




# Weak interspecific interactions in a sagebrush steppe? Conflicting evidence from observations and experiments

PETER B. ADLER <sup>1,5</sup> ANDREW KLEINHESSELINK,<sup>1</sup> GILES HOOKER,<sup>2</sup> JOSHUA B. TAYLOR,<sup>3</sup>  
BRITTANY TELLER,<sup>1</sup> AND STEPHEN P. ELLNER<sup>4</sup>

<sup>1</sup>Department of Wildland Resources and the Ecology Center, Utah State University, 5230 Old Main, Logan, Utah, USA

<sup>2</sup>Department of Biological Statistics and Computational Biology, Cornell University, 1198 Comstock Hall, Ithaca, New York, USA

<sup>3</sup>USDA, Agricultural Research Service, U.S. Sheep Experiment Station, 19 Office Loop, Dubois, Idaho, USA

<sup>4</sup>Department of Ecology and Evolutionary Biology, Cornell University, E145 Corson Hall, Ithaca, New York, USA

**Abstract.** Stable coexistence requires intraspecific limitations to be stronger than interspecific limitations. The greater the difference between intra- and interspecific limitations, the more stable the coexistence, and the weaker the competitive release any species should experience following removal of competitors. We conducted a removal experiment to test whether a previously estimated model, showing surprisingly weak interspecific competition for four dominant species in a sagebrush steppe, accurately predicts competitive release. Our treatments were (1) removal of all perennial grasses and (2) removal of the dominant shrub, *Artemisia tripartita*. We regressed survival, growth, and recruitment on the locations, sizes, and species identities of neighboring plants, along with an indicator variable for removal treatment. If our “baseline” regression model, which accounts for local plant–plant interactions, accurately explains the observed responses to removals, then the removal coefficient should be non-significant. For survival, the removal coefficients were never significantly different from zero, and only *A. tripartita* showed a (negative) response to removals at the recruitment stage. For growth, the removal treatment effect was significant and positive for two species, *Poa secunda* and *Pseudoroegneria spicata*, indicating that the baseline model underestimated interspecific competition. For all three grass species, population models based on the vital rate regressions that included removal effects projected 1.4- to 3-fold increases in equilibrium population size relative to the baseline model (no removal effects). However, we found no evidence of higher response to removal in quadrats with higher pretreatment cover of *A. tripartita*, or by plants experiencing higher pre-treatment crowding by *A. tripartita*, raising questions about the mechanisms driving the positive response to removal. While our results show the value of combining observations with a simple removal experiment, more tightly controlled experiments focused on underlying mechanisms may be required to conclusively validate or reject predictions from phenomenological models.

**Key words:** coexistence; competition; facilitation; integral projection model; population dynamics; removal experiment; sagebrush steppe.

## INTRODUCTION

Stable coexistence of competitors requires that intraspecific interactions must limit populations more than interspecific interactions (Chesson 2000). When studying populations of species that appear to be coexisting stably, such as populations that have co-occurred in close proximity for extended periods, one should expect intraspecific limitations to be stronger than interspecific limitations. However, interspecific interactions might still be strong and negative. In fact, plant community ecologists often motivate research on coexistence by noting that all plant species compete for water, light, and a few mineral resources (e.g., Silvertown 2004). Resource competition plays a central role in both Tilman’s mechanistic resource competition models (Tilman 1982) and Grime’s competitor-stress tolerator-ruderal strategies (Grime 1979). Neutral theory (Hubbell 2001) goes even further, assuming that interspecific competition is exactly as strong as intraspecific competition. Although a growing list of studies has emphasized the role of interspecific

facilitation in some plant communities (Brooker et al. 2008, He et al. 2013), it is unclear whether this research has overturned the default assumption that plants compete intensely with other species.

Competitive release provides an intuitive, population-level measure of the strength of interspecific competition. Following removal of one or more species from a community, how much will the abundance of the remaining species increase? Large increases in the abundance of the remaining species represent a large competitive release, and imply strong interspecific interaction. Plant ecologists have conducted dozens of removal experiments and have observed almost every possible result. In some cases, competitive release was complete: remaining species fully compensated for removed biomass (Leps 1999, Jutila and Grace 2002). In other cases, some competitive release occurred but the degree of compensation depended on the functional group or dominance of the removed and remaining species (Sala et al. 1989, Belsky 1992, Smith and Knapp 2003). Finally, even negative effects of removals have been observed (Keddy 1989, Gilbert et al. 2009), consistent with the recent facilitation literature. Results from removal experiments do come with caveats. Experiments typically do not run long enough to document long-term outcomes, responses may only be measured at

Manuscript received 31 May 2017; revised 2 March 2018; accepted 20 March 2018. Corresponding Editor: Tom E. X. Miller.

<sup>5</sup>E-mail: peter.adler@usu.edu

one life stage, whereas population responses depend on the whole life cycle, and removal treatments require disturbance, which may cause a variety of impacts unrelated to resource competition (Aarssen and Epp 1990).

An alternative approach to studying competitive release is to fit multispecies population models that estimate species interactions using long-term, observational data, and then simulate population responses to removals. Applying this approach to arid and semiarid plant communities, we found evidence of strong intraspecific limitation and weak interspecific interactions, leading to very stable coexistence and implying little competitive release (Adler et al. 2010, Chu and Adler 2015). In fact, at our sagebrush steppe study site (Adler et al. 2010), our model predicts removal of sagebrush would actually reduce the equilibrium cover of the three dominant grass species very slightly, reflecting the facilitative effect of sagebrush on grass recruitment and very weak competitive effects on growth and survival. Since we only studied the most common, co-occurring species, we expected to find evidence of stable coexistence, but the predicted negative effect of sagebrush removal on grass cover is surprising given the long history of sagebrush eradication to increase grass forage production in sagebrush rangelands (Robertson 1947, Mueggler and Blaisdell 1958, Cook and Lewis 1963). Compared to traditional field experiments, our approach has the advantage of projecting full life-cycle responses over long time periods, but its reliance on observational data sets limits the strength of our inference. A successful experimental test would increase our confidence in the model's projections of competitive release.

The idea that results from observational studies of competition represent hypotheses that should be tested experimentally is not new (Hairston 1980), but such tests remain extremely rare. In fact, independent tests of any kind of quantitative prediction are rare in ecology, spurring recent calls for a greater emphasis on prediction (Mouquet et al. 2015, Petchey et al. 2015, Houlahan et al. 2017). These authors argue that testing predictions will not only advance applied goals such as forecasting ecological impacts of global change, but will also accelerate progress in basic research.

In the spirit of this argument, we conducted a removal experiment to test whether our sagebrush steppe community model, which is based on observational data, underestimates the strength of interspecific competition and competitive release. We studied the responses of three perennial grass species to removal of sagebrush, and the response of sagebrush to removal of all grasses. The primary purpose of the experiment was not to directly measure competitive release, which is difficult to do over a short time period in a semiarid ecosystem with low and variable plant cover, but rather to test whether our model accurately characterizes the interspecific interactions that determine the magnitude of competitive release. We addressed two questions: First, are the survival, growth, or recruitment of our target species significantly different in control vs. removal treatments, after accounting for differences in local species composition? Because our demographic models include the effects of neighborhood crowding, they should be able to predict how experimental removal of neighbors alters demographic rates. If the full effect of experimental removals is stronger than predicted, it means that our statistical models underestimate

the response to decreased crowding. Our second question integrates across the full life cycle: how do short and long-term population-level responses to removals compare to projections from multispecies population models based on observational data? If models based on observational data adequately describe facilitative and competitive interactions, the population dynamics they project should be similar to projections from models that include information from the experimental removal treatments. Obtaining that result would validate the use of long-term data sets and modeling to study interspecific interactions.

## MATERIALS AND METHODS

### *Study site and data set description*

The U.S. Sheep Experiment Station (USSES) is located 9.6 km north of Dubois, Idaho (44.2° N, 112.1° W), 1,500 m above sea level. During the period of data collection (1926–1957), mean annual precipitation was 270 mm and mean temperatures ranged from –8°C (January) to 21°C (July). The vegetation is dominated by the shrub *Artemisia tripartita* and the C<sub>3</sub> perennial bunchgrasses *Hesperostipa comata*, *Pseudoroegneria spicata*, and *Poa secunda*. These four species, the focus of our models, accounted for over 70% of basal cover (grasses) and 60% of canopy cover (shrubs and forbs).

Scientists at the USSES established 26 1-m<sup>2</sup> quadrats between 1926 and 1932. Eighteen quadrats were distributed among four ungrazed exclosures, and eight were distributed in two paddocks grazed at medium intensity spring through fall. All quadrats were located on similar topography and soils. In most years until 1957, all individual plants in each quadrat were mapped using a pantograph (Blaisdell 1958). The historical data set is public and available online (Zachmann et al. 2010). In 2007, we located 14 of the original quadrats, all of which are inside permanent livestock exclosures, and resumed annual mapped censusing using the traditional pantograph method.

We extracted data on survival, growth, and recruitment from the mapped quadrats based on plants' spatial locations. Our approach tracks genets, which may be composed of multiple polygons, as they fragment and/or coalesce. Each mapped polygon is classified as a surviving genet or a new recruit based on its spatial location relative to genets present in previous years (Lauenroth and Adler 2008). We modeled vital rates using data from 22 year-to-year transitions between 1926 and 1957, and nine year-to-year transitions from 2007 to 2016. Only four quadrats were observed for the first two transitions in the 1920s, while at least 14 quadrats were observed for all subsequent transitions.

All data and R code necessary to reproduce our analysis are available on the Dryad Digital Repository (see *Data Availability*).

### *Removal experiment*

In spring 2011, in a large exclosure containing six of the historical permanent quadrats, we established an additional 16 quadrats for the removal treatments. New quadrats were selected to be similar to historical plots, but also needed to

have enough of the focal species present so that removal experiments could be conducted. We evaluated potential locations for the new quadrats near the original quadrats and then used only those locations with at least 5% canopy cover of *Artemisia tripartita* and at least 3% combined basal cover of the three focal perennial grasses. We rejected locations falling on hill slopes, areas with greater than 20% bare rock, or with over 10% cover of the woody shrubs *Purshia tridentata* or *Amelanchier utahensis*. We mapped the new quadrats in June 2011 and then implemented the removal treatments in September 2011. We randomly assigned 8 of the 16 new quadrats to the sagebrush removal treatment, and the remaining 8 to the grass removal treatment. We cut sagebrush stems at ground level and applied a 5% solution of Roundup Pro (Glyphosate, Monsanto, St. Louis, MO, USA) to the cut stems. In the grass removal plots, we painted the same herbicide solution on all perennial grasses. We removed plants 50 cm beyond the quadrat boundaries to minimize edge effects. The sagebrush removal was virtually 100% effective; grass removals required additional herbicide applications in April and May of each year, with fewer grasses remaining each year. Appendix S1: Fig. S1 shows photographs of example quadrats before and after treatment. We repeated mapped censuses of all plots each June from 2012 to 2016.

Although the primary focus of our analysis is comparison of observed removal responses with responses predicted by our population models, we also conducted a more traditional statistical analysis of differences in cover trends on control vs. removal plots. The response variable for this analysis was year-to-year change in cover at the quadrat scale, calculated as  $\log(\text{Cover}_{t+1}/\text{Cover}_t)$ . We tested for the effect of removal treatment on this measure of cover change with a linear mixed effects model implemented in the `lme4` package (Bates et al. 2015) of R, including year and quadrat as random effects. We considered the removal treatment effect to be significant when the 95% credible interval on the estimated coefficient did not overlap zero. Because variance in residuals of plots with low cover was high for two species, we explored fitting by weighted least squares, to account for the apparent heteroscedasticity. The results of these tests were qualitatively similar, so we opted to only report the simpler analysis.

#### Overview of statistical models for vital rates

Our regression models for survival, growth, and recruitment follow previous work (Adler et al. 2010, Chu and Adler 2015). The current analysis departs from previous work with the inclusion of a categorical removal treatment effect. Although the three vital rate models differ in functional form (see Appendix S1 for complete model descriptions), all of them have a linear predictor,  $\mu$ , containing the following terms:

$$\mu_{ij} = \gamma_j + \phi_{jg} + \chi_j + \sum_m \omega_{jm} W_{ijm} \quad (1)$$

where  $i$  refers to an individual genet (for survival and growth) or quadrat (for recruitment) of species  $j$ ,  $\gamma$  is a time-dependent intercept (year effect),  $\phi$  is the coefficient for the effect of spatial location ( $g$  represents a group of nearby quadrats

located within one pasture or grazing exclosure),  $\chi$  is the removal treatment effect,  $W_{ijm,t}$  is the crowding on species  $j$  exerted by neighbors of species  $m$  at the individual (survival and growth) or quadrat (recruitment) scale, and  $\omega_{jm}$  is an interaction coefficient that determines the impact of crowding by species  $m$  on the focal species  $j$ . For plants in removal plots in 2012, we set crowding for removed neighbor species to zero, rather than using the pre-treatment (2011) data.

If the conspecific and heterospecific crowding effects, the  $\omega_{jm}$ s, explain demographic rates in removal treatment plots as well as they do in control plots, then there will be no need for the removal treatment effect, and the estimated coefficient will be small and not significantly different from zero. However, if demographic rates in removal plots differ from control plots even after accounting for crowding effects, then the removal treatment effect will be significant. For example, we might find that, holding crowding constant, grasses grow more quickly in the sagebrush removal plots than in the control plots. This result would indicate that predictions from a “baseline model,” a model including all estimated parameters except for the removal treatment effect, underestimates competitive release. We refer to the model that includes the estimated removal treatment effect as the “removal model.”

Note that the “baseline” and “removal” models are not estimated separately; we estimate all parameters together, but when we generate predictions from the baseline model, we simply set the removal treatment effects to zero. We chose this over an alternative approach of fitting a model only to data from control plots and then making and evaluating predictions for the removal plots. The principal advantage of our approach is that the significance test (confidence intervals) for the removal effect is well defined; fitting on one data set and testing on another would require developing non-standard null distributions to account for variation in both data sets. In addition, the alternative approach would return essentially the same results because adding data from the removal plots had negligible effects on other model parameters. Correlations between the parameters of a model fit only to control plot data and our full model (leaving aside the removal effects) were stronger than 0.999. Why don't we find strong interactions between removal treatments and the estimated competition coefficients? The answer is that because our treatments involved complete, not just partial, removal of target species, their effects can be captured with an intercept term, as illustrated in Appendix S1: Fig. S2. As a result, if, for example, one of the grasses responds more positively to removal of *A. tripartita* than our baseline model predicts, we cannot determine whether this reflects a poorly estimated competition coefficient, or some other effect of the removal treatment unrelated to competitor densities.

#### Individual-based models

The vital rate regressions provide a qualitative test of removal treatment effects: Does removal have an effect on a vital rate that goes above and beyond the effects of species interactions that are explicit in the model? However, the population-level effects of removal integrate the effects of the removal treatment on all three vital rates. We used an individual-based model (IBM) to predict changes in population size (cover) from one year to the next in removal and

control quadrats, and compared these predictions with observed changes.

To simulate changes in cover in each quadrat from year  $t$  to year  $t + 1$ , we initialized the IBM with the actual configuration of the four focal species observed in year  $t$  (genet sizes and locations). We then projected the model forward one year, applying the random year effects corresponding to year  $t$  and the spatial random effects corresponding to the quadrat of interest. We used the survival regression to predict the survival probability of each genet, and the growth regression to predict the expected size (on arithmetic scale) of each genet at time  $t + 1$  (note that the expected size on arithmetic scale is not  $\exp(\text{expected log size})$ , because of Jensen's inequality). Expected quadrat-level cover at  $t + 1$  is expected area of new recruits, plus the product of survival probability and expected new size summed over all genets present at time  $t$ . Because we do not project forward beyond one time step, we do not need to assign locations to the new recruits. For control plots, we used the baseline version of the model with all removal treatment effects set to zero. For removal treatment plots, we generated predictions from both the baseline and removal versions of the model, with and without treatment effects set to zero, respectively.

Although all the parameters in our model are estimated with uncertainty, we chose not to propagate parameter uncertainty through these simulations because we wanted to maximize our ability to detect the population-level consequences of a small removal effect. We previously found that interspecific competition is weak relative to the effects of interannual environmental variability (Tredennick et al. 2017). Including parameter uncertainty along with random year effects would conflate parameter uncertainty with environmental variability, overstating the magnitude of natural fluctuations and making it more difficult to show that competitive release is larger than the observationally-based model predicts. Therefore, we used estimated means of all parameters, including those with confidence intervals that overlapped zero. Based on the same rationale, we used a deterministic version of the model (e.g., recruitment is the  $\lambda$  of Appendix S1: Eq. S5, rather than a Poisson random variable with mean  $\lambda$ ). Specifically, we used the recruitment regression to calculate the expected number of new recruits at time  $t + 1$ , which we then multiplied by the observed mean size of recruits for the focal species to get expected cover of new recruits.

Another reason we might observe small effects of the removal treatment is the limited power of our experiment. Perhaps large uncertainty in the estimates of the removal effects themselves makes it difficult to reject the null hypothesis represented by our baseline model. To address this possibility, we also simulated a version of the model in which we replaced the point estimate of each removal treatment effect with either the 5% or 95% credible limit for the treatment effect, whichever was farthest from zero. This version of the model shows the population-level changes we would expect given maximum removal effects that are consistent with our data.

#### Integral projection models

To quantify the long-term population-level consequences of the removal effects, we built and simulated an integral

projection model (IPM). The population of species  $j$  is represented by a density function  $n_j(u, t)$ , which gives the density of genets of size  $u$  at time  $t$  ("size" is natural log of genet area, so that  $n_j(u, t)h$  is the number of genets whose area is between  $\exp(u)$  and  $\exp(u + h)$  for  $h \ll 1$ ). The size distribution function at time  $t + 1$  is given by

$$n_j(v, t + 1) = \int_{L_j}^{U_j} k_j(v, u, \vec{n}) n_j(u, t) du \quad (2)$$

where the kernel  $k_j(v_j, u_j, \vec{n}_j)$  describes all possible transitions from size  $u$  to  $v$ , and  $\vec{n}$  is the set of size-distribution functions for all species in the community. The integral is evaluated over a range of sizes  $[L_j, U_j]$  that extends past the range of observed sizes.

As in our previous studies of this system (Adler et al. 2010, 2012), the kernel is constructed from the fitted survival ( $S$ ), growth ( $G$ ), and recruitment ( $R$ ) models

$$k_j(v, u, \vec{n}) = S_j(u, \vec{n})G_j(v, u, \vec{n}) + R_j(v, u, \vec{n}) \quad (3)$$

(see Appendix S1 for a complete description). The IPM uses a mean-field approximation for neighborhood competition that captures the essential features of spatially explicit local interactions; in Appendix S1: Section 1.5 we explain how the spline competition kernels were used to compute neighborhood competition as a function of individual size for the IPM.

We simulated the IPM to estimate equilibrium abundances under three different scenarios. In all of these simulations, we set the spatial random effects to zero, meaning that we are simulating dynamics on a hypothetical, average site. To incorporate temporal environmental variation, at each time step, we randomly selected one of the 30 sets of random year effects for the survival, growth, and recruitment models ("kernel selection" in the terminology of Metcalf et al. 2015). For each scenario, we initialized all species with very low cover and ran the model for 500 time steps (years) to allow the community to reach its equilibrium (i.e., its steady-state pattern of fluctuations in response to environmental variability). We calculated the average steady-state cover for each species by simulating an additional 2,000 time steps and averaging each species' cover over this period. The first of the three scenarios featured all four species and our baseline model for control plots (no removal effects). For the second scenario, we simulated our experimental removals by setting either the cover of *A. tripartita* to zero or the cover of all three grass species to zero, and then simulating the dynamics of the remaining species as under the first scenario. This simulation projects competitive release following removal according to the baseline model (no removal effects). For the third scenario, we again set *A. tripartita* or grass cover to zero, but this time we projected the model using the vital rate regressions that include removal effects. If our baseline model successfully predicts competitive release observed in the removal experiment, then the results of the second and third scenarios should be similar. To the extent that the baseline model underestimates (or overestimates) the actual competitive release, the results of the second and third scenarios will differ. As with the IBM,

we repeated these simulations using maximum removal treatment effects represented by parameter values at the 5% or 95% credible limit farthest from zero.

RESULTS

During the period of our experiment, annual water-year (October–September) precipitation fell within the 5–95% quantiles of the long-term mean, although more years fell below the mean than above it (Appendix S1: Fig. S4a). However, annual mean temperatures were consistently warm relative to the long-term mean, and two years were warmer than the 95% quantile (Appendix S1: Fig. S4b).

Cover trends

Trends in cover by species and removal treatment were variable. For example, grass removal had little effect on the cover of *A. tripartita* (Fig. 1), but shrub removal did appear to cause an increase in the cover of *P. spicata*. However, when we tested the effects of removal treatments on year-to-year population growth rates ( $\log(\text{Cover}_{t+1}/\text{Cover}_t)$ ) at the quadrat level, none of the removal treatment effects were significantly different from zero (Appendix S1: Table S1). For *P. spicata*, the 95% credible interval only just contained zero

(95% CI:  $-0.01, 0.38$ ), providing marginal evidence for a positive effect of removal on population growth. Fitting with weighted least squares to address the high variance in residuals for plots with very low cover indicated even less evidence for a positive response of *P. spicata* to removal. However, these treatment comparisons do not directly address our questions about whether responses to removal were adequately predicted by models fit to observational data.

Vital rates

Our vital rate models take into account the effects of local species composition and should therefore be able to explain the effect of removals on survival, growth, and recruitment without any additional parameters. As in our previous studies using only the historical data, the neighborhood competition coefficients in the fitted models imply that interspecific competition is much weaker than intraspecific competition, in many cases by one or two orders of magnitude (Appendix S1: Table S13). The intraspecific interaction effects are large and have credible intervals that do not overlap zero, while the interspecific coefficients are small and the majority have credible intervals that do overlap zero. If the removal treatment effects that we added to these models are statistically significant, it would indicate an effect of removals either larger or

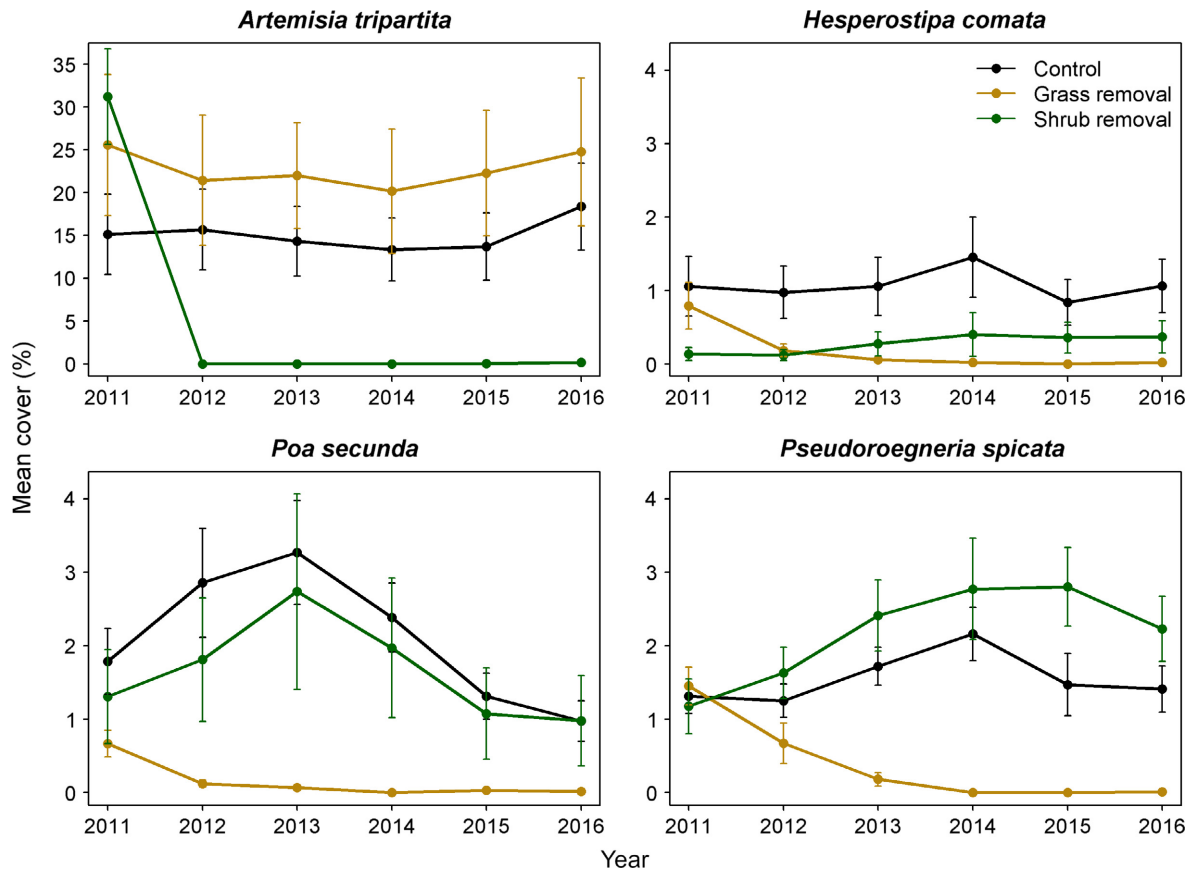


FIG. 1. Cover trends by treatment for the four dominant species. Note the difference in y-axis scale between the canopy cover of *A. tripartita* and the basal cover of the dominant grasses. “Shrub removal” refers to removal of *A. tripartita*, and “Grass removal” refers to removal of all perennial grasses. The cover of species in treatments from which they are removed (e.g., *A. tripartita* cover in shrub removal plots) is included to show values at the start of the experiment in 2011 and the efficacy of the removal treatments. Source file: treatment\_trends\_removals.r.

smaller than the effect predicted by a model based on data from the control plots. For the survival models, the 95% credible intervals on the removal effects overlapped zero for all four species (Fig. 2; Appendix S1: Tables S2–S5), meaning that the survival probabilities of plants in removal plots can be predicted based on data from control plots. For the growth models, the 95% credible intervals on the removal effects were positive for *P. secunda* and *P. spicata*, meaning that individuals of these species grew more rapidly in removal plots than controls even after accounting for neighborhood interactions (Fig. 2; Appendix S1: Tables S6–S9). At the recruitment stage, only *A. tripartita* showed a significant response to removals, and it was negative (Fig. 2; Appendix S1: Table S12).

The positive removal effects on *P. secunda* and *P. spicata* growth suggest that models based on observational data may underestimate competitive release. We conducted additional analyses to further explore this hypothesis. First, we asked whether increases in grass growth following removal were greatest in quadrats or neighborhoods where we removed the most *A. tripartita*. We did this by taking residuals from the fitted growth model that included removal effects, and plotting the residuals for plants in removal quadrats against the pre-treatment, quadrat-level cover of *A. tripartita*. No species showed the positive relationship that would be expected (Fig. 3). In fact, *Poa secunda* residuals were significantly lower in quadrats with high pre-treatment cover of *A. tripartita*: individuals in high pre-treatment cover quadrats grew more slowly than predicted by the growth model with removal effects. We repeated this analysis at the individual plant level for grasses present in 2011. We calculated pre-treatment crowding from *A. tripartita* for each of those individuals, then plotted their growth residuals against crowding for all years in which each plant survived. We found no significant relationships between growth residuals and pre-treatment crowding from *A. tripartita* (Appendix S1: Fig. S5).

Second, we evaluated the possibility that removal effects might be transient. We removed aboveground plant material, but roots of the removed plants remained in the ground potentially creating a flush of mineral nutrients as they decompose. We tested for a transient removal effect by adding a year-by-treatment interaction. For both *P. secunda* and *P. spicata*, we found no sign of a decrease in the removal effect over time (Appendix S1: Tables S10 and S11). In fact, for *P. secunda*, the removal effects increased over time.

Third, responses to removals might be poorly predicted by models based on observational data if the removals reduce plant densities (crowding) below the range of historical variation, or if the removals result in combinations of neighborhood-scale densities that deviate from naturally occurring patterns of correlation. However, the distribution of crowding observed within control plots easily spans the crowding values observed in removal plots (Appendix S1: Fig. S6), and we failed to find evidence for strong interspecific covariances in crowding (Appendix S1).

Finally, we evaluated pre-existing differences in community composition in removal and control plots. In selecting removal plots, our criteria required candidate plots to include at least a certain amount of our target removal species. Because we adopted the historical plots as our controls, the same criteria were not applied (a control plot might not contain any *A. tripartita*). To quantify resulting differences in pre-treatment (2011) composition, we compared neighborhood crowding at the individual plant level (the  $W_{jms}$  of the growth regressions) across treatments. *A. tripartita* crowding and crowding by all subdominant shrub and grass species appeared greater in removal plots than in the controls, whereas crowding by *H. comata* was lower in removals than controls, though it was absent from most *P. spicata* neighborhoods in both treatments (Appendix S1: Fig. S7). These pre-existing differences in composition raise the possibility that plots selected for the removal treatments might

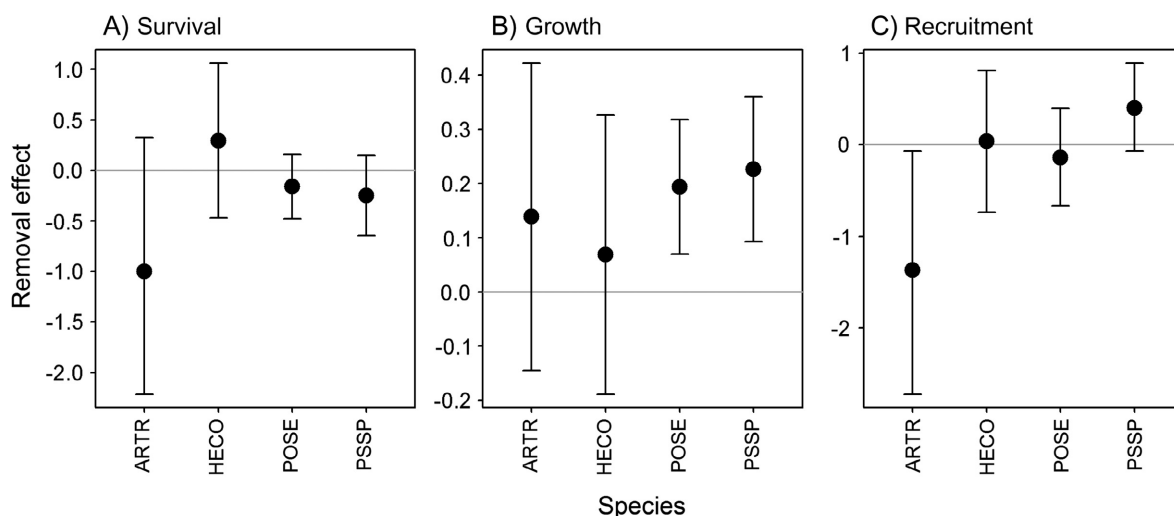


FIG. 2. Effects of removal treatments on the (A) survival, (B) growth, and (C) recruitment of the four dominant species. Circles show the estimated coefficient of the removal effect, bars show 95% credible intervals. The removal effect shows response of *A. tripartita* (ARTR) to perennial grass removal and response of the perennial grasses to *A. tripartita* removal, after accounting for local neighborhood interactions. The other species codes are HECO, *Hesperostipa comata*; POSE, *Poa secunda*, and PSSP, *Pseudoroegneria spicata*. Source file: treatment\_test\_figure.r.

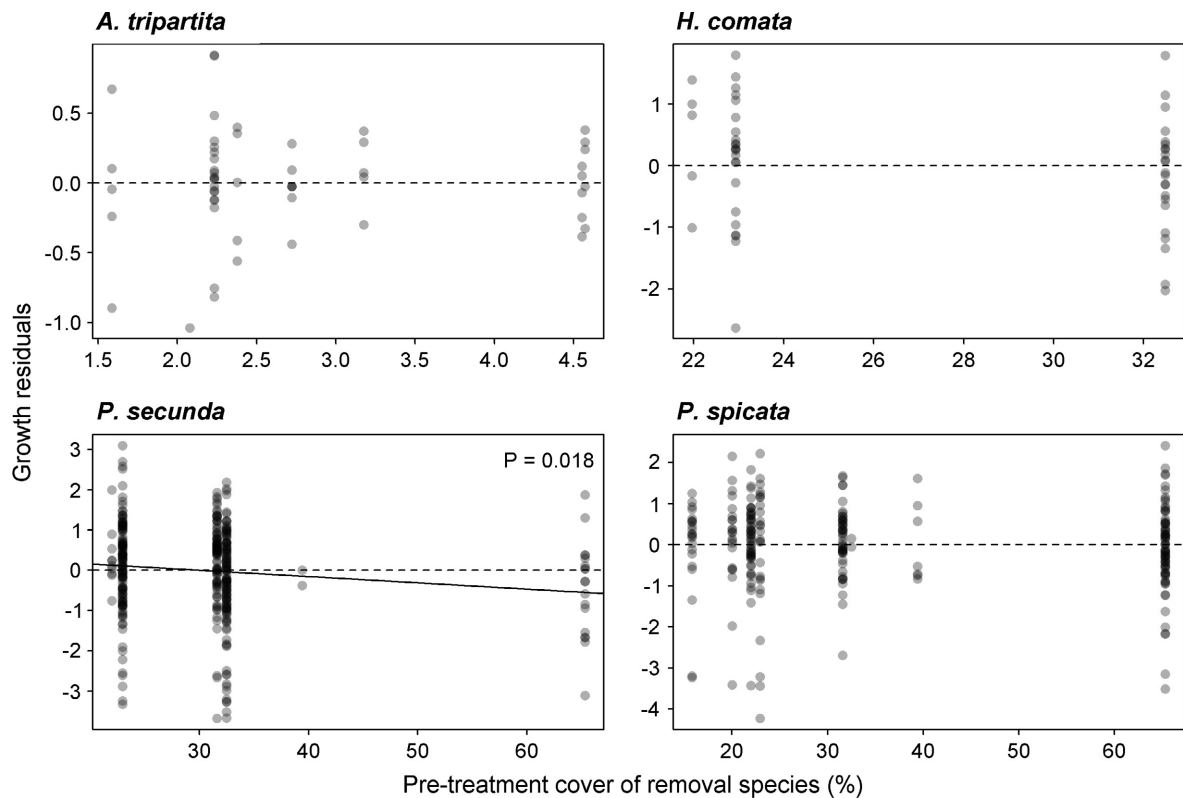


FIG. 3. The relationship between residuals from the fitted growth model and quadrat-level pre-treatment cover of the removal species (all grasses removed in quadrats where we studied *A. tripartita*, and *A. tripartita* removed in the quadrats where we studied the grasses). Linear regressions are shown only when statistically significant. Source file: growthgrowth\_residuals\_by\_quad.R.

differ from the control plots in some unobserved covariate that could influence plant composition and performance. On the other hand, some of the control plots that currently contain zero *A. tripartita* had high *A. tripartita* in the 1900s, and vice versa. If abiotic variation, such as soil depth, were the primary control of variation in *A. tripartita* cover, such large changes in cover over time would be unlikely.

#### Population-level responses

To quantify the short-term population-level consequences of removals, we used an IBM to project the cover of each species in each quadrat one year ahead using versions of the model with and without removal treatment effects. Plotting observed vs. predicted cover for each quadrat in each year (Fig. 4) shows points clustered around the 1:1 line, indicating reasonable accuracy and little bias for both control and removal quadrats. For *H. comata*, differences in predicted cover of removal plots between the baseline and removal models were very small. For the other three species, the removal model predicted higher cover in removal plots than the baseline model predicted.

Time series of predicted and observed cover, averaged over quadrats, show similar patterns. Predicted population size closely tracked observed population size (Fig. 5) in most years for all species, but the difference between predictions from the baseline and removal models varied among species. For *H. comata*, differences in predictions from the two models were virtually indistinguishable. For *A. tripartita*,

*P. secunda* and especially *P. spicata*, the model with removal effects consistently predicted higher population growth than the baseline model (Fig. 5). Using the version of the model with maximum removal treatment effects, the patterns were qualitatively similar, with a somewhat greater contrast between predictions of the baseline and removal models for *P. secunda* and *P. spicata* (Appendix S1: Fig. S8).

To quantify the long-term population-level consequences of permanent competitor removals, we used an IPM to project the long-term average cover of each species. The IPM predicts little effect of *A. tripartita* removal on the cover of any grass species when we simulate the baseline model with no removal treatment coefficients (compare gray and blue boxes in Fig. 6). In fact, projected average cover of all three grass species declines slightly following *A. tripartita* removal. Using a model with removal treatment coefficients, the IPM projects that *A. tripartita* removal will increase equilibrium cover, relative to the baseline scenario, by a factor of 1.4, 1.5, and 3.0 for *H. comata*, *Poa secunda*, and *P. spicata*, respectively (Fig. 6). Running a version of the model with maximum removal treatment effects (removal coefficients set to the 95% credible limit furthest from zero), resulted in a much larger projected increase in *H. comata*, reflecting its small sample size and corresponding high uncertainty, but little change in the projections for *Poa secunda* and *P. spicata* (Appendix S1: Fig. S9).

The equilibrium cover of *A. tripartita* simulated by the IPM (5.4%) is consistent with observed cover during the historical data collection period (6.6% from 1931 to 1956), but

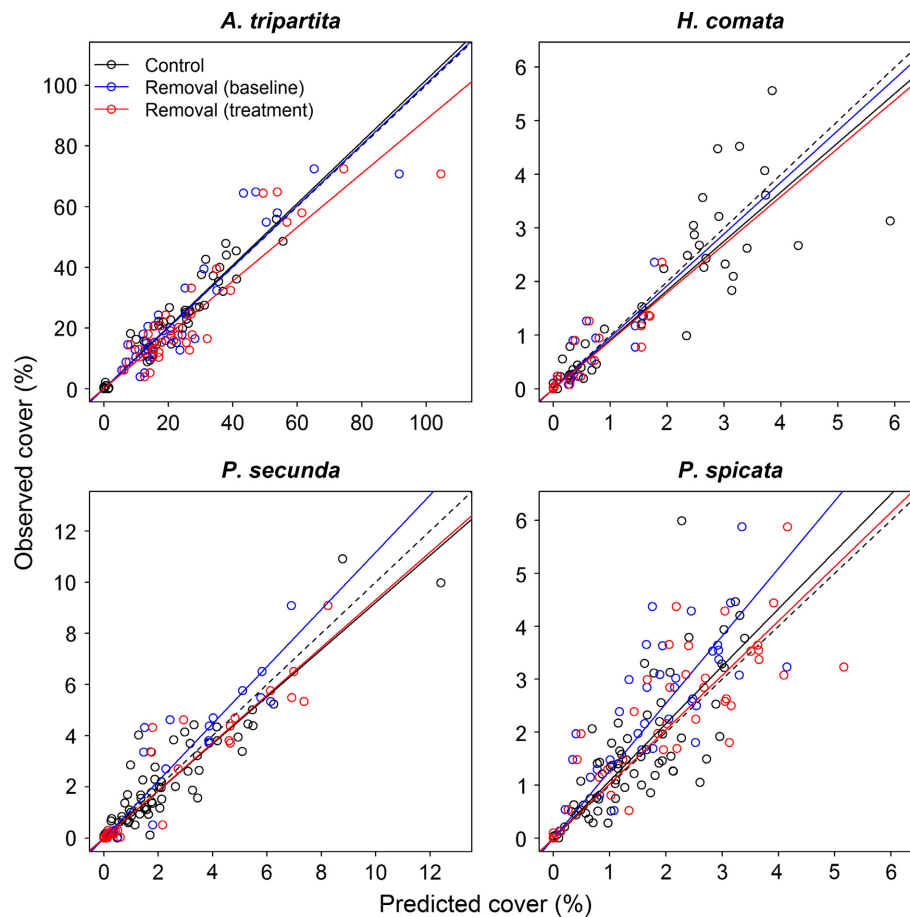


FIG. 4. The relationship between observed and predicted cover for each quadrat, year, and species. The 1:1 line is dashed, the solid black line shows the trend for control plots with predictions from the baseline IBM model (black symbols), the blue line shows the trend for removal plots with predictions from the baseline model (blue symbols), and the red line shows the trend for removal plots with predictions from the removal model (red symbols). Source file: `ibm\summarize_sims1step.r`.

is low compared to contemporary observed cover (15–25%, Fig. 1). Could the small competitive release of the grasses following simulated *A. tripartita* removal simply reflect the low simulated *A. tripartita* cover? To explore this possibility, we increased *A. tripartita* cover in the baseline simulation by increasing its average fecundity. The increased fecundity led to a doubling of *A. tripartita* average cover, but had essentially zero effect on the cover of the grasses (changes of <0.1% cover), implying that, even if shrub cover were higher, its removal would still have little effect on the grasses.

The baseline model predicts more dramatic competitive release of *A. tripartita* following grass removal: mean cover increases from 4% to over 11%. Adding removal treatment effects dramatically alters the outcome, with *A. tripartita* decreasing toward extinction. However, these outcomes both are driven by parameters that are estimated with very high uncertainty (Appendix S1: Table S12). The baseline model predicts an increase following grass removal because of the relatively strong negative effects of *Poa secunda* and *H. comata* on *A. tripartita* recruitment (Appendix S1: Table S12). The estimated effect of *Poa* is  $-0.50$ , but the standard deviation of the estimate is 0.28 and the 95% credible interval stops just short of zero; the 95% credible interval for the *H. comata* effect includes values less than one-half of

the point-estimate. The grass removal treatment effect, which leads to predicted extinction, has a credible interval that broadly overlaps zero (Appendix S1: Table S12). *A. tripartita* recruitment parameters have exceptionally high uncertainty, presumably because *A. tripartita* recruitment is episodic (see *Discussion*), and this produces extreme uncertainty about its long-term response to grass removal.

## DISCUSSION

Our population model based on historical, observational data revealed weak interspecific interactions (Adler et al. 2010, Chu and Adler 2015). As a result, the model predicts very little competitive release among the four dominant species in a sagebrush steppe. Because the model was based solely on observational data, we were suspicious of its ability to accurately describe the outcome of interspecific interactions, so we conducted a removal experiment to test it. Our approach represents a strong test of a surprising prediction for two reasons. First, the prediction was already published, so the experiment gave us a very real opportunity to be wrong in public. Second, testing an observationally based model with an experimental manipulation, not just additional observations, further increases the potential to



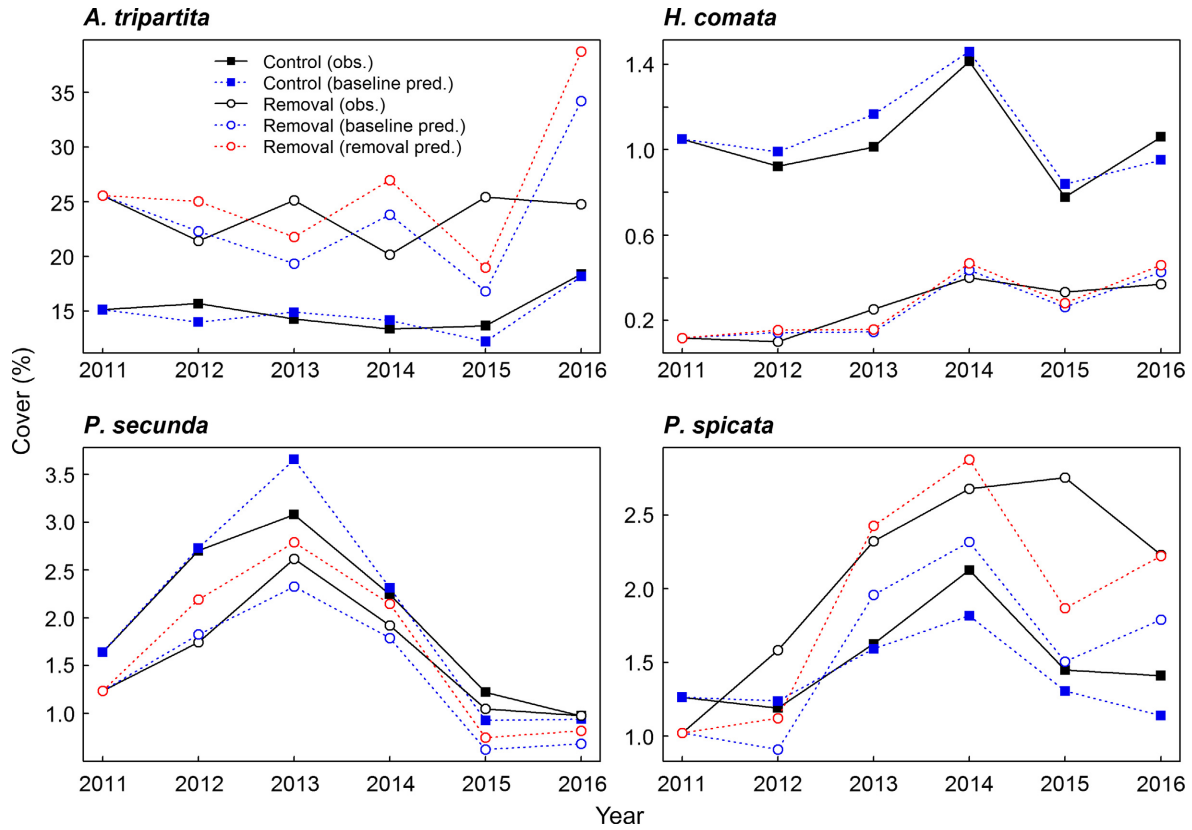


FIG. 5. Observed and predicted cover for each of the four modeled species. Solid black lines show observed cover averaged across control (closed symbols) and removal treatment (open symbols) quadrats. Blue lines and symbols show one-step-ahead predictions from the baseline IBM (no removal treatment coefficients), averaged across quadrats. Red lines and symbols show one-step-ahead predictions from an IBM that includes removal treatment coefficients. For *H. comata*, lines showing predictions for the baseline and removal effect models are virtually indistinguishable. Source file: `ibm\summarize_sims1step.r`.

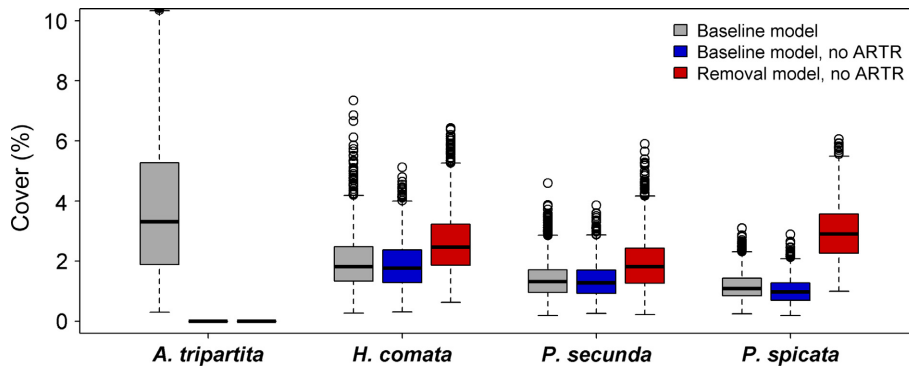


FIG. 6. Equilibrium cover of the four dominant species simulated by the IPM. Boxplots show interannual variation reflecting random year effects. Gray boxes show cover simulated by a model with *A. tripartita* present and no removal treatment coefficients, blue boxes show results from the same model but with *A. tripartita* cover set to zero (a species removal), and red boxes show results from a model with *A. tripartita* set to zero and removal treatment coefficients included. Source file: `ipm\IPM-figures.r`.

confront the model with a novel situation outside the historical range of variation. We expected this study to provide clear evidence for the success or failure of our previous model’s predictions. Unfortunately, our results do not provide that clarity. In the sections below, we first summarize the evidence that our experiment revealed greater competitive release than our baseline model predicted, then explain weaknesses in the experimental results that complicate the

interpretation, and end by discussing general lessons for the study of interspecific interactions.

*Evidence for greater competitive release than predicted*

We evaluated response to removals in three ways: vital rates, short-term population growth rates, and changes in long-term equilibrium cover. Two species, *Poa secunda* and

*P. pseudoroegneria*, showed significant, positive responses to *A. tripartita* removal at the level of vital rates. Individuals of these species grew 21% and 24% faster, respectively, in removal plots than controls, after accounting for variation in local neighborhood composition. In other words, our baseline model appeared to underestimate the competitive release of these species following *A. tripartita* removal.

The elevated growth of *Poa secunda* and *P. pseudoroegneria* individuals translated into greater short-term population growth rates. For these species, the individual-based model with removal effects predicted higher cover than the baseline version of the model. The differences in predicted cover of roughly 20% (Fig. 5) are comparable to the differences in the individual growth rates, reflecting the fact that at these short time scales changes in individual size influence cover more than survival or recruitment. For *A. tripartita*, the magnitude of the difference between predictions from the baseline and removal effects models was similar, further emphasizing the short-term influence of growth compared to survival and recruitment, which responded negatively to removals in this species.

Long-term projections using the integral projection model showed even larger responses to removal. For the three grasses, projected equilibrium cover in the absence of *A. tripartita* is 1.4- to 3-fold higher for the removal effects model compared to the baseline model. While the absolute changes in basal cover are small, these relative changes are substantial, and seem to present compelling evidence that competitive release is stronger than predicted by our baseline model.

#### *Weaknesses in the experimental evidence*

Although the growth regressions for *Poa secunda* and *P. pseudoroegneria* and the long-run projections for all three grass species suggest stronger competitive release than our baseline model predicted, a variety of other evidence makes it harder to confidently reject those baseline model predictions. Only 2 of the 12 vital rate tests indicated greater competitive release than predicted. Granted, small samples sizes contributed to greater uncertainty for *A. tripartita* and *H. comata*, but even for *Poa secunda* and *P. pseudoroegneria* the increases in growth were small relative to other sources of variation. For example, using quadrat-level cover data, we could not detect treatment differences (Appendix S1: Table S1).

The strongest evidence that competitive release of grasses following *A. tripartita* removal was greater than predicted comes from the long-term projections of the IPM. These projections effectively make removal treatment effects permanent; the removal effects on vital rates observed during the experiment are locked in for the entire 2,500 time step simulation. Furthermore, these projections ignore parameter error because we wanted to highlight the population consequences of the removal effects. Given these strong assumptions, the projections should be viewed with some caution.

For *A. tripartita*, the long-term IPM projections actually show extinction following grass removal. However, this result primarily reflects the influence of recruitment parameters that are highly uncertain (see *Results*). This high uncertainty results from *A. tripartita* recruitment being low and very episodic (roughly half of the total recruits in the 22 yr of historical data were observed in two of those years, and

only 28 *A. tripartita* recruits have been observed since we resumed annual censuses in 2007). Because of this uncertainty, we put more weight on the direct observations and short-term projections showing little effect of grass removal on *A. tripartita* populations.

Another reason we view the growth results with suspicion is that increases in growth were no larger in individual neighborhoods or quadrats with higher pre-treatment *A. tripartita* cover. If *A. tripartita* removal is releasing grasses from competition, shouldn't this effect be larger for grass genets in areas with high *A. tripartita* cover? The absence of this relationship (Fig. 3; Appendix S1: Fig. S5) raises questions about the mechanism by which *A. tripartita* removal promotes grass growth. Perhaps *A. tripartita*'s roots extend far enough beyond its canopy to compete for resources with grasses located in interspaces. While this might explain the lack of the relationship at the neighborhood scale, it seems unlikely to explain the quadrat-level results. An alternative explanation involves pre-treatment differences in composition between control plots and the *A. tripartita* removal plots, which had higher cover of *A. tripartita* (see 2011 means in Fig. 1) and higher cover of subdominant grasses and shrubs (Appendix S1: Fig. S7). These pre-treatment differences could reflect subtle differences in edaphic conditions that could promote *P. spicata* growth. However, since the 1900s, many quadrats have undergone dramatic changes in *A. tripartita* cover, arguing against viewing *A. tripartita* cover as a sensitive indicator of edaphic conditions. A final possibility is that our removal treatment influenced grass growth through some other mechanism, unrelated to competition. For example, loss of shrub cover might benefit grasses by reducing small mammal herbivory (Orrock et al. 2010). Without a better understanding of the mechanism driving increased grass growth in the shrub removal quadrats, we are hesitant to dramatically adjust our prior understanding gained from models fit to decades of observational data.

#### *Lessons for studying interspecific interactions*

A variety of approaches are available for studying interspecific interactions among plants, and each has its advantages and disadvantages. Observational datasets have the advantage of realism: they reflect all the processes and interactions that drive community dynamics across a range of spatial and temporal scales. Their disadvantage is weak inference. Distinguishing between correlation and causation is challenging, and simulations of a model fit to observational data may rely heavily on extrapolation. Experiments, such as response surface designs (Inouye 2001), offer an alternative approach to studying interactions at the population scale. Their chief advantage is strong inference about causality. But they sacrifice realism: experiments are typically implemented at fine spatial and temporal scales that exclude potentially important processes. Finally, we can study interactions mechanistically by describing patterns of resource use (Silvertown et al. 1999, Dyzinski and Tilman 2007). Mechanistic studies have the virtue of focusing directly on the processes of interest, but rarely address more than one limiting resource at a time.

Our previous work was based purely on observational data, so we were particularly concerned about the

extrapolation issue. Specifically, our model projects the performance of individual plants growing in the absence of neighbors. Although the observational data set seems to include such situations (Appendix S1: Fig. S6), the data-driven competition kernels we use to calculate neighborhood crowding are quite narrow: they are heavily weighted toward nearby neighbors and essentially ignore the influence of more distant neighbors (Appendix S1: Fig. S3). However, if our competition kernels underestimate the strength of long-range interactions, then our model might underestimate the true intrinsic growth rates that would occur when both near and distant neighbors are removed. We conducted the removal experiment to address this kind of concern, with the hope of gaining the advantages offered by multiple, complementary approaches. The positive responses of *Poa secunda* and *P. pseudoroegneria* to removals suggests our observational approach might have underestimated intrinsic growth rates after all. But removal experiments fall into a gray area between an observational study and a randomized, controlled experiment. They are an experimental manipulation, but they rely on natural spatial variation in composition; we can only remove plants from locations where they naturally occur. The reliance on natural variation weakens inference, leaving nagging questions about pre-existing differences among quadrats. We could have avoided those concerns by conducting a response surface experiment, planting our study species in randomly selected locations.

We also learned that partial removals, rather than complete removal treatments, could have helped to distinguish between competition-mediated and other effects of removal treatments (e.g. Appendix S1: Fig. S2). Partial removals, or experimental gradients in densities, would give us inference about the competition coefficients themselves, whereas our design lumped all treatment effects into the intercept terms. Complete removals might be adequate for testing the presence of competition, but do not work as well for quantifying per capita competitive effects. Furthermore, while our experiment focused on a few interspecific interactions, ultimately it is the difference in the strength of intra- and interspecific interactions that determines the stability of coexistence and the degree to which the dynamics of co-occurring species are coupled. Next time we will also include intraspecific removal treatments.

Our difficulty interpreting our results also made us appreciate the value of mechanistic insight. Even if the removal experiment had perfectly validated the predictions of our phenomenological baseline model, we suspect we would still be asking ourselves, how can interspecific interactions really be so weak? The only way to rigorously answer that question is to study the factors—resources, natural enemies, environmental constraints—limiting population growth.

Although our results point to limitations of models based on long-term observational data, they also give some cause for optimism about population forecasting. Crone et al. (2013) found that matrix projection models for plants often failed to forecast population trends due to environmental differences between the data collection and forecast periods. By including random year-effects to account for year-to-year environmental variation, our population models performed quite well at forecasting between-year population changes (Fig. 4).

## CONCLUSION

Experimental removal of *A. tripartita* led to greater release of grasses than our previous model predicted. However, the mechanisms driving this release remain unclear, and make us hesitant to conclude that our model based on observational data underestimates the strength of interspecific competition. The results do increase our interest in testing model predictions using even more tightly controlled experiments as well as direct study of resource-use.

## ACKNOWLEDGMENTS

P. B. Adler and A. R. Kleinhesselink designed the experiment and supervised data collection with the help of J. B. Taylor; P. B. Adler, S. P. Ellner, G. Hooker, and B. Teller analyzed data and developed models; P. B. Adler wrote the first draft of the manuscript; and all authors contributed to editing. Funding was provided by NSF grants DEB-1353078 and DEB-1054040 to P. B. Adler and DEB-1353039 to S. P. Ellner and G. Hooker, and by the Utah Agricultural Experiment Station, which approves this contribution as journal paper number 9033. The USDA-ARS Sheep Experiment Station generously provided access to historical data and the field experiment site. We thank Robin Snyder, Tom Miller, and two anonymous reviewers for suggestions that improved the manuscript.

## LITERATURE CITED

- Aarssen, L. W., and G. A. Epp. 1990. Neighbour manipulations in natural vegetation a review. *Journal of Vegetation Science* 1:13–30.
- Adler, P. B., S. P. Ellner, and J. M. Levine. 2010. Coexistence of perennial plants: an embarrassment of niches. *Ecology Letters* 13:1019–1029.
- Adler, P. B., H. J. Dalgleish, and S. P. Ellner. 2012. Forecasting plant community impacts of climate variability and change: When do competitive interactions matter? *Journal of Ecology* 100:478–487.
- Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67:1–48.
- Blaisdell, J. P. 1958. Seasonal development and yield of native plants on the upper Snake River plains and their relation to certain climatic factors. Technical Bulletin No. 1190, USDA, Washington, DC.
- Belsky, A. J. 1992. Effects of grazing, competition, disturbance and fire on species composition and diversity in grassland communities. *Journal of Vegetation Science* 3:187–200.
- Brooker, R. W., et al. 2008. Facilitation in plant communities: the past, the present, and the future. *Journal of Ecology* 96:18–34.
- Chesson, P. 2000. Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics* 31:343–366.
- Chu, C., and P. B. Adler. 2015. Large niche differences emerge at the recruitment stage to stabilize grassland coexistence. *Ecological Monographs* 85:373–392.
- Cook, C. W., and C. E. Lewis. 1963. Competition between big sagebrush and seeded grasses on foothill ranges in Utah. *Journal of Range Management* 16:245–250.
- Crone, E. E., et al. 2013. Ability of matrix models to explain the past and predict the future of plant populations. *Conservation Biology* 27:968–978.
- Dybzinski, R., and D. Tilman. 2007. Resource use patterns predict long-term outcomes of plant competition for nutrients and light. *American Naturalist* 170:305–318.
- Gilbert, B., R. Turkington, D. Srivastava, A. E. O. J. Schmitz, and E. D. L. DeAngelis. 2009. Dominant species and diversity: linking relative abundance to controls of species establishment. *American Naturalist* 174:850–862.
- Grime, J. P. 1979. *Plant strategies and vegetation processes*. John Wiley & Sons, New York, New York, USA.
- Hairston, N. G. 1980. The experimental test of an analysis of field distributions: competition in terrestrial salamanders. *Ecology* 61:817–826.

- He, Q., M. D. Bertness, and A. H. Altieri. 2013. Global shifts towards positive species interactions with increasing environmental stress. *Ecology Letters* 16:695–706.
- Houlahan, J. E., S. T. McKinney, T. M. Anderson, and B. J. McGill. 2017. The priority of prediction in ecological understanding. *Oikos* 126:1–7.
- Hubbell, S. P. 2001. *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, New Jersey, USA.
- Inouye, B. D. 2001. Response surface experimental designs for investigating interspecific competition. *Ecology* 82:2696–2706.
- Jutila, H. M., and J. B. Grace. 2002. Effects of disturbance on germination and seedling establishment in a coastal prairie grassland: a test of the competitive release hypothesis. *Journal of Ecology* 90:291–302.
- Keddy, P. A. 1989. Effects of competition from shrubs on herbaceous wetland plants: a 4-year field experiment. *Canadian Journal of Botany* 67:708–716.
- Lauenroth, W. K., and P. B. Adler. 2008. Demography of perennial grassland plants: survival, life expectancy and life span. *Journal of Ecology* 96:1023–1032.
- Leps, J. 1999. Nutrient status, disturbance and competition: an experimental test of relationships in a wet meadow copy. *Journal of Vegetation Science* 10:219–230.
- Metcalf, C. J. E., S. P. Ellner, D. Z. Childs, R. Salguero-Gmez, C. Merow, S. M. McMahon, E. Jongejans, and M. Rees. 2015. Statistical modelling of annual variation for inference on stochastic population dynamics using Integral Projection Models. *Methods in Ecology and Evolution* 6:1007–1017.
- Mouquet, N., et al. 2015. Review: Predictive ecology in a changing world. *Journal of Applied Ecology* 52:1293–1310.
- Mueggler, W. F., and J. P. Blaisdell. 1958. Effects on associated species of burning, rotobating, spraying, and railing sagebrush. *Journal of Range Management* 11:61–66.
- Orrock, J. L., R. D. Holt, and M. L. Baskett. 2010. Refuge-mediated apparent competition in plant–consumer interactions. *Ecology Letters* 13:11–20.
- Petchey, O. L., et al. 2015. The ecological forecast horizon, and examples of its uses and determinants. *Ecology Letters* 18:597–611.
- Robertson, J. H. 1947. Responses of range grasses to different intensities of competition with sagebrush (*Artemisia tridentata* Nutt.). *Ecology* 28:1–16.
- Sala, O. E., R. A. Golluscio, W. K. Lauenroth, and A. Soriano. 1989. Resource partitioning between shrubs and grasses in the Patagonian steppe. *Oecologia* 81:501–505.
- Silvertown, J. 2004. Plant coexistence and the niche. *Trends in Ecology and Evolution* 19:605–611.
- Silvertown, J., M. Dodd, D. Gowing, and J. Mountford. 1999. Hydrologically defined niches reveal a basis for species richness in plant communities. *Nature* 400:61–63.
- Smith, M. D., and A. K. Knapp. 2003. Dominant species maintain ecosystem function with non-random species loss. *Ecology Letters* 6:509–517.
- Tilman, D. 1982. *Resource competition and community structure*. Princeton University Press, Princeton, New Jersey, USA.
- Tredennick, A. T., C. de Mazancourt, M. Loreau, and P. B. Adler. 2017. Environmental responses, not species interactions, determine synchrony of dominant species in semiarid grasslands. *Ecology* 98:971–981.
- Zachmann, L., C. Moffet, and P. B. Adler. 2010. Mapped quadrats in sagebrush steppe: long-term data for analyzing demographic rates and plant–plant interactions. *Ecology* 91:3427.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at <http://onlinelibrary.wiley.com/doi/10.1002/ecy.2363/supinfo>

## DATA AVAILABILITY

Data associated with this study are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.96dn293>.