

**Supplementary Materials: Eco-evolutionary dynamics and collective dispersal:
implications for salmon metapopulation robustness**

Justin D. Yeakel^{1,2,*}, Jean P. Gibert¹, Peter A. H. Westley³, & Jonathan W. Moore⁴

¹School of Natural Sciences, University of California, Merced, Merced CA, USA

²The Santa Fe Institute, Santa Fe NM, USA

³College of Fisheries and Ocean Sciences, University of Alaska, Fairbanks, Fairbanks AK, USA

⁴Earth2Oceans Research Group, Simon Fraser University, Vancouver BC, Canada

*To whom correspondence should be addressed: jdyeakel@gmail.com

We aim to understand how straying between populations may affect phenotypic evolution, and how in turn this evolutionary change may lead to changes in ecological dynamics as they unfold. To do so, we build upon ideas from classic papers (Lande 1976, most notably). In what follows, we show 1) how the application of Lande's equation for the case of two populations linked by dispersal leads to complex equations of trait evolution that are not easily modeled, 2) that this complex trait distribution can be approximated as a single Gaussian distribution, and 3) under what conditions this approximation is appropriate.

I. Phenotypic evolution from one to two sites

Trait evolution in one site To keep track of the evolution of a phenotypic trait x with probability density function p and mean \bar{x} and variance σ^2 , we need a few basic ingredients. First, we define the mean fitness, \bar{W} of the population of individuals bearing the trait x as:

$$\bar{W} = \int p(x)W(x) dx. \quad (1)$$

We also define the mean phenotype in the population as

$$\bar{x} = \int xp(x) dx, \quad (2)$$

and the mean phenotype after selection (but before reproduction), \bar{x}_w Falconer (1967), as:

$$\bar{x}_w = \frac{1}{\bar{W}} \int xp(x)W(x) dx. \quad (3)$$

Then, we can assess how the change in the phenotype x occurs over time following

$$\Delta x = x(t+1) - x(t), \quad (4)$$

which can be rewritten as:

$$\Delta x = h^2(\bar{x}_w - x(t)), \quad (5)$$

where h^2 is the heritability of the trait, and the difference $\bar{x}_w - x(t)$ is the classical "selection differential" Falconer (1967), Lande (1976). In other words, 5 is none other than the Breeder's equation.

To understand how x changes over time, we first need to understand how the mean fitness of the population changes with a change in the mean trait value or $\frac{\partial \bar{W}}{\partial \bar{x}}$ (i.e. the shape of the "adaptive landscape"). Using 1 we have:

$$\frac{\partial \bar{W}}{\partial \bar{x}} = \frac{\partial}{\partial \bar{x}} \int p(x)W(x) dx. \quad (6)$$

Using Leibniz rule we can pass the derivative under the integral sign,

$$\frac{\partial \bar{W}}{\partial \bar{x}} = \int \frac{\partial p(x)}{\partial \bar{x}} W(x) dx, \quad (7)$$

25 and assuming p is Gaussian (i.e. of the form $\frac{1}{\sqrt{2\pi}\sigma^2}e^{-\frac{1}{2}\frac{(x-\bar{x})^2}{\sigma^2}}$), we calculate the derivative of p with respect to \bar{x} ,
 26 which leads to:

$$\frac{\partial \bar{W}}{\partial \bar{x}} = \int \left[\frac{x - \bar{x}}{\sigma^2} \right] p(x) W(x) dx. \quad (8)$$

We can now expand in 8 to get,

$$\frac{\partial \bar{W}}{\partial \bar{x}} = \frac{1}{\sigma^2} \int x p(x) W(x) dx - \frac{\bar{x}}{\sigma^2} \int p(x) W(x) dx, \quad (9)$$

which making use of equations 1 and 3, becomes $\frac{\partial \bar{W}}{\partial \bar{x}} = \frac{\bar{W}}{\sigma^2} \bar{x}_w - \frac{\bar{x}}{\sigma^2} \bar{W}$, which can be factored as $\frac{\partial \bar{W}}{\partial \bar{x}} = \frac{\bar{W}}{\sigma^2} (\bar{x}_w - \bar{x})$. Using 5, we can rearrange this expression to obtain an equation that related the change in the average phenotype from one time step to the next to the amount of heritable variation, $\sigma^2 h^2$, and the adaptive landscape $\frac{\partial \bar{W}}{\partial \bar{x}}$:

$$\Delta \bar{x} = \frac{\sigma^2 h^2}{\bar{W}} \frac{\partial \bar{W}}{\partial \bar{x}} = \sigma^2 h^2 \frac{\partial \ln \bar{W}}{\partial \bar{x}}. \quad (10)$$

Equation 10 (Lande 1976's equation 7), is a staple of evolutionary biology. It is also possible to use 10 and 4 to write a recurrence relationship that allows to simulating evolutionary change over time in a trait x as:

$$\bar{x}(t+1) = \bar{x}(t) + \sigma^2 h^2 \frac{\partial \ln \bar{W}}{\partial \bar{x}}. \quad (11)$$

27 **Expanding the single site approach to two sites linked by dispersal** In what follows we show how Lande's
 28 approach can be used to calculate the magnitude of the phenotypic change between two populations with one trait with
 29 two different means in each site. We will then 1) show why it is not necessarily useful to use the obtained expression
 30 in a simulation context, and 2) introduce an approximation that retains the features important for exploring how
 31 trait and population dynamics impact each other in the case of dispersing populations. In Section II, we will explore
 32 how well the approximation reproduces the underlying trait dynamics. Specifically, we will discuss in what cases it is
 33 exact and for which cases the approximation becomes problematic, such that it should be avoided.

34 We explore the case of two populations that exchange migrants, with a normally distributed trait x that controls
 35 recruitment (as explained in the main text), with each population having a different mean, μ_i and μ_j , and same
 36 variance σ^2 . For simplicity, we will focus on a local population i relative to the connected population j .

37 When individuals from population j stray into population i , and individuals of population i stray into population
 38 j , the phenotypic distribution of trait x in population i is a mixture distribution, $p(x)$, of the form:

$$p(x) = \omega_i g(x, \mu_i) + (1 - \omega_i) g(x, \mu_j). \quad (12)$$

39 where $g(\cdot)$ is the Gaussian probability density function, and ω_i is the proportion of the mixed population that is
 40 composed of individuals local to site i relative to those that have dispersed into site i from site j . This distribution
 41 has a mean $\bar{x} = \omega_i \mu_i + (1 - \omega_i) \mu_j$. As before, we define the mean fitness as in equation 1, only that now p is a
 42 mixture distribution. Let us use Lande's approach to obtain an expression for the change in the mean phenotype of
 43 the mixture distribution from one time step to the next. We thus ask, what is the change in \bar{W} with a change in
 44 \bar{x} ? We can use Lande's approach Lande (1976) to derive equation 7, replacing $p(x)$ by the mixture (equation 12) to
 45 obtain

$$\frac{\partial \bar{W}}{\partial \bar{x}} = \int \frac{\partial(\omega_i g(x, \mu_i) + (1 - \omega_i)g(x, \mu_j))}{\partial \bar{x}} W(x) dx. \quad (13)$$

To obtain the derivative of equation 13, we introduce a change in variables, which is provided by the fact that $\bar{x} = \omega_i \mu_i + (1 - \omega_i) \mu_j$:

$$\begin{cases} \mu_i = \frac{\bar{x} - (1 - \omega_i) \mu_j}{\omega_i} \\ \mu_j = \frac{\bar{x} - \omega_i \mu_i}{1 - \omega_i} \end{cases}. \quad (14)$$

We can thus replace equation 14 into 13,

$$\frac{\partial \bar{W}}{\partial \bar{x}} = \int \frac{\partial(\omega_i g(x, \frac{\bar{x} - (1 - \omega_i) \mu_j}{\omega_i}) + (1 - \omega_i)g(x, \frac{\bar{x} - \omega_i \mu_i}{1 - \omega_i}))}{\partial \bar{x}} W(x) dx, \quad (15)$$

46 which, after expanding and taking the derivatives with respect to \bar{x} becomes:

$$\begin{aligned} \frac{\partial \bar{W}}{\partial \bar{x}} &= \int \omega_i \left[\frac{x - \frac{\bar{x} - (1 - \omega_i) \mu_j}{\omega_i}}{\omega_i \sigma^2} \right] g(x, \frac{\bar{x} - (1 - \omega_i) \mu_j}{\omega_i}) W(x) dx \\ &\quad + \int (1 - \omega_i) \left[\frac{x - \frac{\bar{x} - \omega_i \mu_i}{1 - \omega_i}}{(1 - \omega_i) \sigma^2} \right] g(x, \frac{\bar{x} - \omega_i \mu_i}{1 - \omega_i}) W(x) dx. \end{aligned} \quad (16)$$

To simplify notation, at this point we can come back to the original variables μ_i and μ_j . We then expand and collect terms to get:

$$\begin{aligned} \frac{\partial \bar{W}}{\partial \bar{x}} &= \frac{1}{\sigma^2} \left[\int x g(x, \mu_i) W(x) dx - \mu_i \int g(x, \mu_i) W(x) dx \right. \\ &\quad \left. + \int x g(x, \mu_j) W(x) dx - \mu_j \int g(x, \mu_j) W(x) dx \right]. \end{aligned} \quad (17)$$

47 Using equations 1 and 3, we can define the following quantities,

$$\bar{W}_i = \int g(x, \mu_i) W(x) dx \quad (18)$$

$$\bar{W}_j = \int g(x, \mu_j) W(x) dx \quad (19)$$

$$\bar{x}_{w,i} = \frac{1}{\bar{W}} \int x g(x, \mu_i) W(x) dx \quad (20)$$

$$\bar{x}_{w,j} = \frac{1}{\bar{W}} \int x g(x, \mu_j) W(x) dx, \quad (21)$$

then replace them into equation 17:

$$\begin{aligned} \frac{\partial \bar{W}}{\partial \bar{x}} &= \frac{1}{\sigma^2} [\bar{x}_{w,i} \bar{W}_i - \mu_i \bar{W}_i + \bar{x}_{w,j} \bar{W}_j - \mu_j \bar{W}_j] \\ &= \frac{1}{\sigma^2} [\bar{W}_i (\bar{x}_{w,i} - \mu_i) + \bar{W}_j (\bar{x}_{w,j} - \mu_j)]. \end{aligned} \quad (22)$$

Using equation 5 and rearranging terms we get:

$$\overline{W}_i \Delta \mu_i + \overline{W}_j \Delta \mu_j = \sigma^2 h^2 \frac{\partial \overline{W}}{\partial \bar{x}}, \quad (23)$$

with $\overline{W} = \omega_i \overline{W}_i + (1 - \omega_i) \overline{W}_j$. Equation 23 relates how changes in either μ_i or μ_j lead to changes in fitness, and vice-versa. However, it doesn't tell us anything as to how straying may lead to evolutionary change, or how that may feedback onto ecological dynamics. Because of this, we will once again use the change of variables in 14, to rewrite 23 as:

$$\overline{W}_i \left(\frac{\Delta \bar{x} - (1 - \omega_i) \Delta \mu_j}{\omega_i} \right) + \overline{W}_j \left(\frac{\Delta \bar{x} - \omega_i \Delta \mu_i}{1 - \omega_i} \right) = \sigma^2 h^2 \frac{\partial \overline{W}}{\partial \bar{x}}, \quad (24)$$

which can be simplified and rearranged to obtain an expression for the change in the mean phenotype of the mixture distribution,

$$\Delta \bar{x} = \left[\sigma^2 h^2 \frac{\partial \overline{W}}{\partial \bar{x}} + \left(\frac{(1 - \omega_i)}{\omega_i} \overline{W}_i \Delta \mu_j + \frac{\omega_i}{1 - \omega_i} \overline{W}_j \Delta \mu_i \right) \right] \frac{\omega_i (1 - \omega_i)}{(1 - \omega_i) \overline{W}_i + \omega_i \overline{W}_j}. \quad (25)$$

By multiplying 25 by $\frac{\overline{W}}{\overline{W}}$, we can further rewrite the equation as:

$$\Delta \bar{x} = \left[\sigma^2 h^2 \frac{\partial \ln \overline{W}}{\partial \bar{x}} + \frac{1}{\overline{W}} \left(\frac{(1 - \omega_i)}{\omega_i} \overline{W}_i \Delta \mu_j + \frac{\omega_i}{1 - \omega_i} \overline{W}_j \Delta \mu_i \right) \right] \frac{\omega_i (1 - \omega_i) \overline{W}}{(1 - \omega_i) \overline{W}_i + \omega_i \overline{W}_j}, \quad (26)$$

which has a similar form to Lande's equation (10), but shows the explicit dependence of the change in the mean of the mix on the proportion of straying individuals. A symmetrical expression can be derived for the other site.

While some understanding can be gained using equation 26, the expression does not easily allow us to simulate the change in \bar{x} over time because the change in μ_i and μ_j is difficult to track in closed form. Moreover, this equation only holds for the first time step, and as new dispersing individuals arrive and the trait distribution is again updated by the influx of individuals from the other site, the mix becomes a mix of mixes, such that the analytical expressions presented above no longer hold.

This presents a tradeoff: we can either track changes in phenotype exactly for a single generation, or we can track changes in phenotype across multiple generations using an approximation. Given that our central goal is to gain an understanding of the feedback between evolutionary and ecological processes in salmon populations with dispersal in explicit space, we simplify the complexities created by phenotypic evolution by assuming that, as new migrants arrive, the mixture distribution can be locally approximated by a Gaussian distribution with a mean equal to that of the mix. Below, we show why this approximation is appropriate for describing the phenotypic distributions under most conditions relevant to our questions. Moreover, based on this assumption, we are able to use Lande's equation (here equation 10), without violating the original assumptions that led to its current form.

II. Gaussian approximation for the mixed trait distribution

We assume a Gaussian distribution to approximate the more complex trait distribution that results when two populations with two distinct trait distributions are linked by gene flow and reproduction. Here and throughout, we refer to the resident population as i and the connected population as j , though both populations stray simultaneously. The Gaussian approximation is assumed to have a mean that is equal to the mean of the mixed trait distribution

$\omega_i \mu_i + (1 - \omega_i) \mu_j$ with a constant variance σ^2 , equal to the original trait variance prior to mixing. We employ this approximation because a mixed trait distribution that remixes and recombines (through reproduction) at each time-step becomes complicated very quickly, as we shall demonstrate from a mechanistic perspective. *Importantly, because the evolutionary framework tracks only the mean and not the variability through time, this approximation is exact with respect to the first moment.* Accordingly, the only attribute that is approximate is the assumption of normality. Below we will show that a constant variance of σ^2 for the mixed trait distribution is a good approximation in most relevant cases, and becomes increasingly accurate when 1) generations become less overlapping (i.e. the case of semelparity), and 2) with density-dependent straying ratios.

The trait distribution at $t = 0$. At time $t = 0$, it is assumed that individuals from two separate populations have not yet mixed. Thus, the mean trait values for both populations are equal to the optimal trait values that accord to their respective habitats $\mu_i(t = 0) = \theta_i$, and that their trait values are distributed as $X_i = x_i \sim g(x_i, \mu_i, \sigma) = \mathcal{N}(\mu_i, \sigma)$. Here and henceforth we use upper-case to denote random variables, and lower-case to denote specific values of random variables.

The trait distribution at $t = 1$. At time $t = 1$, dispersal of individuals from site j to site i , and vice versa, has begun. The proportion of individuals that disperse to and from each population is given by the stray ratio m . The calculation of the mixed distribution that results from dispersal and subsequent recruitment incorporates the following processes: 1) dispersal of strays into the local population from the remote population, and vice versa; 2) recruitment of juveniles that are produced from this mixed distribution (under the assumption of random mating); 3) mortality of individuals in the adult class. These events occur in the order in which they are listed.

The mixed trait distribution for site i immediately following dispersal is simply a mixed Gaussian with weights determined by the proportion of the population that are residents i and the proportion that are strays j , such that

$$p(x_{\text{disperse}}) = \omega_i g(x_i, \mu_i, \sigma) + (1 - \omega_i) g(x_j, \mu_j, \sigma), \quad (27)$$

where the proportion of the population i composed of *resident* individuals is

$$\omega_i = \frac{(1 - m)N_i}{(1 - m)N_i + mN_j}, \quad (28)$$

and the proportion of population i composed of strays is $(1 - \omega_i)$. Once individuals in both habitats mix, they mate without preference. Recruits are assumed to have trait values intermediate to their parents such that $x_r = (x_1 + x_2)/2$ where trait values $X_1 = x_1 \sim p(x_{\text{disperse}})$ and $X_2 = x_2 \sim p(x_{\text{disperse}})$. We thus assume that parental traits are drawn independently from the mixed distribution $p(x_{\text{disperse}})$, conditional on the requirement that $x_2 = 2x_r - x_1$, which we get when rearranging the equation for x_r . Incorporating this condition and that traits are independently drawn from the mix of residents and strays yields the probability density for the trait value of recruits,

$$q(x_r) = \int_{-\infty}^{\infty} 2p(x_1)p(2x_r - x_1)dx_1. \quad (29)$$

The mean of the recruit trait distribution is calculated to be

$$E\{q(x_r)\} = \omega_i \mu_i + (1 - \omega_i) \mu_j, \quad (30)$$

and the variance is

$$\text{Var}\{q(x_r)\} = \frac{m^2 N_j^2 (4\mu_j^2 + \sigma^2) + (m - 1)^2 N_i^2 (4\mu_i^2 + \sigma^2) - (m - 1)mN_i N_j (\mu_i^2 + 6\mu_i \mu_j + \mu_j^2 + 2\sigma^2)}{2(-mN_i + mN_j + N_i)^2}. \quad (31)$$

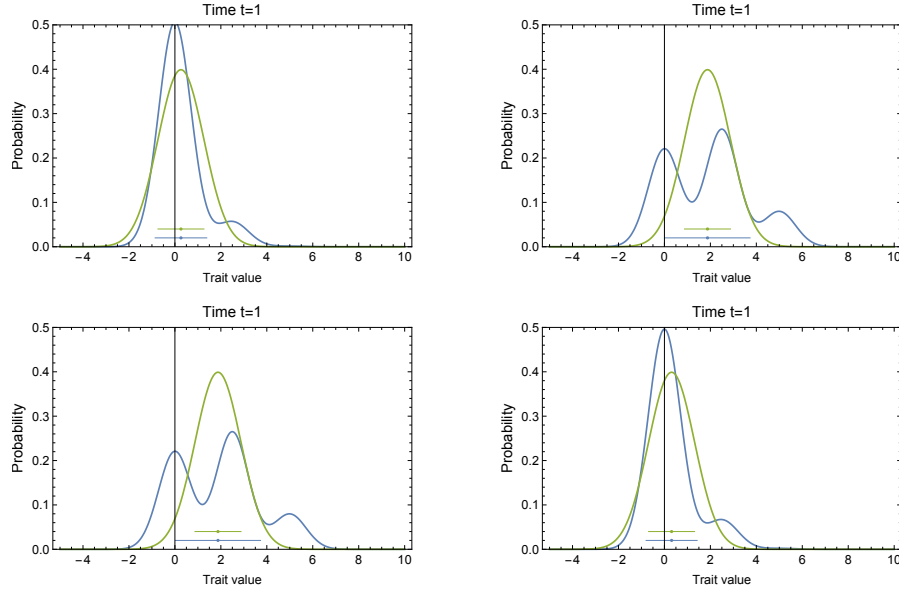


Figure S1: Clockwise from the upper-left. The calculated probability distribution function (PDF) for individuals in population i after immigration from population j , recruitment, and mortality (blue) vs. the Gaussian approximation used for the dynamical system (green), with a) low stray ratio, large straying population, overlapping generations: $m = 0.1$, $Z = 0.5$; b) maximal stray ratio, overlapping generations: $m = 0.5$, $Z = 0.5$; c) maximal stray ratio, non-overlapping generations: $m = 0.5$, $Z = 1000$; d) density-dependent stray ratio, non-overlapping generations: $m = m_0(1 - N_i/(C + N_i))$, $Z = 1000$, where $m_0 = 0.2$ and $C = 1000$. The mean and variability of both distributions is shown by a point (mean) and line (σ) at the base of each distribution.

The relative contribution of trait values from resident parents, stray parents, and recruits depends on the numbers of each within the mixed population. The number of residents is $(1 - m)N_i$, and the number of strays is mN_j . However, because adult mortality occurs after reproduction, we must adjust these numbers to account for only those who survive, which is given by e^{-Z} . As $Z \rightarrow \infty$, the population becomes semelparous, as the adult class gives way to new recruits at each timestep. Accordingly, the number of adult residents and strays is given by $(1 - m)N_i e^{-Z}$ and $mN_j e^{-Z}$, respectively. In contrast, the number of recruits is given by the combined contribution of resident and straying individuals to recruit production as

$$N_r = \int_{-\infty}^{\infty} r_{\max} e^{\frac{(x_r - \theta_i)^2}{2\tau^2}} q(x_r) dx_r ((1 - m)N_i + mN_j) e^{-\beta((1 - m)N_i + mN_j)}, \quad (32)$$

where θ_i is the optimal trait value in site i , r_{\max} is the maximum recruitment rate, and τ is the strength of selection.

The proportion of individuals from the resident, straying, and recruitment classes following mortality at site i are then c_i , c_j , and c_r respectively, where

$$\begin{aligned} c_i &= \frac{(1 - m)N_i e^{-Z}}{(1 - m)N_i e^{-Z} + mN_j e^{-Z} + N_r} \\ c_j &= \frac{mN_j e^{-Z}}{(1 - m)N_i e^{-Z} + mN_j e^{-Z} + N_r} \\ c_r &= \frac{N_r}{(1 - m)N_i e^{-Z} + mN_j e^{-Z} + N_r}. \end{aligned} \quad (33)$$

where $c_i + c_j + c_r = 1$. The trait value distribution for time $t = 1$ is then calculated as a mix of trait distributions

from 1) surviving residents, 2) surviving strays, and 3) their offspring, such that

$$g_{\text{next}}(x_i) = c_i g(x_i) + c_j g(x_j) + c_r q(x_r). \quad (34)$$

The expectation of this distribution is

$$\mathbb{E}\{g_{\text{next}}(x_i)\} = \omega_i \mu_i + (1 - \omega_i) \mu_j. \quad (35)$$

Importantly, $\mathbb{E}\{g_{\text{next}}(x_i)\}$ is exactly equal to the trait mean for the original mix, as well as the trait mean for recruits. Thus, a Gaussian approximation distributed as $\mathcal{N}(\omega_i \mu_i + (1 - \omega_i) \mu_j, \sigma)$ will not alter the dynamics under the assumption of constant variance in Eq 4. Moreover, the mixed distribution described above becomes more complex at the next timestep, as two multi-modal trait distributions then mix and recombine through recruitment. One could make the argument that a Gaussian approximation does not represent the multi-modality in trait values inherent to a mixed distribution, and thus underestimates variance. Although the Gaussian approximation exactly captures the expectation, we next explore to what extent the approximation does and does not capture the variability expressed in $g(x_{\text{next}})$.

The Gaussian approximation $\mathcal{N}(\omega_i \mu_i + (1 - \omega_i) \mu_j, \sigma)$ will generally underestimate the true variance of the mix, but not by much. If the straying ratio is low, (ca. 0.0 to 0.2), the Gaussian approximation is very similar to the actual distribution (Fig. S1a). As the straying ratio increases, so does the bimodality of the actual distribution, and this magnifies our underestimate of variance (Fig. S1b). The relevant dynamics observed in our model generally occur between $m = 0.05$ and $m = 0.2$, where estimates of variance are similar. Decreasing the overlap between generations also leads to greater similarity in the actual distribution and the Gaussian approximation, and the potential bias introduced by bimodality will be less for cases of non-overlapping generations, or semelparity (Fig. S1c). Most importantly, density-dependent straying increases the accuracy of our Gaussian approximation by actively suppressing the straying ratio (Fig. S1d).

Because the Gaussian approximation is exact to first order, and generally provides a good approximation of realized variance, we suggest that the use of Lande's formulation for the change in mean trait values through time under the assumption of constant variance is appropriate. Because our estimate of variance is generally an underestimate, the rate of recruitment will be a slight overestimate, and the timescale of trait evolution will generally be slower. Changes in the rate of recruitment (Fig. S14) and heritability do not have an impact on our qualitative findings, and this suggests any bias that our approximation introduces by underestimating true variance is minimal.

Under these assumptions, the governing equations for the dynamic system are then written as

$$\begin{aligned} N_i(t+1) = & \\ & ((1-m)N_i(t) + mN_j(t)) e^{-Z} \\ & + R_i[\omega_i \mu_i(t) + (1 - \omega_i) \mu_j(t), \theta_i] ((1-m)N_i + mN_j) \\ & \times e^{-\beta((1-m)N_i(t) + mN_j(t))}, \end{aligned} \quad (36)$$

where the trait distribution for recruits is also approximated by the Gaussian $\text{pr}(x_i) \approx \mathcal{N}(\omega_i \mu_i(t) + (1 - \omega_i) \mu_j(t), \sigma^2)$,

where the mean is the analytical expectation. This gives the rate of recruitment as

$$\begin{aligned}
 R_i[\omega_i\mu_i(t) + (1 - \omega_i)\mu_j(t), \theta_i] &= \\
 \int_{-\infty}^{\infty} r_{\max} \exp\left\{\frac{(x_i - \theta_i)^2}{2\tau^2}\right\} \text{pr}(x_i, \omega_i\mu_i(t) + (1 - \omega_i)\mu_j(t), \sigma^2) dx_i + \tilde{P}_i \\
 &= \frac{r_{\max}\tau}{\sqrt{\sigma^2 + \tau^2}} \exp\left\{-\frac{(\theta_i - (\omega_i\mu_i(t) + (1 - \omega_i)\mu_j(t)))^2}{2(\sigma^2 + \tau^2)}\right\} + \tilde{P}_i.
 \end{aligned} \tag{37}$$

The mean of the trait distribution then changes as

$$\begin{aligned}
 \mu_i(t+1) &= \omega_i\mu_i(t) + (1 - \omega_i)\mu_j(t) + h^2\sigma^2 \frac{\partial}{\partial \mu'} \ln(R_i[\mu', \theta_i]), \\
 &= \omega_i\mu_i(t) + (1 - \omega_i)\mu_j(t) + h^2\sigma^2 \left(\frac{\theta_i - \omega_i\mu_i - (1 - \omega_i)\mu_j}{\sigma^2 + \tau^2} \right),
 \end{aligned} \tag{38}$$

given $\mu' = \omega_i\mu_i(t) + (1 - \omega_i)\mu_j(t)$. Under the condition that $\sigma = 1$ and $\tau = 1$ as is set in the manuscript, this simplifies to

$$\mu_i(t+1) = \omega_i\mu_i(t) + (1 - \omega_i)\mu_j(t) + h^2 (\theta_i - \omega_i\mu_i - (1 - \omega_i)\mu_j). \tag{39}$$

112 D. S. Falconer, *Introduction to Quantitative Genetics* (Oliver and Boyd LTD, Edinburgh, 1967).

113 R. Lande, *Evolution* **30**, 314 (1976).

II. Additional Figures

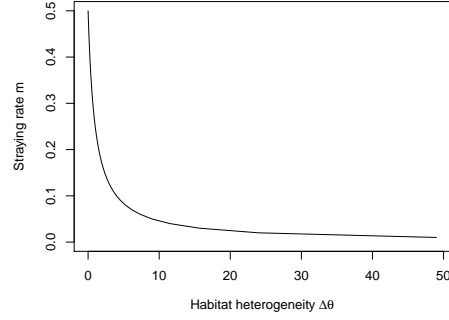


Figure S2: In some cases, habitat heterogeneity may be assumed to determine the rate of straying, if for example: 1) sites are distributed over greater spatial distances, where habitat differences are assumed to be greater between more distant sites, or 2) individuals have behaviors promoting dispersal between habitats that are more similar. To examine such cases, we use the relationship $m = 0.5(1 + \Delta\theta)^{-1}$ where maximum straying is assumed to occur at $m = 0.5$ (perfect mixing).

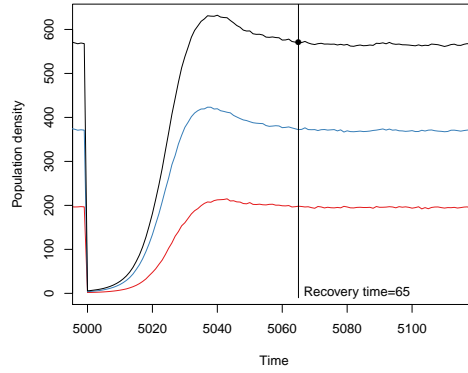


Figure S3: Example of the numerical procedure used to estimate recovery time. After a disturbance is introduced, the recovery time is calculated by measuring the point in time where N_T (in black), which is the aggregate of both populations (blue, red) settles to within one standard deviation of the new equilibrium N_T^* .

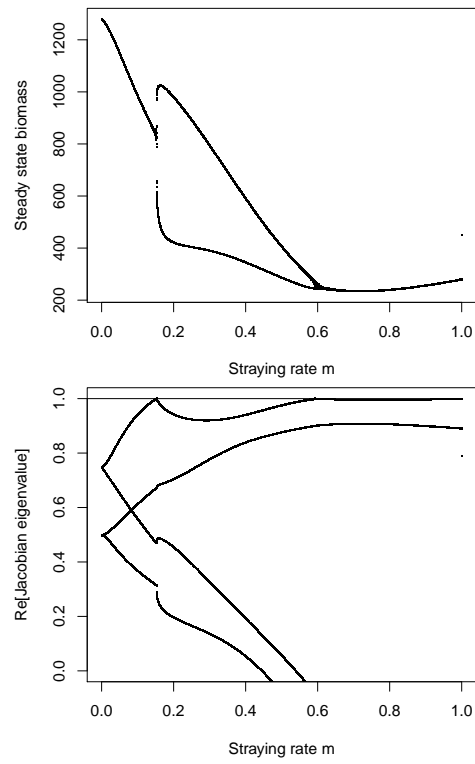


Figure S4: The real parts of the four eigenvalues for the Jacobian matrix of the 4-dimensional system. The cusp bifurcation occurs when the dominant eigenvalue crosses the unit circle at $+1$.

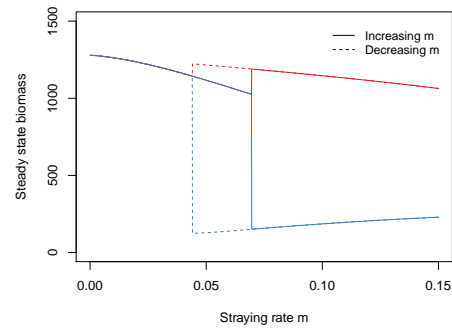


Figure S5: Increasing the straying rate results in the transition from a single steady-state for both populations to a dominant and subordinate states. If the straying rate is subsequently lowered, the single steady-state is not easily obtained, which is the hallmark of hysteresis.

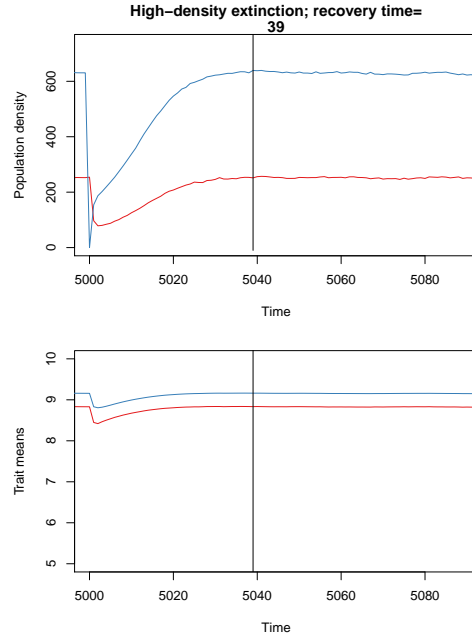


Figure S6: Extinction of high-density population with a high straying rate $m = 0.4$ and low trait heritability $h^2 = 0.2$ (see figure 3a). Black line marks the calculated point of recovery post-perturbation. Trait optima are $\theta_1 = 10$ (blue population trajectory) and $\theta_2 = 5$ (red population).

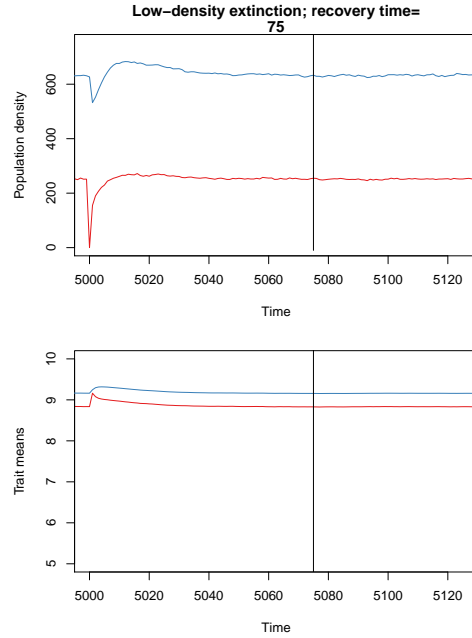


Figure S7: Extinction of low-density population with a high constant straying rate $m = 0.4$ and low trait heritability $h^2 = 0.2$ (see figure 3a). Black line marks the calculated point of recovery post-perturbation. Trait optima are $\theta_1 = 10$ (blue population trajectory) and $\theta_2 = 5$ (red population).

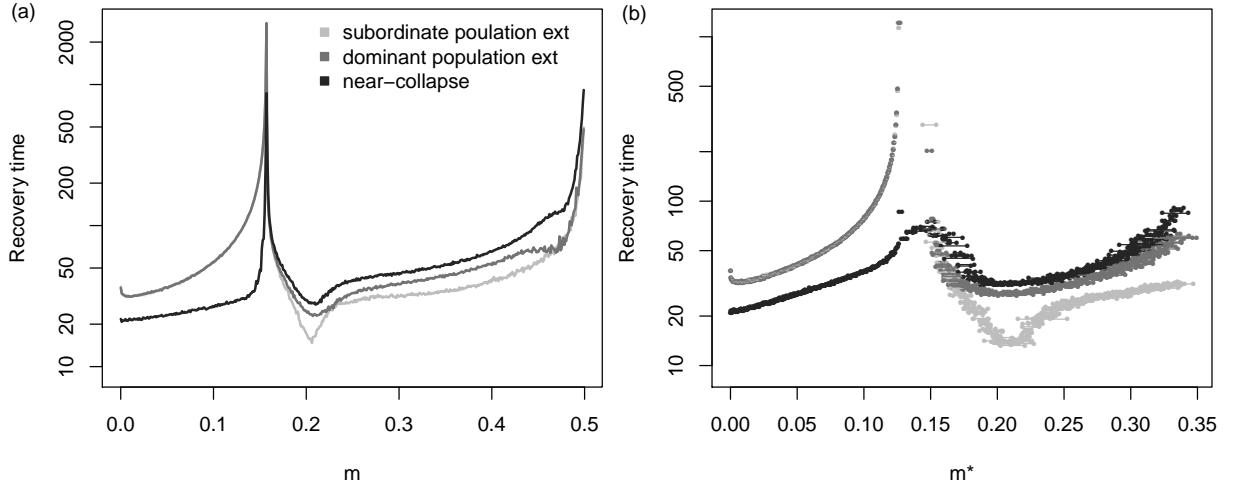


Figure S8: Recovery time of N_T following the extinction of either the low-density (light gray) or high-density (gray) population, or the near-collapse of both (dark gray) assuming (a) constant straying rates m and (b) density-dependent straying rates (evaluated at the steady-state m^*) with trait heritability $h^2 = 0.8$. If m is density-dependent, in the alternative stable state regime there are two straying rates observed: one each for the low- and high-density populations, respectively, which are linked by a horizontal line.

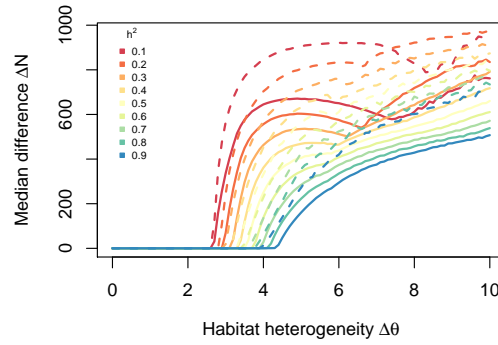


Figure S9: Median difference in population densities taken over the straying rate as a function of habitat heterogeneity $\Delta\theta$. Solid lines are for constant m ; dashed lines are for density-dependent m .

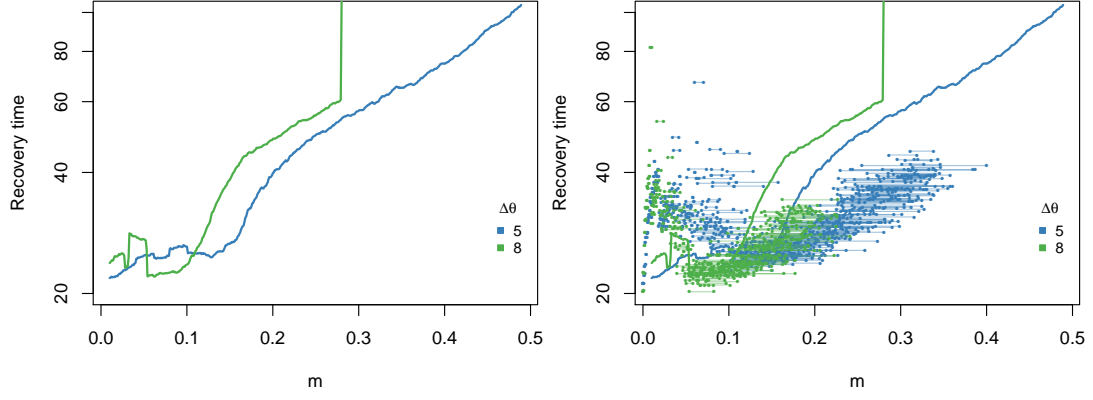


Figure S10: (a) Recovery time after near collapse of both populations as a function of straying rate m and habitat heterogeneity $\Delta\theta$. (b) The same as (a) but including recovery times when straying is density-dependent evaluated at m^* , shown by linked point pairs. Recovery times for systems with density-dependent straying are longer when straying is low and shorter when straying is high, mirroring the change in portfolio effects with respect to density-dependent straying shown in figure 5.

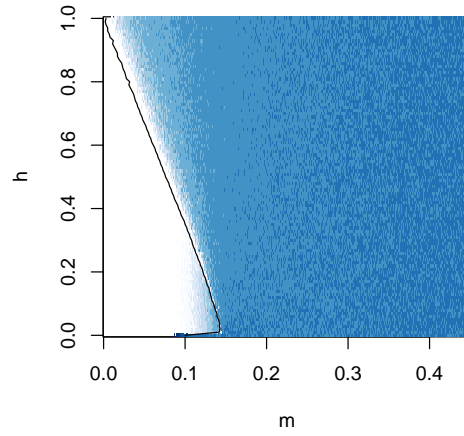


Figure S11: Portfolio effects as a function of straying rate m and trait heritability h^2 when the rate of straying is $m = 0.5(1 + \Delta\theta)^{-1}$. Alternative steady-states emerge for low values of m (left of the cusp bifurcation, denoted by the black line), whereas a single steady-state exists for high m .

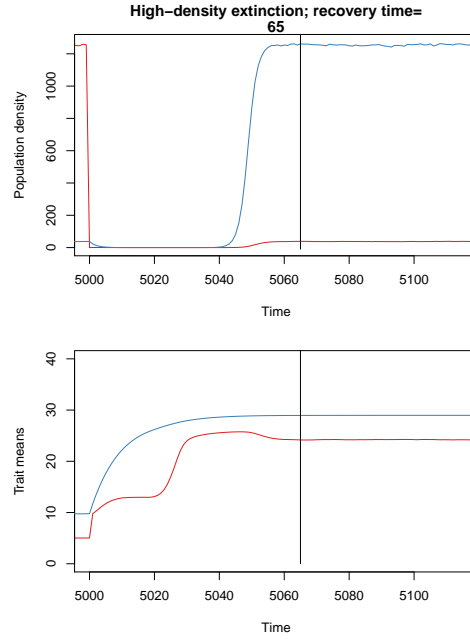


Figure S12: Population inversion where increased differences in trait optima between sites $\Delta\theta$ corresponds to lower rates of straying m . At low rates of straying $m = 0.02$ ($\Delta\theta = 24$), extinction of the dominant population leads to slower-than-expected recovery times because the subordinate population is isolated enough to evolve towards its own trait optimum. In this case, m is less than $m = 0.034$ (denoted by the asterisk in figure 6), such that isolation allows the subordinate population to ‘run away’ from the influence of the dominant population. This leads to a switch in subordinate/dominant states for the two populations. If m is low but greater than 0.034, isolation permits the subordinate population to ‘run away’ from the influence of the dominant population, until it is overwhelmed by the recovering dominant population, and reverts back to its previous trait mean prior to disturbance.

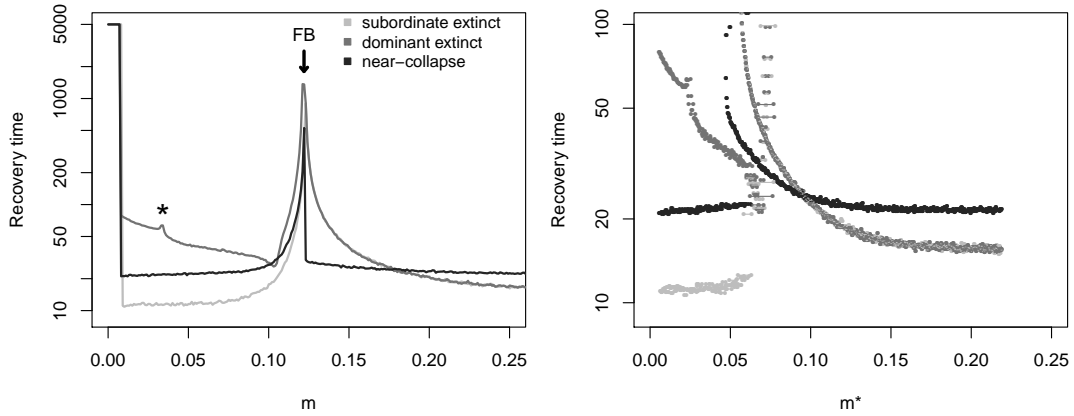


Figure S13: Recovery times for three disturbance types when the straying rate covaries with habitat heterogeneity as $m = 0.5(1 + \Delta\theta)^{-1}$ for constant (a) and density-dependent (b) straying rates. The cusp bifurcation is not as clear in (b) because $\Delta\theta$ is a function of the individual straying rate m_0 , whereas the x-axis in (b) is the straying rate at the steady-state m^* . Despite this difference, the general behavior shown in (a) are the same in (b).

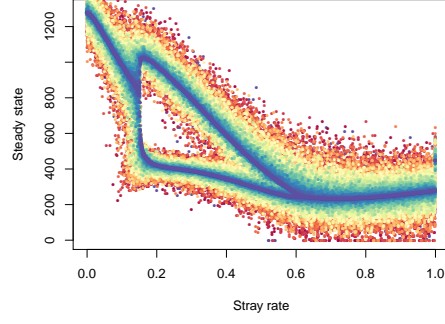


Figure S14: Steady-state densities of both populations as a function of m , where a cusp bifurcation indicates the emergence of alternative steady-states: one in a dominant state and one in a subordinate state. Steady-states for populations with symmetrical values ($\alpha = 0$) in the vital rates r_{\max} and β are shown with cool tones. As the asymmetry among populations between sites increases ($\alpha > 0$), their vital rates diverge, such that the maximal growth at sites 1 and 2 is now $r_{\max}(1) = r_{\max}(1 + \tilde{r}v_1)$ and $r_{\max}(2) = r_{\max}(1 + \tilde{r}v_2)$ where $\tilde{r}v_{1,2}$ are independently drawn from $\text{Normal}(0, \alpha)$ and $r_{\max} = 2$. Similarly the strength of density dependence is calculated at sites 1 and 2 as $\beta(1) = \beta(1 + \tilde{r}v_1)$ and $\beta(2) = \beta(1 + \tilde{r}v_2)$ where $\tilde{r}v_{1,2}$ are independently drawn from $\text{Normal}(0, \alpha)$ and $\beta = 0.001$. Steady-states for populations with increasingly asymmetric values ($\alpha \rightarrow 0.1$) for vital rates r_{\max} and β are shown in warmer tones.

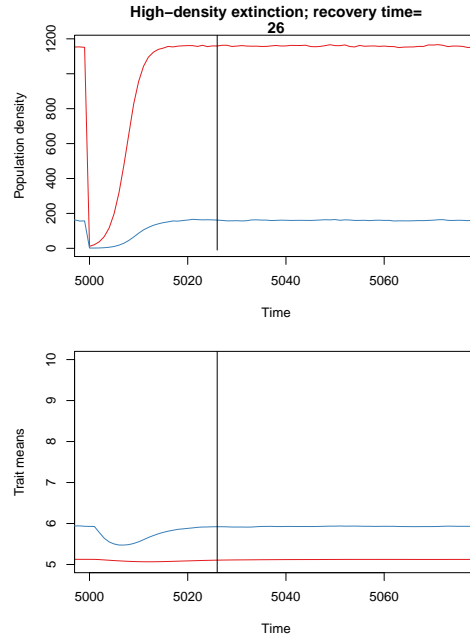


Figure S15: Near collapse of both populations with a low straying rate $m = 0.1$ and low trait heritability $h^2 = 0.2$ (see figure 3a). Black line marks the calculated point of recovery post-perturbation. Trait optima are $\theta_1 = 10$ (blue population trajectory) and $\theta_2 = 5$ (red population).