***Statistical Analyses***

First, we determined whether *Bombus impatiens* has a different phenology than the four most common native bumble bee species in our surveys (*B. impatiens*, *B. melanopygus*, *B. mixtus*, *B. flavifrons*, and *B. vosnesenskii*). To test differences in species phenology, we applyied a Bayesian generalized mixed model (GLMM) with a Poisson distribution to our bumble bee counts. We restricted the count data to workers only, as males and queens exhibit distinct phenological patterns from workers, emerging earlier and later than workers on average, respectively (Williams et al., 2014). We modeled variation in bumble bee counts using a log-link function and included the following predictors: species, year (2022 or 2023), and species-specific effects of Julian date and (Julian date)2. We included a random effect of site:

;

;

.

*B. impatiens* was used as the reference level for each of the species-specific effects. As such, the effect of Julian date represents the difference in the timing of peak abundance of each species relative to *B. impatiens*. The quadratic effect of (Julian date)2 represents the difference in the phenological decay of counts around the date of peak abundance relative to *B. impatiens* (Leong et al., 2016). We interpreted native species to have an earlier phenology relative to *B. impatiens* if the 95% BCI for the species-specific effect of Julian date was less than zero. Similarly, we interpreted native species to have a shorter flight season relative to *B. impatiens* if the 95% BCI for the species-specific effect of (Julian date)2 was less than zero. Julian date and (Julian date)2 were z-score standardized prior to model fitting. We included a species-specific extra-Poisson dispersion parameter in our model () to capture additional variability in counts that may differ among species (Stan Development Team 2023b).

Next, we tested whether *Bombus impatiens* interacts with invasive plant species more frequently than native bumble bees. To do this, we applied a Bayesian generalized mixed model (GLMM) with a Bernoulli distribution to the visitation choices made by bumble bees (interaction with an invasive versus native plant species as a binary outcome). For this analysis, we considered *B. impatiens* and the four most common native bumble bee species listed above. We included all visitation choices that we observed during our surveys, including interactions made by bumble bee workers, queens, and males (3627 total interactions). We classified the flowering plants that bumble bees visited as invasive if they were included in the species list provided by the Invasive Species Council of BC (Invasive species council of BC, n.d.). Several non-native plant species were not classified as invasive. We then modeled the odds of an individual bumble bee choosing an invasive plant over a native plant with species and a species-specific effect of the availability of invasive plants as predictors, and site included as a random effect:

;

.

The availability of invasive plants was calculated as the proportion of flowers on a survey date (and site) that were provided by invasive versus native plants. Availability of invasive plants was z-score standardized prior to model fitting. With *B. impatiens* as the reference level, the species-specific intercepts describe each species’s probability of interacting with an invasive plant species when the scaled availability of invasive plants equals 0 (i.e., when there is an average proportion of invasive plants in the community). The species-specific effect of increasing availability of invasive plants describes the change in the probability of interacting with an invasive plant when the availability of invasive plants increases by one standard deviation. We interpreted native species to have a lower preference for invasive plants relative to *B. impatiens* if the 95% BCI for the species-specific effect on the intercept was less than zero. We interpreted a lower dietary flexibility for using invasive plants if the 95% BCI for the species-specific effect of increasing availability of invasive plants was less than zero.

Both Bayesian GLMMs were written in the probabilistic programming language *Stan* (Stan Development Team 2023b), implemented in R with *rstan* (Stan Development Team 2023a). For the both Bayesian GLMMs we employed best practices for model development and fitting by: using weakly-informative priors to discourage unrealistic parameter values (Table \_); confirming sufficient mixing of chains (Gelman-Rubin R-hat values < 1.05), minimal within-chain autocorrelation (effective sample size / total HMC steps > 0.1), and no divergent transitions (Figures S5 and S6); and running our model for a length of 4,000 HMC steps, discarding the first 2,000 (Stan Development Team 2023b; Gelman et al. 2020). We conducted visual posterior predictive checks to assess model goodness-of-fit.

To test whether *B. impatiens* diet overlaps significantly with the diet of native *Bombus* species, we used non-metric multi-dimensional scaling (NMDS) to assess the similarity in the plant associations for each species of the five most common species of bumble bee from each site across the two years of surveys. Briefly, NMDS uses ordination to create a distance matrix of observations (Bakker, 2024). In ecological studies, the method is used to establish pairwise dissimilarity among community species lists. Here, we defined the set of plants each species of bumble bee was recorded foraging on as the bumble bee ‘community’, with the goal of illustrating the degree of dissimilarity in plant associations among the five most commonly observed bumble bees in my data set. Briefly, we created a matrix for each site, with the bumble bee species at each site as the columns, and the plant community as the rows. We populated the matrix with the abundance of interactions between a particular bee and plant species observed at that site over the two years of surveys. Using the vegan package in R and applying a Bray-Curtis distance metric (v2.6-4; Oksanen et al., 2022), plots were created to visualize the degree of overlap in the plant community observations for each of the five most common species of bumble bee in my data set, with the goal of establishing whether the species had distinct or overlapping foraging preferences. Explain the test STATS TEST HERE?

Data preparation and plotting was conducted in R (v4.2.3; R Core Team 2023).

A graph with different colored lines

Description automatically generated with medium confidence**Figure 1: Phenology differences between B. impatiens and native bumble bee species.** 95% BCI’s for expected abundance by julian date (a) illustrate the phenological differences among the five most common bumble bee species encountered during our surveys. Expected abundances are displayed for the average site, for the year 2022, and without the extra-poisson dispersion that we accounted for with our model. Posterior model estimates indicate that, for all native species, the peak abundance occurs earlier in the year relative to *Bombus impatiens* (b), with 50% BCI’s (thick bars) and 95% BCI’s (thin bars) for the effects of species identity on peak abundance date < 0.

A close-up of a graph

Description automatically generated

**Figure 2: Invasive plant interaction choice differences between B. impatiens and native bumble bee species.** 50% BCI’s for each bumble bee species illustrate the expected probability of interacting with an invasive plant given the availability of invasive plants (a). Colored points indicate the observed interactions used to fit the logistic GLMM model where an interaction with an invasive was treated as a success (y = 1) and an interaction with a non-invasive was treated as a failure (y = 0). Posterior parameter distributions lower than zero indicate that, relative to *B. impatiens*, other bumble bee species are less likely to interact with an invasive plant species when there is an average proportion of invasive plant species available (i.e., when the z-score scaled proportion of invasive plants = 0) (b); or that other bumble bee species have a lower increase in the probability of interacting with an invasive plant when the proportion of invasive plant species increases by one standard deviation above average (c).

A diagram of a space

Description automatically generated

**Figure 3: add figure caption here**

**SUPPLEMENTAL TABLES AND FIGURES (JENS)**

Table \_: Prior distributions used for the phenology model:

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Prior distribution** | **Justification** |
| βspecies | Normal(0, 2) | Weakly-informative prior for species-specific differences in baseline abundance |
| βsite | Normal(0, σsite) | Partial-pooling of site effects informed by within site data and variation among sites (i.e., site treated as a random effect) |
| σsite | Half-Normal(0, 2) | Weakly-informative prior for variation among sites |
| βyear | Normal(0, 1) | Stronger prior for effect of year on overall abundance |
| βdate | Normal(0, 2) | Weakly-informative prior for peak abundance date |
| βdate2[*B. impatiens*] | Normal(0, 2) | Weakly-informative prior for overall shape of phenology for the reference level |
| βdate2[*B. flavifrons*] | Normal(0, 1) | With a stronger prior that species don’t differ drastically in their phenology shape relative to the reference level |
| βdate2[*B. melanopygus*] | Normal(0, 0.25) | With a much stronger prior that *Bombus melanopygus* phenology shape generally follows the shape of the reference level. Due to poorer data availability for this species, a stronger prior was needed to achieve sufficient mixing of chains and an adequate effective sample size. |
| βdate2[*B. mixtus*] | Normal(0, 1) | With a stronger prior that species don’t differ drastically in their phenology shape relative to the reference level |
| βdate2[*B. vosnesenskii*] | Normal(0, 1) | With a stronger prior that species don’t differ drastically in their phenology shape relative to the reference level |
| ε | Normal(0, σε[species[i]]) | Extra-poisson variance is drawn from a distribution with a width that could vary among species |
| σε | Half-Normal(0, 0.5) | Stronger prior to keep ε closer to zero (i.e., that overdispersion doesn’t range towards infinity, rather that the fixed and random effects in the model explain most of the variation in the data) |

Table \_: Prior distributions used for the diet choice model:

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Prior distribution** | **Justification** |
| βspecies | Normal(0, 2) | Weakly-informative prior for species-specific differences in the probability of interacting with an invasive plant when invasive plants are at an average proportion in the plant community. |
| βsite | Normal(0, σsite) | Partial-pooling of site effects informed by within site data and variation among sites (i.e., site treated as a random effect) |
| σsite | Half-Normal(0, 1) | Weakly-informative prior for variation among sites |
| βprop\_invasive | Normal(0, 2) | Weakly-informative prior for species-specific effect of increasing proportion of invasive plants on the probability of interacting with an invasive plant. |

Supplemental Figure (?):

a)

**Phenology model:** for observation *i*, where y[i] is an integer count of a species on some given survey:

y[i] ~ Poisson(lambda[i])

log(lambda[i]) = βspecies[species[i]] + βsite[site[i]] + βyear[year[i]] + βdate[species[i]] \* date[i] + βdate2 [[species[i]]] \* date2[i]+ ε[i]

βsite[site[i]] ~ Normal(0, σsite)

ε[i]~ Normal(0, σε[species[i]])

b)

**Diet choice model:** for observation *i* where y[i] is a binary outcome of whether a bumble bee interacted with an invasive plant y = 1 or not y = 0:

y[i] ~ Bernoulli(theta[i])

logit(theta[i]) = βspecies[species[i]] + βsite[site[i]] + βprop\_invasive[species[i]] \* prop. invasive[i]

βsite[site[i]] ~ Normal(0, σsite)

1. A screenshot of a computer

   Description automatically generated
2. A screenshot of a computer

   Description automatically generated
3. A screenshot of a computer

   Description automatically generated
4. A screenshot of a computer

   Description automatically generated

**Supplemental Figure \_: Phenology model diagnostics.** Traceplots illustrate strong mixing of chains (a). Pairs plots lacking divergent transitions illustrate efficient searching of the full posterior distribution (b). Visual posterior predictive checks for goodness-of-fit indicate that the fitted model provides reasonable estimates of the mean count per survey (red uncertainty bands) in relation to the true mean count per survey (black dots) (c); and reasonable estimates of the maximum count per survey (red uncertainty bands) in relation to the true maximum count per survey (black dots) (d).

1. A screenshot of a computer

   Description automatically generated
2. A screenshot of a computer

   Description automatically generated
3. A screenshot of a computer

   Description automatically generated

**Supplemental Figure \_: Diet choice model diagnostics.** Traceplots illustrate strong mixing of chains (a). Pairs plots lacking divergent transitions illustrate efficient searching of the full posterior distribution (b). Visual posterior predictive checks for goodness-of-fit indicate that the fitted model provides reasonable estimates of the number of interactions with invasive plants (red uncertainty bands) in relation to the true number of interactions with invasive plants (black dots) (c).

A screenshot of a computer

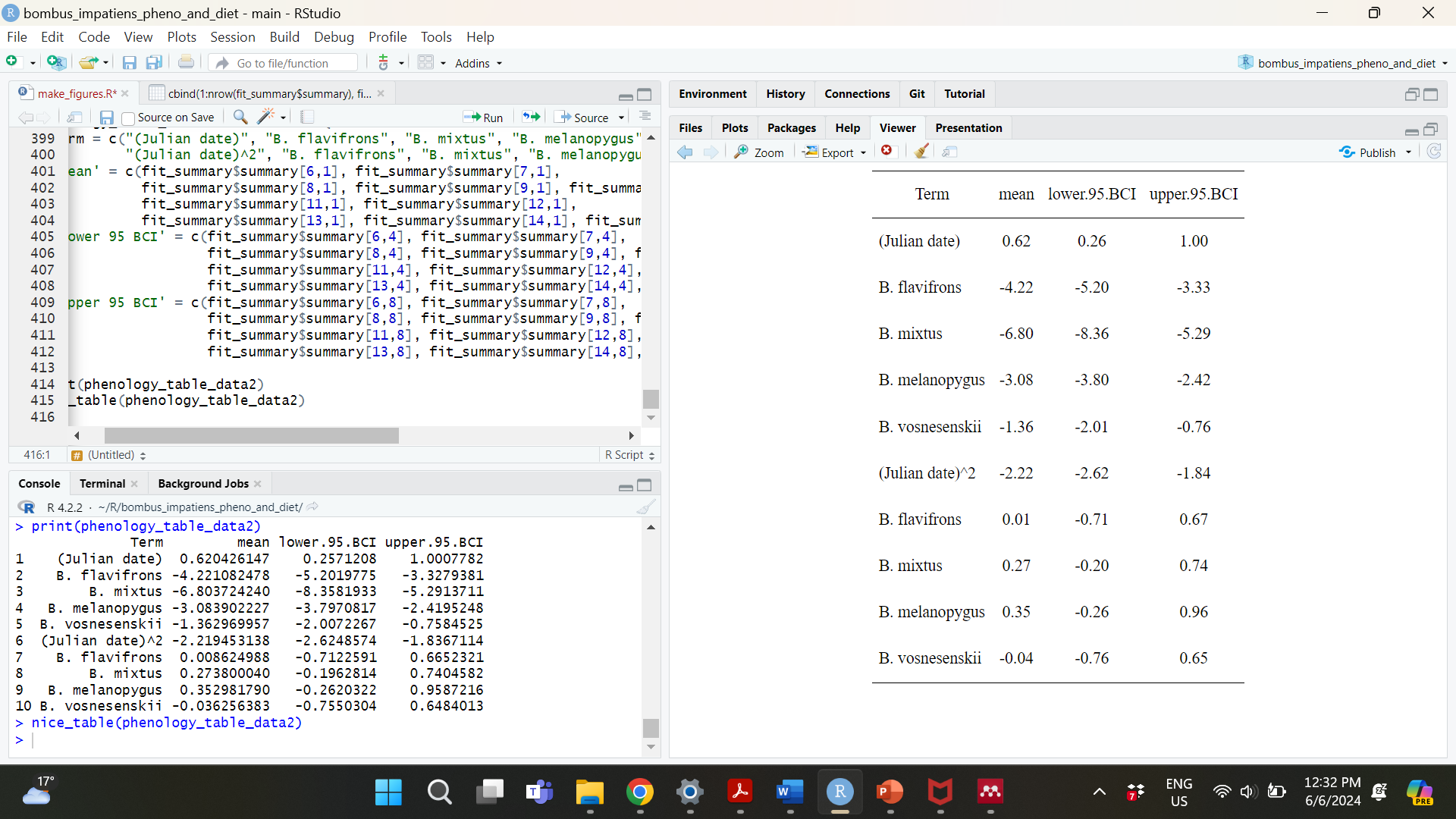
Description automatically generated

Diet choice model parameter estimates.

A screenshot of a computer

Description automatically generated

Phenology parameter estimates (set 1).



Phenology parameter estimates (set 2).