

# Can historical data predict population responses to climate change experiments?

Andrew R. Kleinhesselink<sup>\*a</sup>, Peter B. Adler<sup>a</sup>, Andrew Tredennick<sup>a</sup>, Caitlin Andrews<sup>b</sup> and John Bradford<sup>b</sup>

<sup>a</sup>Department of Wildland Resources and the Ecology Center, Utah State University, Logan, Utah

<sup>b</sup>US Geological Survey, Southwest Biological Science Center, Flagstaff, Arizona

Last compile: December 22, 2016

**Keywords:** Climate change, demographic models, rain-out shelter.

**Running title:** Predicting climate response

**Article type:** Letter

**Authorship statement:** ARK and PBA designed the experiment and supervised data collection. AT helped with the data analysis. JB and CA helped with the soil moisture model.

---

<sup>\*</sup>Corresponding author. Department of Wildland Resources and the Ecology Center, Utah State University, Logan, Utah Email: arklein@aggiemail.usu.edu

## **1 Abstract**

2 Climate is an important driver of population ecology however there have been few tests of whether  
3 observational data correlating population performance with climate variation can be used to predict  
4 how populations respond to climate change experiments. We tested whether longterm observational  
5 data relating the effects of climate variation to the survival, growth and recruitment of four com-  
6 mon plants in a sagebrush steppe could be produce accurate predictions for how each species would  
7 respond to a drought and irrigation experiment. We used rainout shelters and automatic sprin-  
8 klers to manipulate the soil moisture in sixteen new plots established at the same field site as  
9 the longterm observational data. After running the experiment for five years, we analyzed how  
10 the growth, survival and recruitment of each species responded to the experimental drought and  
11 irrigation treatments. We made predictions using two types of models fit to the observational data  
12 collected prior to the experiment: a no climate model that only included the effects of plant size  
13 and competition on plant performance in each plot in each year; and a climate model that also  
14 included the effects of three seasonal soil moisture variables. We then used these models to predict  
15 the performance of each species during the course of the experiment. We also generated one step a  
16 head predictions of population size in each experimental plot using an individual based population  
17 model. We compared predictions made by the the no climate and climate models to the observed  
18 experimental responses. Over the course of the experiment, average cover of the perennial grasses,  
19 *Hesperostipa comata* and *Pseudoroegneria spicata* declined significantly in the drought treatment.  
20 At the level of individual vital rates, experimental drought reduced the survival of *Hesperostipa*  
21 *comata* and *P. spicata* and the growth of the grass *Poa secunda*. In contrast, drought increased  
22 the growth of the shrub *A. tripartita*. The climate model better predicted the response to the  
23 experiment than the no climate model in six out of twelve cases. Across all species and vital rates,  
24 there was a strong positive correlation between the observed effects of the experiment and the  
25 effects predicted by the climate model fit to the observational data. At the population-level, the  
26 climate model made better predictions of the changes in species cover than the no climate model  
27 for *P. secunda* and *P. spicata*. Observational climate data held valuable information for predicting  
28 species' responses to a climate change experiment in this ecosystem. Treatment responses often  
29 matched the direction of predicted responses even when the effects were not significant. We were

30 better able to predict species' responses to the drought treatment than to the control and irrigation  
31 treatments. This suggests that soil moisture is an important factor in predicting the population  
32 dynamics of these species but only when water is truly limiting.

### 33 Introduction

34 Climate is one of the most powerful drivers of changes in species abundance across space and time  
35 (Post and Forchhammer, 2002; Davis and Shaw, 2001; Walther et al., 2002). The effects of climate  
36 on populations and ecosystems are most apparent at the largest scales: climate determines the  
37 distribution of ecosystems (Whittaker, 1975), treelines (Kerner, 2012) and the range limits of  
38 many species (Parmesan and Yohe, 2003; Davis and Shaw, 2001). Understanding and predicting  
39 the effects of climate on populations is an increasingly important goal if we are to anticipate the  
40 effects of climate change on earth's ecosystems (Tredennick et al., 2016; Petchey et al., 2015; Ehrlin  
41 et al., 2016; Teller et al., 2016).

42 Ecologists often resort to one of two methods for predicting the effects of future climate  
43 change on populations and communities: they may use experiments to manipulate aspects of  
44 climate directly and observe the response of populations (Elmendorf et al., 2015; Knapp et al., 2016;  
45 Compagnoni and Adler, 2014); or they may use longterm observational data on species performance  
46 and abundance collected over many years and relate this to ambient annual variation in climate  
47 (Koons et al., 2012; Lunn et al., 2016; Dalgleish et al., 2010; Jenouvrier et al., 2009). The strength of  
48 the experimental approach is in the stronger inference that comes from manipulating some aspects  
49 of climate while controlling for other factors; for instance knowing that loss of snow cover, and not  
50 necessarily changes in soil temperature or moisture are the factors causing a species performance to  
51 change with warming (Compagnoni and Adler, 2014). It also allows for the creation of conditions  
52 that may be more extreme than those observed historically Knapp et al. (2016) but are possible  
53 in the future. However, it is often expensive to control climate at even the smallest scales, and  
54 larger scale climate manipulation is often impossible. Moreover, experimental manipulation can  
55 come with artifacts that may make them less than ideal models for understanding and predicting  
56 the effects of future variation in climate (Wolkovich et al., 2012).

57 Using observational data to predict the effects of climate variation on populations has many  
58 advantages over using experiments. The primary one being the ever increasing availability of long-

59 term ecological data and ever increasingly detailed climate data. Observational studies may also be  
60 the only way to study the effects of climate on large and or migratory species, for which it would  
61 be difficult to manipulate climate (Koons et al., 2012; Jenouvrier et al., 2009; Aubry et al., 2013).  
62 However, there are many potential drawbacks to using essentially correlative relationships between  
63 species performance and climate to predict future species' responses. First, many years of data are  
64 needed to reliably detect climate effects, especially when annual variation in demographic rates is  
65 high (Teller et al., 2016; Gerber et al., 2015). (Teller et al., 2016) estimate that even cutting edge  
66 statistical approaches for fitting relationships between climate and species performance require at  
67 least 20-25 years of independent climate observations before they perform well. Moreover, even  
68 strong correlations between species performance and the climate covariates we choose to include in  
69 our models may not be due to direct causation and may not hold up in future years of observation  
70 (Hilborn, 2016).

71 And yet even overcoming this challenge, predicting the effects of climate change would  
72 face yet another challenge: in many parts of the world, climate change will lead to mean annual  
73 temperatures that will be above the range of historical variation. If species performance responds  
74 non-linearly to temperature, we may be unable to predict how species respond to temperatures  
75 outside the range of observed variation (Doak and Morris, 2010). In addition, climate change will  
76 not just change the average climate, but is also likely to increase the variance in precipitation  
77 and lead to extreme events, which will have their own consequences independent of changes in  
78 climate means (Gherardi and Sala, 2015; Jentsch et al., 2007). Any models based on observations  
79 drawn from the historical range of variation will therefore be extrapolating beyond both the range  
80 of observed averages and variance when used to predict the future (Williams and Jackson, 2007).  
81 Demonstrating that models fit to historical data can accurately predict the effects of experimental  
82 climate manipulations, especially manipulations that generate extreme conditions, would be strong  
83 confirmation that the climate effects they describe are not spurious and will hold even in the novel  
84 conditions of the future Adler et al. (2013).

85 There is abundant evidence that plant performance shows high year to year variation, both  
86 at the level of individual growth, survival and reproduction and total population abundance. Pre-  
87 cipitation is often especially important in driving interannual variation in net primary productivity  
88 among terrestrial plants (Knapp and Smith, 2001; Hsu and Adler, 2014), the annual growth rings

89 in trees (Yang et al., 2014) and smaller plants (Srur and Villalba, 2009; Franklin, 2013), as well as  
90 the germination and reproductive output of many annuals (Venable, 2007).

91 Despite the clear signs that precipitation drives net primary productivity at the ecosystem  
92 level and annual variation in individual growth rate in many plants, there have been relatively few  
93 studies that clearly link observed climate variation to species performance in population models  
94 (Ehrln et al., 2016). Fewer still have tested whether the population-level effects of climate variation  
95 shown in observational data can actually be used to predict the responses of plant populations to  
96 future climate variation. (Adler et al., 2013) showed that population models based on observed  
97 correlations between plant population growth rates and precipitation did have some predictive  
98 power in describing species response to a short-term climate manipulation in a North American  
99 grassland. Three species showed responses to experimentally imposed drought and irrigation that  
100 were well predicted by population models fitted to historical observations. However, another three  
101 species, showed responses to the experimental conditions that were not well predicted by historical  
102 observations. The authors suggested that limited replication in the historical data for two of these  
103 species and changing competitive conditions in the community may have led to the poor predictions.

104 The demography of and competitive interactions between three dominant perennial bunch-  
105 grasses and a shrub in a sagebrush steppe plant community at the US Sheep Experiment Station  
106 near Dubois, Idaho have been described in at least seven different studies since 2010 (Adler et al.,  
107 2010, 2012, 2009, 2016; Chu et al., 2016; Chu and Adler, 2015; Dalgleish et al., 2010). And several of  
108 these studies report significant effects of climate variation on the vital rates and overall population  
109 growth rates of these species (Dalgleish et al., 2010; Adler et al., 2012, 2009; Chu et al., 2016). This  
110 well-studied system offers the ideal opportunity to test whether statistical associations between  
111 annual climate and plant demography in longterm observational data can be used to predict the  
112 responses of plant populations to experimental climate change experiments.

113 In this study, we report how the four dominant plant species at the USSES respond to a  
114 five year drought and irrigation experiment and use the results to address two research questions:  
115 first, how much do the growth, recruitment and survival of our target species differ between the  
116 precipitation manipulation treatments? If our experiment does affect species vital rates we interpret  
117 that as strong evidence that changes in precipitation should have an effect on populations in the  
118 future. Second, can we predict each species' response to the experimental conditions based on

119 how they respond to natural climate variation in the observational data? If models based on  
120 observational data can predict the response of species to this experiment this should give us more  
121 confidence in longterm population monitoring data to predict species response to climate change.

## 122 Methods

### 123 Study site and data set description

124 The U.S. Sheep Experiment Station (USSES) is located at Dubois, Idaho (44.2°N, 112.1°W), 1500  
125 m above sea level. During the period of data collection (1926–2016), mean annual precipitation  
126 was 270 mm and mean temperatures ranged from -8°C (January) to 21°C (July). The vegetation  
127 is dominated by a shrub, *Artemisia tripartita*, and three perennial C3 grasses: *Pseudoroegneria*  
128 *spicata*, *Hesperostipa comata*, and *Poa secunda*. These dominant species account for over 70% of  
129 basal cover and 60% of canopy cover at this site.

130 Scientists at the USSES established 26 1-m<sup>2</sup> quadrats between 1926 and 1932. Eighteen  
131 quadrats were distributed among four ungrazed exclosures, and eight were distributed in two pad-  
132 docks grazed at medium intensity spring through fall. All quadrats were located on similar topog-  
133 raphy and soils. In most years until 1957, all individual plants in each quadrat were mapped using  
134 a pantograph (?). The historical data set is public and available online (?). In 2007, we located  
135 14 of the original quadrats, all of which are inside permanent livestock exclosures, and resumed  
136 annual mapped censusing using the traditional pantograph method. Daily temperature and pre-  
137 cipitation has been monitored throughout this period at a climate station located at the USSES  
138 headquarters (station id: GHCND:USC00102707) which located within 2 km of the research plots.  
139 We downloaded daily and monthly tmin, tmax, and precipitation data from the National Climate  
140 Data Centers online database.

141 We extracted data on survival, growth, and recruitment from the mapped quadrats based  
142 on plants' spatial locations. Our approach tracks genets representing individual plants. For the  
143 shrub, each genet is associated with the basal position of a stem. For the bunchgrasses, each genet  
144 represents a spatially distinct polygon in the mapped quadrat. These genets may fragment and/or  
145 coalesce over time. Each mapped polygon is classified as a surviving genet or a new recruit based  
146 on its spatial location relative to genets present in previous years (Lauenroth and Adler, 2008). We

<sup>147</sup> modeled vital rates using data from 21 year-to-year transitions between 1929 and 1957, and four  
<sup>148</sup> year-to-year transitions from 2007 to 2011.

## <sup>149</sup> Precipitation experiment

<sup>150</sup> In spring 2011, we selected locations for an additional 16 quadrats for the precipitation experiment.  
<sup>151</sup> We located these in a large exclosure containing six of the historical permanent quadrats. We  
<sup>152</sup> avoided plots falling on hill slopes, areas with greater than 20% bare rock, or with over 10% cover  
<sup>153</sup> of the woody shrubs *Purshia tridentata* or *Amelanchier utahensis*. New plots were established  
<sup>154</sup> in pairs, and one plot per pair was randomly assigned to either the precipitation reduction or  
<sup>155</sup> the precipitation addition treatment. We mapped the quadrats in June, 2011 and then built the  
<sup>156</sup> rainfall shelters and set-up the irrigation systems in the fall of 2011. We used a rain-out shelter and  
<sup>157</sup> automatic irrigation design described in (Gherardi and Sala, 2013). Each rain-out shelter covered  
<sup>158</sup> an area of 2.5 by 2 m and consisted of transparent acrylic shingles held up 1.5 to 1 m over the plot  
<sup>159</sup> to channel 50% of incoming rainfall off of the plot and into 75 l reservoirs. The collected water  
<sup>160</sup> was pumped out of reservoirs and sprayed onto paired irrigation treatment plots. Pumping was  
<sup>161</sup> initiated automatically with float switches that were triggered when water levels in the reservoirs  
<sup>162</sup> were approximately 20 l, or equivalently irrigation was triggered once for every 6 mm of rainfall  
<sup>163</sup> collected. We disconnected the irrigation pumps in late fall each year and re-connected them in  
<sup>164</sup> April. The drought shelters remained in place throughout the year.

<sup>165</sup> We monitored soil moisture and air temperature in four of the precipitation experiment plot  
<sup>166</sup> pairs using Decagon Devices (Pulman Washington) 5TM and EC-5 soil moisture sensors and 5TE  
<sup>167</sup> temperature sensors. We installed two soil moisture sensors in each monitored plot at 5 cm and  
<sup>168</sup> two at 25 cm deep in the soil. Air temperature was measured underneath the roofing of the shelter  
<sup>169</sup> at 30 cm above ground. For each pair of manipulated plots, we also installed sensors in a nearby  
<sup>170</sup> area to measure ambient rainfall and temperature. Data were logged automatically every four  
<sup>171</sup> hours. We augmented automatic monitoring of the climate in these plots with direct measurements  
<sup>172</sup> of soil moisture with a handheld EC-5 soil moisture sensor at six points around all 16 plots on  
<sup>173</sup> 6/6/2012, 4/29/2015, 5/7/2015, 6/9/2015 and 5/10/2016. We analyzed these spot measurements  
<sup>174</sup> for significant treatment effects on soil moisture using a linear mixed effects model with the *lmer*  
<sup>175</sup> package in *R*, with plot, plot group, and date as random effects in the model ((Bates et al., 2015)).

176 We conducted a simple statistical to determine the net effect of the experimental treatments  
177 on cover in the experiment. First we calculated the log change in cover for each of the four focal  
178 species in each quadrat from from the start of the experiment in spring prior to manipulation, to  
179 the last year of the experiment. Log change in cover was defined as ,  $\log(Cover_{2016}/Cover_{2011})$   
180 where  $Cover_{2016}$  is the cover of each species in 2011 and  $Cover_{2015}$  is cover in 2011. We tested for  
181 the effect of precipitation treatment on this measure with a linear model in *R*.

## 182 Soil moisture modeling

183 We expected that our precipitation manipulation experiment would affect plants by altering avail-  
184 able soil moisture during the growing season. Because we do not have direct soil moisture measures  
185 for each year of observed plant cover in the historical record, we used the SOILWAT soil mois-  
186 ture model to estimate daily soil moisture at the USSES from 1925 to the present (?). We used  
187 an enhanced version of soilwat that has recently been developed for use in semi-arid shrubland  
188 ecosystems (?). SOILWAT uses daily weather data, ecosystem specific vegetation properties and  
189 site specific soil properties to estimate water balance processes. SOILWAT specifically estimates  
190 rainfall interception by vegetation, evaporation of intercepted water, snow melt and snow redistri-  
191 bution, infiltration into the soil, percolation through the soil, bare-soil evaporation, transpiration  
192 from each soil layer, and drainage. We parameterized SOILWAT with the generic sagebrush steppe  
193 vegetation parameters and site specific soil texture and bulk density data. We used daily weather  
194 data collected at the USSES from 1925 until the present as weather forcing data for the SOILWAT  
195 predictions.

196 We averaged daily soil moisture predictions from SOILWAT from upper 40 cm of soil and  
197 then averaged these seasonally to serve as the covariates in the vital rate regressions for each species.  
198 Because we did not monitor soil moisture directly in all control, drought and irrigation plots, we  
199 used a model to describe the average treatment effects on soil moisture during the course of the  
200 experiment. To do this we first averaged observed soil moisture data by date and plot and then  
201 standardized these by the mean and standard deviation of the control soil moisture conditions  
202 observed within each plot group. We then found the difference between the soil moisture in the  
203 treated plots and the ambient conditions. We then modeled these treatment effects as a function of  
204 season and whether a day was rainy or dry. We expected that our drought and irrigation treatments

205 might be more effective during rainy weather than during dry weather. Rainy days were defined as  
206 any days when any precipitation was recorded and average temperatures were above 3 degrees C.  
207 The day immediately following rainfall was also classified as rainy. We fit this model using the *lmer*  
208 package in *R* (?) with random effects for plot group and date. We then used this model to predict  
209 the treatment effects on soil moisture for the entire study period from the ambient soil moisture  
210 values predicted from the SOILWAT model described above. These adjusted soil moisture values  
211 reflected the average season and rainfall dependent effects of the experimental treatments on soil  
212 moisture and could be used as covariates for predicting the effects of our manipulation on each  
213 species demographic rates.

## 214 Overview of the analysis

215 Our analysis consists of two separate datasets and three different categories of vital rate models.  
216 We refer to the first dataset as the observational data. It consists of all the historical data collected  
217 from 1925 to 1957 as well as the modern data collected from the same plots from 2007 to 2010.  
218 These data record the response of plants in each plot to the ambient climate variation. We refer  
219 to the second dataset as the experimental data. It consists of the data collected from 2011 to 2016  
220 from the 16 new experimentally manipulated plots, as well as from 14 of the original historical plots  
221 which serve as ambient climate controls.

222 In order describe the effects of the experimental treatments on each vital rate, we fit "treat-  
223 ment" models. The treatment models included parameters for the effects of the drought and ir-  
224 rigation treatments on each vital rate. We fit these models to both datasets together, all the  
225 experimental and all the observational data. We did this because we wanted to focus our predic-  
226 tions on the effects of the experimental treatments on the vital rates, rather than any differences  
227 between the datasets in the effects of crowding and plant size on the vital rates.

228 Next in order to test how well we could predict the responses in the experimental plots, we fit  
229 two classes of models to the observational dataset only. Thus when we used these models to predict  
230 the response of each species to the five years of experimental data we are generating predictions  
231 for out of sample data. We refer to the first type of model as the "no climate" models. These  
232 models include parameters for the effects of competition on each vital rate and the size dependence  
233 of survival and growth but they do not include climate or treatment effects. Lastly, we fit the

234 "climate" models which include the effects of annual variation in soil moisture on each vital rate.  
 235 The no climate model provides us a baseline by which to measure the accuracy of the predictions  
 236 from the climate model. Because much of the variation in growth, survival and recruitment in this  
 237 system can be explained by plant size and competition, we expect that these two models will make  
 238 similar predictions for individual plant performance in the experiment. However, if the climate  
 239 model makes more accurate predictions than the no climate model, this indicates that the climate  
 240 parameters it includes contain useful information for prediction.

## 241 Statistical models of vital rates

242 All three categories of models described above follow the same basic structure and differ only in  
 243 how they treat climate and treatment effects and (Adler et al., 2010; Chu and Adler, 2015). We  
 244 model the survival probability of an individual genet as a function of genet size, the neighborhood-  
 245 scale crowding experienced by the genet from both conspecific and heterospecific genets, temporal  
 246 variation among years, and permanent spatial variation among groups of quadrats ('group'; here  
 247 means a set of nearby quadrats located within one pasture or grazing exclosure). In this analysis  
 248 we only include crowding from the four main focal species.

249 Formally, we modeled the survival probability,  $S$ , of genet  $i$  in species  $j$ , group  $g$ , and from  
 250 time  $t$  to  $t + 1$  as

$$\text{logit}(S_{ijg,t}) = \varphi_{jg}^S + \gamma_{j,t}^S + \beta_{j,t}^S u_{ij,t} + \langle \boldsymbol{\omega}_{j,t}^S, \mathbf{W}_{ij,t} \rangle \quad (1)$$

251 where  $\varphi$  is the spatial group dependent intercept,  $\gamma$  is a year-effect,  $\beta$  is year-dependent coefficient  
 252 that represents the effect of log genet size,  $u$ , on survival in year  $t$ .  $\boldsymbol{\omega}$  is a vector of interaction  
 253 coefficients which determine the impact of crowding,  $\mathbf{W}$ , by each species on the focal species. The  
 254 vector  $\mathbf{W}$  includes crowding from the four dominant species, *A. tripartita*, *P. spicata*, *H. comata*,  
 255 and *Poa secunda*.  $\langle \mathbf{x}, \mathbf{y} \rangle$  denotes the inner product of vectors  $\mathbf{x}$  and  $\mathbf{y}$ , calculated as `sum(x*y)` in  
 256 R. This model is the no climate model for survival.

257 In the treatment model, a new term is added to the above model,  $\mathbf{T}\chi_j^S$  where  $\chi$  is a vector  
 258 of treatment effect coefficients for each experimental treatment level  $h$  on the survival rate, and  $\mathbf{T}$   
 259 is a design matrix indicating the treatment level of each observation in the data. The design matrix  
 260 also includes terms for the interaction between plant size  $u$  and the treatment effects which allow  
 261 the effect of each treatment to vary with plant size.

In the climate model, the above term is replaced with  $\mathbf{C}\xi_j^S$ , where  $\xi$  gives a vector of coefficients describing the effects of a set of soil moisture covariates  $\mathbf{C}$  in treatment  $h$  and year  $t$  on the survival rate of species  $j$ .  $\mathbf{C}$  can include interaction effects between plant size,  $u$ , and the soil moisture covariates allowing the effects of soil moisture to vary with plant size.

Our growth model has a similar structure. The change in genet size from time  $t$  to  $t + 1$ , conditional on survival, is given by:

$$u_{ijg,t+1} = \varphi_{jg}^G + \gamma_{j,t}^G + \chi_{jh}^G + \beta_{j,t}^G u_{ij,t} + \langle \boldsymbol{\omega}_{j,t}^G, \mathbf{W}_{ij,t} \rangle + \varepsilon_{ij,t}^G. \quad (2)$$

As in the survival regression above, parameters describing the treatment effects on growth are added in the treatment model,  $\mathbf{T}\chi_j^G$ , where  $\chi$  is a treatment effect describing the effect of experimental treatment  $h$  on growth, including treatment by size interactions.

Similarly, in the climate model, the above term is replaced with  $\mathbf{C}\xi_j^G$ , where  $\xi$  is a vector of coefficients describing the effects of soil moisture covariates in the matrix  $\mathbf{C}$  for treatment  $h$  and year  $t$  on growth of species  $j$ . Again this can include interactions between soil moisture and plant size  $u$ .

Although the main focus of the current analysis the effects of soil moisture, we also modeled the effects of inter- and intra-specific competition in our vital rate models. We model the crowding experienced by a focal genet as a function of the distance to and size of neighbor genets. These effects are well described in previous work (Teller et al., 2016; Adler et al., 2016). Briefly, we model the crowding experienced by genet  $i$  of species  $j$  from neighbors of species  $m$  as the sum of neighbor areas across a set of concentric annuli,  $k$ , centered at the plant,

$$w_{ijm,k} = F_{jm}(d_k)A_{i,k} \quad (3)$$

where  $F_{jm}$  is the competition kernel (described below) for effects of species  $m$  on species  $j$ ,  $d_k$  is the average of the inner and outer radii of annulus  $k$ , and  $A_{i,m,k}$  is the total area of genets of species  $m$  in annulus  $k$  around genet  $i$ . The total crowding on genet  $i$  exerted by species  $m$  is

$$W_{ijm} = \sum_k w_{ijm,k}. \quad (4)$$

284 Note that  $W_{ijj}$  gives intraspecific crowding. The  $W$ 's are then the components of the  $\mathbf{W}$  vectors  
285 introduced as covariates in the survival (1) and growth (2) regressions.

286 We assume that competition kernels  $F_{jm}(d)$  are non-negative and decreasing, so that distant  
287 plants have less effect than close plants. Otherwise, we let the data dictate the shape of the kernel  
288 by fitting a spline model using the methods of Teller et al. (2016). We used data from all historical  
289 plots and contemporary control-treatment plots to estimate the competition kernels and these are  
290 described in more detail in (Adler et al., 2016).

291 Once we had estimated the competitions kernels, we used them to calculate the values of  $\mathbf{W}$   
292 for each individual, and fit the full survival and growth regressions, which include the interspecific  
293 interaction coefficients,  $\boldsymbol{\omega}$ . All genets in a quadrat were included in calculating  $W$ , but plants  
294 located within 5 cm of quadrat edges were not used in fitting.

295 We model recruitment at the quadrat level rather than at the individual genet level because  
296 the mapped data do not allow us to determine which recruits were produced by which potential  
297 parent plants. We assume that the number of individuals,  $y$ , of species  $j$  recruiting at time  $t + 1$  in  
298 the location  $q$  follows a negative binomial distribution:

$$y_{jq,t+1} = \text{NegBin}(\lambda_{jq,t+1}, \theta) \quad (5)$$

299 where  $\lambda$  is the mean intensity and  $\theta$  is the size parameter. In turn,  $\lambda$  depends on the composition  
300 of the quadrat in the previous year:

$$\lambda_{jq,t+1} = C'_{jq,t} \exp \left( \varphi_{jg}^R + \gamma_{j,t}^R + \langle \boldsymbol{\omega}^R, \sqrt{\mathbf{C}'_{q,t}} \rangle \right) \quad (6)$$

301 where the superscript  $R$  refers to Recruitment,  $C'_{jq,t}$  is the ‘effective cover’ ( $\text{cm}^2$ ) of species  $j$  in  
302 quadrat  $q$  at time  $t$ ,  $\varphi$  is a group dependent intercept,  $\gamma$  is a random year effect,  $\boldsymbol{\omega}$  is a vector of  
303 coefficients that determine the strength of intra- and interspecific density-dependence, and  $\mathbf{C}'$  is  
304 the vector of “effective” cover of each species in the community. Following previous work (Adler  
305 et al., 2010), we treated year as a random factor allowing intercepts to vary among years.

306 Because plants outside the mapped quadrat could contribute recruits to the focal quadrat or  
307 interact with plants in the focal quadrat, we estimated effective cover as a mixture of the observed

308 cover,  $C$ , in the focal quadrat,  $q$ , and the mean cover,  $\bar{C}$ , across the spatial location,  $g$ , in which  
309 the quadrat is located:  $C'_{jq,t} = p_j C_{jq,t} + (1 - p_j) \bar{C}_{jq,t}$ , where  $p$  is a mixing fraction between 0 and  
310 1 that was estimated as part of fitting the model.

311 In the treatment model for recruitment, a new term is added to the exponential term in  
312 the equation above,  $\mathbf{T}\boldsymbol{\chi}_j^R$  where  $\boldsymbol{\chi}$  describes the effect of each treatment level on recruitment.

313 Likewise in the climate model this term is replaced by  $\mathbf{C}\boldsymbol{\xi}_j^R$  where the  $\boldsymbol{\xi}$  gives a set of  
314 coefficients for the year, and treatment specific soil moisture covariates in  $\mathbf{C}$ .

315 We fit all vital rate models using Hamiltonian-Markov Chain Monte Carlo (HMC)C  
316 simulations in the programs STAN 10.1 and rStan (). The priors and model code are described more  
317 completely in appendix A. Each model was run for 2,000 iterations and four independent chains  
318 with different initial values for parameters. We discarded the initial 1,000 samples. Convergence  
319 was observed graphically for all parameters, and confirmed by assessing the split  $\hat{R}$  statistic which  
320 at convergence is equal to one () .

321 We fit the treatment models for species survival and growth with and without the size  
322 by treatment interactions in the treatment effect term  $\boldsymbol{\chi}$ . We then judged whether including the  
323 interaction terms improved model fit by comparing the Watanabe-Aikake Information Criteria  
324 (WAIC) scores of each version of the model and retained the version with the lower WAIC score  
325 (?). WAIC are similar to AIC scores and allow for comparison of Bayesian models. Lower WAIC  
326 scores indicate a more parsimonious model. When a treatment model for survival or growth of a  
327 species included a size by treatment effect in  $\boldsymbol{\chi}$ , we also included a size by soil moisture effect in  
328 the  $\boldsymbol{\xi}$  term in the climate model for that species and vital rate. This allowed us to more directly  
329 compare the predictions from the climate model to the effects in the treatment model.

### 330 Selecting soil moisture covariates

331 After generating a time series of predicted daily soil moisture from SOILWAT, we averaged daily  
332 soil moisture across spring, summer and fall seasons in each year. We considered each of the three  
333 seasonal soil moisture variables at three different time periods relative to the demographic transition  
334 from year  $t$  to year  $t + 1$ . Soil moisture in the year between  $t$  and  $t + 1$  is indicated with a "1"  
335 subscript. Soil moisture in the year before  $t$  is indicated with a "0" subscript. And soil moisture  
336 preceding this year is indicated with a "lag" subscript. For example, for the year 2010,  $spring_1$

337 indicates soil moisture in the spring of 2010,  $spring_0$  indicates soil moisture during spring of 2009  
338 and  $spring_{lag}$  indicate soil moisture during spring 2008.

339 We wanted to avoid fitting nine soil moisture covariates (three seasons and three lags each)  
340 for each species and vital rate, so we used only three soil moisture covariates per species and vital  
341 rate. We selected these three by calculating the correlations of each soil moisture variable with  
342 the random year effects from the no climate model fit and then selecting the three soil moisture  
343 variables with the strongest correlations with these year effects. This screening technique has  
344 been used in previous demographic studies at this site (Dalgleish et al., 2010) and is often used in  
345 dendrochronology to screen for potential climate influence on tree-ring growth (Wang et al., 2003).  
346 We felt this approach was justified because we did not make inference on these fitted parameters  
347 until after we validated their ability to predict the out of sample data in the experimental plots.

348 **Predicting cover from individual-based models**

349 The vital rate regressions allow us to evaluate whether soil moisture and the experimental treat-  
350 ments had an effect on species performance. But the population response ultimately depends on  
351 the integrated effects of treatment or soil moisture on all three vital rates. To evaluate whether the  
352 climate models could predict the responses of these species in the drought and irrigation experiment  
353 at the overall population level we used an individual-based model (IBM) to compare observed and  
354 predicted changes in population size from one year to the next.

355 To simulate changes in cover in each quadrat from year  $t$  to year  $t + 1$ , we initialized the  
356 IBM with the observed genet sizes and locations of the four focal species observed in year  $t$  in each  
357 quadrat. For every individual genet in a quadrat, we projected its size and survival probability in  
358 the next year using the growth and survival models and the appropriate crowding and soil moisture  
359 or treatment covariates for that year and quadrat. Likewise we projected the number of new recruits  
360 in the quadrat in the next year using the recruitment model. We calculated the expected cover in  
361 year  $t + 1$  as the total area of new recruits, plus the sum of the predicted area of each existing plant  
362 at time  $t + 1$  multiplied by each plant's expected survival probability from time  $t$  to  $t + 1$ .

363 We generated predictions using 1000 samples from the posterior distributions of each model  
364 parameter which allowed us to carry forward all of the uncertainty of the fitted vital rate models into  
365 our cover predictions. Because we were interested in comparing model predictions to observations,

<sup>366</sup> and were not interested in the effects of demographic stochasticity, we used a deterministic version  
<sup>367</sup> of the models (e.g., recruitment is the  $\lambda$  of (6), rather than a random draw from a negative binomial  
<sup>368</sup> distribution with a mean of  $\lambda$ ).

<sup>369</sup> After generating predictions for each year from the climate and no climate models, we found  
<sup>370</sup> the predicted quadrat-level changes in cover as  $\log(Cover_{t+1}/Cover_t)$ .

<sup>371</sup> **Quantifying predictive accuracy**

<sup>372</sup> We assessed the predictive performance of the climate and no climate by calculating the mean  
<sup>373</sup> square error (MSE) between the predicted and observed responses in the experimental data as,

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - E(y_i|\theta))^2, \quad (7)$$

<sup>374</sup> where  $y_i$  is the outcome of observation  $i$  and  $E(y_i|\theta)$  gives the expected outcome given the pa-  
<sup>375</sup> rameters in the model  $\theta$ . The MSE is easy to interpret, but is not always appropriate for models  
<sup>376</sup> fit with non-normal error structures (Gelman et al., 2014). A more general statistic for assessing  
<sup>377</sup> model predictions is the log pointwise predictive density (lppd) (Gelman et al., 2014). The lppd  
<sup>378</sup> for a given model is defined as,

$$lppd = \sum_{i=1}^n \log \int p(y_i|\theta) p_{post}(\theta) d\theta, \quad (8)$$

<sup>379</sup> where the integral on the right side gives the probability of observing the outcome  $y$  at each data  
<sup>380</sup> point  $i$  given the full posterior distribution of the parameters in the model  $p_{post}(\theta)$ . In practice we  
<sup>381</sup> computed the lppd from the posterior simulations generated by STAN as,

$$\widehat{lppd} = \sum_{i=1}^n \log \left( \frac{1}{S} \sum_{s=1}^S p(y_i|\theta^S) \right), \quad (9)$$

<sup>382</sup> where the summation of  $p(y_i|\theta^S)$  gives the total probability of observing the the actual response  $y_i$   
<sup>383</sup> given the simulated posterior distribution  $\theta^S$  across the full set of model simulations  $S$ . The log of  
<sup>384</sup> this sum is then averaged across the set of all observations  $i$ . Higher lppd scores indicate that the  
<sup>385</sup> model better predicts the observations.

386 In addition, we evaluated whether the climate model predicted treatment effects of similar  
387 direction and magnitude to those observed in the experiment. We did this by extracting the soil  
388 moisture coefficients contained in  $\xi$  for each of the vital rates and then multiplying those by the  
389 appropriate soil moisture covariates for each year and treatment level in the experiment. We then  
390 averaged these across all five years in the experiment to find the average treatment effect predicted  
391 by the climate model. We compared these to the posteriors of the treatment parameters,  $chi$ , from  
392 the treatment model. As a measure of agreement between our predictions and observed response  
393 we calculated the correlation between the predicted and observed treatment effects.

394 We considered the effect of climate covariates or treatment effects to be significant when  
395 the 95% Bayesian credible intervals on the posterior estimates did not overlap zero.

396 All data and R code necessary to reproduce our analysis will be deposited in the Dryad  
397 Digital Repository once the manuscript is accepted. The current version of the computer code  
398 is available at <https://github.com/pbadler/ExperimentTests/tree/master/precip> and the data are  
399 available at <https://bitbucket.org/ellner/driversdata>.

## 400 Results

### 401 Effects on soil moisture

402 Our treatments successfully changed the soil moisture in the experimental plots in the directions  
403 expected (fig. 1). Spring spot measurements of soil moisture from all the plots showed that on  
404 average the drought plots were roughly 50% drier, while irrigated plots were roughly 40% wetter  
405 than ambient conditions (table 1).

406 The continuously recorded soil moisture data also showed treatment effects, but these were  
407 weaker on average than the spot measurements and depended on season and recent rainfall (table  
408 2; fig 2). We saw weaker effects during the spring than during the fall and summer: the drought  
409 plots were about 20-30% drier than ambient in the fall and summer but only 7 to 14% drier during  
410 the spring, while the irrigated plots were 30% wetter during the fall and summer but only 20-25%  
411 wetter during the spring. Treatment differences were slightly larger during rainy periods, especially  
412 in the spring (rainfall effect in table 2). We did not find evidence that the drought shelters and the  
413 irrigation treatments consistently affected air temperature at 30 cm above the plots.

414        The SOILWAT soil moisture model predicted average monthly soil volumetric water content  
415        of between 10 ml/ml and 15 ml/ml each month, with the month of April being the wettest and the  
416        month of July, August and September being the driest on average. Annual variation in seasonal  
417        soil moisture for each year was positively correlated with seasonal precipitation and negatively  
418        correlated with seasonal temperature. During the course of the experiment, SOILWAT reproduced  
419        much of the daily variation observed in soil moisture recorded by our automatic data loggers, but  
420        the average soil moisture predicted by SOILWAT was about 5 ml/ml higher than the soil moisture  
421        content observed in the field.

422        After adjusting the SOILWAT seasonal soil moisture predictions by the treatment effects,  
423        we found that the soil moisture predicted in the drought plots during the course of the experiment  
424        was well below the historical seasonal averages: the summer of 2012 and 2013, the fall of 2013, and  
425        the spring and winter of 2014 fell below the 5th percentile limit for drought in the historical period  
426        (fig 3). Soil moisture in our irrigation plots was generally above the historical average soil moisture  
427        but conditions never exceeded the 90th percentile for soil moisture in the historical period (fig 3).

#### 428        **Effects on cover and vital rates**

429        The cover of *H. comata* and *P. spicata* fell significantly in the drought plots from 2011 to 2016  
430        (tables 4, 6; fig 4). The cover of *P. secunda* showed a slight decrease in the drought plots and an  
431        increase in the irrigated plots but these changes were not significant (table 5). In contrast to the  
432        grasses, the cover of *A. tripartita* increased slightly in all three treatments (fig 4).

433        Our treatment models fit to the experimental and observational data indicated a variety of  
434        treatment effects on the vital rates of each species. Based on the WAIC scores with and without the  
435        size by treatment effects, we retained size by treatment effects in the growth models for *A. tripartita*  
436        and *P. secunda*, and the survival model for *P. secunda*. For *A. tripartita* we found significant size  
437        by treatment effects of drought: drought had positive effects on plants of average size and smaller  
438        (fig 5), but plants larger than the mean size by more than 1.5 standard deviations grew slightly less  
439        in the drought treatment than in the controls. *A. tripartita* showed the opposite response in the  
440        irrigated plots, (although the irrigation parameters were not significant at the 95% confidence level):  
441        irrigation reduced growth for small plants while irrigation increased growth of plants more than  
442        1.5 standard deviations larger than the mean size. Drought led to a strong (but not significant)

443 decrease in *H. comata* growth, while irrigation had no effect on growth. Like *A. tripartita*, we  
444 saw size by treatment effects on *P. secunda* growth, with the negative effects of drought becoming  
445 greater for larger plants. *P. secunda* showed the opposite response in the irrigation plots with larger  
446 plants showing the largest increase in growth in response to irrigation (although not significant).  
447 *P. spicata* growth was relatively unaffected by the drought and irrigation treatments.

448 Survival of all three grass species (fig 6) decreased in the drought plots. And *P. secunda*  
449 showed a negative size by drought interaction effect: the survival of larger plants was more neg-  
450 atively affected by drought than that of the smaller plants. *A. tripartita* survival was relatively  
451 unaffected by the drought and irrigation treatments.

452 Recruitment in our irrigation plots was significantly less than in control plots for two grass  
453 species *P. secunda* and *P. spicata* (fig 7). However, recruitment was also lower in the drought plots  
454 than in the the control plots (although not significantly so), indicating that the decrease in the  
455 irrigated plots may have not been entirely due to the irrigation itself. The recruitment data for *A.*  
456 *tripartita* were relatively limited, with only 32 new recruits in total observed in all 30 plots over  
457 the course of the five year experiment and we observed no treatment effects.

458 Consistent with previous research most of our demographic models estimated strong nega-  
459 tive intra-specific crowding effects and weaker negative inter-specific crowding effects on the focal  
460 species (appendix) (Adler et al., 2010; Chu et al., 2016; Chu and Adler, 2015; Adler et al., 2016).

## 461 Effects of soil moisture on vital rates

462 We choose three seasonal soil moisture variables for each species' based on their correlation with the  
463 random year effects in the no climate model (table 7). We included size by soil moisture variables  
464 for *A. tripartita* and *P. secunda* based on the treatment response we observed in the experiment.  
465 All three time lags and all three seasons show up in the selected variables. After fitting the vital rate  
466 models with the selected soil moisture variable we observed a trend towards positive soil moisture  
467 effects on growth of all three grasses (fig 8). For *H.comata* the soil moisture of the most recent  
468 summer (*summer*<sub>1</sub>) had a significantly positive effect while the soil moisture during *summer*<sub>0</sub> and  
469 *fall*<sub>lag</sub> were also positive but not significant. For *A. tripartita* *fall*<sub>0</sub> and *summer*<sub>0</sub> had strong  
470 negative effects on growth. There were also strong positive size by climate interaction effects for

471 these variables: soil moisture had a stronger negative effect on small plants and a positive effect  
472 only on the largest plants (fig SI-1).

473 Soil moisture had significant effects on the survival of all four species (fig 9). As for growth  
474 the grasses showed mainly positive effects while *A. tripartita* showed a significant negative effect  
475 of *summer<sub>0</sub>* and a strong negative effect of *spring<sub>0</sub>*. *H. comata* showed a significant positive effect  
476 of *springlag* soil moisture and a strong positive effect of *spring<sub>0</sub>* and *fall<sub>1</sub>*. *P. secunda* showed a  
477 significant positive effect of the previous *spring<sub>0</sub>* and there was an interaction between this effect  
478 and plant size: as plant size increased this effect became more positive. Finally for *P. spicata* there  
479 was a significant positive effect of *springlag* soil moisture on survival.

480 There were only two significant effects of soil moisture on recruitment: *falllag* soil moisture  
481 had a positive effect on *P. secunda*, and *summerlag* soil moisture had a negative effect on *P. spicata*  
482 recruitment (fig 10). Soil moisture of *summer<sub>0</sub>* also had a strong negative effect on *P. spicata*  
483 recruitment.

484 The intra- and interspecific crowding effects estimated in the climate model were similar to  
485 those estimated in the treatment model (appendix).

#### 486 Evaluating the predictions

487 For most models adding climate covariates did not improve our ability to predict species responses  
488 in the experiment (table 8). However, the climate models did improve overall prediction MSE for  
489 growth of *A. tripartita* and growth and survival of *P. secunda* (table 8). In terms of lppd, the  
490 climate model outperformed the no climate model in six out of twelve models: for *A. tripartita*  
491 growth, *H. comata* recruitment, *P. secunda* growth and survival and *P. spicata* recruitment (table  
492 8).

493 When we look at the predictions for each treatment separately we see that climate covariates  
494 improved model predictions more often in the drought treatments than in the control or irrigation  
495 treatments (table SI-1). For all four species, the climate model outperformed the no climate model  
496 for predicting the response of growth to drought in terms of lppd (table SI-1). The climate model  
497 also outperformed the no climate model for predicting irrigation effects on growth for all species  
498 except *H. comata*.

499 Overall our climate models often predicted the correct direction of the drought and irriga-  
500 tion treatments (fig 11). In four cases we both observed and predicted treatment effects significantly  
501 different from zero based on the 95% Bayesian credible interval around the parameter mean: the  
502 drought response of *H. comata* survival (fig SI-6), the drought response of *P. secunda* growth (fig  
503 SI-3), the irrigation response of *P spicata* recruitment (fig ??) and the irrigation response of *P.*  
504 *secunda* recruitment (fig ??). In only one of these cases, for *P. secunda* recruitment, was the  
505 predicted effect in the opposite direction from the observed treatment effect (fig 11). The overall  
506 correlation between the predicted and observed treatment effects for all treatments, species and  
507 vital rates was  $r = 0.54$ , whereas the correlation for the drought treatment effects,  $r = 0.77$ , was  
508 better than for the irrigation effects,  $r = 0.46$ . Also the correlation between the size by climate and  
509 size by treatment effects for *A. tripartita* growth and *P. secunda* growth and survival was much  
510 stronger than the correlation between the intercept parameter estimates (fig 11).

511 Using the vital rate models for each species we generated one step ahead cover predictions  
512 for each quadrat in each year of the experiment. Average cover predicted by the climate model  
513 tended to be lower than the observed cover each year for *A. tripartita* and *P. secunda* (fig 12).  
514 Comparing the overall population growth rates predicted to those observed in the experiment,  
515 we see that the MSE of the climate model was lower than the MSE of the no climate model for  
516 *P. secunda* and *P. spicata* (table 9). The predictions produced by the climate model for these  
517 species were also slightly more correlated with the observations than the predictions produced by  
518 the no climate model (table 9). Considering each treatment and species separately, the predicted  
519 population growth rates for *A. tripartita*, *P. secunda* and *P. spicata* were all consistently lower  
520 than the observed population growth rates (figs ??, 15, 16). The climate model showed lower MSE  
521 for *A. tripartita*, *P. secunda* and *P. spicata* in the irrigation treatment, *P. spicata* in the control  
522 treatment and *H. comata* in the drought treatment (fig 14). However, the correlations between the  
523 predicted and observed log changes in cover did not always show the same pattern as MSE: the  
524 climate model made more strongly correlated predictions with the observations than the no climate  
525 model only for *P. spicata* and *P. secunda* in the control treatment and *P. secunda* and *H. comata*  
526 in the drought treatment.

527 **Discussion**

528 Our experiment showed that observational data on the response of plant populations to interannual  
529 climate variation can indeed help us predict the direction of species responses to experimental  
530 climate manipulations (fig 11). This was true even though adding climate parameters to the  
531 demographic models only improved vital rate predictions for half of the models (fig 8). This should  
532 give us some hope that even when climate effects in demographic models fit to observational data  
533 are weak or not significant, they may contain useful qualitative information on the direction of  
534 climate effects in the future.

535 Although a number of previous studies of the plants in this system have looked the effects  
536 of climate on demography and survival, this is the first study to demonstrate effects of climate  
537 experimentally. The significant decline in *H. comata* cover we report in the experiment (fig 4)  
538 is consistent with the population-level sensitivities to precipitation estimated by previous studies  
539 ((Chu et al., 2016), (Dalgleish et al., 2010), (Adler et al., 2012)). Each of these studies report *H.*  
540 *comata* having the strongest positive population response to increased precipitation among all the  
541 species; this is matched by our finding that this species has the strongest negative response to  
542 experimental drought (fig 5, 6). On the other hand if we had only conducted an irrigation exper-  
543 iment our results may not have shown this consistency with previous work as *H. comata* showed  
544 no positive response to irrigation. Previous work has also generally reported postive population  
545 responses to precipitation for the other grasses, *P. secunda* and *P. spicata* ((Adler et al., 2012),  
546 (Chu et al., 2016), (Dalgleish et al., 2010)). These results are consistent with significant decrease  
547 we saw in *P. spicata* cover and survival and *P. secunda* growth and survival in the drought plots  
548 in our experiment (figs 4, 5, 6). The magnitudes of the negative drought effects were greater than  
549 the magnitudes of the positive irrigation effects on these grasses.

550 In contrast to the effects on the grasses, the effects on *A. tripartita* are more complicated.  
551 The effects of drought and irrigation on *A. tripartita* growth were highly dependent on plant size  
552 (figs 5): drought had a positive effect except possibly for the largest plants (fig SI-1). This result  
553 matches the negative effects of precipitation on this species estimated by (Adler et al., 2012) and  
554 (Chu et al., 2016). These studies also report relative strong indirect effects of precipitation on  
555 *A. tripartita* mediated by its competition with grasses. It is possible that some of the positive

556 effect on *A. tripartita* growth we see in the drought plots is the result of reduced grass competition  
557 these plots (fig 12; (Chu et al., 2016)). However, our growth model includes interspecific crowding  
558 and so should take into account any changes in grass abundance that could be driving a positive  
559 response from *A. tripartita*. So this leaves us with the question of why this species would show  
560 a positive direct response to drought? Although there is some evidence that saturated soils can  
561 be detrimental to big sagebrush (*A. tridentata*), a closely related species (Sturges, 1989), there  
562 was no evidence in our data that soils were saturated with moisture in the ambient or irrigation  
563 treatments (fig 2). In a similar study to ours, (Germino and Reinhardt, 2014) showed that big  
564 sagebrush increased in response to deep soil moisture but actually decreased in response to shallow  
565 soil moisture. Perhaps *A. tripartita* responds similarly and benefits directly from the drier soils.  
566 Another possible explanation is that our drought treatments reduced snow cover in the winter and  
567 early spring, an effect that has been shown to benefit big sagebrush in other ecosystems (Perfors  
568 et al., 2003).

569 One trend that appears to stand out in these results is the relatively asymmetrical effects  
570 of drought and irrigation. Although our irrigation plots raised soil moisture, we saw little positive  
571 effect of added moisture on the same species that were reduced by drought (figs 5, 6). We expected  
572 that in this water limited system, increasing soil moisture would lead to a stronger increase in plant  
573 performance. But perhaps we should not have been surprised by this result. It is possible that  
574 adding water to this system does little to increase grass growth and survival for a number of reasons.  
575 First, these species are presumably able to escape drought, by growing early in the year, or avoid  
576 drought stress later in the year through high water use efficiency (Bazzaz, 1979; Franks, 2011). If  
577 the species in our system are well adapted to the ambient water limitation, it makes some sense  
578 that they would show relatively small responses to increased moisture. Indeed our soil moisture  
579 data generally show a pulse of soil moisture in the spring when many grasses are actively growing  
580 (fig 2). During this window of active growth soil moisture in sagebrush steppe may be quite high  
581 and perhaps other resources become limiting (Caldwell et al., 1987). Moreover, our treatments had  
582 relatively smaller effects on soil moisture during the spring (table ).

583 A similar result was recently reported from another precipitation experiment in a warm  
584 desert grassland: the effect of increased soil moisture on grass productivity in that system showed a  
585 concave down pattern of response to increased soil moisture (Gherardi and Sala, 2015). Increasing

586 precipitation had a weaker effect on annual grass productivity than decreasing precipitation by  
587 an equal amount. A pattern like this is qualitatively similar to our finding significant decreases in  
588 either the cover or performance of the three grasses in drought plots, but not seeing any statistically  
589 significant increase in performance in the irrigated plots.

590 Reflecting on the possible mechanisms by which soil moisture could have affected these  
591 species may also shed some light on our successes and failures at prediction. The drought effects we  
592 observed on the three grasses were often stronger than the effects we predicted, while the irrigation  
593 responses observed were often weaker than predicted (figs SI-2, SI-6, SI-3, SI-7, SI-4, SI-8). If  
594 we had fit our growth and survival models with a non-linear function for soil moisture, perhaps  
595 informed by more mechanistic understanding of water limitation in this system, we may have made  
596 more accurate predictions of the drought and irrigation effects (Ehrlin et al., 2016).

597 Out of all the climate effects we predicted and observed, we only observed four cases where  
598 vital rate predictions and observations were both significantly different from zero (fig 11). In three  
599 of these cases, we successfully predicted the direction of the treatment effects. However, for *P.*  
600 *secunda* recruitment we predicted a response to irrigation that was significantly positive and yet  
601 observed a significantly negative response (figs 11). This is arguably our greatest error in prediction.  
602 But there may be some mitigating factors that can excuse our error. First we note that we observed  
603 decreases in recruitment in both the drought and irrigation plots for *P. secunda* and also for *P.*  
604 *spicata* (fig 7). So while we observed a significant decrease in *P. secunda* recruitment in the irrigated  
605 plots this may have been due to underlying differences in the experimental plots from the historical  
606 control plots. Because recruitment is modeled at the quadrat level we also had fewer observations  
607 and predictions to work with for these recruitment models than we had for the other vital rates.

608 Using the IBM, we demonstrated that the climate model predictions of the vital rates could  
609 be used to generate population-level predictions that were also better than the no climate models in  
610 some cases: the climate model improved one step ahead predictions of cover changes of *P. spicata*  
611 and *P. secunda* (table 9). In the drought treatment our one step ahead cover predictions for *H.*  
612 *comata* and *P. secunda* were also better than the no climate model. In this analysis, we either  
613 used the climate model, or the no climate model to generate predictions for the IBM. In theory  
614 we could have refined our predictions by climate model to predict all three vital rates in the IBM.  
615 Improvements of our cover predictions may have been possible if we only included climate models

616 in the IBM when those models improved vital rate predictions—for instance, the *A. tripartita* IBM  
617 would only include the climate model for growth but not for survival and recruitment.

618 Mixed success at predicting species responses to high and low moisture availability has in-  
619 teresting implications. On the one hand it is reassuring that observational data captures important  
620 climate effects. On the other hand, will knowing that some species' responses are unpredictable,  
621 but not knowing which ones, discourage ecologists from making any predictions at all? Among  
622 plants, perhaps improved ecophysiological models or trait data can be used as a screen to generate  
623 hypothesis about which species are more likely to be affected by climate and in which ways. In our  
624 system, for instance, it may make sense that the grasses showed a stronger response to the drought  
625 treatment than the a woody shrub. Although these grasses are adapted to the arid conditions  
626 that characterize the sagebrush steppe, they thrive during the brief window in spring and early  
627 summer when the soil moisture and temperatures are warmer. *A. tripartita* on the other hand  
628 grows throughout the summer and has deeper roots than the grasses, these traits may buffer it  
629 from some of the annual variability in soil moisture experienced by the grasses.

630 Will our predictions be useful for the predicting the effects of climate change in the future?  
631 Because these species compete, one could argue that our predictions for any one species in this  
632 community will only be as good as the predictions we make for their competitors, pathogens and  
633 predators (Tylianakis et al., 2008). For instance, while we observed weak treatment effects on *A.*  
634 *tripartita* in our experiment (fig 4) could drought ultimately lead to a decline in the cover of its  
635 competitors which will cause it to increase in abundance in the future? While this is possible for  
636 some species, research at our site and several other grasslands suggest that species are more limited  
637 by intra-specific competition than inter-specific competition Chu et al. (2016); Adler et al. (2016).  
638 In theory, this should mean that we can predict species short term responses to climate change  
639 with single species models ((Levine et al., 2010; ?; Adler et al., 2012)).

640 Our results give us more confidence that historical observational data can be used to de-  
641 tect and predict the demographic effects of climate change among sagebrush steppe plants. This  
642 should encourage more researchers to try and use observational data to predict the future in both  
643 experimental and natural settings (Houlahan et al., 2016). Nevertheless, our success at predicting  
644 the short-term response of two species to a small-scale climate manipulation is not likely to be very  
645 reassuring to applied ecologists and resource managers wishing to make accurate predictions about

646 the effects of climate change in large complex systems. Clearly more work is needed to learn how  
647 best to meet the challenge of prediction in ecological systems. Towards that goal, perhaps the best  
648 way forward is to conduct more tests like this one.

649 **Acknowledgements**

650 Funding was provided by NSF grants DEB-1353078, and DEB-1054040 to PBA, an NSF GRFP  
651 award to AK, and by the Utah Agricultural Experiment Station (get journal paper number). The  
652 USDA-ARS Sheep Experiment Station generously provided access to historical data and the field  
653 experiment site. Laureano Gherardi provided us with valuable advice in constructing the rain-out  
654 and irrigation experiment.

655 **References**

- 656 2015a. Stan: A C++ Library for Probability and Sampling.
- 657 2015b. Stan Modeling Language User's Guide and Reference Manual, Version 2.10.0.
- 658 Adler, P. B., K. M. Byrne, and J. Leiker. 2013. Can the past predict the future? Experimental  
659 tests of historically based population models. *Global Change Biology*, **19**:1793–1803.
- 660 Adler, P. B., H. J. Dagleish, and S. P. Ellner. 2012. Forecasting plant community impacts of  
661 climate variability and change: when do competitive interactions matter? *Journal of Ecology*,  
662 **100**:478–487.
- 663 Adler, P. B., S. P. Ellner, and J. M. Levine. 2010. Coexistence of perennial plants: an embarrassment  
664 of niches. *Ecology Letters*, **13**:1019–1029.
- 665 Adler, P. B., J. HilleRisLambers, and J. M. Levine. 2009. Weak effect of climate variability on  
666 coexistence in a sagebrush steppe community. *Ecology*, **90**:3303–3312.
- 667 Adler, P. B., A. R. Kleinhesselink, G. Hooker, B. Taylor, B. J. Teller, and S. P. Ellner. 2016. Weak  
668 interspecific interactions in a sagebrush steppe: Evidence from observational data, population  
669 models, and a removal experiment.
- 670 Aubry, L. M., R. F. Rockwell, E. G. Cooch, R. W. Brook, C. P. H. Mulder, and D. N. Koons.  
671 2013. Climate change, phenology, and habitat degradation: drivers of gosling body condition  
672 and juvenile survival in lesser snow geese. *Global Change Biology*, **19**:149–160.
- 673 Bates, D., M. Mchler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using  
674 lme4. *Journal of Statistical Software*, **67**:1–48.
- 675 Bazzaz, F. A. 1979. The Physiological Ecology of Plant Succession. *Annual Review of Ecology and*  
676 *Systematics*, **10**:351–371.
- 677 Caldwell, M. M., J. H. Richards, J. H. Manwaring, and D. M. Eissenstat. 1987. Rapid shifts in  
678 phosphate acquisition show direct competition between neighbouring plants. *Nature*, **327**:615–  
679 616.

- 680 Chu, C. and P. B. Adler. 2015. Large niche differences emerge at the recruitment stage to stabilize  
681 grassland coexistence. *Ecological Monographs*.
- 682 Chu, C., A. R. Kleinhesselink, K. M. Havstad, M. P. McClaran, D. P. Peters, L. T. Vermeire,  
683 H. Wei, and P. B. Adler. 2016. Direct effects dominate responses to climate perturbations in  
684 grassland plant communities. *Nature communications*, **7**.
- 685 Compagnoni, A. and P. B. Adler. 2014. Warming, soil moisture, and loss of snow increase *Bromus*  
686 *tectorum*s population growth rate. *Elementa: Science of the Anthropocene*, **2**:000020.
- 687 Dalgleish, H. J., D. N. Koons, M. B. Hooten, C. A. Moffet, and P. B. Adler. 2010. Climate influences  
688 the demography of three dominant sagebrush steppe plants. *Ecology*, **92**:75–85.
- 689 Davis, M. B. and R. G. Shaw. 2001. Range Shifts and Adaptive Responses to Quaternary Climate  
690 Change. *Science*, **292**:673–679.
- 691 Doak, D. F. and W. F. Morris. 2010. Demographic compensation and tipping points in climate-  
692 induced range shifts. *Nature*, **467**:959–962.
- 693 Ehrlin, J., W. F. Morris, T. von Euler, and J. P. Dahlgren. 2016. Advancing environmentally explicit  
694 structured population models of plants. *Journal of Ecology*, **104**:292–305.
- 695 Elmendorf, S. C., G. H. R. Henry, R. D. Hollister, A. M. Fosaa, W. A. Gould, L. Hermanutz,  
696 A. Hofgaard, I. I. Jnsdttir, J. C. Jorgenson, E. Lvesque, B. Magnusson, U. Molau, I. H. Myers-  
697 Smith, S. F. Oberbauer, C. Rixen, C. E. Tweedie, and M. Walker. 2015. Experiment, monitoring,  
698 and gradient methods used to infer climate change effects on plant communities yield consistent  
699 patterns. *Proceedings of the National Academy of Sciences*, **112**:448–452.
- 700 Franklin, R. S. 2013. Growth response of the alpine shrub, *Linanthus pungens*, to snowpack and  
701 temperature at a rock glacier site in the eastern Sierra Nevada of California, USA. *Quaternary*  
702 *International*, **310**:20–33.
- 703 Franks, S. J. 2011. Plasticity and evolution in drought avoidance and escape in the annual plant  
704 *Brassica rapa*. *New Phytologist*, **190**:249–257.

- 705 Gelman, A., J. Hwang, and A. Vehtari. 2014. Understanding predictive information criteria for  
706 Bayesian models. *Statistics and Computing*, **24**:997–1016.
- 707 Gerber, B. D., W. L. Kendall, M. B. Hooten, J. A. Dubovsky, and R. C. Drewien. 2015. Optimal  
708 population prediction of sandhill crane recruitment based on climate-mediated habitat limita-  
709 tions. *Journal of Animal Ecology*, **84**:1299–1310.
- 710 Germino, M. J. and K. Reinhardt. 2014. Desert shrub responses to experimental modification of  
711 precipitation seasonality and soil depth: relationship to the two-layer hypothesis and ecohydro-  
712 logical niche. *Journal of Ecology*, **102**:989–997.
- 713 Gherardi, L. A. and O. E. Sala. 2013. Automated rainfall manipulation system: a reliable and  
714 inexpensive tool for ecologists. *Ecosphere*, **4**:1–10.
- 715 Gherardi, L. A. and O. E. Sala. 2015. Enhanced precipitation variability decreases grass- and  
716 increases shrub-productivity. *Proceedings of the National Academy of Sciences*, **112**:12735–  
717 12740.
- 718 Hilborn, R. 2016. Correlation and Causation in Fisheries and Watershed Management. *Fisheries*,  
719 **41**:18–25.
- 720 Houlahan, J. E., S. T. McKinney, T. M. Anderson, and B. J. McGill. 2016. The priority of prediction  
721 in ecological understanding. *Oikos*, pages n/a–n/a.
- 722 Hsu, J. S. and P. B. Adler. 2014. Anticipating changes in variability of grassland production due  
723 to increases in interannual precipitation variability. *Ecosphere*, **5**:1–15.
- 724 Jenouvrier, S., H. Caswell, C. Barbraud, M. Holland, J. Strve, and H. Weimerskirch. 2009. De-  
725 mographic models and IPCC climate projections predict the decline of an emperor penguin  
726 population. *Proceedings of the National Academy of Sciences*, **106**:1844–1847.
- 727 Jentsch, A., J. Kreyling, and C. Beierkuhnlein. 2007. A new generation of climate-change experi-  
728 ments: events, not trends. *Frontiers in Ecology and the Environment*, **5**:365–374.
- 729 Knapp, A. K., M. L. Avolio, C. Beier, C. J. W. Carroll, S. L. Collins, J. S. Dukes, L. H. Fraser,  
730 R. J. Griffin-Nolan, D. L. Hoover, A. Jentsch, M. E. Loik, R. P. Phillips, A. K. Post, O. E.

- 731 Sala, I. J. Slette, L. Yahdjian, and M. D. Smith. 2016. Pushing precipitation to the extremes  
732 in distributed experiments: recommendations for simulating wet and dry years. *Global Change  
733 Biology*.
- 734 Knapp, A. K. and M. D. Smith. 2001. Variation Among Biomes in Temporal Dynamics of Above-  
735 ground Primary Production. *Science*, **291**:481–484.
- 736 Koons, D. N., P. Terletzky, P. B. Adler, M. L. Wolfe, D. Ranglack, F. P. Howe, K. Hersey, W. Pas-  
737 kett, and J. T. d. Toit. 2012. Climate and density-dependent drivers of recruitment in plains  
738 bison. *Journal of Mammalogy*, **93**:475–481.
- 739 Krner, C. 2012. Alpine Treelines: Functional Ecology of the Global High Elevation Tree Limits.  
740 Springer Science & Business Media.
- 741 Lauenroth, W. K. and P. B. Adler. 2008. Demography of perennial grassland plants: survival, life  
742 expectancy and life span. *Journal of Ecology*, **96**:1023–1032.
- 743 Levine, J. M., A. K. McEachern, and C. Cowan. 2010. Do competitors modulate rare plant response  
744 to precipitation change? *Ecology*, **91**:130–140.
- 745 Lunn, N. J., S. Servany, E. V. Regehr, S. J. Converse, E. Richardson, and I. Stirling. 2016.  
746 Demography of an apex predator at the edge of its range impacts of changing sea ice on polar  
747 bears in Hudson Bay. *Ecological Applications*, pages n/a–n/a.
- 748 Parmesan, C. and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across  
749 natural systems. *Nature*, **421**:37–42.
- 750 Perfors, T., J. Harte, and S. E. Alter. 2003. Enhanced growth of sagebrush (*Artemisia tridentata*)  
751 in response to manipulated ecosystem warming. *Global Change Biology*, **9**:736–742.
- 752 Petchey, O. L., M. Pontarp, T. M. Massie, S. Kfi, A. Ozgul, M. Weilenmann, G. M. Palamara, F. Al-  
753 termatt, B. Matthews, J. M. Levine, D. Z. Childs, B. J. McGill, M. E. Schaeppman, B. Schmid,  
754 P. Spaak, A. P. Beckerman, F. Pennekamp, and I. S. Pearse. 2015. The ecological forecast horizon,  
755 and examples of its uses and determinants. *Ecology Letters*, **18**:597–611.

- 756 Post, E. and M. C. Forchhammer. 2002. Synchronization of animal population dynamics by large-  
757 scale climate. *Nature*, **420**:168–171.
- 758 Srur, A. M. and R. Villalba. 2009. Annual growth rings of the shrub *Anarthrophyllum rigidum*  
759 across Patagonia: Interannual variations and relationships with climate. *Journal of Arid Envi-*  
760 *ronments*, **73**:1074–1083.
- 761 Sturges, D. L. 1989. Response of Mountain Big Sagebrush to Induced Snow Accumulation. *Journal*  
762 *of Applied Ecology*, **26**:1035–1041. ArticleType: research-article / Full publication date: Dec.,  
763 1989 / Copyright 1989 British Ecological Society.
- 764 Teller, B. J., P. B. Adler, C. B. Edwards, G. Hooker, and S. P. Ellner. 2016. Linking demography  
765 with drivers: climate and competition. *Methods in Ecology and Evolution*, **7**:171–183.
- 766 Tredennick, A. T., M. B. Hooten, and P. B. Adler. 2016. Do we need demographic data to forecast  
767 plant population dynamics? *Methods in Ecology and Evolution*, pages n/a–n/a.
- 768 Tylianakis, J. M., R. K. Didham, J. Bascompte, and D. A. Wardle. 2008. Global change and species  
769 interactions in terrestrial ecosystems. *Ecology Letters*, **11**:1351–1363.
- 770 Venable, D. L. 2007. Bet Hedging in a Guild of Desert Annuals. *Ecology*, **88**:1086–1090.
- 771 Walther, G.-R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J.-M. Fromentin,  
772 O. Hoegh-Guldberg, and F. Bairlein. 2002. Ecological responses to recent climate change. *Nature*,  
773 **416**:389–395.
- 774 Wang, J., P. M. Rich, and K. P. Price. 2003. Temporal responses of NDVI to precipitation and  
775 temperature in the central Great Plains, USA. *International Journal of Remote Sensing*, **24**:2345–  
776 2364.
- 777 Whittaker, R. H. 1975. Communities and ecosystems. Macmillan, New York, second edition edition.
- 778 Williams, J. W. and S. T. Jackson. 2007. Novel climates, no-analog communities, and ecological  
779 surprises. *Frontiers in Ecology and the Environment*, **5**:475–482.
- 780 Wolkovich, E. M., B. I. Cook, J. M. Allen, T. M. Crimmins, J. L. Betancourt, S. E. Travers, S. Pau,  
781 J. Regetz, T. J. Davies, N. J. B. Kraft, T. R. Ault, K. Bolmgren, S. J. Mazer, G. J. McCabe,

- 782 B. J. McGill, C. Parmesan, N. Salamin, M. D. Schwartz, and E. E. Cleland. 2012. Warming  
783 experiments underpredict plant phenological responses to climate change. *Nature*, **485**:494–497.
- 784 Yang, B., C. Qin, J. Wang, M. He, T. M. Melvin, T. J. Osborn, and K. R. Briffa. 2014. A 3,500-year  
785 tree-ring record of annual precipitation on the northeastern Tibetan Plateau. *Proceedings of the*  
786 *National Academy of Sciences*, **111**:2903–2908.

## Tables

Table 1: Spring soil moisture

	Model 1
(Intercept)	8.81*** (1.54)
TreatmentDrought	-3.97*** (0.45)
TreatmentIrrigation	3.26*** (0.45)
AIC	3191.87
BIC	3222.92
Log Likelihood	-1588.93
Num. obs.	624
Num. groups: plot	24
Num. groups: PrecipGroup	8
Num. groups: date	5
Var: plot (Intercept)	0.45
Var: PrecipGroup (Intercept)	0.23
Var: date (Intercept)	11.24
Var: Residual	8.90

\*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$

Table 2: soil moisture model

	Model 1
(Intercept)	-0.57*** (0.16)
TreatmentIrrigation	1.23*** (0.03)
rainfallrainy	-0.05 (0.03)
seasonspring	0.27*** (0.02)
seasonsummer	0.15*** (0.02)
seasonwinter	0.25*** (0.02)
TreatmentIrrigation:rainfallrainy	0.18*** (0.03)
TreatmentIrrigation:seasonspring	-0.23*** (0.03)
TreatmentIrrigation:seasonsummer	-0.26*** (0.03)
TreatmentIrrigation:seasonwinter	-0.33*** (0.03)
rainfallrainy:seasonspring	-0.23*** (0.04)
rainfallrainy:seasonsummer	-0.07 (0.04)
rainfallrainy:seasonwinter	-0.07 (0.07)
AIC	14581.58
BIC	14695.49
Log Likelihood	-7274.79
Num. obs.	9133
Num. groups: simple_date	1596
Num. groups: PrecipGroup	8
Var: simple_date (Intercept)	0.00
Var: PrecipGroup (Intercept)	0.19
Var: Residual	2.50

\*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$

Table 3: Treatment effects on log cover change for extitA. tripartita from 2011 to 2016. Intercept gives control effects.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.2835	0.4627	-0.61	0.5487
TreatmentDrought	0.9378	0.7743	1.21	0.2434
TreatmentIrrigation	0.8882	0.7743	1.15	0.2682

Table 4: Treatment effects on log cover change for extitH. comata from 2011 to 2016. Intercept gives control effects.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.3982	0.2587	1.54	0.1548
TreatmentDrought	-2.9951	0.5784	-5.18	0.0004
TreatmentIrrigation	-0.1219	0.4953	-0.25	0.8105

Table 5: Treatment effects on log cover change for extitP. secunda from 2011 to 2016. Intercept gives control effects.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.7247	0.4613	-1.57	0.1298
TreatmentDrought	0.0273	0.8208	0.03	0.9737
TreatmentIrrigation	1.1459	0.7797	1.47	0.1552

Table 6: Treatment effects on log cover change for extitP. spicata from 2011 to 2016. Intercept gives control effects.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.0188	0.2124	0.09	0.9303
TreatmentDrought	-0.8851	0.3780	-2.34	0.0287
TreatmentIrrigation	0.1453	0.3780	0.38	0.7044

Table 7: Selected climate variables for each vital rate model for each species. Correlations and p-values between the chosen variables and the intercept of year effects model are shown. For ARTR growth and POSE growth and survival, the correlations between the year effects on size and the soil moisture variables are also given. "f" = fall, "su" = summer, "sp" = spring. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

vital_rate	species	climate variable	Int. cor.	p val.	Size cor.	Size p. val.
growth	ARTR	su.0	-0.49	0.02	0.26	0.26
growth	ARTR	f.0	-0.28	0.23	0.40	0.08
growth	ARTR	sp.1	0.17	0.45	-0.33	0.14
growth	HECO	su.1	0.69	0.00		
growth	HECO	su.0	0.50	0.02		
growth	HECO	f.lag	0.37	0.10		
growth	POSE	f.lag	0.31	0.17	-0.11	0.64
growth	POSE	su.lag	0.29	0.20	-0.20	0.38
growth	POSE	sp.1	0.26	0.25	-0.20	0.38
growth	PSSP	f.lag	0.34	0.13		
growth	PSSP	su.lag	0.25	0.27		
growth	PSSP	f.0	-0.22	0.34		
recruitment	ARTR	su.lag	-0.32	0.16		
recruitment	ARTR	su.0	-0.26	0.25		
recruitment	ARTR	sp.1	0.22	0.34		
recruitment	HECO	su.lag	-0.31	0.18		
recruitment	HECO	su.0	-0.30	0.18		
recruitment	HECO	f.lag	0.19	0.40		
recruitment	POSE	sp.1	0.49	0.02		
recruitment	POSE	f.lag	0.34	0.13		
recruitment	POSE	f.1	0.32	0.16		
recruitment	PSSP	su.lag	-0.52	0.02		
recruitment	PSSP	su.0	-0.48	0.03		
recruitment	PSSP	sp.0	0.30	0.19		
survival	ARTR	su.0	-0.60	0.00		
survival	ARTR	sp.0	-0.41	0.06		
survival	ARTR	su.1	-0.40	0.07		
survival	HECO	sp.0	0.44	0.04		
survival	HECO	sp.lag	0.43	0.05		
survival	HECO	f.1	0.33	0.14		
survival	POSE	sp.0	0.44	0.04	0.22	0.34
survival	POSE	sp.1	0.27	0.23	-0.46	0.04
survival	POSE	f.lag	-0.00	0.99	0.30	0.19
survival	PSSP	sp.0	0.36	0.11		
survival	PSSP	sp.lag	0.34	0.13		
survival	PSSP	su.1	0.26	0.26		

Table 8: Comparison of model predictions from climate model and year effects model for each species and vital rate. Two prediction scores are reported, MSE and lppd. Lower MSE indicates improved predictions whereas higher lppd indicates improved predictions. Instances where the climate model outperformed the random year effects model are marked with "\*\*\*\*" in the last column. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

species	vital_rate	score	climate model	year effects model	diff	improved
ARTR	growth	lppd	-186.08	-189.17	3.09	***
ARTR	growth	MSE	0.55	0.55	-0.00	***
ARTR	recruitment	lppd	-79.05	-77.81	-1.24	
ARTR	recruitment	MSE	169.86	9.47	160.39	
ARTR	survival	lppd	-37.55	-36.42	-1.13	
ARTR	survival	MSE	0.06	0.06	0.00	
HECO	growth	lppd	-475.54	-454.36	-21.18	
HECO	growth	MSE	1.26	1.18	0.09	
HECO	recruitment	lppd	-149.43	-151.60	2.17	***
HECO	recruitment	MSE	288.16	235.57	52.59	
HECO	survival	lppd	-158.99	-147.06	-11.93	
HECO	survival	MSE	0.12	0.11	0.01	
POSE	growth	lppd	-1823.71	-1831.56	7.85	***
POSE	growth	MSE	1.72	1.73	-0.01	***
POSE	recruitment	lppd	-260.05	-257.46	-2.59	
POSE	recruitment	MSE	45.40	37.32	8.08	
POSE	survival	lppd	-698.06	-718.60	20.54	***
POSE	survival	MSE	0.14	0.14	-0.00	***
PSSP	growth	lppd	-1232.93	-1237.92	4.99	***
PSSP	growth	MSE	1.51	1.51	0.00	
PSSP	recruitment	lppd	-271.34	-273.94	2.60	***
PSSP	recruitment	MSE	79.09	42.68	36.41	
PSSP	survival	lppd	-332.47	-307.26	-25.21	
PSSP	survival	MSE	0.11	0.10	0.01	

Table 9: MSE of predicted log cover changes and correlations between log cover changes predicted and observed. Predictions for the cover changes in the experimental plots were generated either from the year effects or the climate models. Instances where the climate model made better predictions than the year effects model are indicated with the "\*\*\*\*". ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

	species	stat	year effects model	climate model	diff	improved
1	ARTR	cor	0.48	0.19	-0.29	
2	ARTR	MSE	0.30	0.30	0.00	
3	HECO	cor	0.29	0.22	-0.07	
4	HECO	MSE	0.49	0.57	0.07	
5	POSE	cor	0.45	0.53	0.07	***
6	POSE	MSE	0.42	0.41	-0.01	***
7	PSSP	cor	0.36	0.38	0.03	***
8	PSSP	MSE	0.39	0.39	-0.01	***

## Figures

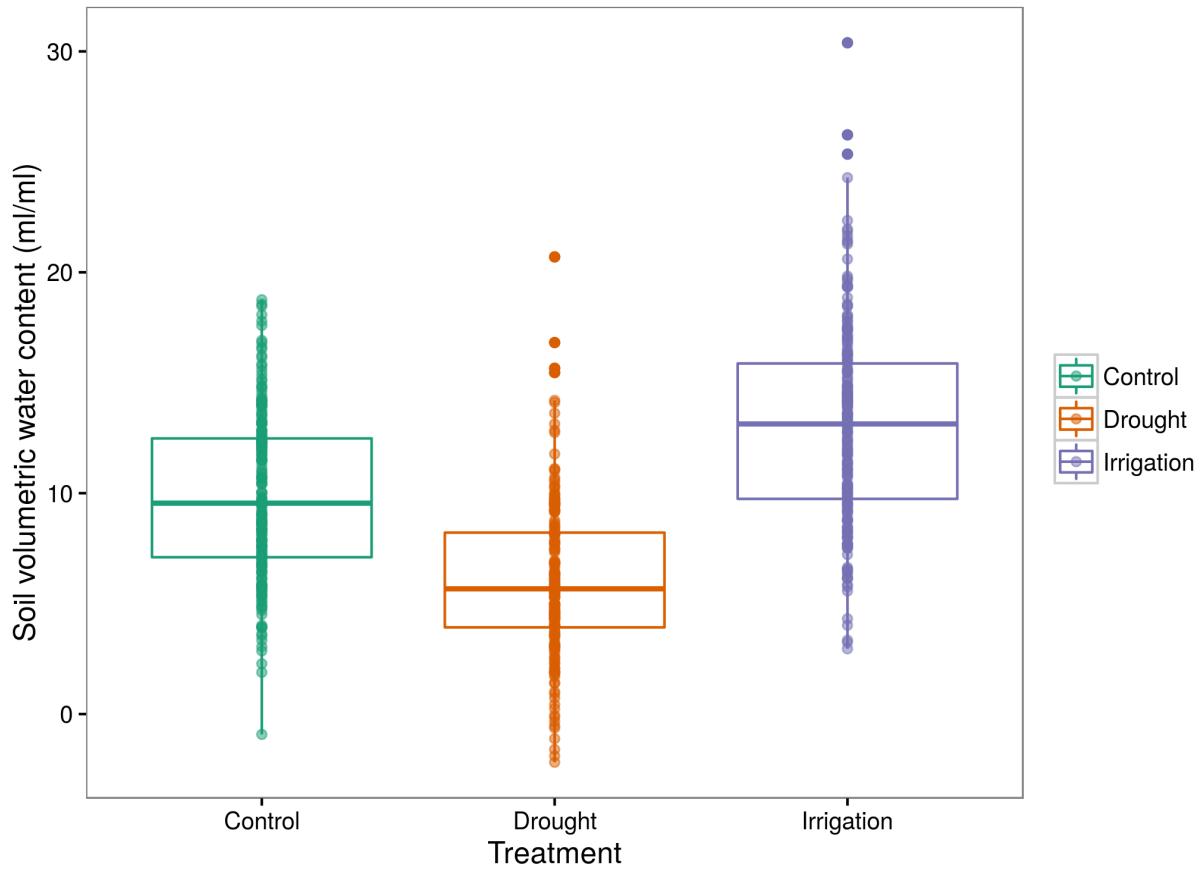


Figure 1: Soil moisture in the upper 5 cm of drought and irrigated plots compared to ambient controls. Soil moisture was measured at six locations around each plot at five different dates during the spring. Control plots were nearby areas of experiencing ambient soil moisture. Box plots show the median soil moisture and the interquartile range. Dots show individual soil moisture measurements. Readings of volumetric soil moisture less than zero were occasionally obtained in very dry soil.

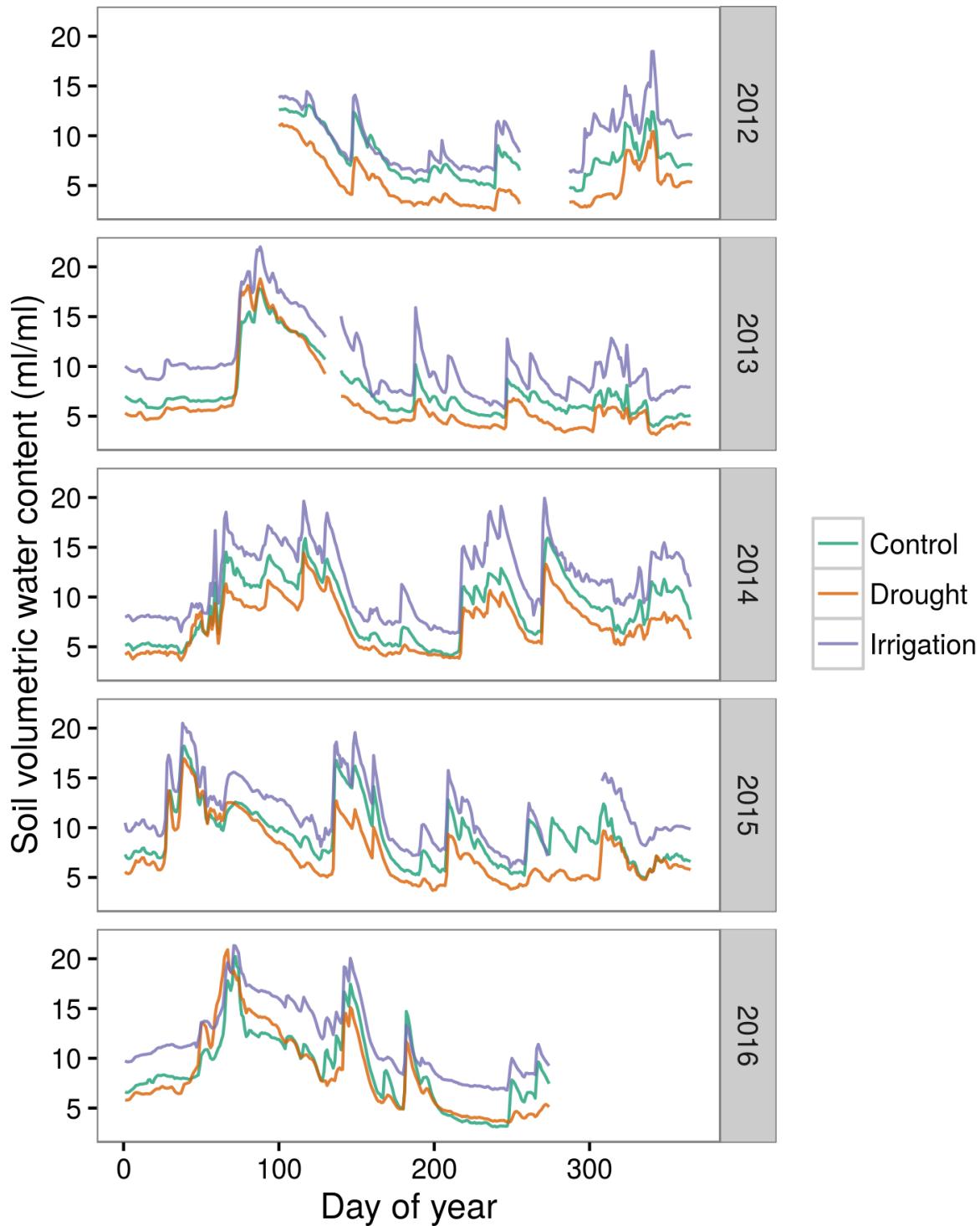


Figure 2: Average soil moisture in the control, drought, and irrigation treatments during each year of the experiment. Soil moisture was monitored in four drought plots, four irrigated plots and four ambient control plots. Two sensors were installed at 5 cm depth at each plot and two at 25 cm and data was logged every 2 hours.

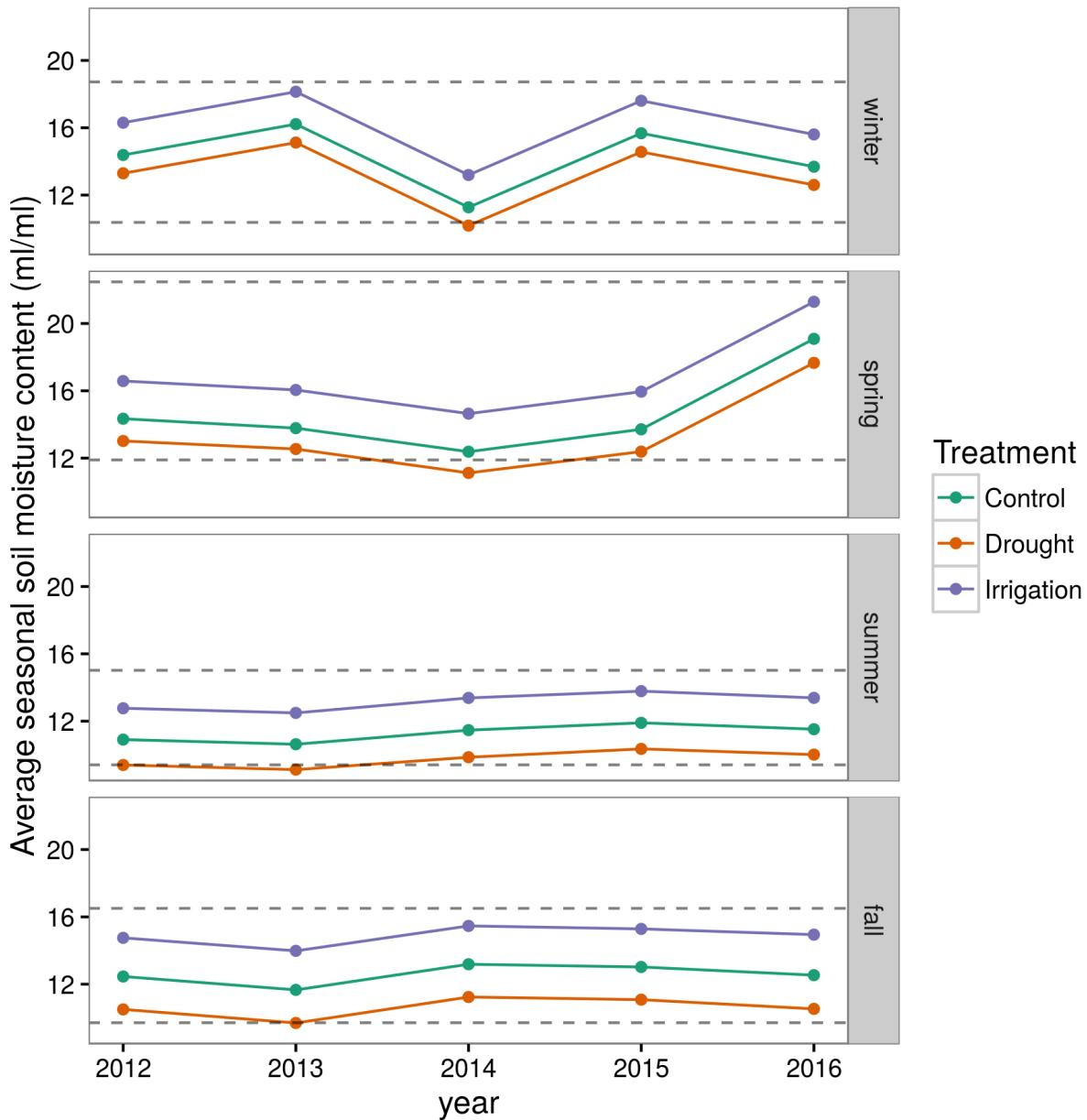


Figure 3: Average seasonal soil moisture in the control, drought, and irrigation treatments during each year of the experiment. The dashed gray lines give the 5th percentile and 95th percentile limits for seasonal soil moisture in the historical record (1929 to 2010).

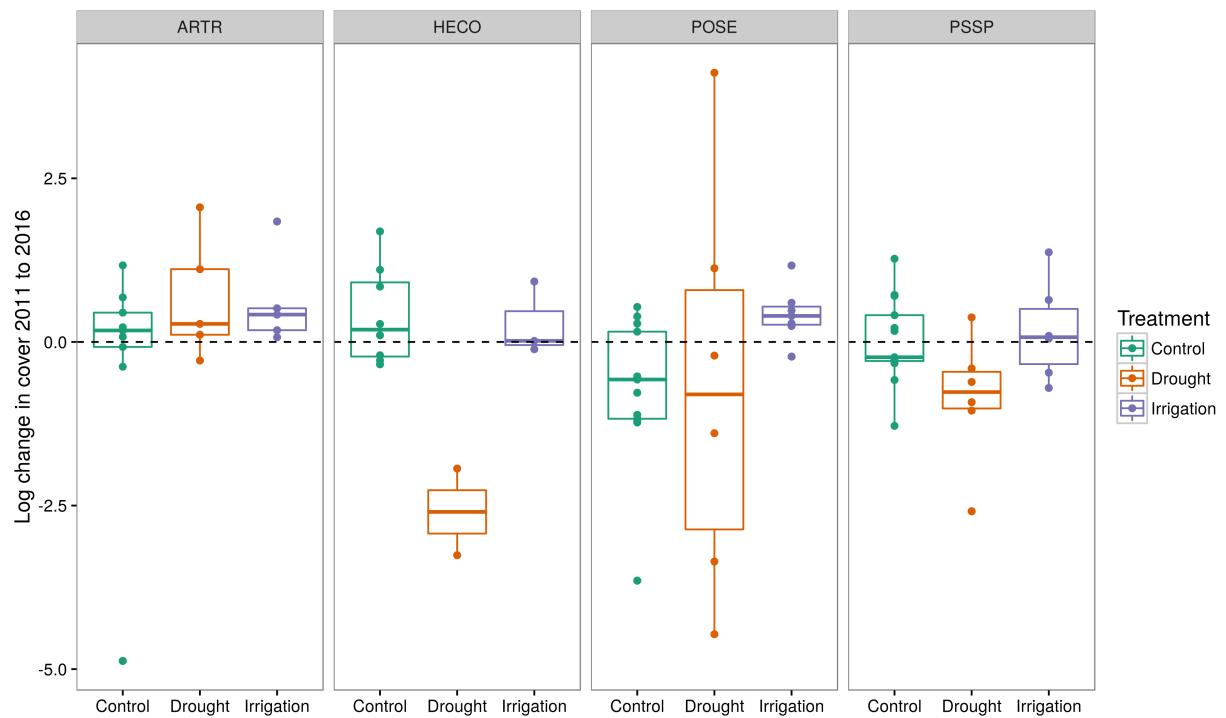


Figure 4: Log change in cover in each of the experimental plots from the pre-treatment monitoring in 2011 to the last year of the experiment in 2016. Box plots show the median cover change and the interquartile range. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

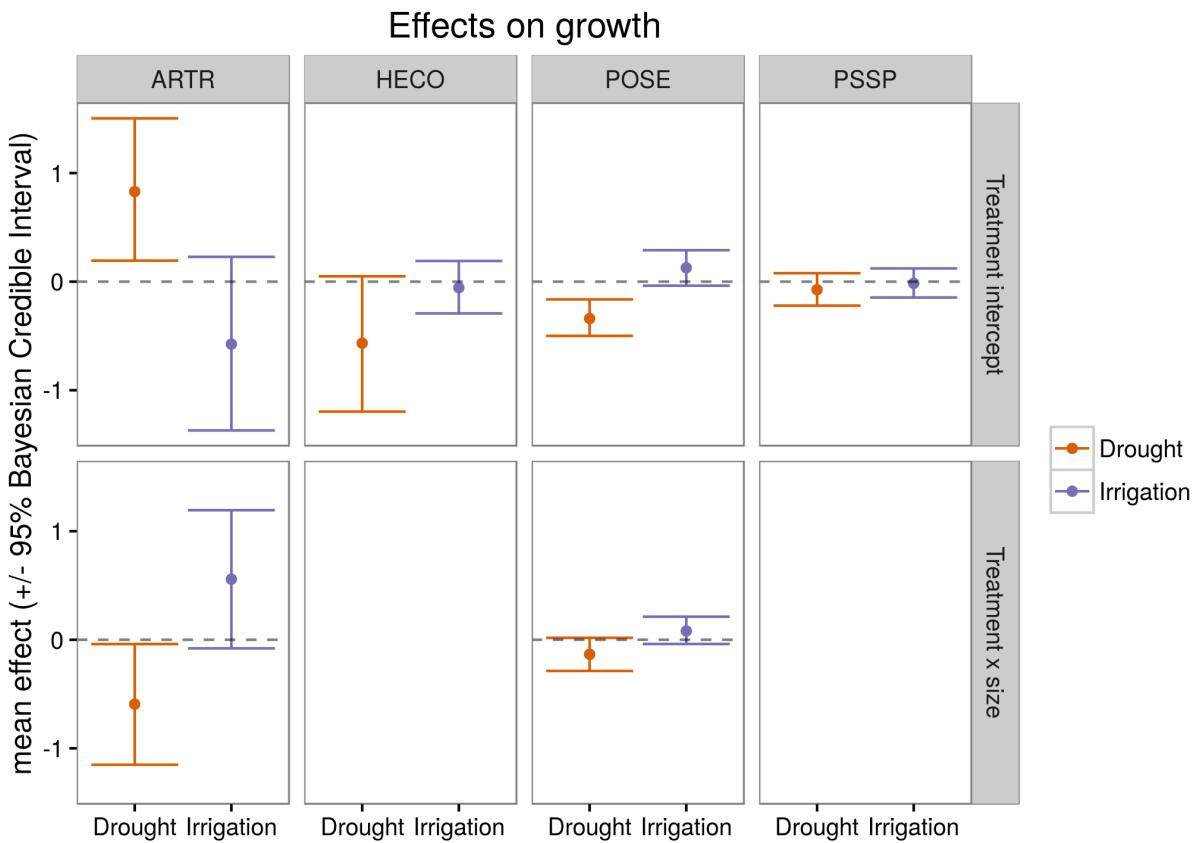


Figure 5: Parameter estimates for the effects of treatment on growth of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. Size by treatment interactions were only fit for ARTR, and POSE. Plant size was centered on mean size and scaled by its standard deviation. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

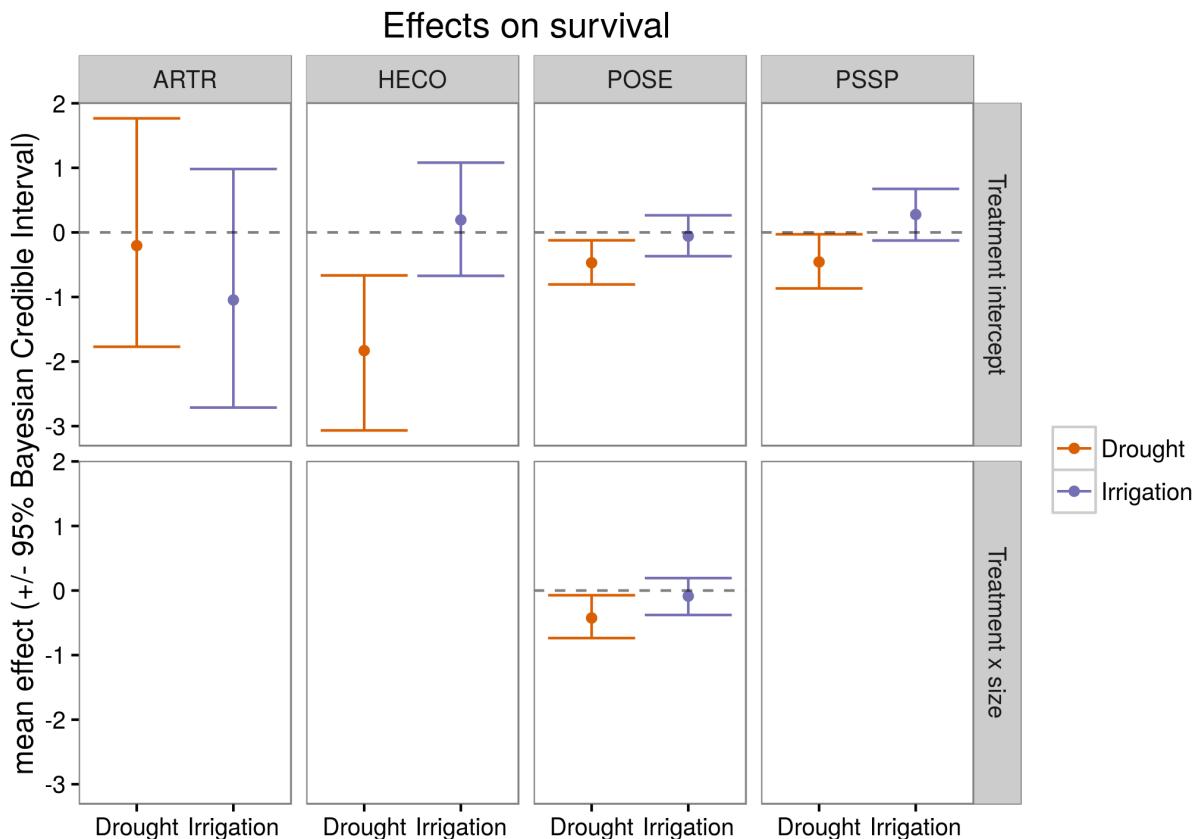


Figure 6: Parameter estimates for the effects of treatment on survival of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. Size by treatment interactions were only fit for POSE. Plant size was centered on mean size and scaled by its standard deviation. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

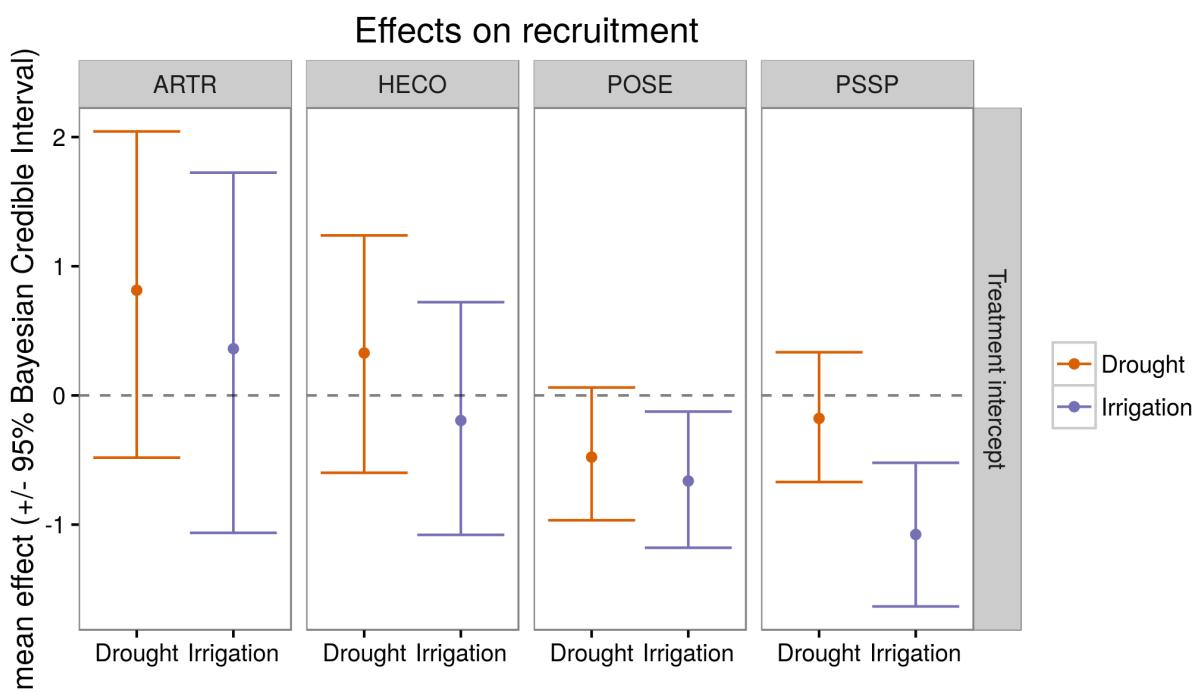


Figure 7: Parameter estimates for the effects of treatment on recruitment of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

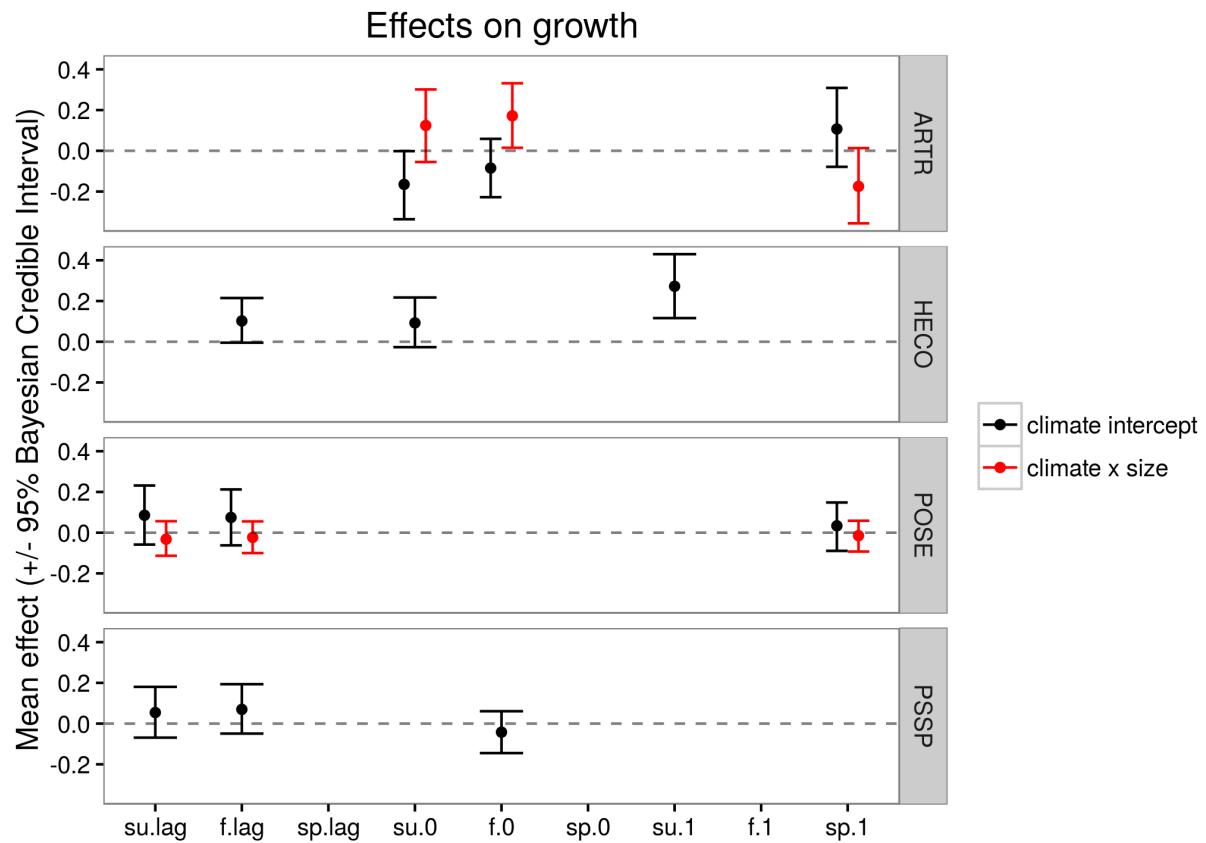


Figure 8: Parameter estimates for the selected seasonal soil moisture covariates on the growth of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. Dark blue parameters show size x climate interaction effects. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

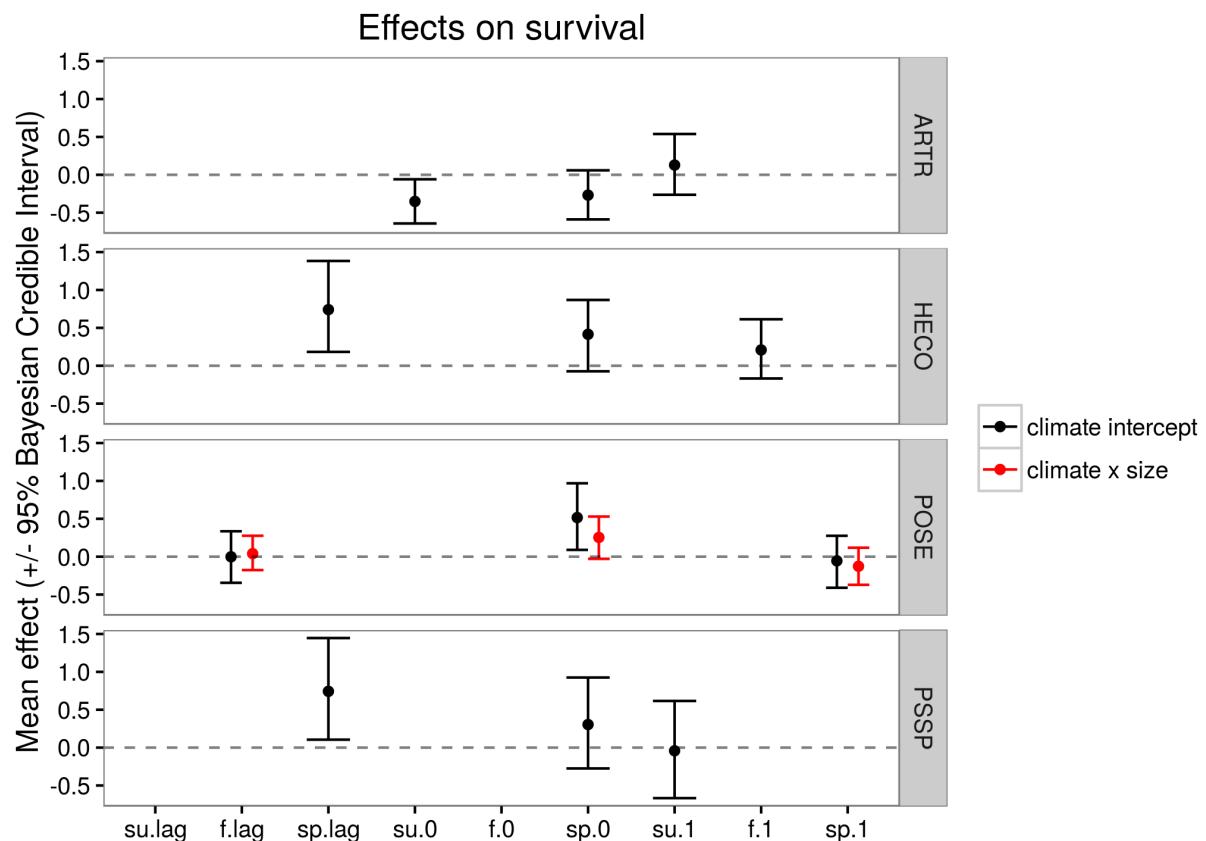


Figure 9: Parameter estimates for the selected seasonal soil moisture covariates on the survival of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. Dark blue parameters show size x climate interaction effects. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

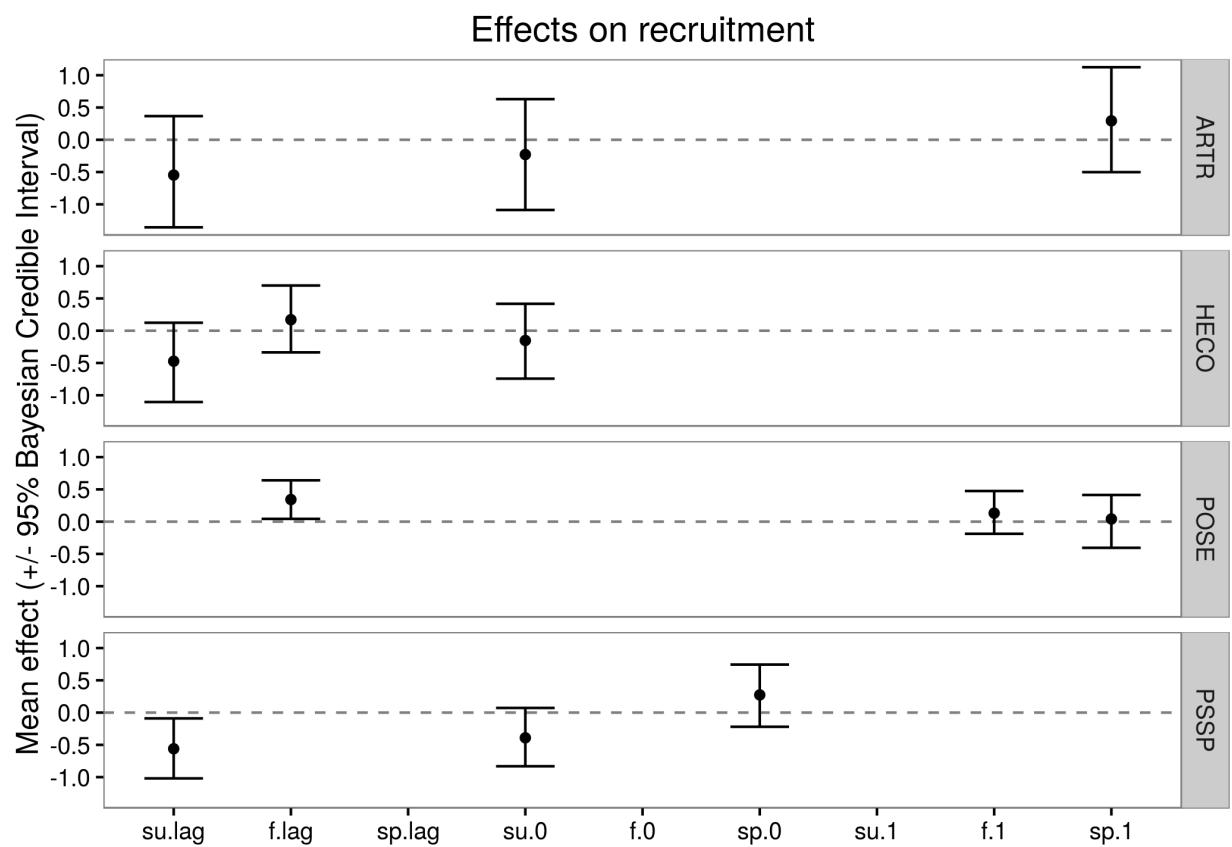


Figure 10: Parameter estimates for the selected seasonal soil moisture covariates on the survival of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. Dark blue parameters show size x climate interaction effects. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

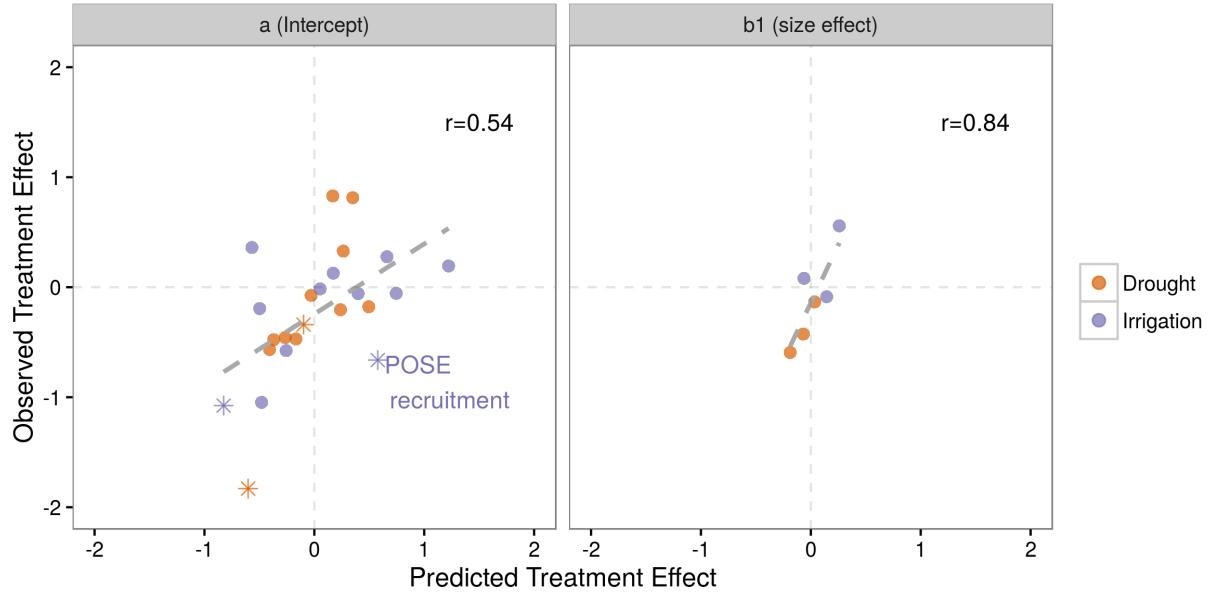


Figure 11: The treatment effects predicted by the climate model compared to the treatment effects observed for the intercept parameters (left side) and size by climate/treatment effects (right side). The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational data but do not include the five years of the experiment (2011 to 2016). Treatment parameters that were both observed and predicted to be significantly different from zero are shown with the “\*” symbol. The correlation between predicted and observed parameters is given on each panel. *P. secunda* recruitment was predicted to be positively affected by the irrigation treatment but was in fact negatively affected. The other significant effects were in the correct direction. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero.

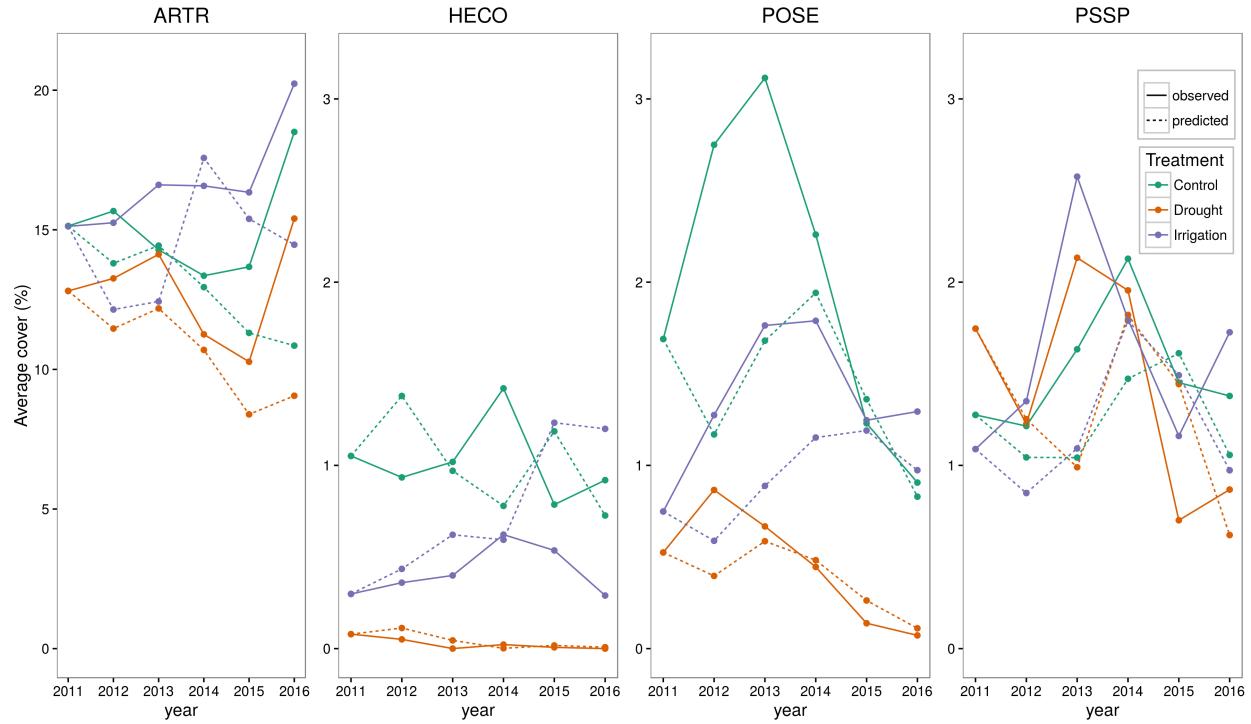


Figure 12: Observed average cover per quadrat in the experimental and control plots and one step ahead cover predictions from the climate model. Cover predictions for each year are generated from the IBM based on the observed distribution of plants in each quadrat in the current year. Quadrat cover was not predicted for the first year of the experiment in 2011. Note the different cover scales for ARTR and the three grass species. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

### *A. tripartita*

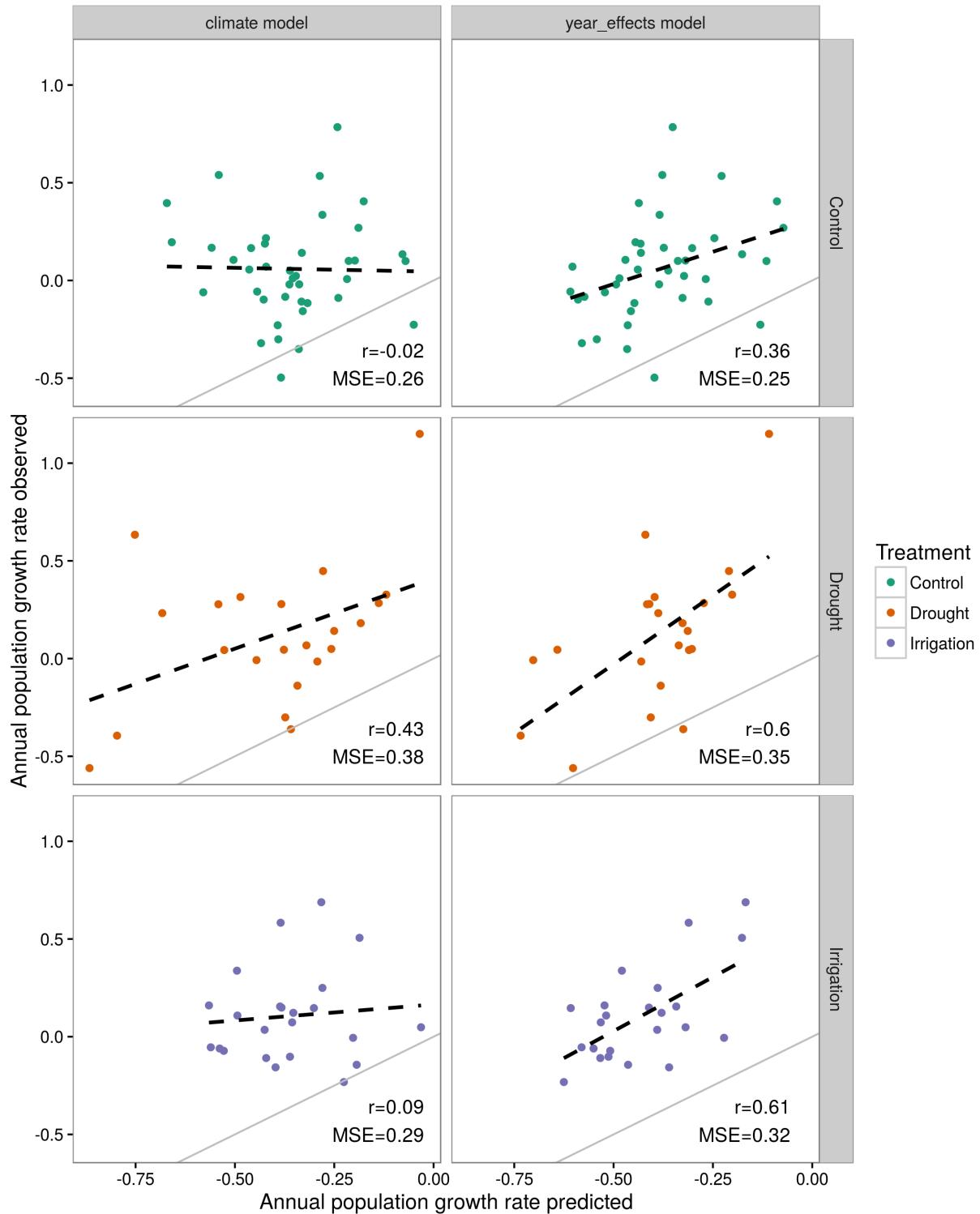


Figure 13: Observed and predicted one step ahead log change in *A. tripartita* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations and MSE are shown for each treatment and model. Gray line shows 1:1 line.

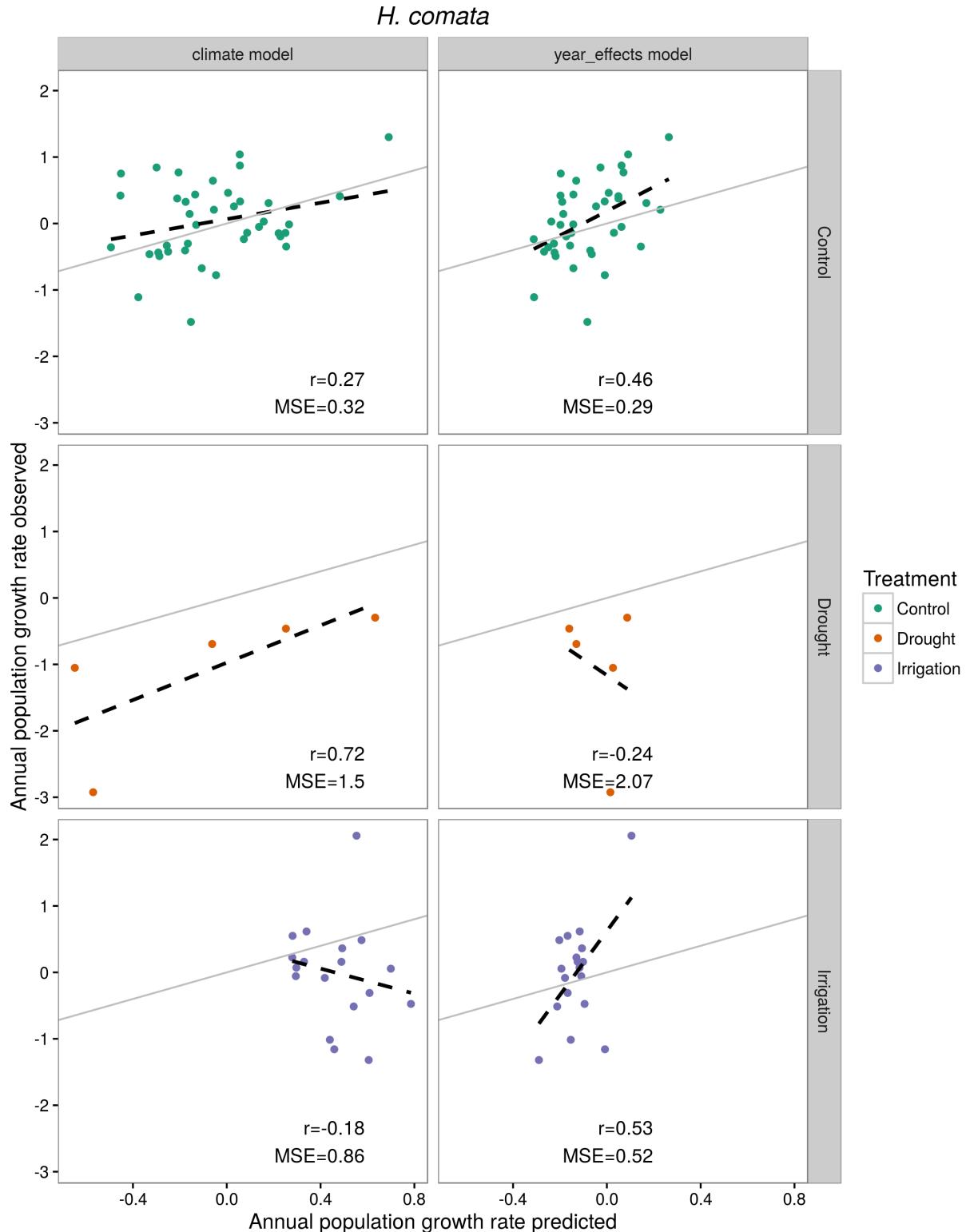


Figure 14: Observed and predicted one step ahead log change in *H. comata* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations and MSE are shown for each treatment and model. Gray line shows 1:1 line.

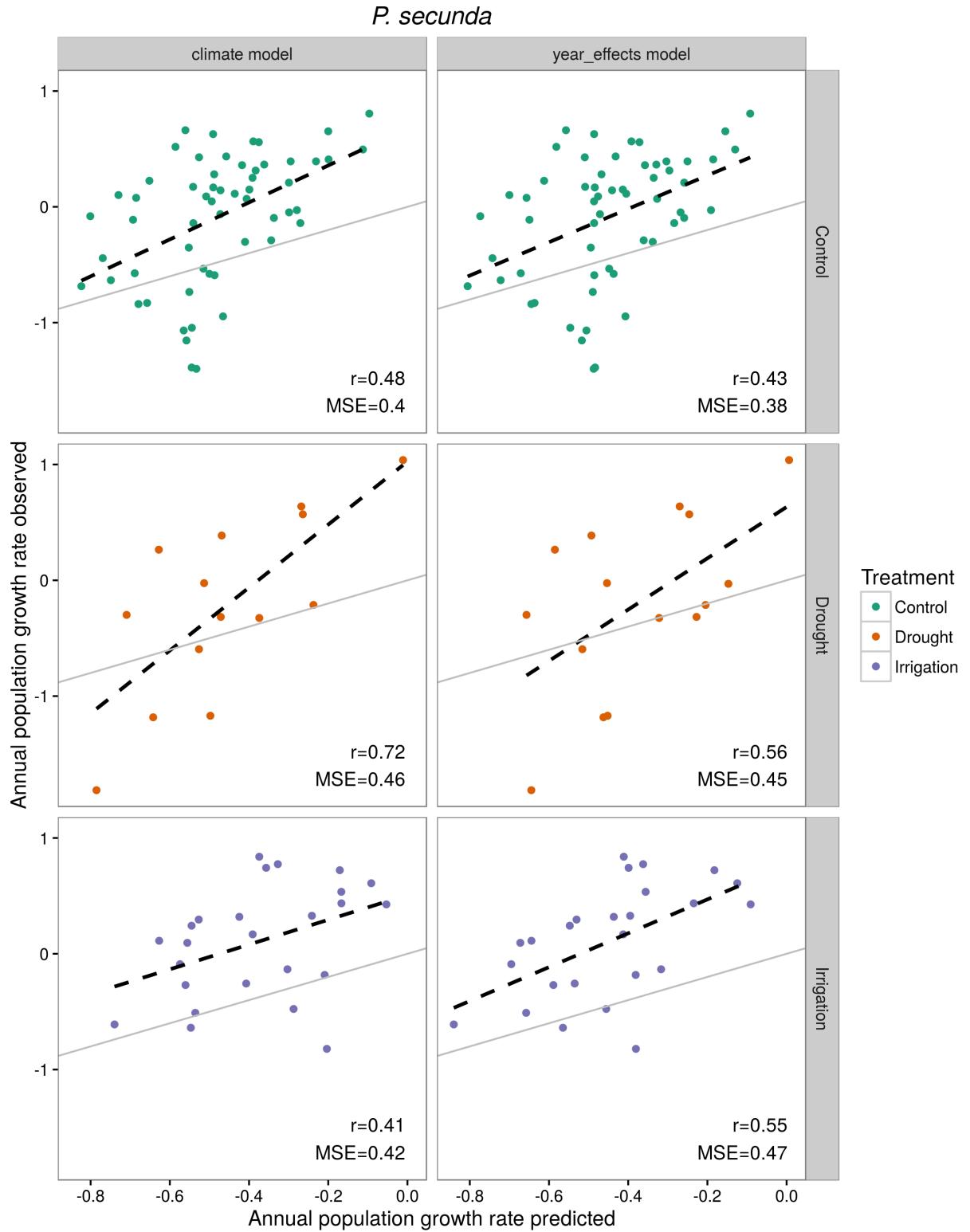


Figure 15: Observed and predicted one step ahead log change in *P. secunda* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations and MSE are shown for each treatment and model. Gray line shows 1:1 line.

*P. spicata*

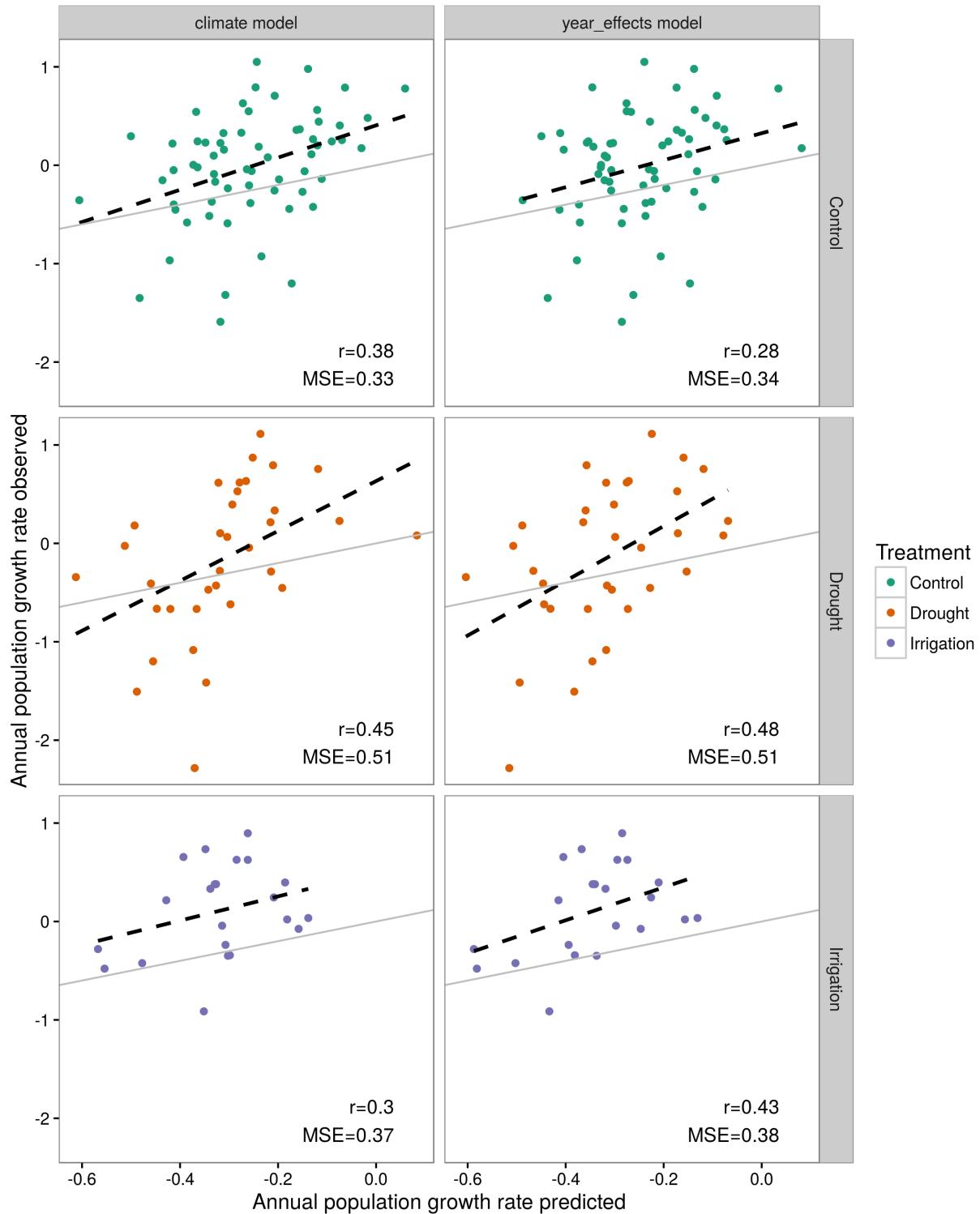


Figure 16: Observed and predicted one step ahead log change in *P. spicata* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations and MSE are shown for each treatment and model. Gray line shows 1:1 line.

## Supporting Information

Kleinheisselink et al., “Predicting climate response”

## Supplementary Methods

### Interspecific covariance in local crowding

We explored interspecific covariance in local crowding experienced by individual plants, by regressing the  $W$  values exerted by one neighbor species, the response variable, against the  $W$  values of all other species, the independent variables. Because some  $W = 0$ , we conducted two separate regressions. First, using all  $W$ ’s, we fitted a generalized linear model with a logit link function to evaluate whether the probability that the focal species’  $W = 0$  is influenced by the value of other species’  $W$ ’s. In this model, the dependent variable is a Bernoulli variate coding for the zero or non-zero value of the focal species’ crowding, and the independent variables are the  $W$ ’s for all other species. Second, for the set of records in which the focal species has  $W > 0$ , we performed a linear regression, where the focal species’  $W$  is the dependent variable, and the other species’  $W$ ’s are the independent variables. We repeated these regressions for each focal species. Due to large samples size, interspecific  $W$  values were often statistically significant predictors of intraspecific. However, they explained very little variance. The maximum reduction in deviance for the generalized linear regressions and  $R^2$  for the linear regressions were both less than 8%. The R code for this analysis is included as ..\Wdistrib\exploreSurvivalWs.r.

### Mean field approximation of local crowding for the IPM

Adler et al. (2010) developed a mean field approximation for local crowding when the competition kernels are all Gaussian functions,  $F_{jm}(d) = e^{-\alpha_{jm}d^2}$ . The approximation is explained in the online SI to Adler et al. (2010) and in section 5.3 of ?. Here we explain how that approximation was modified for the IPMs in this paper, which used fitted nonparametric competition kernels.

For  $j \neq m$  (between-species competition), overlap between individuals is allowed. The mean field approximation is that from the perspective of any focal plant in species  $j$ , individuals of species  $m$  are distributed at random in space, independent of each other and of their size.

Consider the region between the circles of radius  $x$  and  $x + dx$  centered on a focal genet of species  $j$ . The area of this annulus is  $2\pi x dx$  to leading order for  $dx \approx 0$ . A species  $m$  genet in the annulus puts competitive pressure  $F_{jm}(x)$  times its area on the focal genet. The expected total competitive pressure from all such genets is therefore is  $F_{jm}(x)2\pi x dx$  times the expected fractional cover of species  $m$  in the annulus (fractional cover is the total area of species  $m$  genets, as a fraction of the total area). The excepted fractional cover  $C_m$  of species  $m$  in the annulus equals its fractional cover in the habitat as a whole, because of the assumption of random distribution spatial distributions. We therefore have  $C_m = \int e^u n_m(u, t) du / A$  where  $A$  is the total area of the

habitat. The total expected competitive pressure on a species- $j$  genet due to species  $m$  is then

$$W_{jm} = \int_0^\infty C_m F_{jm}(x) 2\pi x \, dx = C_m \left[ 2\pi \int_0^\infty x F(x) \, dx \right]. \quad (\text{SI.1})$$

The quantity in square brackets is a constant (that is, it only depends on what the kernel function is) so it can be computed once and for all for each kernel used in the IPM. The integral is finite because all fitted kernels fall to zero at a finite distance from the focal plant.

Our kernel fitting method only uses competition kernel values at the “mid-ring” distances halfway between the inner and outer radii of a series of annuli around each focal plant, scaled so that the value at the innermost mid-ring distance equals 1. In the IPM we defined the kernel at other distances by linear interpolation between values at mid-ring distances, except that for the innermost ring a kernel value of 1 was specified at the outer radius of the ring and at distance  $x = 0$ .

Now consider within-species competition. We assume that conspecifics cannot overlap. Genet shapes are irregular, but we nonetheless implement the no-overlap rule by assuming that a genet of log area  $u_i$  is a circle of radius  $r_i$  where  $\pi r_i^2 = e^{u_i}$ . The no-overlap rule is then that the centroids of two conspecific individuals must be separated by at least the sum of their radii.

For any one focal genet, the no-overlap restriction on its neighbors’ locations affects only a negligibly small part of the habitat. The expected cover of individuals in the places where they can occur (relative to one focal plant) is thus assumed to equal their expected locations in the habitat as a whole.

Let  $C_m(u)$  be the total cover of species  $m$  genets of radius  $r$  or smaller,

$$C_m(r) = \int_L^{\log(\pi r^2)} e^z n_m(z, t) \, dz. \quad (\text{SI.2})$$

A focal genet of radius  $r$  cannot have any conspecific neighbors centered at distances less than  $r$ . It can have a neighbor centered at distance  $x > r$  if that neighbor’s radius is no more than  $x - r$ . Adding up the expected cover of all such possible neighbors for a focal genet of radius  $r$ ,

$$W_{mm}(r) = 2\pi \int_r^\infty F_{mm}(x) x C_m(x - r) \, dx \quad (\text{SI.3})$$

This integral is again finite and computable because the kernels  $F$  fall to 0 at finite  $x$ .

Table SI-1: Comparison of model predictions from climate model and year effects model for each species and vital rate and treatment. Two prediction scores are reported, MSE and lppd. Lower MSE indicates improved predictions whereas higher lppd indicates improved predictions. Instances where the climate model outperformed the random year effects model are marked with “\*\*\*” in the last column. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

	species	vital_rate	Treatment	score	climate model	year effects model	diff	improved
1	ARTR	growth	Control	lppd	-107.53	-108.40	0.87	***
2	ARTR	growth	Control	MSE	0.57	0.56	0.01	
3	ARTR	growth	Drought	lppd	-39.71	-40.01	0.30	***
4	ARTR	growth	Drought	MSE	0.53	0.49	0.03	
5	ARTR	growth	Irrigation	lppd	-38.85	-40.75	1.91	***
6	ARTR	growth	Irrigation	MSE	0.50	0.57	-0.07	***
7	ARTR	recruitment	Control	lppd	-32.61	-31.05	-1.56	
8	ARTR	recruitment	Control	MSE	61.88	10.59	51.29	
9	ARTR	recruitment	Drought	lppd	-28.93	-26.52	-2.41	
10	ARTR	recruitment	Drought	MSE	523.95	11.62	512.33	
11	ARTR	recruitment	Irrigation	lppd	-17.51	-20.24	2.73	***
12	ARTR	recruitment	Irrigation	MSE	4.72	5.34	-0.62	***
13	ARTR	survival	Control	lppd	-24.19	-23.06	-1.13	
14	ARTR	survival	Control	MSE	0.07	0.07	0.00	
15	ARTR	survival	Drought	lppd	-5.48	-5.34	-0.13	
16	ARTR	survival	Drought	MSE	0.04	0.04	-0.00	***
17	ARTR	survival	Irrigation	lppd	-7.88	-8.01	0.13	***
18	ARTR	survival	Irrigation	MSE	0.06	0.06	-0.00	***
19	HECO	growth	Control	lppd	-377.77	-369.46	-8.31	
20	HECO	growth	Control	MSE	1.11	1.09	0.02	
21	HECO	growth	Drought	lppd	-8.74	-10.97	2.23	***
22	HECO	growth	Drought	MSE	3.12	4.35	-1.23	***
23	HECO	growth	Irrigation	lppd	-89.03	-73.93	-15.10	
24	HECO	growth	Irrigation	MSE	2.02	1.47	0.55	
25	HECO	recruitment	Control	lppd	-93.35	-93.47	0.12	***
26	HECO	recruitment	Control	MSE	613.41	499.11	114.29	
27	HECO	recruitment	Drought	lppd	-25.39	-27.56	2.17	***
28	HECO	recruitment	Drought	MSE	1.94	2.33	-0.38	***

Continued on next page

	species	vital_rate	Treatment	score	climate model	year effects	model	diff	improved
29	HECO	recruitment	Irrigation	lppd	-30.69		-30.56	-0.12	
30	HECO	recruitment	Irrigation	MSE	5.21		7.62	-2.41	***
31	HECO	survival	Control	lppd	-124.59		-112.70	-11.90	
32	HECO	survival	Control	MSE	0.12		0.11	0.01	
33	HECO	survival	Drought	lppd	-17.89		-20.82	2.93	***
34	HECO	survival	Drought	MSE	0.22		0.26	-0.05	***
35	HECO	survival	Irrigation	lppd	-16.51		-13.54	-2.97	
36	HECO	survival	Irrigation	MSE	0.09		0.07	0.02	
37	POSE	growth	Control	lppd	-1117.20		-1117.29	0.08	***
38	POSE	growth	Control	MSE	1.50		1.51	-0.00	***
39	POSE	growth	Drought	lppd	-254.17		-257.32	3.16	***
40	POSE	growth	Drought	MSE	2.66		2.69	-0.03	***
41	POSE	growth	Irrigation	lppd	-452.34		-456.95	4.61	***
42	POSE	growth	Irrigation	MSE	1.87		1.90	-0.03	***
43	POSE	recruitment	Control	lppd	-127.64		-128.73	1.09	***
44	POSE	recruitment	Control	MSE	35.20		44.82	-9.61	***
45	POSE	recruitment	Drought	lppd	-60.34		-63.15	2.80	***
46	POSE	recruitment	Drought	MSE	23.60		33.88	-10.28	***
47	POSE	recruitment	Irrigation	lppd	-72.06		-65.58	-6.48	
48	POSE	recruitment	Irrigation	MSE	85.04		27.63	57.41	
49	POSE	survival	Control	lppd	-366.49		-380.49	14.00	***
50	POSE	survival	Control	MSE	0.12		0.12	-0.00	***
51	POSE	survival	Drought	lppd	-175.29		-187.88	12.59	***
52	POSE	survival	Drought	MSE	0.21		0.22	-0.01	***
53	POSE	survival	Irrigation	lppd	-156.28		-150.23	-6.05	
54	POSE	survival	Irrigation	MSE	0.13		0.13	0.00	
55	PSSP	growth	Control	lppd	-627.07		-625.63	-1.44	
56	PSSP	growth	Control	MSE	1.38		1.36	0.02	
57	PSSP	growth	Drought	lppd	-292.90		-297.61	4.71	***
58	PSSP	growth	Drought	MSE	1.82		1.85	-0.03	***
59	PSSP	growth	Irrigation	lppd	-312.96		-314.67	1.71	***
60	PSSP	growth	Irrigation	MSE	1.51		1.51	-0.00	***
61	PSSP	recruitment	Control	lppd	-123.81		-125.23	1.42	***
62	PSSP	recruitment	Control	MSE	30.59		34.79	-4.20	***
63	PSSP	recruitment	Drought	lppd	-87.34		-86.58	-0.77	
64	PSSP	recruitment	Drought	MSE	222.16		54.58	167.58	

Continued on next page

	species	vital_rate	Treatment	score	climate model	year effects	model	diff	improved
65	PSSP	recruitment	Irrigation	lppd	-60.18		-62.12	1.94	***
66	PSSP	recruitment	Irrigation	MSE	20.90		44.57	-23.67	***
67	PSSP	survival	Control	lppd	-148.69		-140.15	-8.54	
68	PSSP	survival	Control	MSE	0.10		0.09	0.01	
69	PSSP	survival	Drought	lppd	-83.26		-78.09	-5.17	
70	PSSP	survival	Drought	MSE	0.11		0.11	0.01	
71	PSSP	survival	Irrigation	lppd	-100.52		-89.02	-11.50	
72	PSSP	survival	Irrigation	MSE	0.13		0.11	0.01	

## Additional Figures

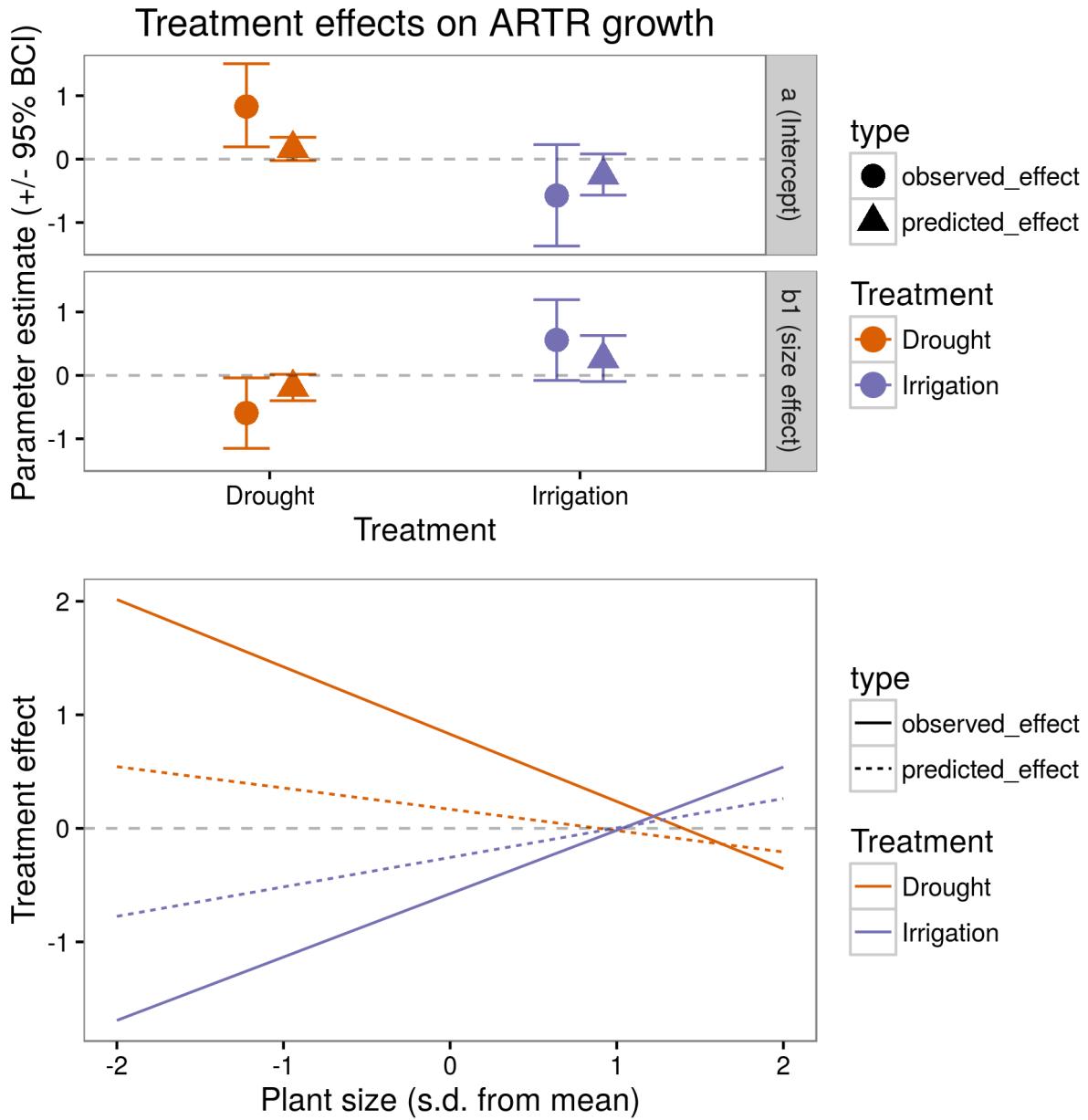


Figure SI-1: Comparison of treatment effects predicted and observed for *A. tripartita* growth. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational data but do not include the five years of the experiment (2011 to 2016). We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero.

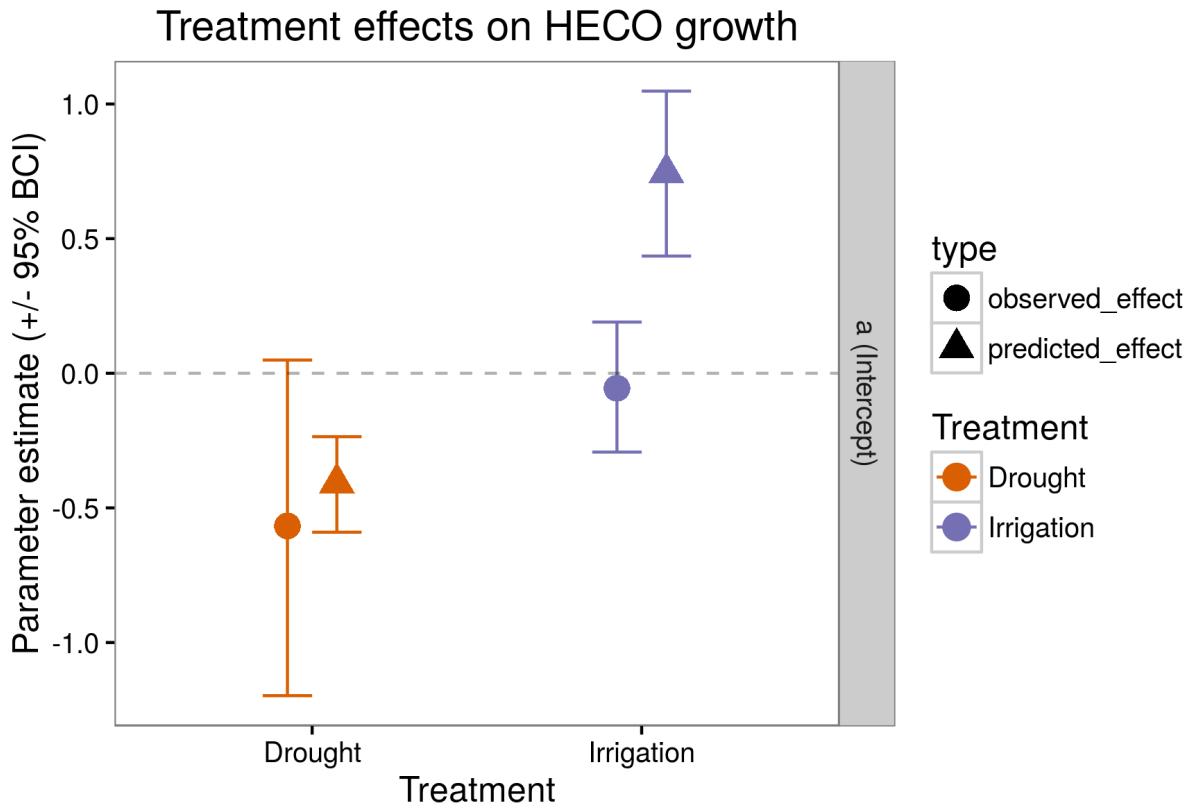


Figure SI-2: Comparison of treatment effects predicted and observed for *H. comata* growth. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational data but do not include the five years of the experiment (2011 to 2016). We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero.

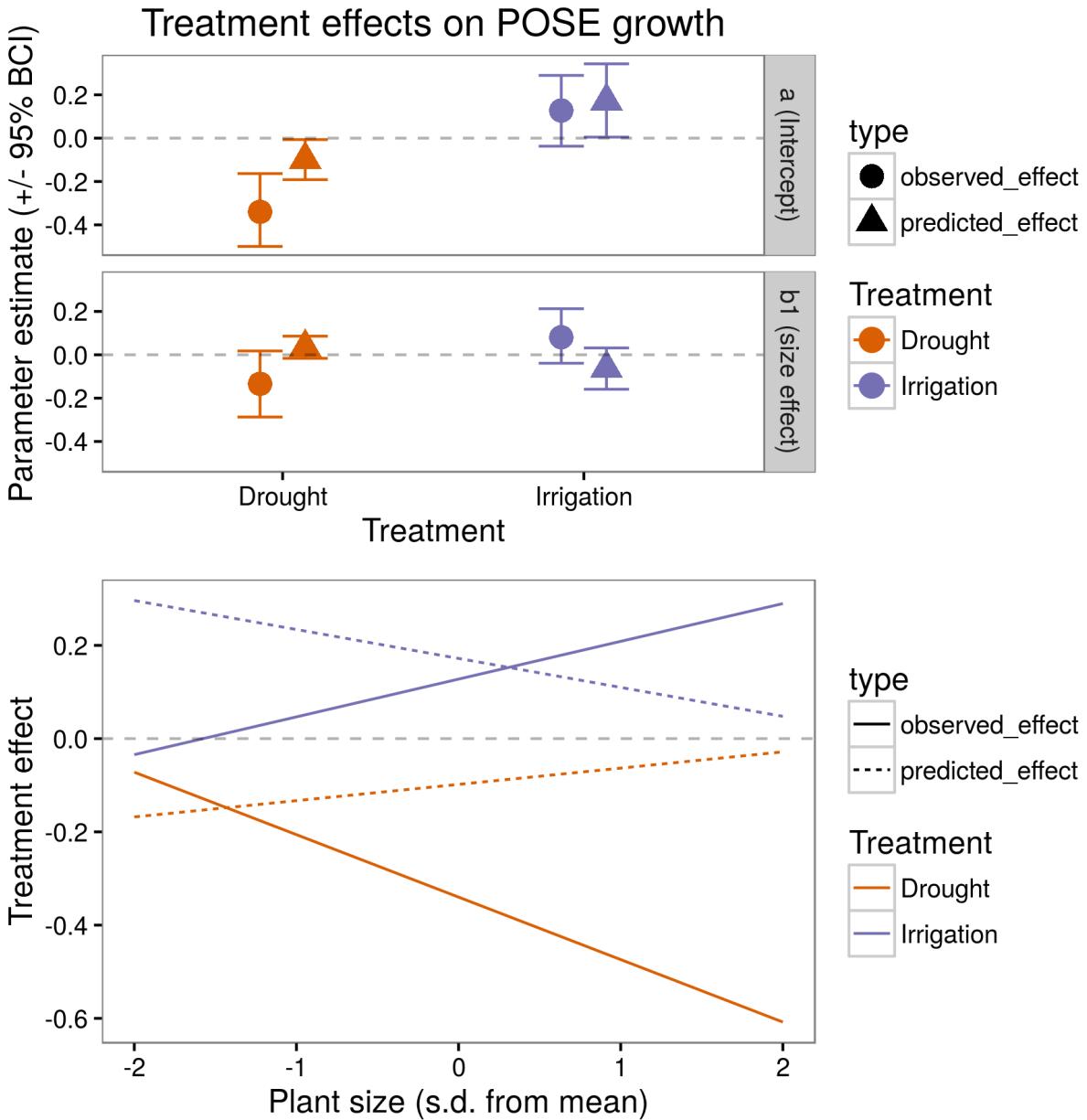


Figure SI-3: Comparison of treatment effects predicted and observed for *P. secunda* growth. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational data but do not include the five years of the experiment (2011 to 2016). We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero.

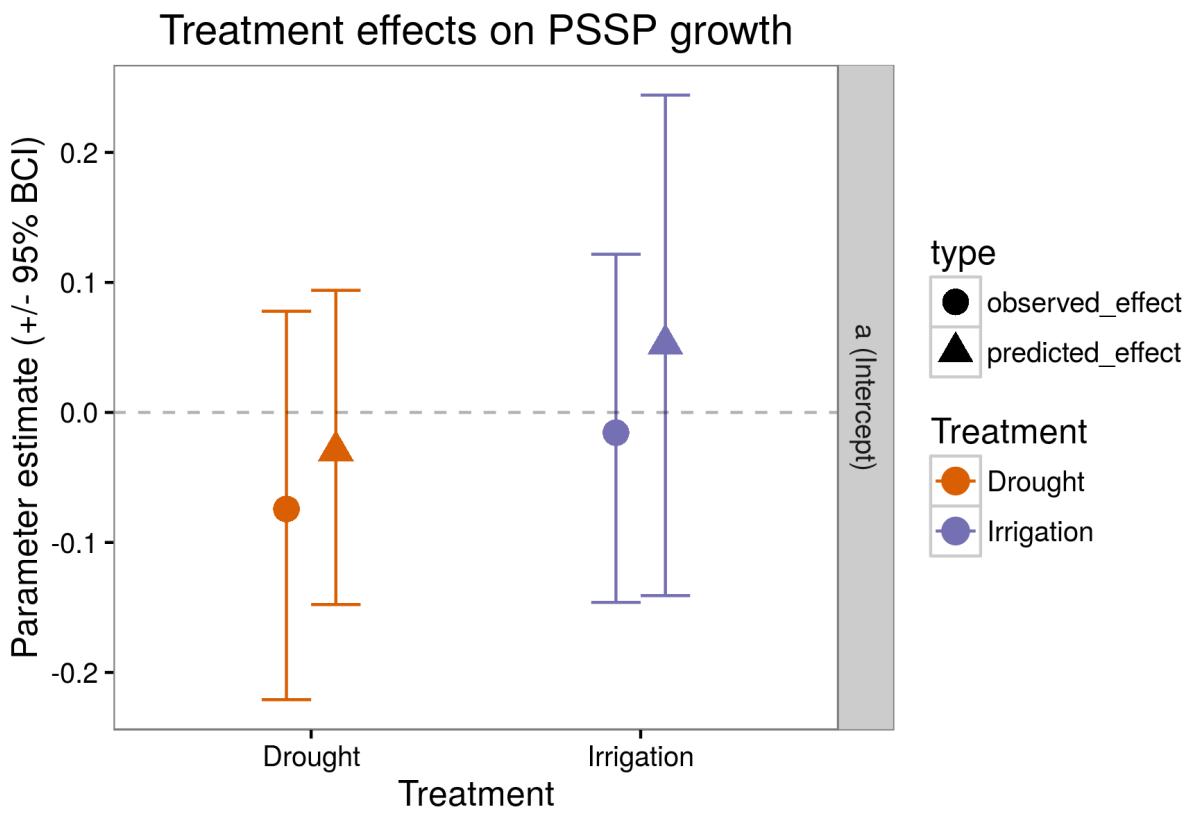


Figure SI-4: Comparison of treatment effects predicted and observed for *P. spicata* growth. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational data but do not include the five years of the experiment (2011 to 2016). We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero.

### Treatment effects on ARTR survival

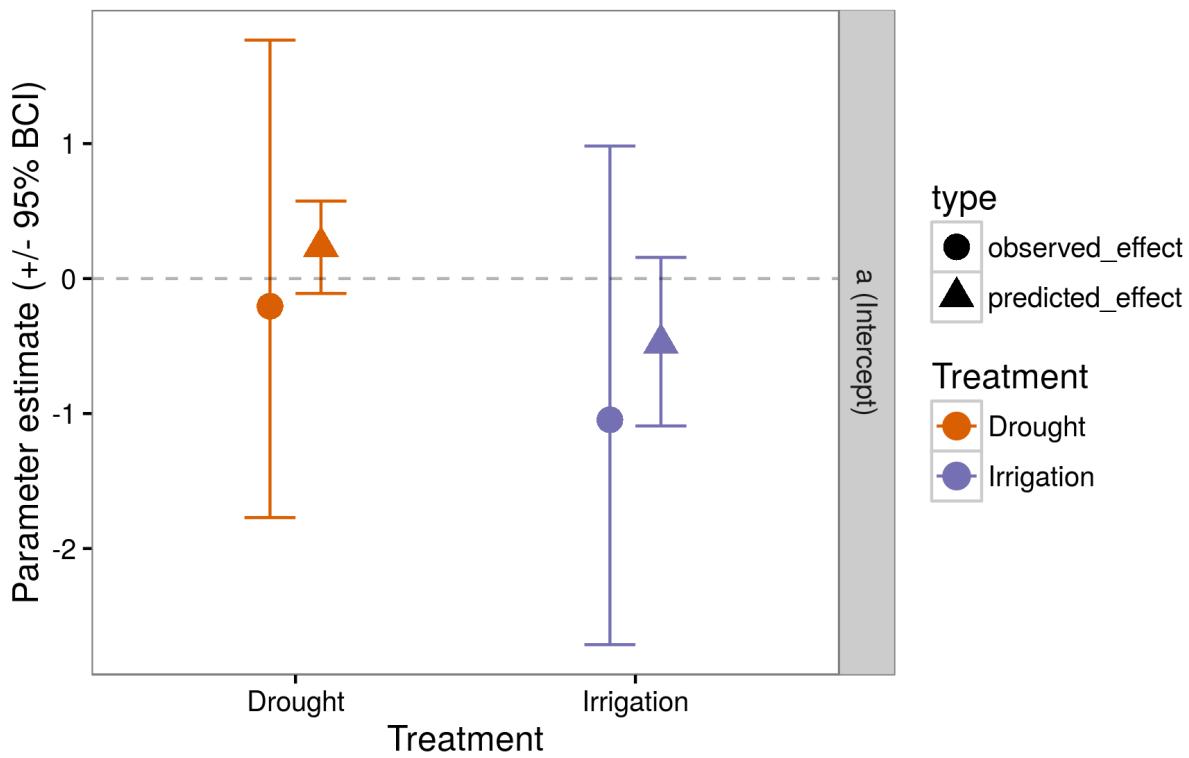


Figure SI-5: Comparison of treatment effects predicted and observed for *A. survival* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational data but do not include the five years of the experiment (2011 to 2016). We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero.

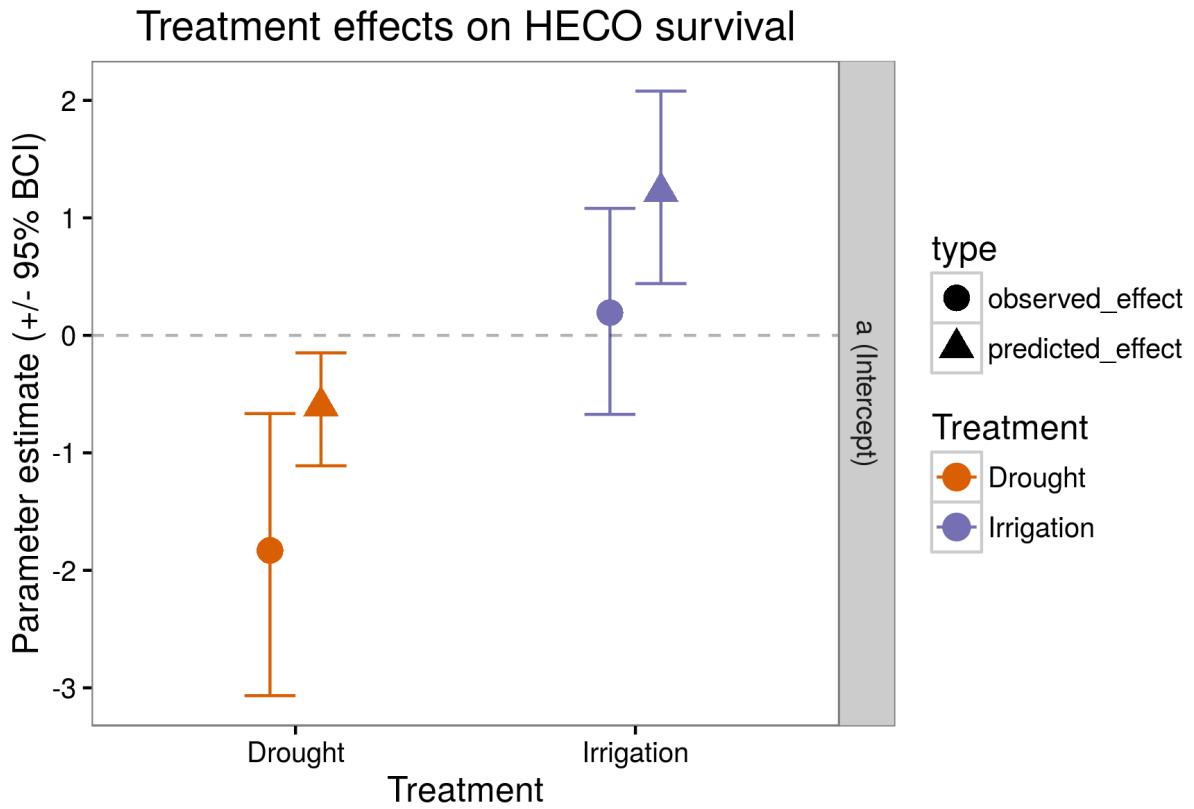


Figure SI-6: Comparison of treatment effects predicted and observed for *H. comata* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational data but do not include the five years of the experiment (2011 to 2016). We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero.

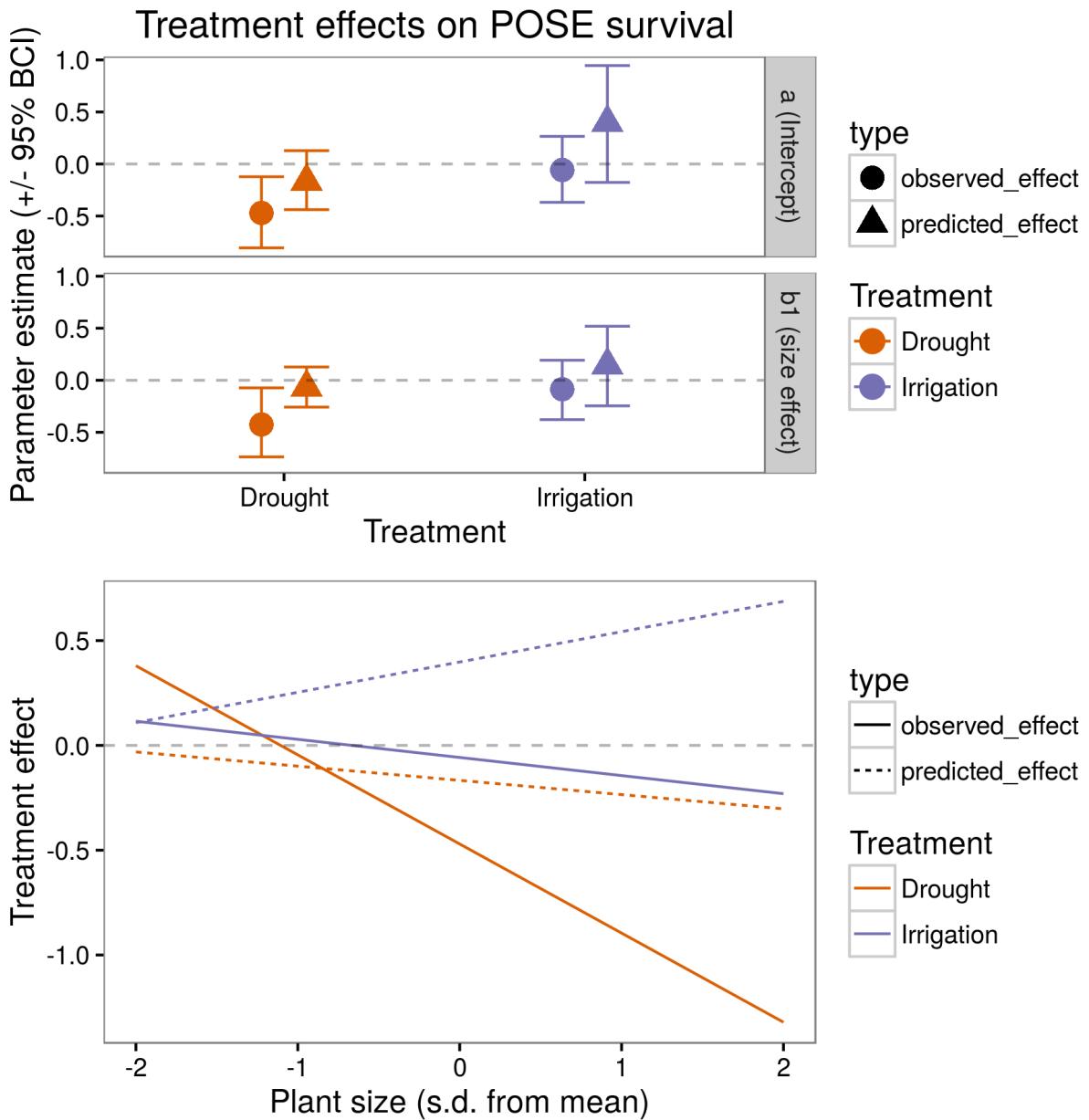


Figure SI-7: Comparison of treatment effects predicted and observed for *P. secunda* survival. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational data but do not include the five years of the experiment (2011 to 2016). We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero.

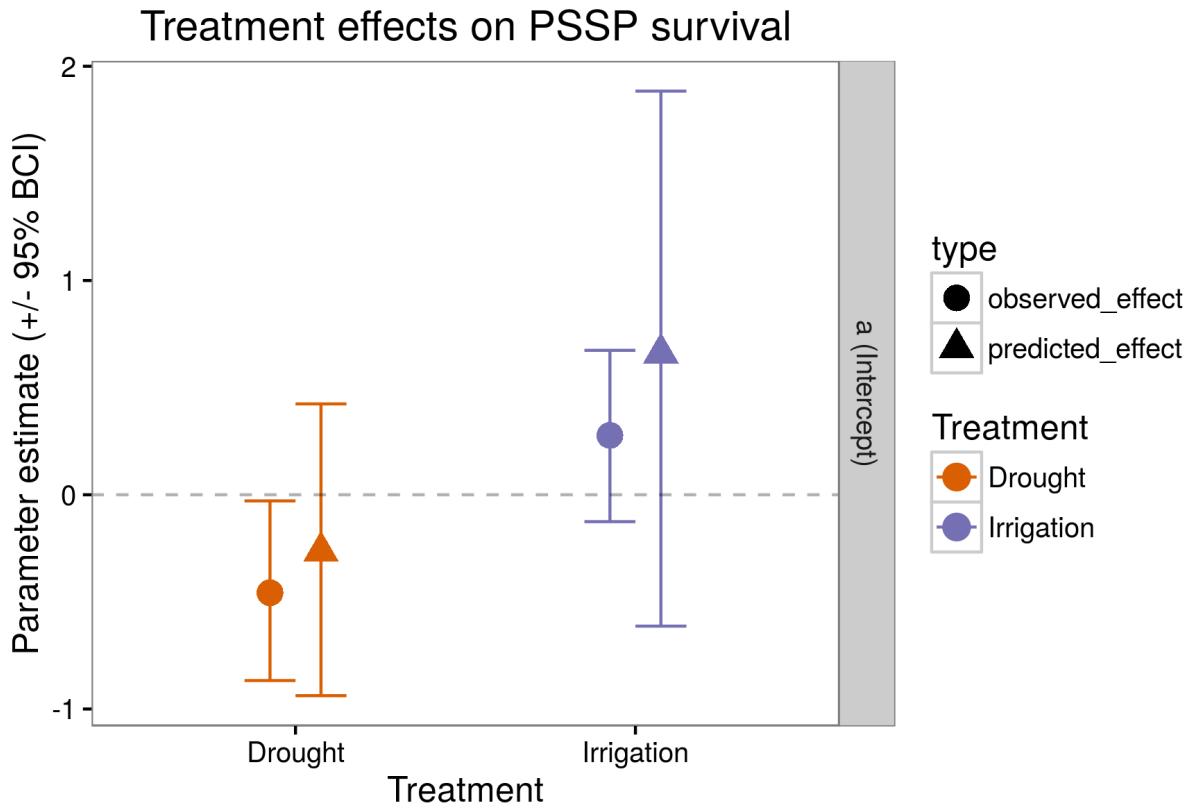


Figure SI-8: Comparison of treatment effects predicted and observed for *P. spicata* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational data but do not include the five years of the experiment (2011 to 2016). We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero.