Title: Fine scale spatial variation in life time fitness

Authors: Shaun R. Coutts[1], Pedro Quintana-Ascencio[3], ..., Roberto Salguero-Gómez[1,2], Dylan Childs [1]

1. Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, UK.

2. Evolutionary Demography Laboratory. Max Planck Institute for Demographic Research. Rostock, DE-18057, Germany.

3. Florida State University

Introduction

The scale at which important aspects of population performance vary has been a foremost question in ecology since its beginnings [some review REF]. The scale at which we measure population performance has important ramifications for the inferences we can make from our observations [CRONE REF]. For example populations with low average lifetime reproductive output are at higher risk of extinction [REF]. However, that average performance may hide wide variation in performance between individuals or locations. If there are some spots or individuals that consistently perform much better than others then the overall extinction risk may be much lower than that indicated by the average population performance since the species may persist in the most favourable micro-sites [KENDALL, DEER REF, WHITE I THINK].

microsites and micro habitats, source sink at small scales driving over all population performance and local extinction risk.

R0 as an important measure of individual performance.

challenges of finding the right scale to measure at.

Even if the broad over-all environment changes to be less favourable for a species, that

species may still persist in favourable micro-habitats [CERGO REF]. Finding the relevant scale at which to measure the effect of the environment on population/individual performance is difficult because it may require the collection and analysis of data at a very fine spatial scale.

We use dataset of *hypericum* demography with individual locations measured to centimetre accuracy to test the spatial scale at which individual level performance varies.

Methods

Study site and species

For Pedro to fill in

Talk about patches of *hypericum* existing in gaps of a shrubby vegetation matrix and refer to map. Patches (or gaps, depending how you view it) will be important later when talking about scale.

demographic and location data

For Pedro to fill in.

Statistical analysis

R_0 depends on survival and seed production, seed production in turn is affected by plant size, and so growth indirectly affects R_0. We model these three components independently, this implicitly assumes survival, seed production and growth do not trade-off against each other. We model these three vital rates using a non-spatial site level model and a spatial errors model. If the non-spatial site level model explains as much of the variation in R_0

as the spatial errors model it suggests factors that impact the whole site uniformly (such as broad scale climate and time since fire) are the most important drivers of individual level fitness. If the spatial model explains the most variation in observed R_0 it suggests that environmental factors that vary within the site are important drivers of individual fitness. All models are fit using Hamiltonian MCMC in stan, using the 'rstan' interface [REF].

The three components of R_0 are modelled slightly differently since survival and probability of reproduction are binary variables and size is a continuous variable. For the binary variables we use a generalized linear modelling frame work

$$v_i \sim B(p_i) \tag{1a}$$

$$p_i = \frac{1}{1 + e^{\eta_i}} \tag{1b}$$

where v_i is the binary vital rate survival or reproduction for the i^{th} observation, and is assumed to be drawn from the Bernoulli distribution $B(\cdot)$. The linear predictor, η_i, for the site scale model includes no spatial effects and so only captures variation across the whole 200m by 500m site.

$$\eta_i = y_j + \beta_h h_i^t \tag{2}$$

The linear predictor for the i^{th} observed survival or flowering, η_i, has a year specific intercept, y_j, where each observation i is made in year $j \in \{2002, 2003, 2004, 2005, 2006, 2007\}$. We use a weak prior on the effect of year, with $y_j \sim N(0, 20)$. Year captures both year effects and a known effect of time since fire. The linear effect of the height of the i^{th} individual is β_h and h_i^t is the height of individual i at time t, where $t \in \{-1, 0\}$ is the height in the current year (for reproduction) or the previous year (for survival).

To capture variation in vital rates over space we include a spatial error term in the linear predictor. The spatial error term is computationally challenging to fit to the full set of spatial locations, q. We follow the approach of (Viana *et al.*, 2013) and fit the spatial error term to a smaller, more computationally tractable, set of knot locations, q^*, and

then interpolate the error term from those knot locations to each data point. However, at the resolutions required to capture the dominant spatial patterns in our data we could not use the interpolation process used by (Viana *et al.*, 2013), instead we used a much simpler nearest neighbour interpolation.

The first step in this process is to construct q^*, the set of knot locations. Our aim was to have the minimum number of knot locations, that also minimized the distance between every data point and its nearest knot (so that nearest neighbour interpolation is effective), but gave good coverage of the data with a small distance between knots. We employed a simple heuristic approach. First we set every unique observation location as a knot location. Then, in turn, for each knot location we took the average coordinate across all neighbouring knot locations within 50cm of the target knot and add it to q^* as a new aggregate knot. We then removed all the original knot locations from the set q^*. This resulted in 575 knot locations in q^*, with good coverage of the data (Figure 1a). 75% of knots had another knot within 82cm and 50% had at least one other knot within 61cm, giving us sub 1m resolution. All observations had a knot location within 56cm, and 75% have a knot within 25cm of their location.

The linear predictor for the spatial model is

$$\eta_i = y_j + \beta_h h_i^t + G(q_i) \tag{3}$$

where

$$G(q_i) = \overline{G}(q_i^*) \tag{4}$$

is the spatial error term for the i^{th} observation and q_i^* is the nearest knot location to observation location q_i (simple nearest neighbour interpolation). The spatial error term can be thought of as a random effect that varies continuously over space. This effect is assumed to be drawn from a multivariate-normal distribution

$$\overline{G}(q^*) \sim MVN(\mathbf{0}, \Omega(q^*; \alpha_s, \sigma_s^2)). \tag{5}$$

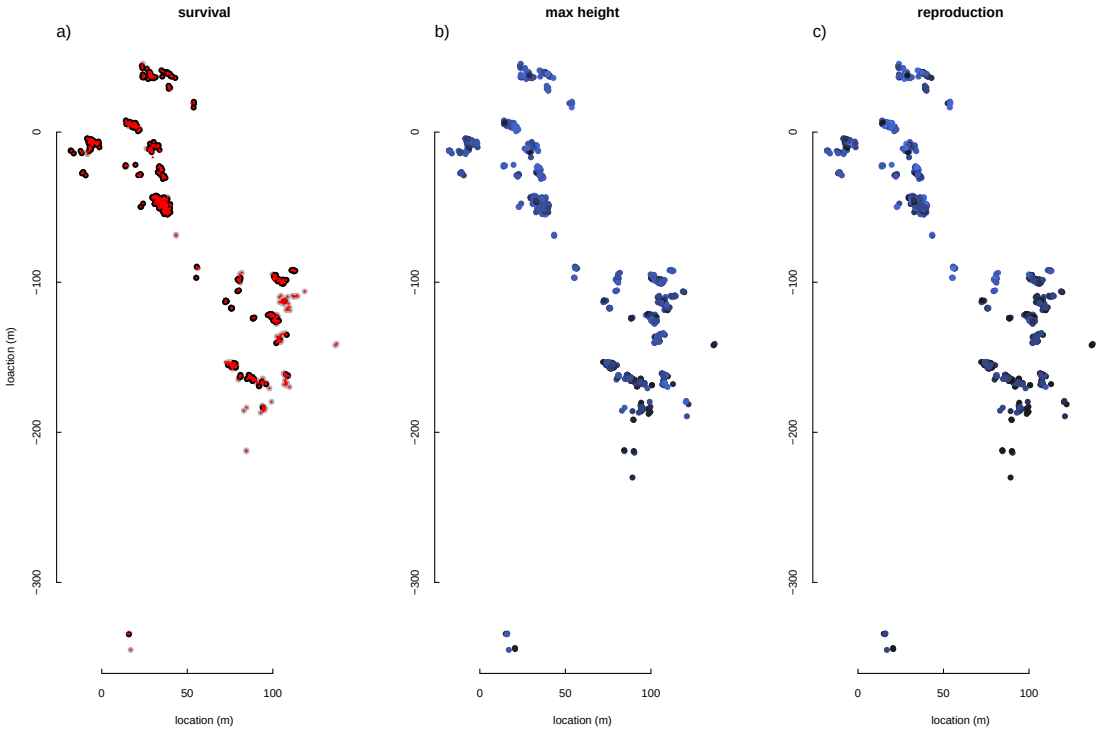

Figure 1: Each point shows the location of an observation for the vital rates survival (a), maximum height reached, a proxy for growth (b) and reproduction (c). In (a) black points show locations where mortality was observed, grey indicates the individual tag was lost or the study ended before the individual died. Red crosses show the knot locations q^*. Note that most knot locations sit directly over an observation, on in the center of a dense cluster. In (b) and (c) lighter purples show greater max height and number of reproductive events, respectively.

We call the total number of observations (and thus the length of q) Q, and the total number of knots Q^*. $\mathbf{0}$ is a vector of 0's Q^* elements long, setting the mean of each draw to 0. The covariance matrix between knot locations, q^*, is $\Omega(q^*; \alpha_s, \sigma_s^2)$. The covariance is modelled as a exponentially decreasing function of distance so that the covariance at the mth row and kth column is

$$\omega_{m,k} = \sigma_s^2 \exp(-\alpha_s d(m,k)) \tag{6}$$

where $d(m,k)$ is the Euclidean distance between knot locations m and k. σ_s^2 is the variance, drawn from a weakly informative Cauchy prior, $\sigma_s^2 \sim Cauchy(0,5)$, truncated at 0.00001 (positive part used). The rate that covariance between knot points decays with distance is controlled by $\alpha_s = 1/\alpha^*$, where α^* is sampled from the weak prior

$\alpha^* \sim Cauchy(0, 5)$, truncated at 0.1 and 1000. We truncate these priors just above 0 to improve numerical stability of the sampling. The lower limit of truncation results in $\alpha_s = 10$, a distance decay rate where points just 30cm apart are independent. Since this distance is smaller than the resolution of our knot points it is pointless to sample lower. The upper truncation limit ($\alpha_s = 0.001$) implies that all knot locations co-vary, i.e. a site level effect, and so there is little point in sampling higher.

We model height using a slightly different model as it is a continuous, rather than binary variable. We assume height is drawn from a normal distribution

$$h_{i,j} \sim N(\mu_{i,j}, \sigma) \tag{7}$$

where σ is the standard deviation on the error term and drawn from an uninformative prior, and for the site level model

$$\mu_{i,j} = y_j + \beta_j h_{i,j-1} \tag{8}$$

with $\mu_{i,j}$ the mean predicted height for individual i in year j, y_j a year specific intercept, β_j a year specific linear effect of the previous height of an individual (i.e. its growth rate) and $h_{i,j-1}$ is the height of individual i in the previous year. Both y_j and β_j are drawn from an uninformative prior for the height model.

We include the spatial effect for growth in an similar manner to the survival and reproduction model, with

$$\mu_{i,j} = y_j + \beta_j h_{i,j-1} + G(q_i) \tag{9}$$

where $G(q_i)$ is given in Eq. 4.

The number of fruiting bodies (assumed to be proportional to seed production) produced by each individual was not observed at the study site during the study period. Instead we use data from 15 sites (including the study site), located ???? in a similar ecological system (predro to fill in), across 9 years (1994 – 2003) were used to estimate the number

of fruiting bodies based on plant height, with a separate intercept chosen for each study site (common slope on height estimated across sites).

$$\ln(f_i + 1) \sim N(\mu_i^f, \sigma_f) \tag{10a}$$

$$\mu_i^f = \beta_0^g + \beta_f h_i \tag{10b}$$

where f_i is the predicted number of seeds produced by individual i, and σ_f is the variance around that prediction, β_f is the effect of height (h_i) on seed number and β_0^g is the intercept for site g. These parameters were estimated on the log scale to improve the normality and homoscedasticity of the model residuals, hence the the prediction f_i is the anti-log.

Individual level performance and life history strategy

Even if the vital rates show strong spatial structure, if those vital rates tend to show negative co-variance with each other integrated measures of demographic performance may show much less spatial structure. Alternately vital rates might show positive covariance (i.e. and location that is good for growth is also good for survival and reproduction) then variation in integrated measures of demographic performance may be even more strongly spatially structured than the underlying rates.

One of the most important measures of individual demographic performance is reproductive output over time. To measure this we calculate R_{j0}^J, the number of fruiting bodies an individual is expected to produce between years $j0$ and J. We assume R_{j0}^J is proportional to the number of seeds produced by an individual. R_{j0}^J is similar to a cohort specific R_0, but since survival and growth in particular are strongly influenced by time since fire we do not extrapolate reproductive performance outside the observation period of the data. We put the models Eq. 1, 7 and 10 into an integral projection modelling framework to calculate per-individual seed productivity over the study period (R_{j0}^J). Because we

calculate R^J_{j0} at an individual level (as opposed to the population level measure R_0), distributions over size are interpreted as uncertainty around the size of an individual, rather than variation in size within a population.

We start by defining the frequency distribution of height z of individual i in year j, at location q_i as $n(z, j, q_i)$. Because we are interested in the performance of individuals the distribution over height represents our uncertainty over how much an individual will grow in a given year, rather than variation within a population, as is more usual. This means that in the first year $n(z, j, q_i)$ is a probability density that should integrate to 1 since we know an individual has to have some size. In addition at initial year j_0 this distribution will be a single value (i.e. a distribution with 0 variance), since we choose a height to start tracking an individual at. In order to propagate uncertainty in growth through time we need to integrate over size and incorporate size and year dependent survival, so that

$$n(Z, j+1, q_i) = \int_Z n(z, j, q_i) S(z, j, q_i) K(z, j, q_i) \mathrm{d}z \tag{11}$$

where $n(Z, j+1, q_i)$ is the frequency distribution of height over the whole domain of height Z, $S(z, j, q_i)$ is the survival of an individual of size z in year j at location q_i, and is adapted from Eq. 1 to be

$$S(z, j, q_i) = \frac{1}{1 + \exp(-\eta_s(z, j, q_i))} \tag{12a}$$

$$\eta_s(z, j, q_i) = y^s_j + \beta^s_h z + G_s(q_i) \tag{12b}$$

where y^s_j, β^s_h and $G_s(q_i)$ are estimated in Eq's 1 and 3, when the model Eq. 1 is fit to survival data. The growth kernel $K(z, j, q_i)$ is adapted from Eq. 7, and is

$$K(z, j, q_i) \sim N(\mu_k(z, j, q_i), \sigma_k) \tag{13a}$$

$$\mu_k(z, j, q_i) = y^k_j + \beta_j z + G_k(q_i) \tag{13b}$$

where y^k_j, β_j, σ_k and $G_k(q_i)$ are estimated by fitting Eq. 7, to height data.

Note that for each height z in domain Z, $K(z, j, q_i)$ is a normal distribution ($N(\cdot)$). As a result the integration in Eq. 11 is adding the normal distributions $K(z, j, q_i)$ produced by each value of z, which means that the variance of $n(z, j, q_i)$ will increase as we project it forward in time. That is, the further forward we project the fate of individual $n(z, j, q_i)$, the less certain we are about that fate. There is also another important aspect of uncertainty in the forward projection of $n(z, j, q_i)$, which is that there is uncertainty around the estimates y_j^s, β_h^s, y_j^k, β_j, σ_k, $G_k(q_i)$ and $G_s(q_i)$. We detail how we incorporate that uncertainty into the explanation of how we use Eq. 11 to calculate life expectancy (L_E), age of first reproduction (R_a), expected number of lifetime reproductive events ($E[R_n]$) and expected life time seed production ($E[R_0]$).

One aspect of life history we are interested in is how long an individual can be expected to live given they start life in initial year j_0. The probability that an individual we start tracking in year j_0, location q_i, will survive until year j is

$$\Pr(S|j, j_0, q_i) = \int_Z n(z, j, q_i) \mathrm{d}z \tag{14}$$

We calculate $n(Z, j, q_i)$ in each year by choosing location q_i and year j_0 to start tracking an individual, then we recursively calculate Eq. 11. We define the life expectancy of an individual as their age in the first year in which there is a less than 5% chance that the individual is still alive, thus,

$$L_E = j - j_0 \text{ such that } \Pr(S|j, j_0, q_i) \leq 0.05 \tag{15}$$

To incorporate the uncertainty in estimated growth rate and survival we repeatedly draw an instance of parameter set $\theta = \{y_j^s, \beta_h^s, y_j^k, \beta_j, \sigma_k, G_k(q_i), G_s(q_i)\}$ from the posterior distributions of these parameters obtained from their fitting. A value of L_E is calculated for each draw of θ to build a distribution of L_E that reflects both the uncertainty around growth rate, and uncertainty in parameter estimation.

Age of fist reproduction incorporates probability of flowering into Eq. 14. The probability

that an individual that we begin tracking in year j_0 flowers in year j depends on the growth of that individual and the uncertainty in that growth. Thus, the probability of flowering at least once before year j^*, after starting life in year j_0 is

$$\Pr(R|j^*, j_0, q_i) = 1 - \prod_{j=j_0}^{j^*} (1 - \Pr(R|j, q_i)) \tag{16}$$

We define the age of first reproduction (R_a) as the age by which there is a 95% chance an individual will have flowered,

$$R_a = j^* - j_0, \text{ where } j^* \text{ is the smallest value such that } \Pr(R|j^*, j_0, q_i) \geq 0.95 \tag{17}$$

The probability that an individual in location q_i will flower in year j is

$$\Pr(R|j, q_i) = \int_Z \frac{n(z, j, q_i) R(z, j, q_i) \mathrm{d}z}{\Pr(S|j, j_0, q_i)} \tag{18}$$

where the probability of survival until year j, location q_i ($\Pr(S|j, j_0, q_i)$, Eq. 14) acts as a normalizing constant so that the distribution $n(Z, j, q_i)$ integrates to 1 and can be treated as a probability density. We obtain $n(Z, j, q_i)$ by starting with an individual in year j_0, location q_i and recursively calculating Eq. 11. The probability that an individual of size z will flower in year j, location q_i is adapted from Eq. 1 to be

$$R(z, j, q_i) = \frac{1}{1 + \exp(-\eta_r(z, j, q_i))} \tag{19a}$$

$$\eta_r(z, j, q_i) = y_j^r + \beta_h^r z + G_r(q_i) \tag{19b}$$

where y_j^r, β_h^r and $G_r(q_i)$ are estimated in Eq's 1 and 3, when the model Eq. 1 is fit to reproduction data. Parameter uncertainty in estimated age of fist reproduction is incorporated by drawing samples from the posterior distributions of parameters $\theta = \{y_j^s, \beta_h^s, y_j^k, \beta_j, \sigma_k, G_k(q_i), G_s(q_i), y_j^r, \beta_h^r, G_r(q_i)\}$, recalculating R_a for each draw to build a distribution of R_a. Note that parameter set θ includes parameters for growth, survival and probability of flowering, since all these vital rates contribute to R_a.

The expected number of reproductive events for an individual over the period j_0 to J is

$$R_n = \Pr(R|j_0, q_i) + \sum_{j=j_0+1}^{j=J} \Pr(S|j, j_0, q_i)\Pr(R|j, q_i) \tag{20}$$

The final year we have observations for is 2007, thus we set $J = 2007$. $\Pr(R|j, q_i)$ is defined in Eq. 18. Recall that $\Pr(R|j, q_i)$ uses $\Pr(S|j, j_0, q_i)$ (Eq. 14) as a normalizing constant, thus Eq 20 can be written more simply as

$$R_n = \sum_{j=j_0}^{j=J} \int_Z n(z, j, q_i)R(z, j, q_i)\mathrm{d}z \tag{21}$$

We used the same posterior sampling process to build a distribution of R_n as outlined in Eq. 25.

Finally we included the estimated number of seeds produced by an individual of height z into Eq. 21 to get the expected number of seeds produced over time period j_0 to J

$$R_0 = \sum_{j=j_0}^{j=J} \int_Z n(z, j, q_i)R(z, j, q_i)f(z)\mathrm{d}z \tag{22}$$

where $f(z)$ is adapted from Eq. 10

$$f(z) \sim \exp[N(\mu_f(z), \sigma_f)] - 1 \tag{23a}$$

$$\mu_f(z) = \beta_0^f + \beta_h^f z + \delta_{g^*}^f \tag{23b}$$

Where $\delta_{g^*}^f$ is the random intercept for site g^* where the vital rates were mapped and measured.β_0^f, β_h^f, $\delta_{g^*}^f$ and σ_f are estimated in Eq. 10. Note that $f(z)$ is a distribution, and so the integration in Eq. 22 adds these distributions together across all values for z, propagating our uncertainty in $f(z)$ to our uncertainty around R_0.

ALTERNATE STRATEGY Even if the vital rates show strong spatial structure, if those vital rates tend to show negative co-variance with each other integrated measures of demographic performance may show much less spatial structure. Alternately vital rates might

show positive covariance (i.e. and location that is good for growth is also good for survival and reproduction) then variation in integrated measures of demographic performance may be even more strongly spatially structured than the underlying rates.

One of the most important measures of individual demographic performance is reproductive output over time. To measure this we calculate R^J_{j0}, the number of fruiting bodies an individual is expected to produce between years $j0$ and J. We assume R^J_{j0} is proportional to the number of seeds produced by an individual. R^J_{j0} is similar to a cohort specific R_0, but since survival and growth in particular are strongly influenced by time since fire we do not want to extrapolate reproductive performance outside the observation period of the data. Also, because we calculate R^J_{j0} at an individual level (as opposed to the population level measure R_0), distributions over size are interpreted as uncertainty around the size of an individual, rather than variation in size within a population.

R^J_{j0} incorporates size, survival and probability of flowering, since fruit production is conditional on all three processes. We begin by calculating fruit production in the first year $j0$. We assume individual i is alive in year $j0$ and that we are certain about its height. Thus, fruit production in the first year is

$$R(z_0, j0, q_i)f(z_0) \tag{24}$$

where the probability of flowering given height z, year j and location q_i is

$$R(z, j, q_i) = \frac{1}{1 + \exp(-\eta_r(z, j, q_i))} \tag{25a}$$

$$\eta_r(z, j, q_i) = y^r_j + \beta^r_h z + G_r(q_i) \tag{25b}$$

where y^r_j, β^r_h and $G_r(q_i)$ are estimated in Eq's 1 and 3, when the model Eq. 1 is fit to reproduction data. Number of fruits per plant of height z is

$$f(z) \sim lnN(\mu_f, \sigma_f) \tag{26a}$$

$$\mu_f(z) = \beta_0^{f*\text{“}} + \beta_h^f h_i \tag{26b}$$

where β_0^{f*} is the intercept for site g^* where the vital rates were mapped and measured, β_h^f is the effect of height on fruit number and and σ_f is the variance of the underlying exponentiated normal distribution estimated in Eq. 10.

To get the fruit production in subsequent years survival and growth need to be incorporated, so that

$$R_{j0}^J = R(z_0, j0, q_i)f(z_0) + \sum_{j=j0+1}^{J} K(z', j, q_i)S(z, j, q_i)R(z, j, q_i)f(z) \tag{27}$$

where the growth kernel $K(z', j, q_i)$ gives the height on an individual in year j given its location q_i and previous height z'. $K(z, j, q_i)$ is adapted from Eq. 7, and is

$$K(z, j, q_i) \sim N(\mu_k(z, j, q_i), \sigma_k) \tag{28a}$$

$$\mu_k(z, j, q_i) = y_j^k + \beta_j^k z + G_k(q_i) \tag{28b}$$

where y_j^k, β_j, σ_k and $G_k(q_i)$ are estimated by fitting Eq. 7, to height data. The probability that an individual at location q_i with height z in year j survives is

$$S(z, j, q_i) = \frac{1}{1 + \exp(-\eta_s(z, j, q_i))} \tag{29a}$$

$$\eta_s(z, j, q_i) = y_j^s + \beta_h^s z + G_s(q_i) \tag{29b}$$

where y_j^s, β_h^s and $G_s(q_i)$ are estimated in Eq's 1 and 3, when the model Eq. 1 is fit to survival data.

There are multiple levels of uncertainty in our estimation of R_{j0}^J. Both $f(z)$ and $K(z', j, q_i)$ are randomly distributed variables, where the distribution represents our uncertainty over how growth and survival respond to height, year and location. There is also uncertainty around the fitted parameter values used to estimate $R(z, j, q_i)$, $S(z, jq_i)$, $f(z)$ and $K(z, j, q_i)$.

In order to propagate both types of uncertainty into R^J_{j0} we use a simulation process. We estimate total fruit production for the cohort starting in 2002 until 2007 (R^{07}_{02}). There was a major fire at the site in 2001 killing most individuals. By starting our observation period in 2002 we capture the effect of important post fire changes in demography. To include uncertainty in parameter estimates we draw a parameter set

$$\theta = \{y^s_j, \beta^s_h, G_s(q_i), y^k_j, \beta^k_j, \sigma_k, G_k(q_i), y^r_j, \beta^r_h, G_r(q_i), \beta^f_0, \beta^f_h\} \tag{30}$$

from the posterior distribution of each parameter. For each individual in 2002 we know their height and the year, giving us z_0 and $j0$. Then to simulate the fate of that individual through time and determine the number of fruit it produces we calculate $R^J_{j0}|\theta$ by drawing random variates from the distributions $f(z)$ (lognormal), $Bernoulli(R(z, j, q_i))$, $Bernoulli(S(z, j, q_i))$, and $K(z', j, q_i)$ (normal distribution) under parameter set θ. Note that aside from the first year z' will be a randomly drawn height from the previous year. We repeat this process ?1000 times, drawing a new parameter set θ every time. Thus, R^J_{j0} for each location will be a distribution that incorporates our uncertainty in growth rate, fruit production and parameter uncertainty.

Testing for spatial structure in measures of individual performance and life history strategy

We are interested in whether or not R^J_{j0}, which integrates different vital rates, show a similar spatial structure to the vital rates themselves. Complicating this task is that R^J_{j0} is a distribution that represents our uncertainty in both vital rates and the parameters used to estimate those vital rates. The simplest way to approach this is to simply ignore that uncertainty and only use the modes of those distributions. An alternative is to compare the distributions of demographic measures built using the spatially structured models to those built in the same way, but with the non-spatial site level models, where the spatial term $G(q_i)$ is dropped from all vital rate models. We can then compare the cumulative

density function of each of these distributions to see how far the spatial distribution lies to the left or right of the non-spatial model so that

$$\Delta M = \int_m M \mathrm{d}m - \int_m M' \mathrm{d}m \tag{31}$$

Where ΔM is the difference between the CDF of spatial demographic measure (M) and the non-spatial demographic measure (M').

Results

Year (and by proxy time since fire) effected both survival and growth. Following a fire in 2001 survival decreased until 2006, while the intercept for 2006 and 2007 are the same (y_j in Figure 2, survival). Growth showed a similar pattern, with growth rate (β_j) for 2003 (and to a lesser extent 2004) being much higher than estimated growth rate in subsequent years (gr_j in Figure 2, growth). [ASIDE: these effects look a lot like Bethan's time since fire splines for a different species in same region/system]. Year had much less effect on reproduction, with no large differences in intercept for probability of flowering between years (y_j in Figure 2, reproduction).

The variance of the spatial effect ('spp_sd' [NOTE I need to fix the names in the figure]) suggests that sub-site level spatial variation in vital rates is of a similar magnitude to the year effects (Figure 2). The scale of the Gaussian predictive process was similar for all three vital rates, with correlation between knots falling to 0.5 by roughly 5m (Figure 3g–i). This suggests that similar, small scale, environmental proccesses are influencing survival, reproduction and growth. This scale of spatial correlation indicates that these vital rates can vary within patches as much as between them (Figure 3d–f). However, the maps of the magnitude and direction of the Gaussian predictive process (Figure 3a–f) show that while each vital rates varies over a similar spatial scale, they do so in different ways. For example take the patch at the bottom right of Figure 3d–f. For survival (Figure

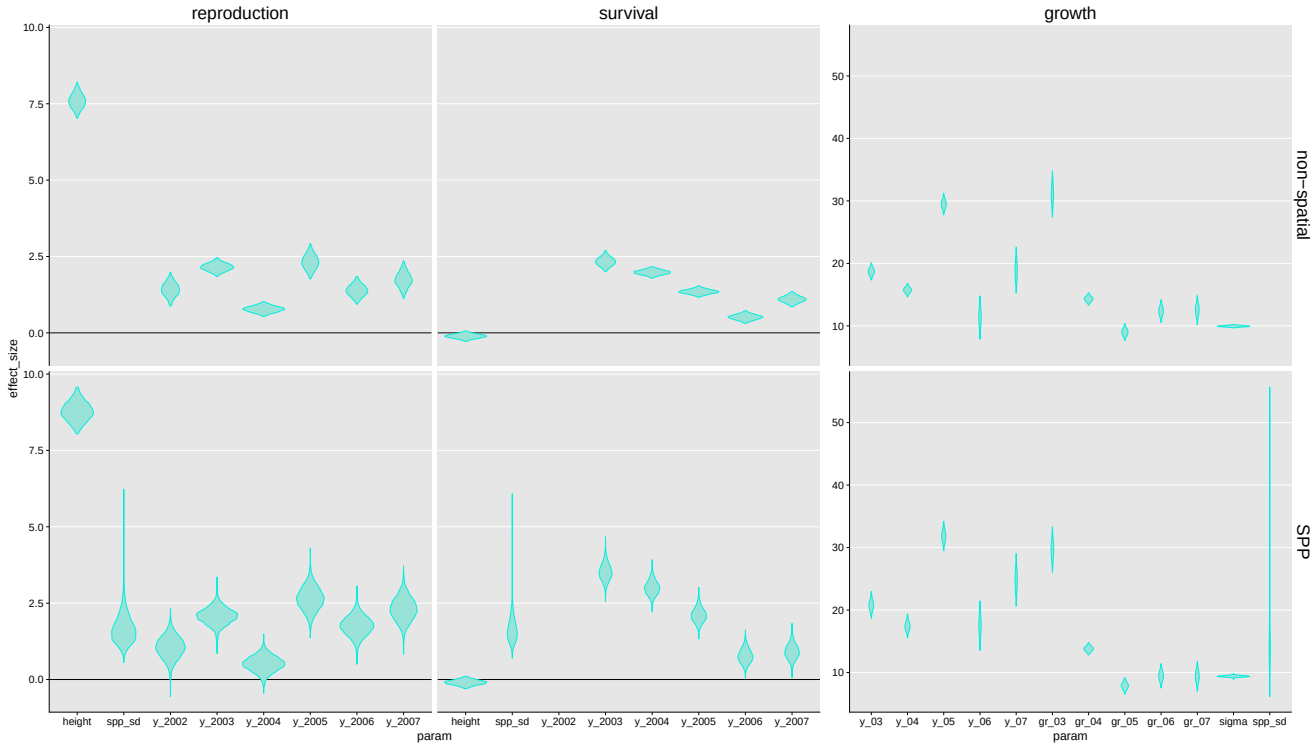

Figure 2: Effect sizes of year (y_j) and height (height and 'gr_year') on survival, probability of flowering and height. Note that the slope of height is multiplied by the mean height to put it on the same scale as the intercepts, and can be interpreted as the effect of height on an average sized individual. 'spp_sd' indicates how large the spatial effect is and sigma is the overall error term in the height model. Note effect sizes for survival and reproduction are on the logit scale and on the arithmetic scale for growth.

3d) this whole patch is red-ish, meaning it is worse than average for survival (after year and height effects are controlled for), with some parts of the patch being much worse (deep red) and some parts only slightly worse (pale red to white). For reproduction (Figure 3e) this whole patch has a higher than average probability of flowering (blues), again with some parts of the patch being better than others (deeper vs paler blues). For growth this patch has a more complex structure (Figure 3f). One end is worse than average for growth (again after controlling for year and height effects), while the other has higher than average growth rates (blue to red gradient). Similar patterns can be seen in several other patches.

This raises the possibility that even though there is spatial structure in individual vital rates, these differences in could cancel out so that R_0 has less spatial structure than any single vital rate. The alternative is that these vital rates re-enforce each other so that

the spatial structure becomes more extreme. Will have to test this by simulating R_0 for different plants under the non-spatial and spatial models to generate distributions of R_0 for each plant. Then map the overlap between spatial and non-spatial distributions. If spatial effects cancel then would expect high over lap across all sites, if they re-enforce should see patches where the R_0 is much higher and much lower for the spatial vs non-spatial distributions. (I need to think and this a bit more).

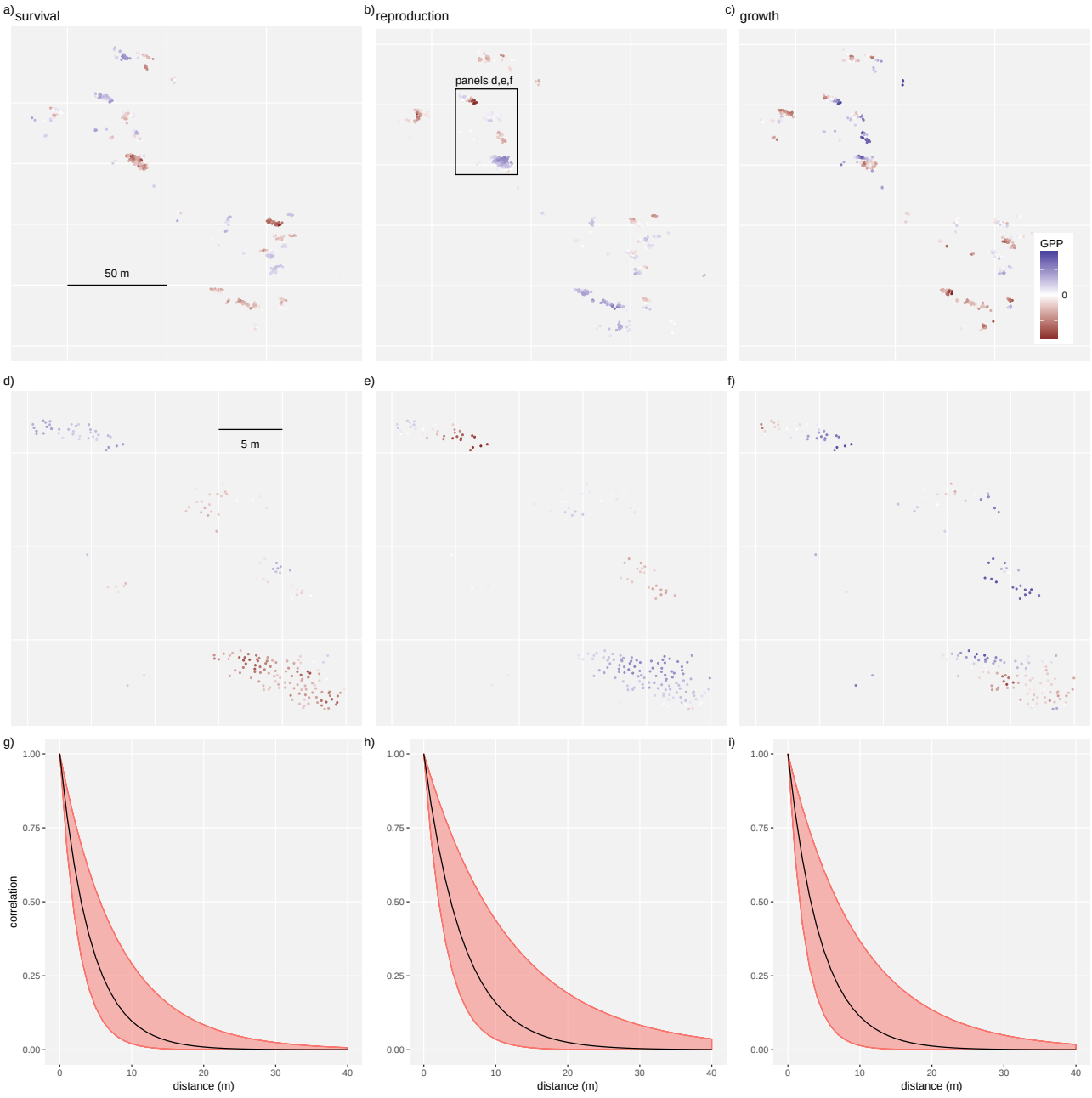

Figure 3: Gaussian predictive process (GPP, i.e. the spatial error term) for each knot point in q^* for each vital rate for the whole site (a–c) and for a smaller section of the site to reveal within patch structure (d–f). Spatial correlation curves (g–i) show the scale at which knots co-vary.

References

Viana, M., Jackson, A.L., Graham, N. & Parnell, A.C. (2013) Disentangling spatio-temporal processes in a hierarchical system: a case study in fisheries discards. *Ecography*, **36**(5), 569–578.