Supplementary Material for

Metapopulation persistence can be inferred from incomplete surveys

Chuliang Song[1,2*,†], Marie-Josée Fortin[2], Andrew Gonzalez[1]

[1] Department of Biology, Quebec Centre for Biodiversity Science, McGill University, Montreal, Canada

[2] Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Canada

[†] Present addreess: Department of Ecology and Evolutionary Biology, Princeton University, Princeton, US

Contents

A Derivation of the analytic estimate **S1**

B Illustration of patch distributions **S2**

C Illustration of dispersal kernel **S3**

D Empirical metapopulation **S4**

E Predicting the effects of patch removal **S5**

F Effects of biased sampling **S6**

A Derivation of the analytic estimate

Here we derive the analytic-based estimate (Eqn. 7). Recall that the connectivity matrix M is given by

$$M_{ij} = A_i^e A_j^\omega f(d_{ij}/\xi). \tag{S1}$$

The connectivity matrix M can be decomposed as $M = A \odot D$, where $A_{ij} = A_i^e A_j^\omega$ and $D_{ij} = f(d_{ij}/\xi)$, and $\odot$ is the Hadamard product of matrix. If the connectivity matrix M is close to the assumption of random matrix, then we can use the asymptotic estimate of leading eigenvector (Eqn. 5) to estimate the capacity $\hat{\lambda}(M)$ (Füredi & Komlós, 1981):

$$\hat{\lambda}(M) = n\overline{M} + \frac{\sigma_M^2}{\overline{M}} + O\left(\frac{1}{\sqrt{n}}\right) \tag{S2}$$

$$= n\overline{A \odot D} + \frac{\sigma_{A\odot D}^2}{\overline{A \odot D}} + O\left(\frac{1}{\sqrt{n}}\right) \tag{S3}$$

If we further assume that patch area and distance are uncorrelated, then we have

$$\hat{\lambda}(M) = n\overline{A \odot D} + \frac{\sigma_{A\odot D}^2}{\overline{A \odot D}} + O\left(\frac{1}{\sqrt{n}}\right) \tag{S4}$$

$$= n\overline{D} \cdot \overline{A} + \frac{\sigma_{A\odot D}^2}{\overline{D} \cdot \overline{A}} + O\left(\frac{1}{\sqrt{n}}\right) \tag{S5}$$

$$= n\overline{D} \cdot \overline{A} + \frac{\left(\sigma_D^2 + \overline{D}^2\right)\left(\sigma_A^2 + \overline{A}^2\right) - \overline{D}^2\overline{A}^2}{\overline{D} \cdot \overline{A}} + O\left(\frac{1}{\sqrt{n}}\right), \tag{S6}$$

where the second equation follows from $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ for independent X and Y, and the third equation follows from $\sigma_{XY} = (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2\mu_Y^2$ for independent X and Y.

The asymmetry in the off-diagonal elements originates from the heterogeneous patch area A. The level of dependency can be measured by

$$\frac{\lambda(A)}{k\overline{A} + \sigma_A^2/\overline{A}}, \tag{S7}$$

which is the ratio between asymptotically estimated eigenvalue and true eigenvalue of A.

Combined Eqns. S6 and S7, we have that

$$\hat{\lambda}(M) = \underbrace{\left(n\overline{D} \cdot \overline{A} + \frac{\sigma_A^2}{\overline{A}} + \frac{\sigma_D^2}{\overline{D}} + \frac{\sigma_D^2\sigma_A^2}{\overline{D} \cdot \overline{A}}\right)}_{\text{estimate under independency}} \cdot \underbrace{\frac{\lambda(A)}{k\overline{A} + \sigma_A^2/\overline{A}}}_{\text{correct for dependency}} \tag{S8}$$

B Illustration of patch distributions

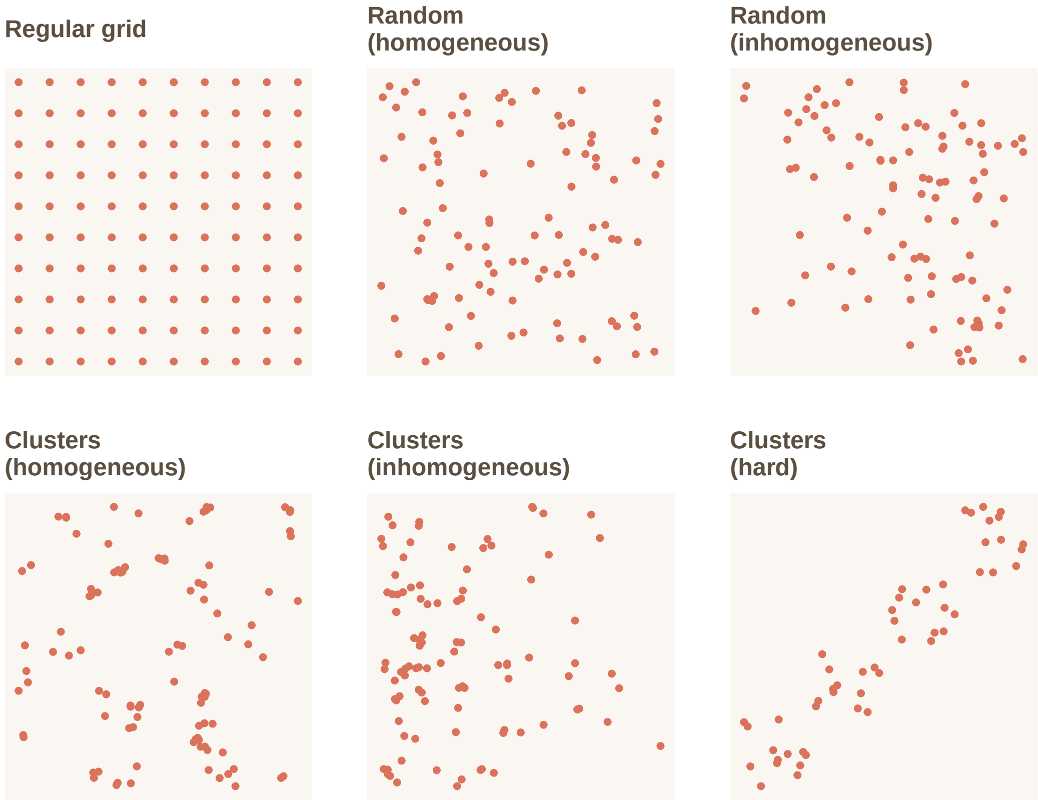

Figure S1: Each panel denotes a different patch distributions. The details for patch simulations can be found in Table 1.

C Illustration of dispersal kernel

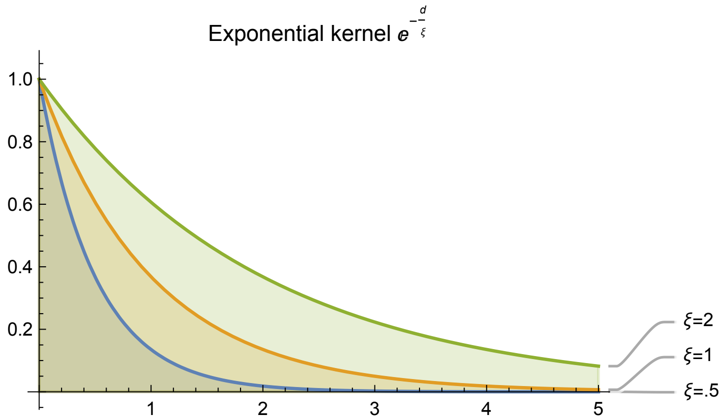

Figure S2: Illustration of the exponential kernel $e^{-d/\xi}$. Three lines correspond to three different characteristic scales ($\xi = 0.5, 1, 2$).

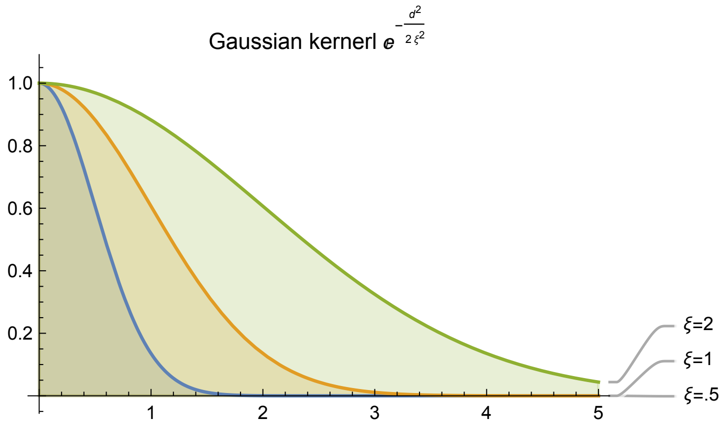

Figure S3: Illustration of the Gaussian kernel $e^{-d^2/(2\xi^2)}$. Three lines correspond to three different characteristic scales ($\xi = 0.5, 1, 2$).

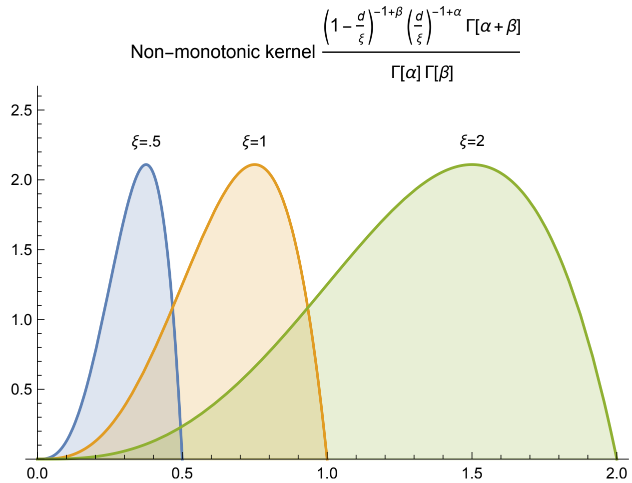

Figure S4: Illustration of the non-monotonic kernel $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\left(\frac{d_{ij}}{\xi}\right)^{\alpha-1}\left(1-\frac{d_{ij}}{\xi}\right)^{\beta-1}$. We chose $\alpha = \beta = 2$. Three lines correspond to three characteristic scales ($\xi = 0.5, 1, 2$).

D Empirical metapopulation

Here we provide some summary statistics of the empirical dataset (Huang *et al.*, 2020).

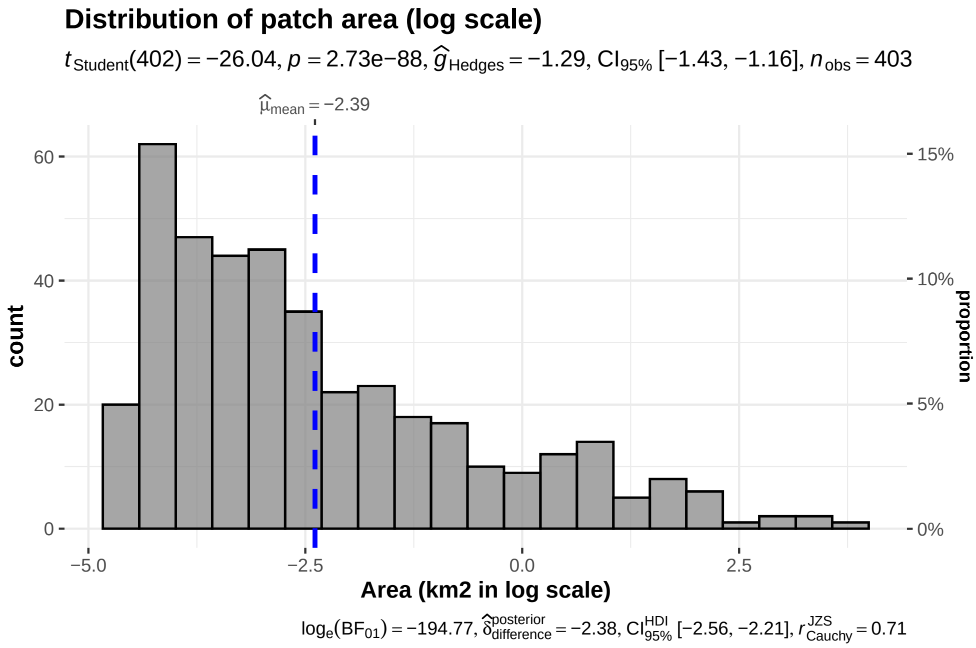

Figure S5: Distribution of patch area in the empirical dataset. The horizontal axis shows the area size in log scale (unit: km^2). The mean area size is 0.09 km^2. The vertical axis shows the frequency of area size.

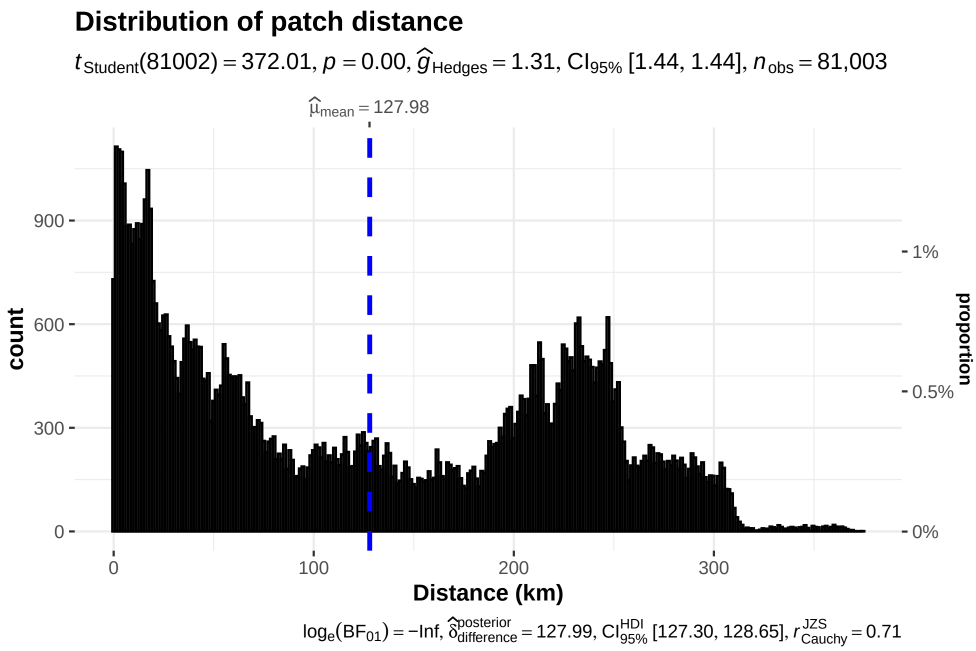

Figure S6: Distribution of patch distance in the empirical dataset. The horizontal axis shows the pairwise patch distance (unit: km). The mean patch distance is 127.98 km. The vertical axis shows the frequency of patch distance.

E Predicting the effects of patch removal

In the main text, we have studied how to predict metapopulation capacity of the whole metapopulation from sub-populations. Here we consider the inverse question of predicting metapopulation capacity of sub-populations from the whole metapopulation.

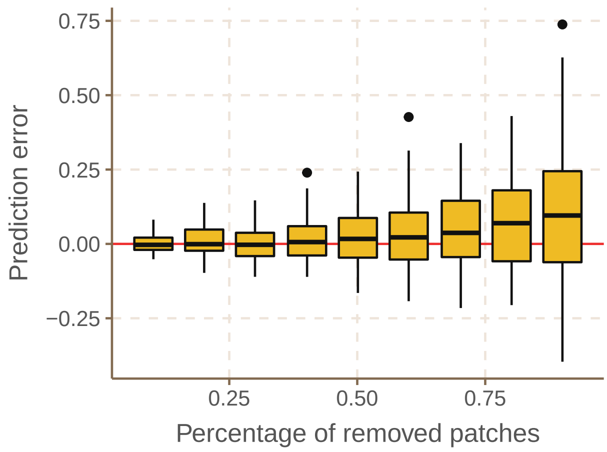

Figure S7: Predicting metapopulation capacity of sub-populations in the empirical dataset. We used all the parameters provided by Huang *et al.* (2020). We used the analytic-based predictor as we only have one sample (i.e., the whole connectivity matrix). The horizontal axis denotes the percentage of removed patches. For example, 75% denotes that 75% of all patches are removed and only 25% of patches remain. The vertical axis denotes the prediction error ($(\hat{\lambda} - \lambda)/\lambda$). Unsurprisingly, because of the law of small numbers, the variance of prediction error increases as more patches are removed. The mean of the prediction error increases as well (this is because the analytic-based estimator is a conservative estimator).

F Effects of biased sampling

Here, we provide a simple example on biased sampling. Suppose we have a bias to sample large patches. Specifically, the probability of sampling a patch i is proportional to A_i^ρ, where A_i is the area of the patch i and ρ is the sampling bias. If $\rho = 0$, then we sample all patches with equal probability (i.e., unbiased). The larger ρ is, we sample large patches with larger probability.

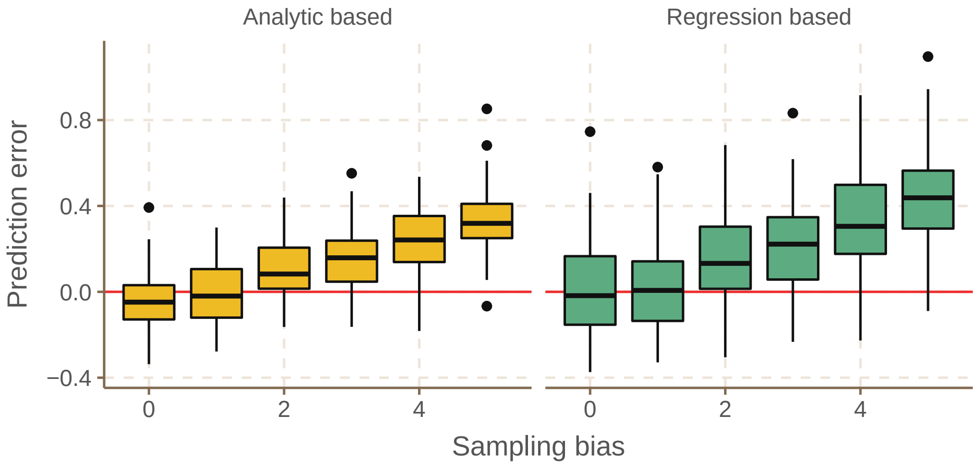

Figure S8: Effects of sampling bias on prediction of metapopulation capacity in the empirical dataset. We used all the parameters provided by Huang *et al.* (2020). The horizontal axis denotes the sampling bias ρ. The vertical axis denotes the prediction error $((\hat{\lambda} - \lambda)/\lambda)$. The red line denotes the zero prediction error. When there is no sampling bias ($\rho = 0$), the distributions of prediction error are the same as Figure 3. When the sampling bias is small ($\rho < 1$), we do not see a strong effect on the prediction error. However, the changes in prediction error become obvious and keep increasing with a higher level of sampling bias.

References

Füredi, Z. & Komlós, J. (1981). The eigenvalues of random symmetric matrices. *Combinatorica*, 1, 233–241.

Huang, R., Pimm, S. L. & Giri, C. (2020). Using metapopulation theory for practical conservation of mangrove endemic birds. *Conservation Biology*, 34, 266–275.