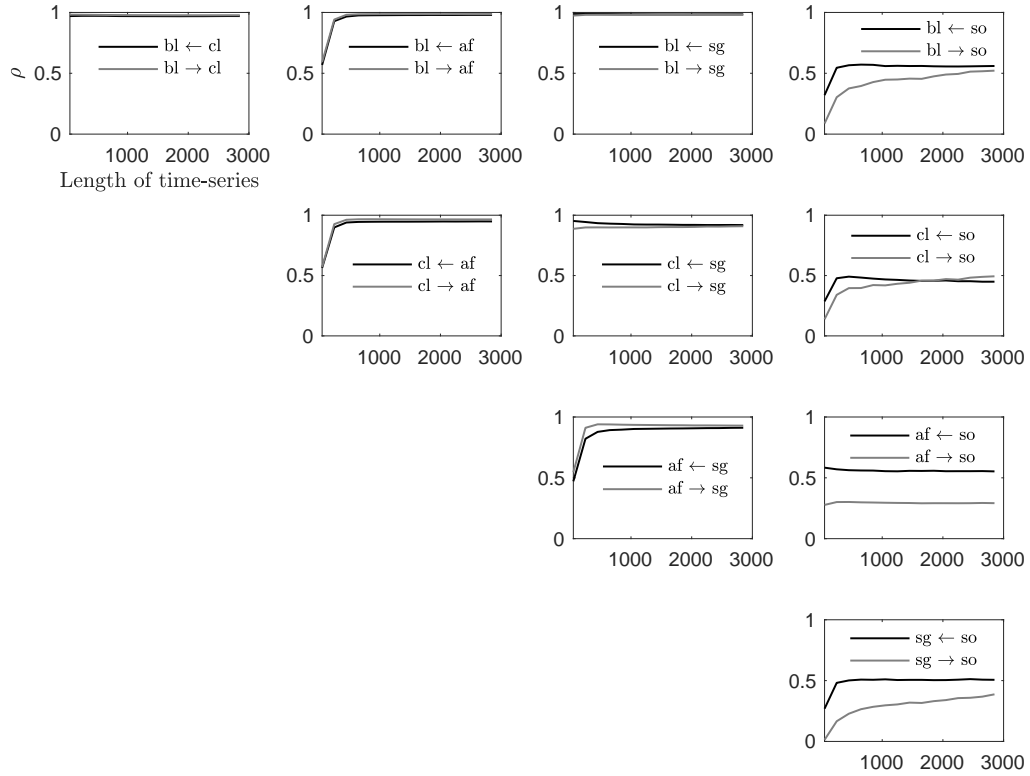


Figure 3. Cross maps for the static model which assumes three prey (brown lemming, collared lemming, and snow goose) and two predators (arctic fox and snowy owl). Only correlation coefficients, ρ , which are significantly different from zero ($P < 0.05$) are shown (see Supplementary Information section 3). A causal relationship is indicated if ρ increases as the length of the time-series increases and if ρ converges to a fixed value. Species 1 \leftarrow species 2 means that species 2 causes species 1 if the above criteria are met. Here bl=brown lemming, cl=collared lemming, af=arctic fox, sg=snow goose, and so=snowy owl. If a causal relationship is not present, then ρ can be taken to represent the degree of correlation between the variables.

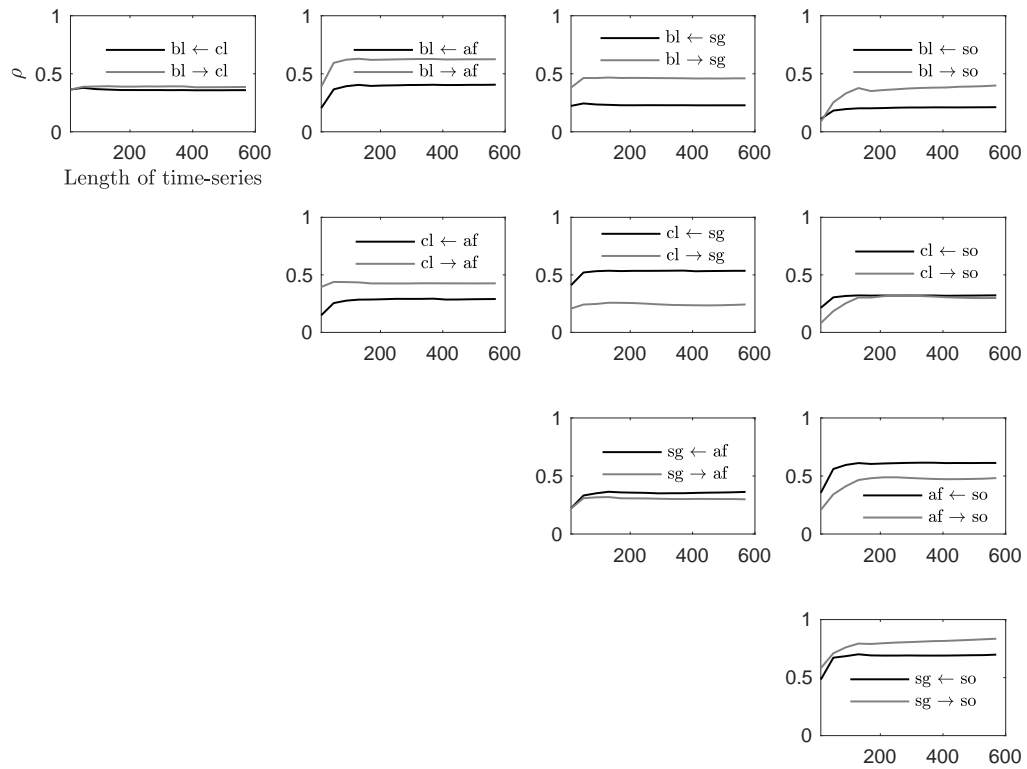


Figure 4. Cross maps for the summer snapshot model obtained from concatenating the time-series in the multi-season model which correspond to summer months. Only correlation coefficients, ρ , which are significantly different from zero ($P < 0.05$) are shown (see Supplementary Information section 3). A causal relationship is indicated if ρ increases as the length of the time-series increases and if ρ converges to a fixed value. Species 1 \leftarrow species 2 means that species 2 causes species 1 if the above criteria are met. Here bl=brown lemming, cl=collared lemming, af=arctic fox, sg=snow goose, and so=snowy owl. If a causal relationship is not present, then ρ can be taken to represent the degree of correlation between the variables.

References

1. P. Legagneux, G. Gauthier, D. Berteaux, J. Bêty, M.-C. Cadieux, F. Bilodeau, E. Bolduc, L. McKinnon, A. Tarroux, J.-F. Therrien, *et al.*, "Disentangling trophic relationships in a high arctic tundra ecosystem through food web modeling," *Ecology*, vol. 93, no. 7, pp. 1707–1716, 2012.
2. C.-W. Chang, M. Ushio, and C.-h. Hsieh, "Empirical dynamic modeling for beginners," *Ecological research*, vol. 32, no. 6, pp. 785–796, 2017.
3. H. Ye, A. Clark, E. K. Deyle, and G. Sugihara, "redm: an r package for empirical dynamic modeling and convergent cross mapping," 2018.
4. M. F. Aftab, M. Hovd, and S. Sivalingam, "Convergent cross mapping (ccm) based approach for isolating the source of plant-wide disturbances," in *2017 IEEE Conference on Control Technology and Applications (CCTA)*, pp. 1492–1498, IEEE, 2017.