Our problem

We write per capita reproductive success as the product of seedling survival to fruiting, fruits per plant, and seeds per fruit (all per capita):

$$\text{per capita RS} = \sigma F \phi \tag{1}$$

We want to understand the contribution of variance in each component of reproductive success to the overall variance in reproductive success.

Background

Venable 2007 calculates the "risk associated with germination. . . as the geometric standard deviation of per capita reproductive success." In that paper, the geometric standard deviation of per capita reproductive success is expressed as:

$$\text{geometric SD(per capita RS)} = \exp\big(SD(\ln(\text{per capita RS}))\big) \tag{2}$$

That's what I've used to calculate the standard deviation in per capita reproductive success so far. When I was thinking and reading about how to decompose the variance of per capita reproductive success, I learned that this expression for the geometric standard deviation is derived from a result involving the log-normal distribution (Kirkwood 1979). I had to learn a bit more about this to figure out how/whether I could decompose the variance for this. In the rest of this, I give the geometric mean, geometric variance, and geometric standard deviation (de Carvalho 2016). I then write out what I think should be the geometric covariance, but note that I have not found a good reference for the expression.

Geometric mean

The geometric mean μ_g of $x_1, x_2, \ldots, x_n$ is:

$$\mu_g = (x_1 x_2 \ldots x_n)^{1/n}$$
$$\ln(\mu_g) = \frac{1}{n} \ln(x_1 x_2 \ldots x_n) \tag{3}$$

We can rewrite this to show that the log of the geometric mean is equal to the arithmetic mean of the logs:

$$\ln(\mu_g) = \frac{1}{n} \big(\ln(x_1) + \ln(x_2) + \cdots + \ln(x_n) \big)$$
$$\ln(\mu_g) = \frac{1}{n} \sum_{i=i}^{n} \ln(x_i) \tag{4}$$
$$\mu_g = \exp \left(\frac{1}{n} \sum_{i=i}^{n} \ln(x_i) \right)$$

Geometric variance

The log of the geometric variance is equal to the arithmetic variance of the logs of $x_1, x_2, \ldots, x_n$:

$$\ln(\sigma_g^2) = \frac{1}{n-1} \sum_{i=i}^{n} \big(\ln x_i - \ln \mu_g \big)^2 \tag{5}$$

We use the geometric mean μ_g to calculate the geometric variance:

$$\text{geometric Var}(\boldsymbol{x}) = \exp \frac{\sum_{i=i}^{n} \big(\ln x_i - \ln \mu_g \big)^2}{n-1}$$
$$= \exp \frac{\sum_{i=i}^{n} \big(\ln \frac{x_i}{\mu_g} \big)^2}{n-1} \tag{6}$$

2

Geometric standard deviation

The log of the geometric standard deviation is equal to the arithmetic standard deviation of the logs of $x_1, x_2, \ldots, x_n$:

$$\ln(\sigma_g) = \sqrt{\frac{1}{n-1} \sum_{i=i}^{n} \left(\ln x_i - \ln \mu_g \right)^2} \tag{7}$$

We use the geometric mean μ_g to calculate the geometric standard deviation:

$$
\begin{aligned}
\text{geometric SD}(\boldsymbol{x}) &= \exp \sqrt{\frac{\sum_{i=i}^{n} \left(\ln x_i - \ln \mu_g \right)^2}{n-1}} \\
&= \exp \sqrt{\frac{\sum_{i=i}^{n} \left(\ln \frac{x_i}{\mu_g} \right)^2}{n-1}}
\end{aligned} \tag{8}
$$

Geometric co-variance

I have not been able to find a reference for this expression. But I think that by extension, the log of the geometric covariance is equal to the arithmetic covariance of the logs of $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$:

$$\ln cov(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{n-1} \sum_{i=i}^{n} \left(\ln x_i - \ln \mu_{x,g} \right) \times \left(\ln y_i - \ln \mu_{y,g} \right) \tag{9}$$

We use the geometric means $\mu_{x,g}$ of $\boldsymbol{x}$ and $\mu_{y,g}$ of $\boldsymbol{y}$ to calculate the geometric standard

deviation:

$$\text{geometric Cov}(\boldsymbol{x}, \boldsymbol{y}) = \exp \frac{\sum\limits_{i=i}^{n} \left(\ln \frac{x_i}{\mu_{x,g}} \right) \times \left(\ln \frac{y_i}{\mu_{y,g}} \right)}{n-1}$$

$$= \exp \frac{\sum\limits_{i=i}^{n} \left(\ln x_i - \ln \mu_{x,g} \right) \times \left(\ln y_i - \ln \mu_{y,g} \right)}{n-1} \tag{10}$$

Variance decomposition

The arithmetic variance of the sum of random variables is the sum of the variance of each random variable plus two times the covariance of the random variables. The variance and covariance terms are additive. The arithmetic variance of the sum of random variables X and Y is written as:

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y) \tag{11}$$

I think the log of the geometric variance of the sum of random variables X and Y should be the following. Here I use σ to denote the arithmetic variance and σ_g^2 to denote the geometric variance:

$$\ln \left(\sigma_g^2(X+Y) \right) = \sigma^2(\ln X) + \sigma^2(\ln Y) + 2 \times \text{geom cov}(\ln X, \ln Y) \tag{12}$$

The geometric variance of the sum of random variables is multiplicative. I think this

shows up when rewriting the equation above.

$$\ln\left(\sigma_g^2(X+Y)\right) = \sigma^2(\ln X) + \sigma^2(\ln Y) + 2 \times \text{cov}(\ln X, \ln Y)$$

$$\sigma_g^2(X+Y) = \exp\left(\sigma^2(\ln X) + \sigma^2(\ln Y) + 2 \times \text{cov}(\ln X, \ln Y)\right)$$

$$= \exp\left(\sigma^2(\ln X)\right) \exp\left(\sigma^2(\ln Y)\right) \exp\left(2 \times \text{cov}(\ln X, \ln Y)\right)$$

$$= \exp\left(\sigma^2(\ln X)\right) \exp\left(\sigma^2(\ln Y)\right) \left(\exp \text{cov}(\ln X, \ln Y)\right)^2$$

$$(13)$$

In the last line of the equation above, there are some familiar terms: the geometric variance of X, the geometric variance of Y, and the geometric covariance of X and Y (which is squared rather than multiplied by two).

Back to our problem

We want to understand the contribution of variance in each component of reproductive success to the overall variance in reproductive success. The log of the geometric variance of a quantity is the arithmetic variance of the log of that quantity. If reproductive success is a product of three terms, the log of the geometric variance of reproductive success is the arithmetic variance of the log of the product of those terms.

$$\text{geometric var(per capita RS)} = \exp\left(\text{Var}(\ln(\sigma F \phi))\right)$$

$$= \exp\left(\text{Var}(\ln(\sigma) + \ln(F) + \ln(\phi))\right)$$

$$(14)$$

Using the properties of logarithms lets us rewrite this more, so that we can see that it is the arithmetic variance of the sum of the logs. We can then apply what we know about the properties of variances and exponents (as above) to expand the expression.

$$\text{geometric var(per capita RS)} = e^{\text{Var}(\ln \sigma)} e^{\text{Var}(\ln F)} e^{\text{Var}(\ln \phi)} \left(e^{\text{Cov}(\ln \sigma, \ln F)}\right)^2 \left(e^{\text{Cov}(\ln \sigma, \ln \phi)}\right)^2 \left(e^{\text{Cov}(\ln F, \ln \phi)}\right)^2$$

$$(15)$$

I've verified that this expression returns the correct result by comparing the geometric variance calculated from the values for reproductive success and plotting them against the product of the decomposed terms (see Figure 1).

Interpretation of the decomposition is different than for an arithmetic variance. I'm not sure I have the nuances but here is what I have so far. First, the variance has a minimum value of 1 because of the exponents. This corresponds to an arithmetic variance of 0. Second, covariances have a minimum of 0. Values of 1 for the covariance indicate a lack of covariation; values less than 1 indicate negative covariation; values greater than one indicate positive covariation. I'm not sure how to plot these results because the contributions are multiplicative but I've tried a few things.

First, it seems like the geometric variance in sigma is much larger than any other terms for many sites (see Figure 2). Fecundity has the next highest variance, and then phi. The covariances are mostly positive, though there does seem to be both positive and negative covariances for sigma and fecundity.

Second, the sites vary in how their variances are distributed but some patterns emerge from plotting each sites components individually. Variance in seedling survival to fruiting dominates the other components at roughly half the sites. The other sites see other patterns such as a more even distribution of variances (e.g. Camp 3), equal variance in seedling survival to fruiting and fruits per plant (Lucas Creek West), or decreasing variance when moving from survival to fruits to seeds (e.g. Upper Richbar).

I've included the plots for interannual estimates of seedling survival to fruiting, total fruit per plant equivalents, and seeds per fruit. Looking at those, it seems like high variance in seedling survival to fruiting may be the result of sites with year(s) of very low seedling survival to fruiting. Mill Creek saw three years with survival very close to zero.
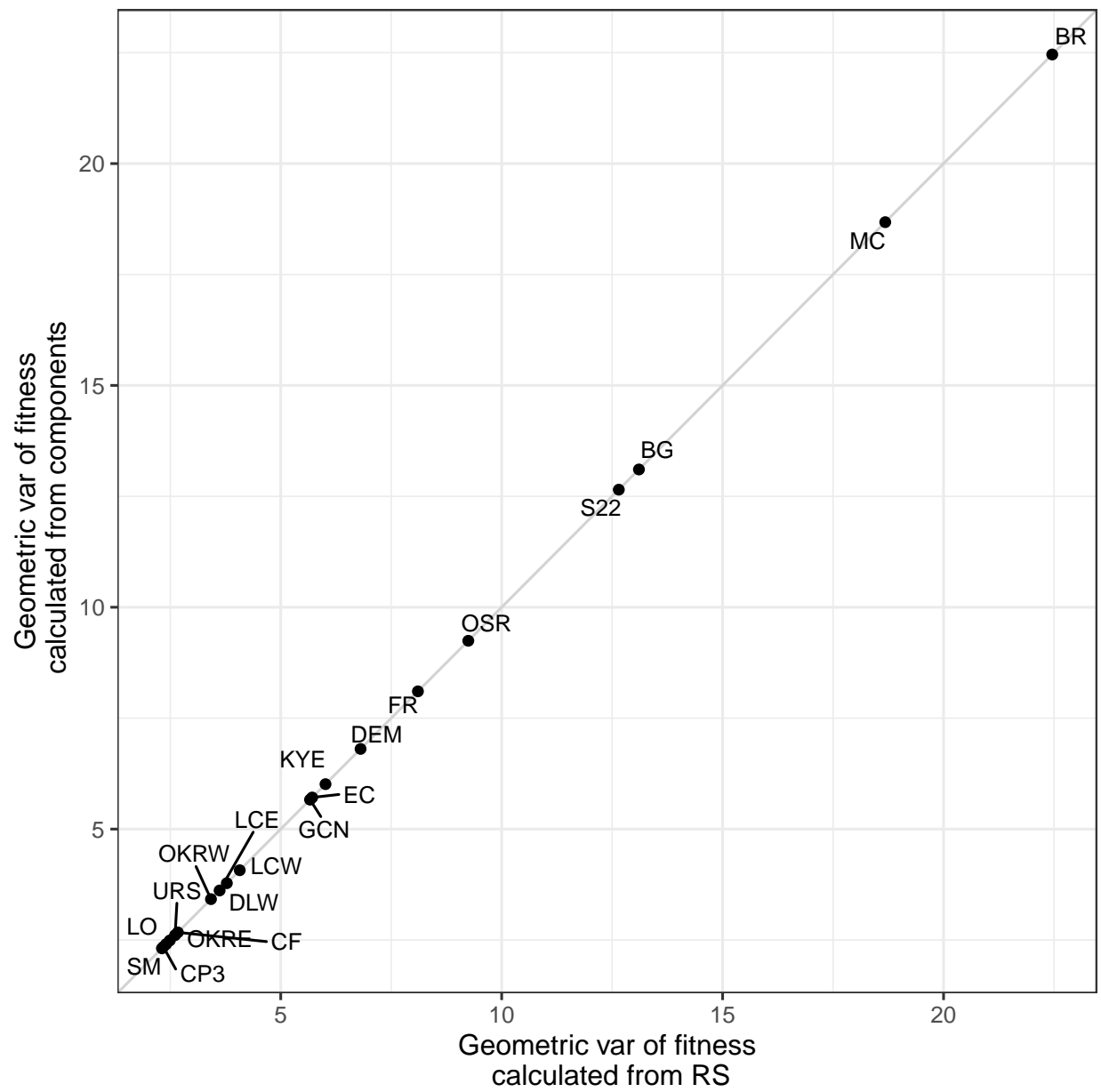
Figure 1: Compares the geometric variance calculated from the values for reproductive success with the product of the decomposed terms.
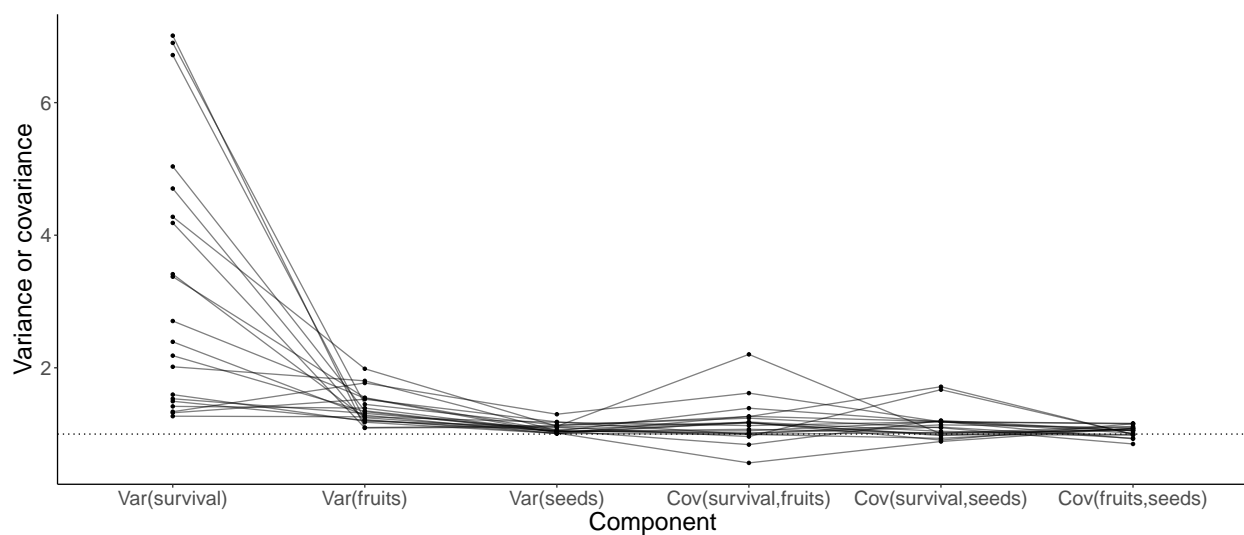
Figure 2: Plots the variance and covariance terms. The covariance terms have not been squared here, as they would be in the contribution to reproductive success - not sure if that's something I need to do?
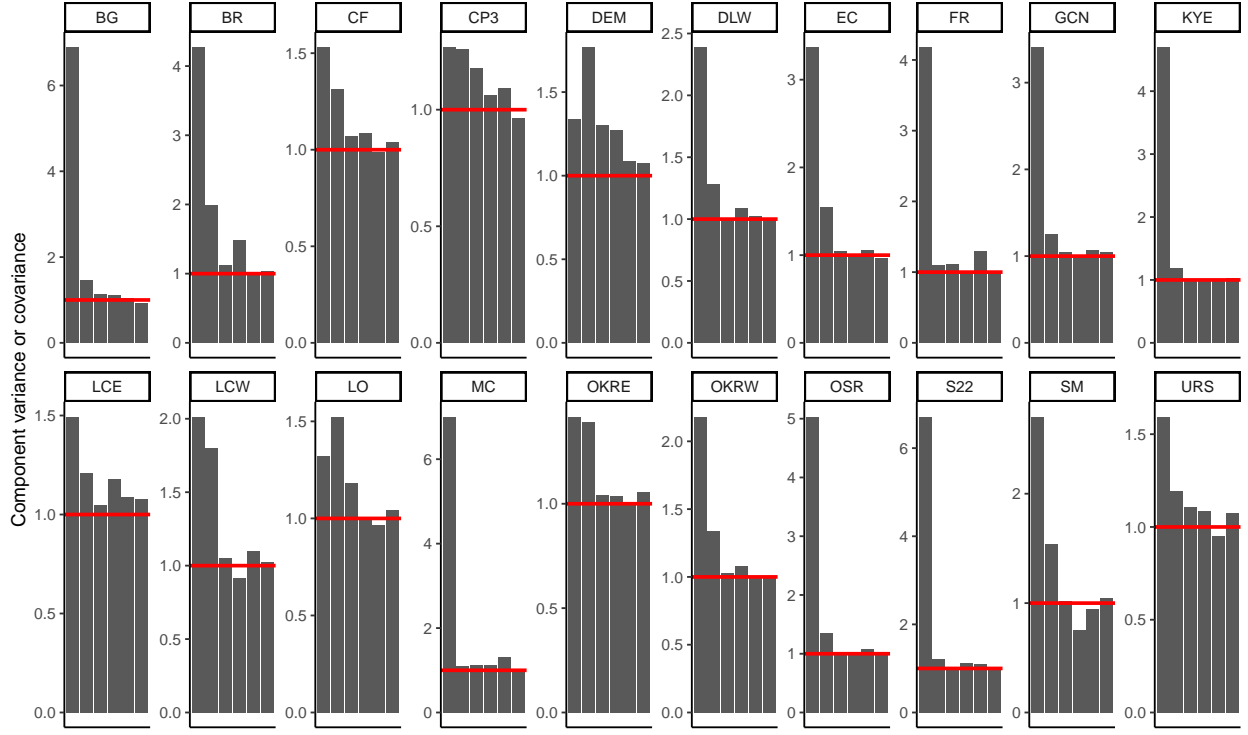
Figure 3: Plots the variance and covariance terms for each site. The covariance terms have not been squared here, as they would be in the contribution to reproductive success - not sure if that's something I need to do? The bars are arranged in the same order as in the previous figure: Var(sigma), Var(F), Var(phi), Cov(sigma,F), Cov(sigma,phi), Cov(F,phi). The y-axis scales vary for each plot so I have included a red line at 1; this is the minimum for the variance terms and the value when there is no covariance.

9

Probability of seedling survival to fruiting

Fruits per plant (total fruit equivalents)