Souther, S., M. Sample, G. Conley, and C. Aslan. 2022. Demographic analysis of the endangered plant *Pectis imberbis* highlights tradeoffs between deer browse and interspecific competition. Natural Areas Journal 42(3).

***Integral Projection Model (IPM) Development—***

*Establishing demographic measurements*

In order to parameterize Integral Project Models (IPMs) for demographic analysis, a continuous state variable, typically a metric of size, must explain variation among individuals in reproduction and survival. In September 2019 (9/17/19), we collected a variety of size measurements on *P. imberbis*, including 1) height of longest stem, 2) summed height of all stems, 3) stem number, 4) leaf number and 5) length of the longest leaf, in order to determine the best predictor of demographic vital rates. We then built statistical models and used the Aikaike Information Criterion (AIC) to compare model performance in predicting floral head number, a measure of reproduction. Models explored single factor predictors (e.g., height), as well as composite factors that represented combinations of state variables (e.g., leaves\*height), and included both linear and polynomial fits. A total of 24 models were tested and height was selected as the best predictor of vital rates (**Table 1**).

*Census timing*

To guide census timing, we measured height at 2 time points, once in June (19th) and again in September (5th) of 2020, to determine whether plant growth was determinant or continuous within a growing season. In June, plants had attained between 42% and 93% of their height as recorded in September, and were on average 69% their size in September, with larger plants having more fully developed. However, floral counts were difficult to ascertain in June, due to the immaturity of floral heads. We again revisited the population in October (16th), but found that floral heads and small seedlings were beginning to senescence at that time, again rendering reproductive counts difficult. For this reason, census data collected in September of both years were used in model development.

**Table 1.** Model rankings produced by AIC scores.

*Model parameterization*

The continuous state variable, total plant height, was used to explain individual variation in growth, survival, and reproduction in order to model population dynamics using Integral Projection Models (IPMs). Variance in residuals describing the relationship describing growth was uniform without transformation of the state variable (Shapiro-Wilk normality test; W = 0.9964. p = 0.1210; **Fig. 1**). In order to parameterize the *P* sub-kernel, height was related to survival and growth and linear and polynomial models compared using the Akaike Information Criterion (AIC) (**Fig. 2**). Based on model comparisons, second degree terms were included for both growth and survival models to form the *P* sub-kernel (**Fig. 3**). We ran diagnostics on the P matrix to determine whether the dimensions of the matrix (bin N = 100) were appropriate and to test for eviction. During the discretization process, individuals with sizes near model limits may be ‘evicted’ from the model if their future size exceeds these limits, with eviction leading to inaccuracies in growth rate estimates. No eviction from the model was detected, so no correction was applied, and models were unresponsive to bin number alteration, thus a bin number of 100 was retained (**Fig. 4**).

Parameterization of the F matrix was more complex, since we were only able to gather reproductive data from a subsample of individuals in year 1. However, a complete census was possible in year 2 of the study, allowing us to compare models between years to ensure appropriate parameters were selected for F matrix construction based on subsampled reproductive data. Model ranking procedures for both the year 1 subsample and year 2 models selected the same model for likelihood of flowering, which was best predicted by linear models (**Fig. 5**), and seed production, which was best predicted by models including polynomial terms (**Fig. 6**). Using the relationships between height and flowering probability and seed production from data in year 1 for a subset of the population, we were able to estimate reproduction for the rest of the population, using the following equations derived from logistic and Poisson GLM models, respectively:

and

For final seed count estimates, production was contingent on flowering, and a plant was scored as having produced flowers if the likelihood was or exceeded a threshold of 0.75. While these estimates are sources of uncertainty in the model, ultimately seed production is extremely high relative to seedling recruitment, and models were constructed to depend principally on observed seedling numbers, which were accurately assessed. Thus, the final F sub-kernel model consists of a size-based probability of flowering, seed production contingent on flowering derived from floral head counts and the resultant size distribution of observed germinants, which ultimately determined entry of new individuals, arising from seed production the previous year, into the population (**Fig. 7**). Based on P and F sub-kernel evaluations, these final four equations describing the predictive relationship between *P. imberbis* height and survival, growth, probability of flowering, and seed production, along with the size distribution of observed seedlings, form the structure of the IPM model to examine population dynamics (**Fig. 8, 9)**.

*Seed bank dynamics*

We calculated that 541 reproductive individuals produced an estimated 233,294 seeds in year 1. A total of 7 germinants were observed in year 2, dividing the number of germinants by seed number minus the seeds used in the seed cage experiment (233,294-500=232,794) yielded a ‘natural’ germination rate (germination occurring within the population, not within seed cages) of 0.003%. Notably, this is a lower germination rate than that calculated from the seed cage experiment, which estimated a germination rate of 0.4% (2 germinants per 500 seeds planted). For modeling purposes, to best approximate germination rates, we created a germination rate based on all observed germinants, including the two germinants from seed cages (Total germinants = 9), divided by the total seed count. Further studies of seed and seedling dynamics are warranted, though discrepancies in germination rates are unlikely to strongly influence demographic model outcomes, since recruitment rates are low and sensitivity analyses do not indicate that population growth is strongly influenced by this life stage.

**Figures**



**Figure 1**. The relationship between *P. imberbis* size in year 1 and year 2 (Panel 1) and quantile-quantile plot (Panel 2).



**Figure 2**. Comparisons of linear and polynomial growth (Panel 1) and survival (Panel 2) models.



**Figure 3.** Visualization of the P sub-kernel. The dashed line indicates stasis, while individuals that fall above the line grow and below the line shrink. Maximum size can be approximated where the growth model crosses the one-to-one line, indicating that while there is the possibility that individuals will grow beyond that size, in this case ca. 80 cm in height, they will have the tendency to shrink to smaller size classes in subsequent years.



**Figure 4**. A summary of P matrix diagnostics. The first panel (left) shows the range of the state variable (black line) and the distribution of the data, indicated by the grey histogram. The red and blue arrows in the left panel indicate the data range of two additional P matrices, one with a larger size range and one with an increased number of bins, respectively, which are explored in following 2 panels of the figured by comparing with the true P matrix. Eviction would be apparent in the ‘Survival’ and ‘Life expectancy’ panels if an increase in size range resulted in disagreement with the actual P model. This was not the case, however, and no correction for eviction was applied. Similarly, alteration of bin size, or the dimensions of the discretized matrix (bin number = 100) did not alter model output, again suggesting that this parameter was appropriately selected.



**Figure 5**. Comparison of models predicting the likelihood of flowering based on plant height in 2019 (Panel 1) and 2020 (Panel 2).



**Figure 6**. Comparison of models predicting seed production from plant height in 2019 (Panel 1) and 2020 (Panel 2).



**Figure 7**. Visualization of the F sub-kernel. As with standard matrix modeling, each cell of the matrix indicates the transition probability from size at time t to a size at time t + 1. Here, the largest individuals within the population contribute the most to seedling recruitment.



**Figure 8.** The relationship of size to A) size the following year, B) probability of survival, C) probability of flowering and D) seed production. In all cases, size is a predictor of y. For each of the underlying regressions used to build growth, survival, and reproduction data, multiple fits were explored and the model of best fit was chosen as the model with the lowest AIC score.



**Figure 9**. Visualization of the final IPM model, including both P and F sub-kernels.