

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

Andrew R. Kleinhesselink<sup>\*a</sup>, Peter B. Adler<sup>a</sup> and Andrew Tredennick<sup>a</sup>

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

Last compile: December 20, 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.

---

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

<sup>1</sup> **Summary**

- <sup>2</sup> 1. Climate is an important driver of population ecology however there have been few tests of  
<sup>3</sup> whether observational data that links population performance with climate variation can be  
<sup>4</sup> used to predict the responses of populations to experimentally imposed climate conditions.
- <sup>5</sup> 2. Using longterm historical observational data from a sagebrush steppe plant community, previ-  
<sup>6</sup> ous research has reported a wide variety of climate effects on four dominant plant species. We  
<sup>7</sup> tested whether statistical models based on these observational data could be used to predict  
<sup>8</sup> how these species respond to a five year drought and irrigation experiment conducted from  
<sup>9</sup> 2011 to 2016.
- <sup>10</sup> 3. We established sixteen new plots at the same field cite as the original historical data and  
<sup>11</sup> assigned each to either drought or irrigation treatments imposed using rainfall shelters and  
<sup>12</sup> automatic sprinklers. The original plots were also monitored and served as ambient climate  
<sup>13</sup> controls. In order to describe the demographic effects of the climate manipulations, we used  
<sup>14</sup> all data, observational and experimental, to fit a set of statistical models describing each  
<sup>15</sup> species survival, growth and recruitment in response to local competitive interactions and the  
<sup>16</sup> effects of the climate manipulation experiment. Next we fit another set of models to only  
<sup>17</sup> the observational data that used seasonal soil moisture covariates to describe annual variation  
<sup>18</sup> in the vital rates. Finally, we used the observation-based models to predict the effects of  
<sup>19</sup> the experimental treatments on each species vital rates. We also used an individual based  
<sup>20</sup> population model to predict one year ahead changes in population size in the experimental  
<sup>21</sup> plots.
- <sup>22</sup> 4. The experimental drought and irrigation treatments successfully lowered and increased soil  
<sup>23</sup> moisture respectively. Over the course of the experiment, average plot cover of *Hesperostipa*  
<sup>24</sup> *comata* and *Pseudoroegneria spicata* declined significantly, while cover of *Poa secunda* showed a  
<sup>25</sup> trend towards increase in the irrigated plots. At the level of individual vital rates, experimental  
<sup>26</sup> drought reduced growth and survival of *Hesperostipa comata* and *Poa secunda* and the survival  
<sup>27</sup> of *P. spicata*, while drought increased the growth of small *A. tripartita*.

- 28     5. Soil moisture parameters improved model predictions for six out of the twelve models we  
29         examined. Moreover, we observed strong positive correlations between the predicted treatment  
30         effects and the treatment effects we observed.
- 31     6. At the population-level, including the effects of soil moisture improved the cover predictions  
32         made by individual based models for *P. secunda* and *P. spicata*.
- 33     7. *Synthesis*: Observational climate data did hold valuable information for predicting species'  
34         responses to this climate change experiment. We were encouraged that treatment responses  
35         were often in the right direction to predicted responses even when the effect were not significant.  
36         We generally were better able to predict species responses to the drought treatment than to  
37         the control or the irrigation treatment. This suggests that soil moisture is an important factor  
38         in predicting the population dynamics of these species but only when water is truly limiting.

## 39     **Introduction**

40     Climate is widely considered one of the most powerful external forces driving changes in species  
41         abundance across space and time (?). The effects of climate on populations and ecosystems are most  
42         apparent at the largest scales in space and time: climate determines the distribution of ecosystems,  
43         treelines and the the range limits of many species (), while the recent historical and the paleoeco-  
44         logical record shows that long-term climate change leads to changes in species range limits (??).  
45         Understanding and predicting the effects of climate on populations is an increasingly important  
46         goal if we are to understand and predict the effects of climate change on earth's ecosystems.

47             Unfortunately, for many species it is difficult to determine how annual climate variation  
48         affects populations and demographic rates (??). Observational data on species performance across  
49         years with varying climate can provide some information on how climate might affect population  
50         abundance (??). However, many years of data are needed to reliably detect climate effects, es-  
51         pecially when annual variation in demographic rates is high (??). Generating predictions from  
52         observational historical data for the future novel conditions created by climate change is fraught  
53         with risk. Climate change will not only change mean annual temperature and precipitation, but  
54         also affect annual variation in these measures, and possibly the covariance between them (). Models  
55         based on the historical response of populations to annual climate variation, will therefore be extrap-  
56         olating beyond the range of observed climate variation when they are used to generate predictions

57 in the future (). If models fit to historical data can be used to accurately predict the effects of  
58 experimental climate manipulations, especially manipulations that generate extreme variation in  
59 climate, it would be strong confirmation that the climate effects they describe are real and will  
60 hold in the future ?.

61 Many plant species occur across a wide range of climates and individual plants in many  
62 environments must tolerate large fluctuations in seasonal temperature and soil moisture. Unlike  
63 most animals, plants must endure these conditions in place and often become dormant in less  
64 optimal seasons . At first pass, these observation would seem to indicate that annual variation in  
65 climate should have little effect on plant population performance.

66 Nevertheless, there is abundant evidence that plant performance shows high year to year  
67 variation, both at the level of individual growth, survival and reproduction and in terms of total  
68 population abundance often this variation c. In many regions interannual variation in precipi-  
69 tation can be directly linked to variation in net primary productivity (??). The growth rates of  
70 trees are also often tightly linked to annual precipitation, so much so that annual growth rings in  
71 their stems can serve as accurate records of historical climate thousands of years before the present  
72 (?). Likewise, many smaller shrub and sub-shrub species show strong variation in growth in re-  
73 sponse to climate that is recorded in their tissues (??). Among annual plants in desert ecosystems,  
74 germination and reproductive output is tightly linked to interannual precipitation variation (?).

75 Linking annual climate variation to demographic performance in plants is a high priority  
76 for building population models for plants that can predict their future response to climate change  
77 (??). However, despite the clear signs that climate drives net primary productivity at the ecosystem  
78 level and the variation in individual growth rate in trees and other woody species, as well as  
79 the germination and fecundity of annual plants, there are few studies that have connected the  
80 effects of interannual climate variation to population models for plants, and fewer still that have  
81 tested whether these population models can be used to accurately predict the future responses of  
82 plants to short term climate variation. Adler et al. (?) showed that population models based  
83 on historically observed correlations between plant population growth rates and precipitation did  
84 have some predictive power in describing species response to a short-term climate manipulation in a  
85 North American grassland. Three species showed responses to experimentally imposed drought and  
86 irrigation that were well predicted by population models fitted to historical observations. However,

87 another three species, showed responses to the experimental conditions that were not well predicted  
88 by historical observations. The authors suggested that limited replication in the historical data for  
89 two of these species and changing competitive conditions in the community may have led to the  
90 poor predictions.

91 The demography and competitive interactions between three dominant grass species and  
92 a dominant shrub species at in a sagebrush steppe plant community at the US Sheep Experiment  
93 Station near Dubois, Idaho have been described in at least X different studies since 2005. X of  
94 these studies report significant effects of historical climate variation on the vital rates and overall  
95 population growth of these species. Both precipitation and temperature have been shown to have  
96 strong species-specific effects on this system. Although past studies used different statistical mod-  
97 els and methods for variable selection they indicate X. This well-studied system offers the ideal  
98 opportunity to test whether statistical associations between annual climate variation and plant  
99 demography in long-term observational data can be used to predict the future responses of plant  
100 populations to climate change.

101 In this study, we report how the four dominant plant species at the USSES responded to  
102 a five year drought and irrigation experiment and use the results to address two research ques-  
103 tions: first, how much do the growth, recruitment and survival of our target species differ between  
104 the precipitation manipulation treatments? If our experiment does affect species vital rates we  
105 interpret that as strong evidence that variation in precipitation should have an effect on popula-  
106 tions. Second, we test whether statistical models parameterized from observational data only can  
107 accurately predict each species response to the experimental precipitation manipulation? If models  
108 based on observational effectively capture the effects of climate on species performance, they should  
109 also predict the effects of precipitation treatments in the experiment.

## 110 Methods

### 111 Study site and data set description

112 The U.S. Sheep Experiment Station (USSES) is located 9.6 km north of Dubois, Idaho ( $44.2^{\circ}\text{N}$ ,  
113  $112.1^{\circ}\text{W}$ ), 1500 m above sea level. During the period of data collection (1926–1957), mean annual  
114 precipitation was 270 mm and mean temperatures ranged from  $-8^{\circ}\text{C}$  (January) to  $21^{\circ}\text{C}$  (July).

<sup>115</sup> The vegetation is dominated by a shrub, *Artemesia tripartita*, and three perennial C3 grasses:  
<sup>116</sup> *Pseudoroegneria spicata*, *Hesperostipa comata*, and *Poa secunda*. These dominant species account  
<sup>117</sup> for over 70% of basal cover and 60% of canopy cover at this site.

<sup>118</sup> Scientists at the USSES established 26 1-m<sup>2</sup> quadrats between 1926 and 1932. Eighteen  
<sup>119</sup> quadrats were distributed among four ungrazed exclosures, and eight were distributed in two pad-  
<sup>120</sup> docks grazed at medium intensity spring through fall. All quadrats were located on similar topog-  
<sup>121</sup> raphy and soils. In most years until 1957, all individual plants in each quadrat were mapped using  
<sup>122</sup> a pantograph (Blaisdell 1958). The historical data set is public and available online (?). In 2007,  
<sup>123</sup> we located 14 of the original quadrats, all of which are inside permanent livestock exclosures, and  
<sup>124</sup> resumed annual mapped censusing using the traditional pantograph method. Daily temperature  
<sup>125</sup> and precipitation has been monitored throughout this period at a climate station located at the  
<sup>126</sup> USSES headquarters (station id: GHCND:USC00102707) which located within 2 km of the research  
<sup>127</sup> plots. We downloaded daily and monthly tmin, tmax, and precipitation data from the National  
<sup>128</sup> Climate Data Centers online database.

<sup>129</sup> We extracted data on survival, growth, and recruitment from the mapped quadrats based  
<sup>130</sup> on plants' spatial locations. Our approach tracks genets representing individual plants. For the  
<sup>131</sup> shrub, each genet is associated with the basal position of a stem. For the bunchgrasses, each genet  
<sup>132</sup> represents a spatially distinct polygon in the mapped quadrat. These genets may fragment and/or  
<sup>133</sup> coalesce over time. Each mapped polygon is classified as a surviving genet or a new recruit based  
<sup>134</sup> on its spatial location relative to genets present in previous years (?). We modeled vital rates using  
<sup>135</sup> data from 21 year-to-year transitions between 1929 and 1957, and four year-to-year transitions from  
<sup>136</sup> 2007 to 2011.

### <sup>137</sup> Precipitation experiment

<sup>138</sup> In spring 2011, we selected locations for an additional 16 quadrats for the precipitation experiment.  
<sup>139</sup> We located these in a large exclosure containing six of the historical permanent quadrats. We  
<sup>140</sup> avoided plots falling on hill slopes, areas with greater than 20% bare rock, or with over 10% cover  
<sup>141</sup> of the woody shrubs *Purshia tridentata* or *Amelanchier utahensis*. New plots were established  
<sup>142</sup> in pairs, and one plot per pair was randomly assigned to either the precipitation reduction or  
<sup>143</sup> the precipitation addition treatment. We mapped the quadrats in June, 2011 and then built the

144 rainfall shelters and set-up the irrigation systems in the fall of 2011. We used a rain-out shelter  
145 and automatic irrigation design described in (??). Each rain-out shelter covered an area of 2.5 by  
146 2 m and consisted of transparent acrylic shingles held up over the plot to channel 50% of incoming  
147 rainfall off of the plot and into 75 l reservoirs. The collected water was pumped out of reservoirs  
148 and sprayed onto paired irrigation treatment plots. Pumping was initiated automatically with  
149 float switches that were triggered when water levels in the reservoirs were approximately 20 l, or  
150 equivalently irrigation was triggered once for every 6 mm of rainfall collected. We disconnected  
151 the irrigation pumps in late fall each year and re-connected them in April. The drought shelters  
152 remained in place throughout the year.

153 We monitored soil moisture and air temperature in four of the precipitation experiment  
154 plot pairs using Decagon Devices (Pulman Washington) 5TM and EC-5 soil moisture sensors and  
155 5TE temperature sensors. We installed two soil moisture sensors in each monitored plot at 5 cm  
156 and two at 25 cm deep in the soil. Air temperature was measured underneath the roofing of the  
157 shelter at 30 cm above ground. For each pair of manipulated plots, we also installed sensors in a  
158 nearby area to measure ambient rainfall and temperature. Data were logged automatically every  
159 four hours. We augmented automatic monitoring of the climate in these plots with by taking direct  
160 measurements of soil moisture with a handheld EC-5 soil moisture sensor at six points around  
161 each plot on 6/6/2012, 4/29/2015, 5/7/2015, 6/9/2015 and 5/10/2016. We analyzed these spot  
162 measurements for significant treatment effects on soil moisture using a linear mixed effects model  
163 with the *lmer* package in *R* (?), with plot, plot group, and date as random effects in the model.

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

## 170 Soil moisture modeling

171 We expected that our precipitation manipulation experiment would affect plants by altering avail-  
172 able soil moisture during the growing season. Because we do not have direct soil moisture measures

173 for each year of observed plant cover in the historical record, we used the SOILWAT soil mois-  
174 ture model to estimate daily soil moisture at the USSES from 1925 to the present (?). We used  
175 an enhanced version of soilwat that has recently been developed for use in semi-arid shrubland  
176 ecosystems (?). SOILWAT uses daily weather data, ecosystem specific vegetation properties and  
177 site specific soil properties to estimate water balance processes. SOILWAT specifically estimates  
178 rainfall interception by vegetation, evaporation of intercepted water, snow melt and snow redistri-  
179 bution, infiltration into the soil, percolation through the soil, bare-soil evaporation, transpiration  
180 from each soil layer, and drainage. We parameterized SOILWAT with the generic sagebrush steppe  
181 vegetation parameters and site specific soil texture and bulk density data. We used daily weather  
182 data collected at the USSES from 1925 until the present as weather forcing data for the SOILWAT  
183 predictions.

184 We averaged daily soil moisture predictions from SOILWAT from upper 40 cm of soil and  
185 then averaged these seasonally to serve as the covariates in the vital rate regressions for each species.  
186 Because we did not monitor soil moisture directly in all control, drought and irrigation plots, we  
187 used a model to describe the average treatment effects on soil moisture during the course of the  
188 experiment. To do this we first averaged observed soil moisture data by date and plot and then  
189 standardized these by the mean and standard deviation of the control soil moisture conditions  
190 observed within each plot group. We then found the difference between the soil moisture in the  
191 treated plots and the ambient conditions. We then modeled these treatment effects as a function of  
192 season and whether a day was rainy or dry. We expected that our drought and irrigation treatments  
193 might be more effective during rainy weather than during dry weather. Rainy days were defined as  
194 any days when any precipitation was recorded and average temperatures were above 3 degrees C.  
195 The day immediately following rainfall was also classified as rainy. We fit this model using the *lmer*  
196 package in *R* (?) with random effects for plot group and date. We then used this model to predict  
197 the treatment effects on soil moisture for the entire study period from the ambient soil moisture  
198 values predicted from the SOILWAT model described above. These adjusted soil moisture values  
199 reflected the average season and rainfall dependent effects of the experimental treatments on soil  
200 moisture and could be used as covariates for predicting the effects of our manipulation on each  
201 species demographic rates.

202 **Statistical models of vital rates**

203 For each of the vital rates and each species we fit three separate models. First we fit a treatment  
204 model fit to all observations in the historical data as well as the contemporary experiment. This  
205 model included parameters estimating the effects of the drought and irrigation treatments on each  
206 species vital rate. This model was used to describe the basic results of the experiment. Next  
207 we fit two models to the historical observational data only (including the first four years of the  
208 modern data 2007 to 2010). We used these two models to generate predictions for the effects of  
209 the experiment. In the first prediction model, year to year variation in vital rates was treated  
210 as random effect. In the second prediction model, the "climate" model, we included paramaters  
211 that described the effects of year to year variation in seasonal soil moisture. These three models  
212 allow for two meaningful comparisons. First we can compare the predictions made by the two  
213 prediction models directly to the raw data in the experimental plots. In addition, we can compare  
214 the direction and magnitude of the soil moisture parameters estimated from the observational data  
215 by the climate model to the coefficients describing the treatment effects in the treatment model.

216 All three versions of the models of the models above follow the same basic structure and  
217 are developed from previous work (??). We model the survival probability of an individual genet  
218 as a function of genet size, the neighborhood-scale crowding experienced by the genet from both  
219 conspecific and heterospecific genets, temporal variation among years, and permanent spatial vari-  
220 ation among groups of quadrats ('group'; here means a set of nearby quadrats located within one  
221 pasture or grazing exclosure). In this analysis we only include crowding from the four main focal  
222 species.

223 Formally, we modeled the survival probability,  $S$ , of genet  $i$  in species  $j$ , group  $g$ , and  
224 removal treatment  $h$ , from time  $t$  to  $t + 1$  as

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

225 where  $\varphi$  is the spatial group dependent intercept,  $\gamma$  is a year-effect,  $\beta$  is year-dependent coefficient  
226 that represents the effect of log genet size,  $u$ , on survival in year  $t$ .  $\boldsymbol{\omega}$  is a vector of interaction  
227 coefficients which determine the impact of crowding,  $\mathbf{W}$ , by each species on the focal species. The  
228 vector  $\mathbf{W}$  includes crowding from the four dominant species, *A. tripartita*, *P. spicata*, *H. comata*,

229 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  
230 R.

231 To describe the treatment effects in the experiment a new term is added to the above model,

$$\mathbf{T}\chi_j^S \quad (2)$$

232 where  $\chi$  is a vector of treatment effect coefficients describing the effects of each experimental  
233 treatment  $h$  on the survival rate, and  $\mathbf{T}$  is a design matrix indicating the treatment level of each  
234 observation in the data. The design matrix also includes terms for the interaction between plant  
235 size  $u$  and the treatment effects. These interaction terms allow the effect of each treatment to vary  
236 with plant size.

237 In the climate model, the  $\chi$  and  $\xi$  terms are replaced with,

$$\mathbf{C}\xi_j^S \quad (3)$$

238 where  $\xi$  gives a vector of coefficients describing the effects of a set of soil moisture covariates  $\mathbf{C}$   
239 in treatment  $h$  and year  $t$  on the survival rate of species  $j$ .  $\mathbf{C}$  is a vector of seasonal average soil  
240 moisture and can include interaction effects between plant size,  $u$ , and the soil moisture covariates  
241 that allow the effects of soil moisture to vary with plant size.

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

$$u_{ijgh,t+1} = \varphi_{jg}^G + \gamma_{j,t}^G + \chi_{jh}^G + \beta_{j,t}^G u_{ij,t} + \langle \boldsymbol{\omega}_{j,t}^G, \mathbf{W}_{ij,t} \rangle + \varepsilon_{ij,t}^G. \quad (4)$$

244 To capture non-constant error variance in growth, we modeled the variance  $\varepsilon$  about the growth  
245 curve (4) as a nonlinear function of predicted genet size:

$$Var(\varepsilon_{ij,t}^G) = a \exp^{(bu_{ij,t+1})}. \quad (5)$$

246 As in the survival regression above, specific parameters describing the treatment  
 247 effects on growth are added in the treatment model,

$$\mathbf{T}\chi_j^G \quad (6)$$

248 where  $\chi$  is a treatment effect describing the effect of experimental treatment  $h$  on growth, including  
 249 treatment by size interactions.

250 Similarly, in the climate model soil moisture influences the growth equation through these  
 251 terms,

$$\mathbf{C}\xi_j^G \quad (7)$$

252 where  $\xi$  is a vector of coefficients describing the effects of soil moisture covariates in the matrix  $\mathbf{C}$   
 253 for treatment  $h$  and year  $t$  on growth of species  $j$ . Again this can include interactions between soil  
 254 moisture and plant size  $u$ .

255 Although the main focus of the current analysis the effects of soil moisture, we also modeled  
 256 the effects of inter- and intra-specific competition in our vital rate models. We model the crowding  
 257 experienced by a focal genet as a function of the distance to and size of neighbor genets. In previous  
 258 work, we assumed that the decay of crowding with neighbor distance followed a Gaussian function  
 259 (?), but here we use a data-driven approach (??). We model the crowding experienced by genet  
 260  $i$  of species  $j$  from neighbors of species  $m$  as the sum of neighbor areas across a set of concentric  
 261 annuli,  $k$ , centered at the plant,

$$w_{ijm,k} = F_{jm}(d_k)A_{i,k} \quad (8)$$

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

$$W_{ijm} = \sum_k w_{ijm,k}. \quad (9)$$

265 Note that  $W_{jj}$  gives intraspecific crowding. The  $W$ 's are then the components of the  $\mathbf{W}$  vectors  
 266 introduced as covariates in the survival (1) and growth (4) regressions.

267 We assume that competition kernels  $F_{jm}(d)$  are non-negative and decreasing, so that distant  
 268 plants have less effect than close plants. Otherwise, we let the data dictate the shape of the kernel  
 269 by fitting a spline model using the methods of Teller et al. (2016). The shape of  $F_{jm}$  is determined  
 270 by a set of spline basis coefficients  $\vec{b}_{jm}$  and a smoothing parameter  $\eta$  that controls the complexity of  
 271 the fitted kernel. Demographic models such as (1) then have  $\gamma, \varphi, \chi, \beta, \omega, \vec{b}$  and  $\eta$  as parameters  
 272 to be fitted. We implemented this in the statistical computing environment, R, by making the  
 273 spline coefficients and  $\eta$  the arguments of an objective function that computes  $\mathbf{W}$  using the input  
 274 spline coefficients, calls the model-fitting functions `lmer` and/or `glmer` to fit the other parameters  
 275 in the survival and growth regressions, and returns an approximate AIC value and model degrees  
 276 of freedom ( $df$ ) for survival and growth combined. We used the  $\vec{b}$  values at the smoothest local  
 277 minimum of AIC as a function of  $df$ , as in ?. This approach assumes that one measure of crowding  
 278 affects survival and growth. In addition, for fitting the kernels we assumed that survival and growth  
 279 depended only on intraspecific crowding, and thus only fitted the within-species competition kernels  
 280  $F_{jj}$ . Based on previous work (?), we set all  $F_{mj}$  equal to  $F_{jj}$ , meaning that the within-species  
 281 competition kernel for species  $j$  is also used to determine the effect of all other species on species  
 282  $j$ . We used data from all historical plots and contemporary control-treatment plots to estimate the  
 283 competition kernels (?).

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

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

$$y_{jq,t+1} = NegBin(\lambda_{jq,t+1}, \theta) \quad (10)$$

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

$$\lambda_{jq,t+1} = C'_{jq,t} \exp \left( \varphi_{jg}^R + \gamma_{j,t}^R + \langle \boldsymbol{\omega}^R, \sqrt{C'}_{q,t} \rangle \right) \quad (11)$$

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

299 Because plants outside the mapped quadrat could contribute recruits to the focal quadrat or  
 300 interact with plants in the focal quadrat, we estimated effective cover as a mixture of the observed  
 301 cover,  $C$ , in the focal quadrat,  $q$ , and the mean cover,  $\bar{C}$ , across the spatial location,  $g$ , in which  
 302 the quadrat is located:  $C'_{jq,t} = p_j C_{jq,t} + (1 - p_j) \bar{C}_{jg,t}$ , where  $p$  is a mixing fraction between 0 and  
 303 1 that was estimated as part of fitting the model.

304 In the treatment model, a new term is added to the exponential term in the equation above,

$$T\chi_j^R \quad (12)$$

305 where  $\chi$  describes the effect of the treatment levels on recruitment.

306 Likewise in the climate model this term is added,

$$C\xi_j^R \quad (13)$$

307 where the  $\xi$  gives a set of coefficients for the soil year, and treatment specific soil moisture covariates  
 308 in  $\boldsymbol{C}$ .

309 We fit all vital rate models using Hamiltonian-Markov Chain Monte Carlo (HMC) simulations in STAN 6.4 (?) and (?). The priors and model code are described more completely  
 310 in appendix A. Each model was run for 2,000 iterations and four independent chains with different  
 311 initial values for parameters. We discarded the initial 1,000 samples. Convergence was observed  
 312 initial values for parameters. We discarded the initial 1,000 samples. Convergence was observed

313 graphically for all parameters, and confirmed by assessing the split  $\widehat{R}$  statistic which at convergence  
314 is equal to one (?).

315 To assess the effects of the experimental treatments on species survival and growth we fit  
316 treatment models with and without the size by treatment interactions in the treatment effect term  
317 for (2) and growth models (6). We judged whether including interaction terms improved model fit by  
318 comparing the Watanabe-Aikake Information Criteria (WAIC) scores for each version of the model  
319 and retained the version with the lower WAIC score (?). WAIC scores approximate cross-validation  
320 predictive accuracy for a model and like traditional AIC allow for model comparison. Lower WAIC  
321 scores indicate a more parsimonious model. When a treatment model for survival or growth for  
322 a species included a size by treatment effect, we also included a size by soil moisture effect in the  
323 climate model for that species and vital rate. This allowed us to more directly compare the effects  
324 in the experimental data to the effects predicted from the climate model fit to the observational  
325 data.

### 326 Selecting soil moisture covariates

327 After generating a time series of predicted daily soil moisture from the SOILWAT model, we  
328 averaged daily soil moisture across spring, summer and fall seasons in each year. We considered  
329 each of the three seasonal soil moisture variables at three different time periods relative to the  
330 demographic transition from year  $t$  to year  $t + 1$ . Soil moisture in the year between  $t$  and  $t + 1$  is  
331 indicated with a "1" subscript. Soil moisture in the year before  $t$  is indicated with a "0" subscript.  
332 And soil moisture preceding this year is indicated with a "lag" subscript. For example, for the  
333 year 2010,  $spring_1$  would indicate soil moisture in the spring of 2010,  $spring_0$  would indicate soil  
334 moisture during the spring of 2009 and  $spring_{lag}$  would indicate soil moisture during the spring of  
335 2008.

336 In order to select among the nine potential soil moisture covariates (three seasons and three  
337 lags each) for each species and vital rate, we first fit a model with random year effects but without  
338 climate covariates to the observational data up to 2010. We then extracted the mean of the year  
339 effects estimates for each fitted year in the data. These random effects represented unexplained  
340 deviations in the average vital rate in a given year. We then found the correlations between each of  
341 the soil moisture variable and the random year effects. For each vital rate and species, we selected

342 the three covariates with highest correlations with these year effects. This screening technique has  
343 been used in previous demographic studies at this site (?) and is often used in dendrochronology to  
344 screen for potential climate influence on tree-ring growth (?); however, it is subject to the criticism  
345 that it is a form of data dredging (?). Nevertheless, we felt that this approach was justified in this  
346 study because ultimately we did not make inference from these fitted parameters until after we  
347 tested their ability to predict the data in the experimental plots.

348 Once we selected soil moisture covariates for each of the species and vital rates we fit the  
349 climate models including the climate terms for survival 3, growth 7 and recruitment 13. We fit these  
350 models only to the observational data from the historical period and (1928 to 1957) and the first  
351 four years of the modern period (2007 to 2010). Thus in generating predictions for the treatment  
352 effects observed in from 2011 to 2016 we are predicting truly out of sample data.

### 353 Predicting cover from individual-based models

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

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

368 We accounted for the uncertainty in our random year effects when generating predictions,  
369 by drawing random year effects for each predicted year from a normal distribution with a mean  
370 of zero and a standard deviation drawn from the posterior estimate of the standard deviation of

371 the random year effects ???. We generated predictions using the full posterior distributions of each  
 372 model parameter which allowed us to carry forward all of the uncertainty of the fitted vital rate  
 373 models into our cover predictions. Because we were interested in comparing model predictions  
 374 to observations, and were not interested in the effects of demographic stochasticity, we used a  
 375 deterministic version of the models (e.g., recruitment is the  $\lambda$  of (11), rather than a random draw  
 376 from a negative binomial distribution with a mean of  $\lambda$ ).

377 We generated predictions for the full time series of observations, including the five exper-  
 378 imental years, from the year effects model, the climate model and the treatment model. After  
 379 generating predictions for each year, we found the mean cover across all quadrats in each treatment  
 380 level and then calculated the predicted log cover change as  $\log(Cover_{t+1}/Cover_t)$ .

### 381 Quantifying predictive accuracy

382 After fitting the year effects and climate models to the observational data, we generated predic-  
 383 tions from these models for each of the vital rates in the experimental plots. We then assessed  
 384 performance of the climate and null models by calculating the mean square error (MSE) between  
 385 the predicted and observed responses in the experimental data as,

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - E(y_i|\theta))^2, \quad (14)$$

386 where  $y_i$  is the outcome of observation  $i$  and  $E(y_i|\theta)$  gives the expected outcome given the param-  
 387 eters in the model  $\theta$ . The MSE is easy to interpret, but is not always appropriate for models fit  
 388 with non-normal error structures (?). A more general statistic for assessing model predictions is  
 389 the log pointwise predictive density (lppd) (?). The lppd for a given model is defined as,

$$lppd = \sum_{i=1}^n \log \int p(y_i|\theta) p_{post}(\theta) d\theta, \quad (15)$$

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

$$\widehat{lppd} = \sum_{i=1}^n \log \left( \frac{1}{S} \sum_{s=1}^S p(y_i|\theta^S) \right), \quad (16)$$

393 where the summation of  $p(y_i|\theta^S)$  gives the total probability of observing the the actual response  $y_i$   
394 given the simulated posterior distribution  $\theta^S$  across the full set of model simulations  $S$ . The log of  
395 this sum is then averaged across the set of all observations  $i$ . Higher lppd scores indicate that the  
396 model better predicts the observations.

397 In addition, we evaluated whether the climate model predicted treatment effects of similar  
398 direction and magnitude to those observed in the experiment. We did this by extracting the soil  
399 moisture coefficients contained in  $\xi$  for each of the vital rates and then multiplying those by the  
400 appropriate soil moisture covariates for each year and treatment level in the experiment. We then  
401 averaged these across all five years in the experiment to find the average treatment effect predicted  
402 by the climate model. We compared these to the posteriors of the treatment parameters,  $chi$ , from  
403 the treatment model. As a measure of agreement between our predictions and observed response  
404 we calculated the correlation between the predicted and observed parameter values.

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

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

## 411 Results

### 412 Effects on soil moisture

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

417 The continuously recorded soil moisture data also showed treatment effects on soil moisture,  
418 but these effects were weaker on average than the spot measurements and depended on season and  
419 recent rainfall (table 2; fig 2). We saw weaker effects during the spring than during the fall and  
420 summer: the drought plots were about 20-30% drier in the fall and summer but only 7 to 14% drier

421 during the spring, while the irrigated plots were 30% wetter during the fall and summer but only  
422 20-25% wetter during the spring. Treatment differences were slightly larger during rainy periods,  
423 especially in the spring. We did not find evidence that the drought shelters and the irrigation  
424 treatments consistently affected air temperature at 30 cm above the plots.

425 The SOILWAT soil moisture model predicted average monthly soil volumetric water content  
426 of between 10 ml/ml and 15 ml/ml each month, with the month of April being the wettest and the  
427 month of July, August and September being the driest on average. Annual variation in seasonal  
428 soil moisture for each year was positively correlated with seasonal precipitation and negatively  
429 correlated with seasonal temperature. During the course of the experiment, SOILWAT reproduced  
430 much of the daily variation observed in soil moisture recorded by our automatic data loggers, but  
431 the average soil moisture predicted by SOILWAT was about 5 ml/ml wetter than the soil moisture  
432 content recorded by the data loggers.

433 After adjusting the SOILWAT seasonal soil moisture predictions by the treatment effects,  
434 we found that the soil moisture predicted in the drought plots during the course of the experiment  
435 was well below the historical seasonal averages: the summer of 2012 and 2013, the fall of 2013, and  
436 the spring and winter of 2014 fell below the 5th percentile limit for drought in the historical period  
437 3. Soil moisture in our irrigation plots was generally above the historical average soil moisture but  
438 conditions never exceeded the 90th percentile for soil moisture in the historical period.

#### 439 Effects on cover and vital rates

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

444 Our treatment models fit to the experimental and observational data indicated a variety of  
445 treatment effects on the vital rates of each species. After comparing WAIC scores for the growth  
446 models with and without the size by treatment effects, we retained size by treatment interaction  
447 effects in the growth models for *A. tripartita* and *P. secunda*, and the survival model for *P. secunda*.  
448 For *A. tripartita* we found significant size by treatment effects of drought: drought had positive  
449 effects on plants of average size and smaller 5, but plants larger than the mean size by more than 1.5

450 standard deviations grew slightly less in the drought treatment than in the controls. *A. tripartita*  
451 showed the opposite response in the irrigated plots, (although the irrigation parameters were not  
452 technically significant at the 95% confidence level): irrigation reduced growth for small plants while  
453 irrigation increased growth of plants more than 1.5 standard deviations greater than the mean size.  
454 Drought led to a strong (but not significant) decrease in *H. comata* growth, while irrigation had  
455 no effect on growth. Like *A. tripartita*, we saw size by treatment effects on *P. secunda* growth,  
456 with the negative effects of drought becoming greater for larger plants. *P. secunda* showed the  
457 opposite response in the irrigation plots with larger plants showing the largest increase in growth  
458 in response to irrigation (although these effects were technically not significant). *P. spicata* growth  
459 was relatively unaffected by the drought and irrigation treatments.

460 Drought decreased the survival of all three grass species (fig 6). And *P. secunda* showed a  
461 negative size by drought interaction effect, indicating that the survival of larger plants was more  
462 negatively affected by drought than that of the smaller plants, with the smallest plants (plants one  
463 standard deviation smaller than the mean) actually seeing a slightly positive effect of drought. *A.*  
464 *tripartita* survival was relatively unaffected by the drought and irrigation treatments.

465 Recruitment in our irrigation plots was significantly less than in control plots for two grass  
466 species *P. secunda* and *P. spicata* (fig 7). However, recruitment was also lower in the drought plots  
467 than in the the control plots (although not significantly so), indicating that the decrease in the  
468 irrigated plots may have not been entirely due to the irrigation itself. The recruitment data for *A.*  
469 *tripartita* were relatively limited, with only 32 new recruits in total observed in all 30 plots over  
470 the course of the five year experiment.

471 As expected from previous research most of our demographic models estimated strong  
472 negative intra-specific crowding effects and weaker negative inter-specific crowding effects on the  
473 focal species (appendix) (???).

#### 474 Effects of soil moisture on vital rates

475 We choose three seasonal soil moisture variables for each species' climate model based on their  
476 correlation with the year effects in the random year effects model fit to the observational data  
477 (table 7). We included size by soil moisture variables for *A. tripartita* and *P. secunda* based on  
478 the treatment response we observed in the experiment. All three time lags and all three seasons

479 show up in the choosen variables. After fitting the vital rate models with the selected soil moisture  
480 variable we observed a trend towards positive soil moisture effects on growth of all three grasses  
481 8. For *H.comata* the soil moisture of the most recent summer had a significantly positive effect  
482 while the soil moisture of the previous summer and the fall before that were also positive but not  
483 significant. For *A. tripartita* the summer and fall soil moisture of the previous year had strong  
484 negative effects on growth. There were also strong positive size by climate interaction effects for  
485 these variables: soil moisture had a stronger negative effect on small plants and a positive effect  
486 only on the largest plants.

487 Soil moisture had significant effects on the survival of all four species 9. As for growth the  
488 grasses showed mainly positive effects while *A. tripartita* showed a significant negative effect of soil  
489 moisture in the summer of the previous year and a strong negative effect of spring soil moisture of  
490 the previous year. *H. comata* showed a significant positive effect of spring lag soil moisture and a  
491 strong positive effect of spring soil moisture in the previous year. *P. secunda* showed a significant  
492 positive effect of the previous spring's soil moisture and there was a significant interaction between  
493 this effect and plant size: as plant size increased this effect became more strongly positive. Finally  
494 for *P. spicata* there was a significant positive effect of lag spring soil moisture on survival.

495 There were only two significant effects of soil moisture on recruitment: lag fall soil moisture  
496 had a positive effect on *P. secunda*, and lag summer soil moisture had a negative effect on *P. spicata*  
497 recruitment. Soil moisture of the previous year's summer had also had a strong negative effect on  
498 *P. spicata* recruitment.

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

## 501 Evaluating the predictions

502 For most models adding climate covariates did not improve our ability to predict species responses  
503 in the experiment 8. However, the climate models did improve overall prediction MSE for growth of  
504 *A. tripartita* and growth and survival of *P. secunda* (table 8). In terms of lppd, the climate model  
505 outperformed the year effects model for *A. tripartita* growth, *H. comata* recruitment, *P. secunda*  
506 growth and survival and *P. spicata* recruitment. .

507 When we look at the predictions for each treatment separately we see that climate covariates  
508 improved model predictions more often in the drought treatments than in the control or irrigation  
509 treatments SI-1. For all four species, the climate model outperformed the year effects model for  
510 predicting the response of growth to drought in terms of lppd SI-1. The climate model outperformed  
511 the year effects model for predicting irrigation effects on growth for all species except *H. comata*.

512 Overall our climate models often predicted the correct direction of the drought and irriga-  
513 tion treatments 11. In four cases we both observed and predicted treatment effects significantly  
514 different from zero based on the 95% Bayesian credible interval around the parameters: the drought  
515 response of *H. comata* survival (fig SI-6), the drought response of *P. secunda* growth (fig SI-3),  
516 the irrigation response of *P. spicata* recruitment (fig ??) and the irrigation response of *P. secunda*  
517 recruitment (fig ??). In only one of these cases, for *P. secunda* recruitment, was the predicted effect  
518 in the opposite direction from the observed treatment effect 11. The overall correlation between  
519 the predicted and observed treatment effects for all treatments, species and vital rates was  $r =$   
520 0.54, whereas the correlation for the drought treatment effects ( $r = 0.77$ ) was better than for the  
521 irrigation effects ( $r = 0.46$ ). Also for the three models in which we included size by treatment or  
522 size by climate interactions the correlation of these size dependent effects were much stronger than  
523 the intercept parameter estimates 11.

524 Using the vital rate models for each species we generated one year ahead cover predictions  
525 for each quadrat in each year of the experiment. Average cover predicted by the climate model  
526 tended to be lower than the observed cover each year for *A. tripartita* and *P. secunda* (fig 12).  
527 Comparing the overall population growth rates predicted to those observed in the experiment, we  
528 see that the MSE of the climate model was lower than the MSE of the year effects model for  
529 *P. secunda* and *P. spicata* (table 9). The predictions produced by the climate model for these  
530 species were also slightly more correlated with the observations than the predictions produced by  
531 the year effects model (table 9). Considering each treatment and species separately we see that the  
532 predicted population growth rates for *A. tripartita*, *P. secunda* and *P. spicata* were all consistently  
533 lower than the observed population growth rates (figs ??, 15, 16). The climate model showed lower  
534 MSE for *A. tripartita*, *P. secunda* and *P. spicata* in the irrigation treatment, *P. spicata* in the  
535 control treatment and *H. comata* in the drought treatment (figs 14). However, the correlations  
536 between the predicted and observed log changes in cover did not always show the same pattern as

537 MSE: the climate model made more strongly correlated predictions with the observations than the  
538 year effects model only for *P. spicata* and *P. secunda* in the control treatment and *P. secunda* and  
539 *H. comata* in the drought treatment.

## 540 Discussion

541 Our experiment showed that observational data on the response of plant populations to interannual  
542 climate variation can indeed help us predict the direction of species response to experimental climate  
543 manipulations 11. This was true even though adding climate parameters to the demographic models  
544 only improved vital rate predictions for half of the models 8. This should give us some hope that  
545 even when climate effects in demographic models to observational data are weak, they may contain  
546 useful qualitative information on the direction of climate effects in the future.

547 \* discuss biggest successes—astreces in fig 11 \* discuss biggest mistake, *P. secunda* recruit-  
548 ment went in wrong direction.

549 Scaling up to the population level, the climate models only produced better one step a head  
550 predictions of the overall response of the species to the experiment for two species: *P. spicata* and  
551 *P. secunda* (table 9). However adding climate effects produced better population-level predictions  
552 in the drought treatments for *P. secunda* and *H. comata*. Both the survival and the growth of these  
553 species were positively affected by increased soil moisture (fig 8, 9), and so it makes intuitive sense  
554 that we would observe declines in these species cover in the drought treatment (fig 4).

555 \*Among our species, we had the most success predicting the response of the three grass  
556 species and less success predicting the response of the shrub species *A. tripartita*. Why might  
557 ARTR be different?

558 \*We also had more success predicting species response to the drought treatment than to the  
559 irrigation treatment. We also tended to see species respond more strongly to the drought treatment  
560 than they did to the irrigation treatment. This tells us that in some cases, demographic models  
561 based on observational data may be of more use for prediction in extreme conditions than when  
562 conditions are close to average. Our drought treatment likely created extremely low soil moisture  
563 compared to the historical average 3. We hypothesize that this made water the most limiting  
564 resource for plants in these plots. In contrast, in the control and irrigated plots soil moisture may  
565 not have always been limiting during the course of the experiment. It may make sense then that

566 our climate models did not make more accurate predictions in these conditions than a random year  
567 effects model.

568 \*Mixed success at predicting species responses to high and low moisture availability has  
569 interesting implications. On the one hand it is re-assuring that observational data is sometimes  
570 useful. On the other hand, for predictions to be truly useful we would also like more information  
571 to help us sort out why our predictions for some species were good and for others no better than  
572 a null model. In other words, we have little ability to predict when our predictions are likely to  
573 be accurate. Among plants detailed physiological ecology may give us a guide to which factors  
574 and climate conditions are likely to affect which species. In our system, for instance, it may make  
575 sense that the grasses showed a stronger response to the drought treatment than the a woody  
576 shrub. Although these grasses are adapted to the arid conditions that characterize the sagebrush  
577 steppe, they thrive during the brief window in spring and early summer when the soil moisture  
578 and temperatures are warmer. *A. tripartita* on the other hand grows throughout the summer and  
579 generally has deeper roots than the grasses. These traits may help it tolerate the water stress  
580 induced by drought.

581 \*Will our predictions be useful for the longer term? Because these species compete, one  
582 could argue that our predictions for any one species in this community will only be as good as the  
583 predictions we make for their competitors. For instance, while we observed little effect of drought  
584 on *A. tripartita* in our experiment, it is possible that it will eventually respond positively to the  
585 drought treatment as cover of the grass species it competes with declines (fig 4). We know that  
586 grass species compete strongly with *A. tripartita*. However, we also know from previous work in  
587 this community that each species is more limited by intra-specific competition than by inter-specific  
588 crowding. This fact ensures that the direct effects of climate change will generally be greater than  
589 the competition mediated indirect effects of climate change ??.

590 \*Our results give us more confidence that historical observational data can in theory be  
591 used to detect and predict the demographic effects of climate change. This should encourage more  
592 researchers to try and use observational data to predict the future in both experimental and natural  
593 settings. Nevertheless, our success at predicting the short-term response of two species to a small-  
594 scale climate manipulation is not likely to be very reassuring to applied ecologists and resource  
595 managers wishing to make accurate medium to long-term quantitative predictions about the effects

596 of climate change on the species they manage. Clearly more work is needed to distinguish which  
597 predictions are worth having confidence in and which predictions we should have less confidence in.  
598 Towards that goal, perhaps the best way to build confidence in ecological predictions is conduct  
599 more tests like this one.

600 **Acknowledgements**

601 Funding was provided by an NSF GRFP to AK, NSF grants DEB-1353078, and DEB-1054040 to  
602 PBA and by the Utah Agricultural Experiment Station (get journal paper number). The USDA-  
603 ARS Sheep Experiment Station generously provided access to historical data and the field experi-  
604 ment site. Joe and Kevin, Lau, John Bradford, Caitlin (soilwat),

## Tables

Table 1: Treatment effects on spring soil moisture

	Model 1	
(Intercept)	8.81*	[5.78; 11.83]
TreatmentDrought	-3.97*	[-4.84; -3.09]
TreatmentIrrigation	3.26*	[2.39; 4.14]
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	

\* 0 outside the confidence interval

Table 2: Model of treatment effects on soil moisture

	Model 1	
(Intercept)	-0.57*	[-0.89; -0.26]
TreatmentIrrigation	1.23*	[1.18; 1.29]
rainfallrainy	-0.05	[-0.12; 0.01]
seasonspring	0.27*	[0.23; 0.32]
seasonsummer	0.15*	[0.10; 0.19]
seasonwinter	0.25*	[0.21; 0.29]
TreatmentIrrigation:rainfallrainy	0.18*	[0.13; 0.24]
TreatmentIrrigation:seasonspring	-0.23*	[-0.29; -0.16]
TreatmentIrrigation:seasonsummer	-0.26*	[-0.33; -0.20]
TreatmentIrrigation:seasonwinter	-0.33*	[-0.40; -0.27]
rainfallrainy:seasonspring	-0.23*	[-0.31; -0.16]
rainfallrainy:seasonsummer	-0.07	[-0.14; 0.01]
rainfallrainy:seasonwinter	-0.07	[-0.20; 0.07]
AIC	14581.58	
BIC	14695.49	
Log Likelihood	-7274.79	
Num. obs.	9133	
Num. groups: date	1596	
Num. groups: PrecipGroup	8	
Var: date (Intercept)	0.00	
Var: PrecipGroup (Intercept)	0.19	
Var: Residual	2.50	

\* 0 outside the confidence interval

	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 3: Treatment effects on log cover change for *A. tripartita* 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 4: Treatment effects on log cover change for *H. comata* 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 5: Treatment effects on log cover change for *P. secunda* 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 6: Treatment effects on log cover change for *P. spicata* from 2011 to 2016. Intercept gives control effects.

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 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*.

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 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	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	***

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*.

## Figures

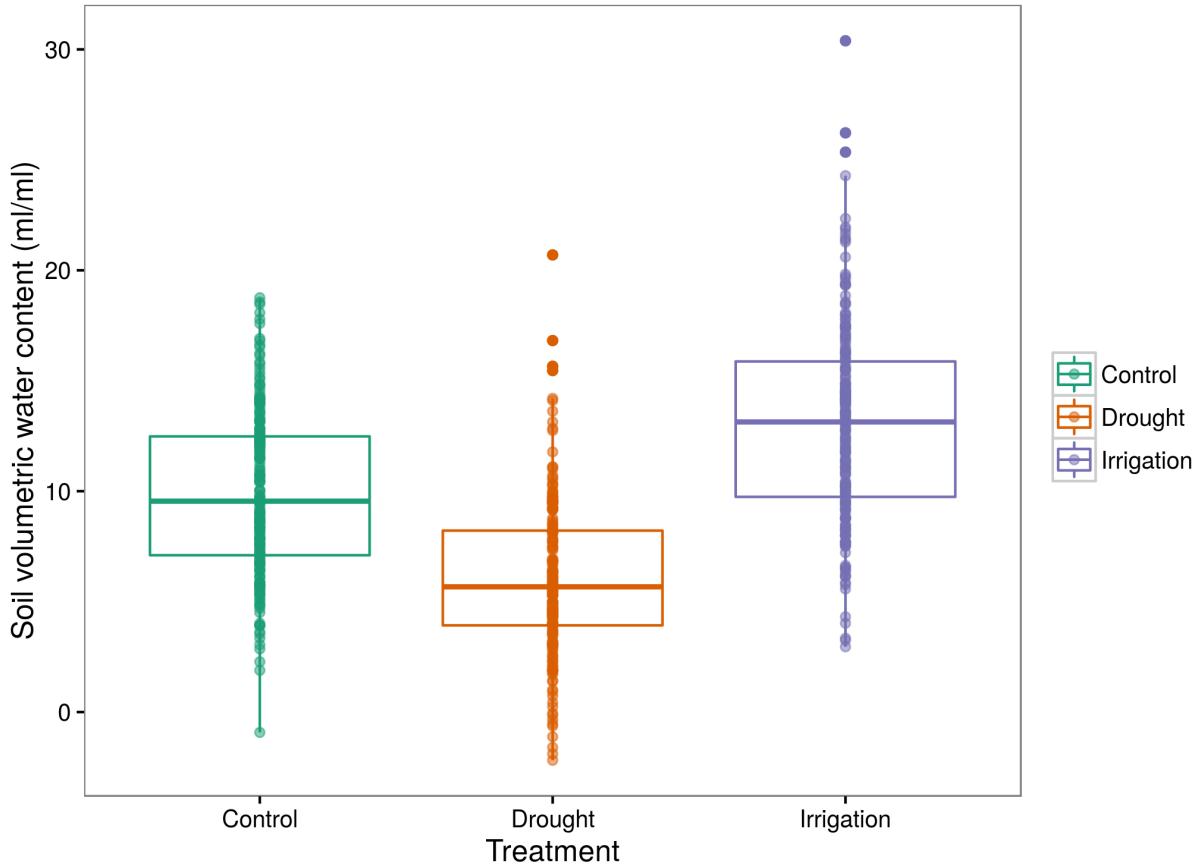


Figure 1: Soil moisture in the upper 5 cm of drought and irrigated plots compared to ambient controls. Soil moisture was measured at six locations around each plot at five different dates during the spring. Control plots were nearby areas of experiencing ambient soil moisture. Box plots show the median 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.

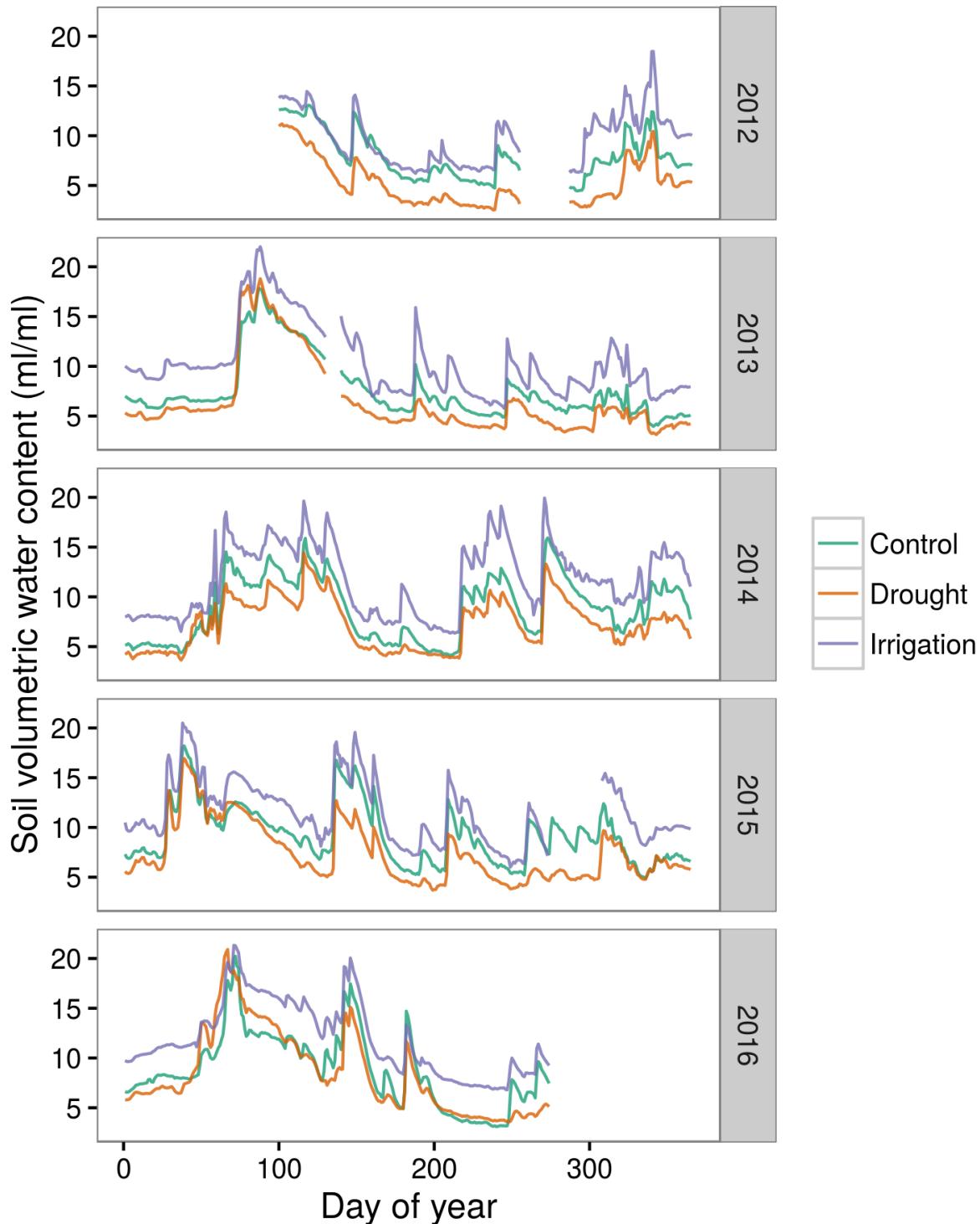


Figure 2: Average soil moisture in the control, drought, and irrigation treatments during each year of the experiment. Soil moisture was monitored in four drought plots, four irrigated plots and four ambient control plots. Two sensors were installed at 5 cm depth at each plot and two at 25 cm and data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

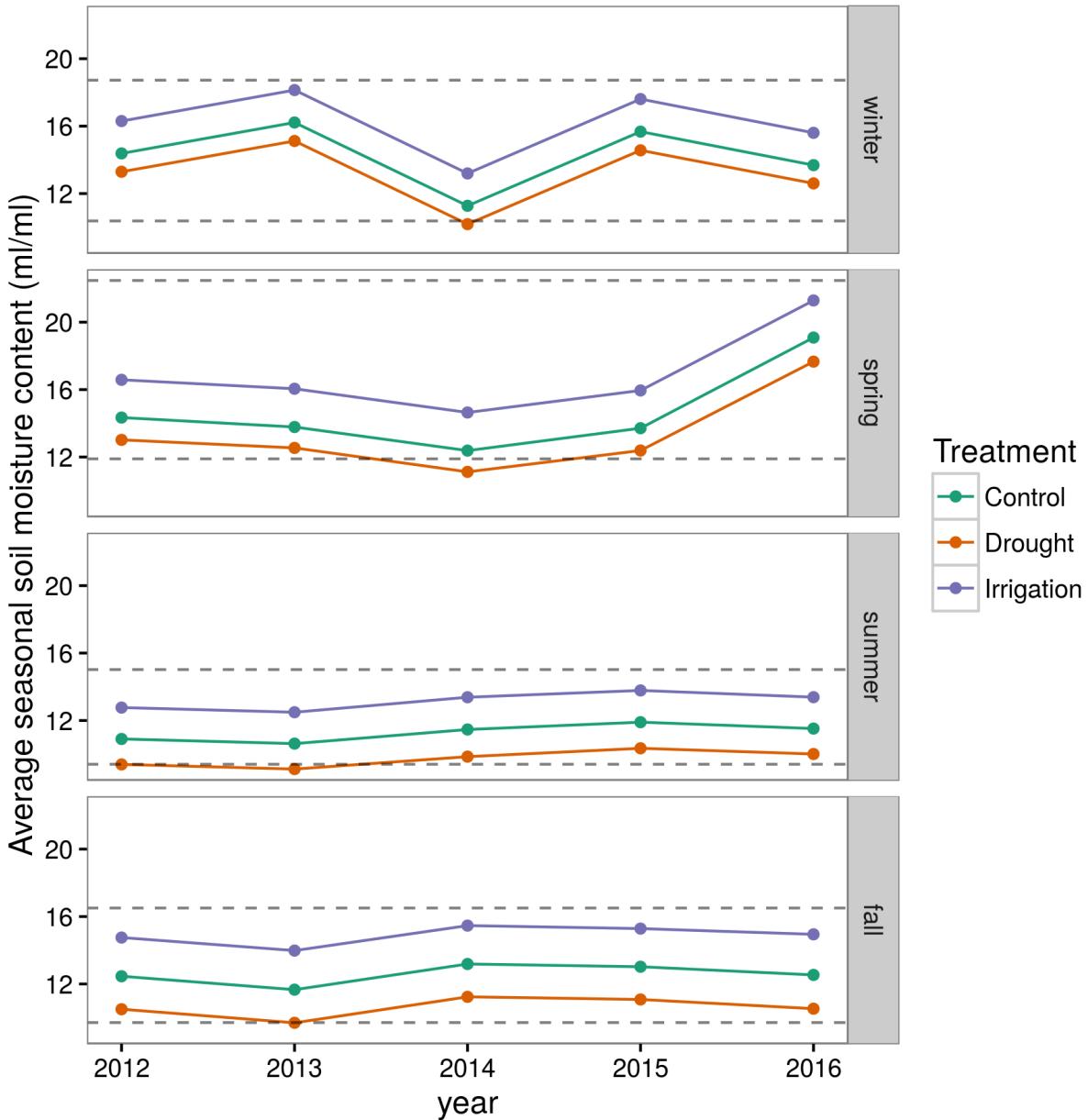


Figure 3: Average seasonal soil moisture in the control, drought, and irrigation treatments during each year of the experiment. The dashed gray lines give the 5th percentile and 95th percentile limits for seasonal soil moisture in the historical record (1929 to 2010).

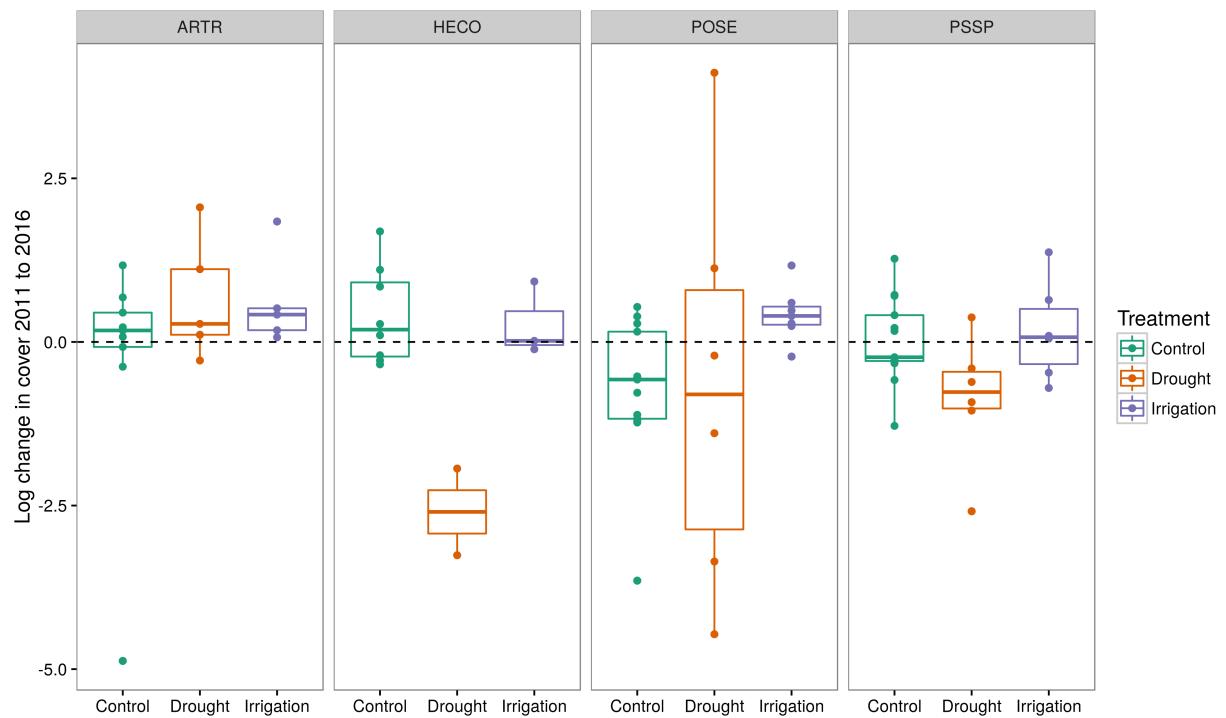


Figure 4: Log change in cover in each of the experimental plots from the pre-treatment monitoring in 2011 to the last year of the experiment in 2016. Box plots show the median cover change and the interquartile range. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

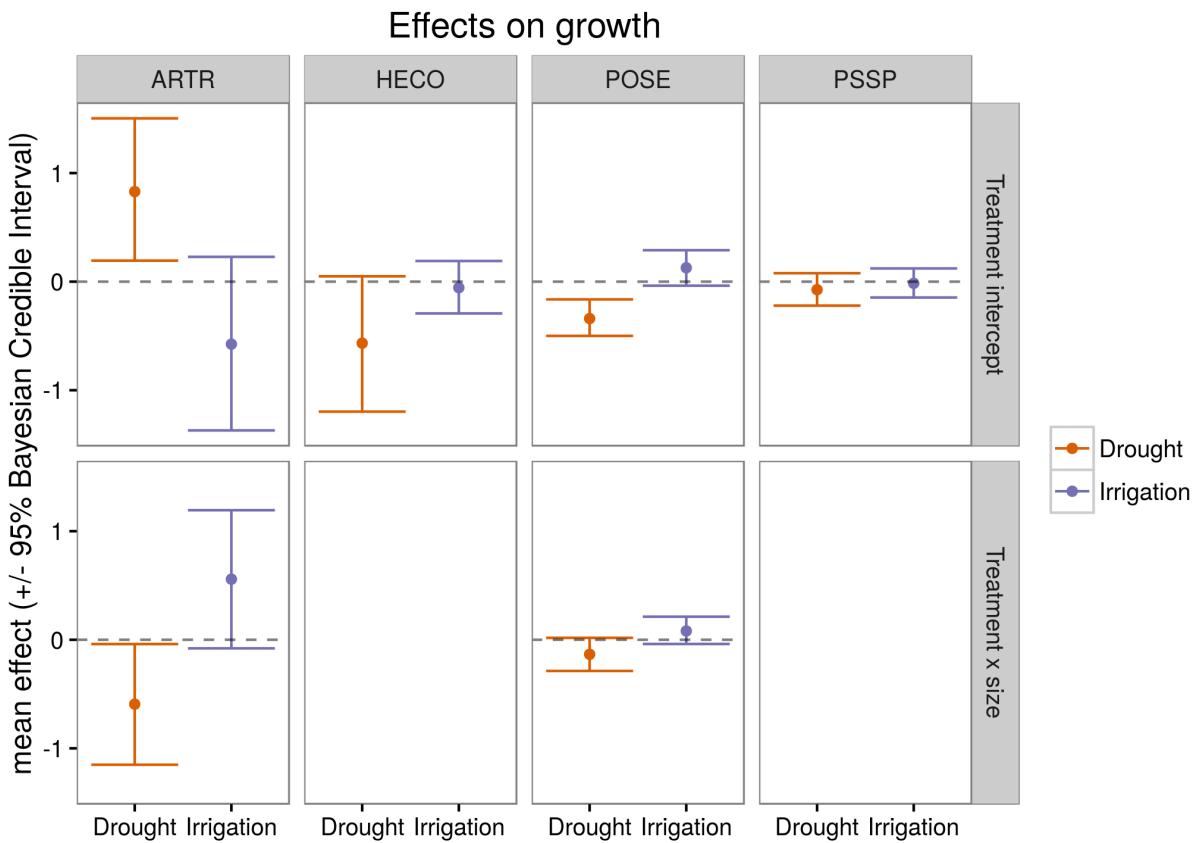


Figure 5: Parameter estimates for the effects of treatment on growth of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. Size by treatment interactions were only fit for ARTR, and POSE. Plant size was centered on mean size and scaled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

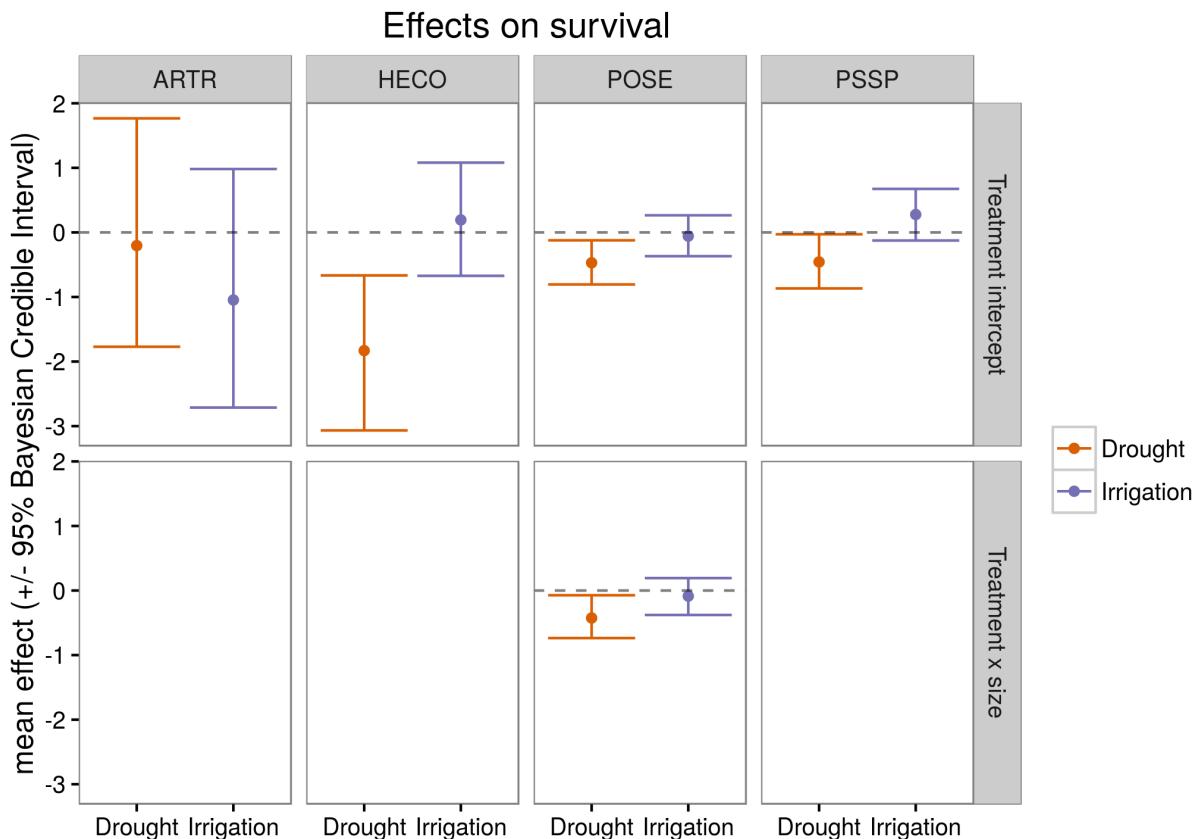


Figure 6: Parameter estimates for the effects of treatment on survival of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. Size by treatment interactions were only fit for POSE. Plant size was centered on mean size and scaled by its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

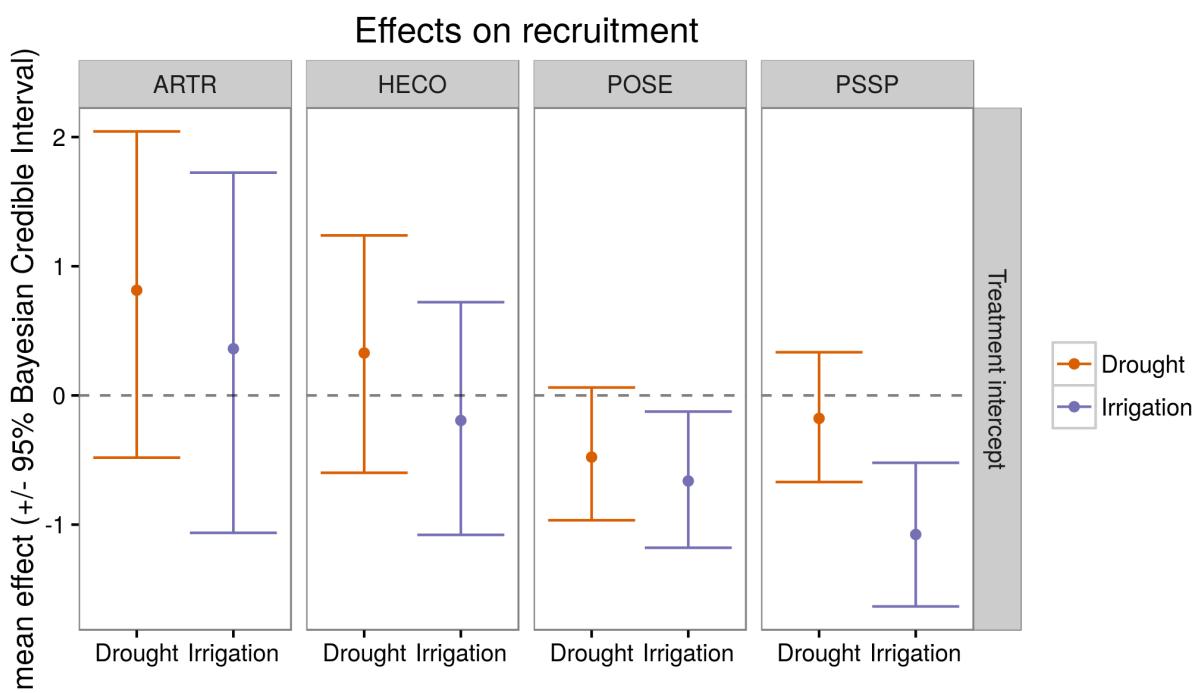


Figure 7: Parameter estimates for the effects of treatment on recruitment of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

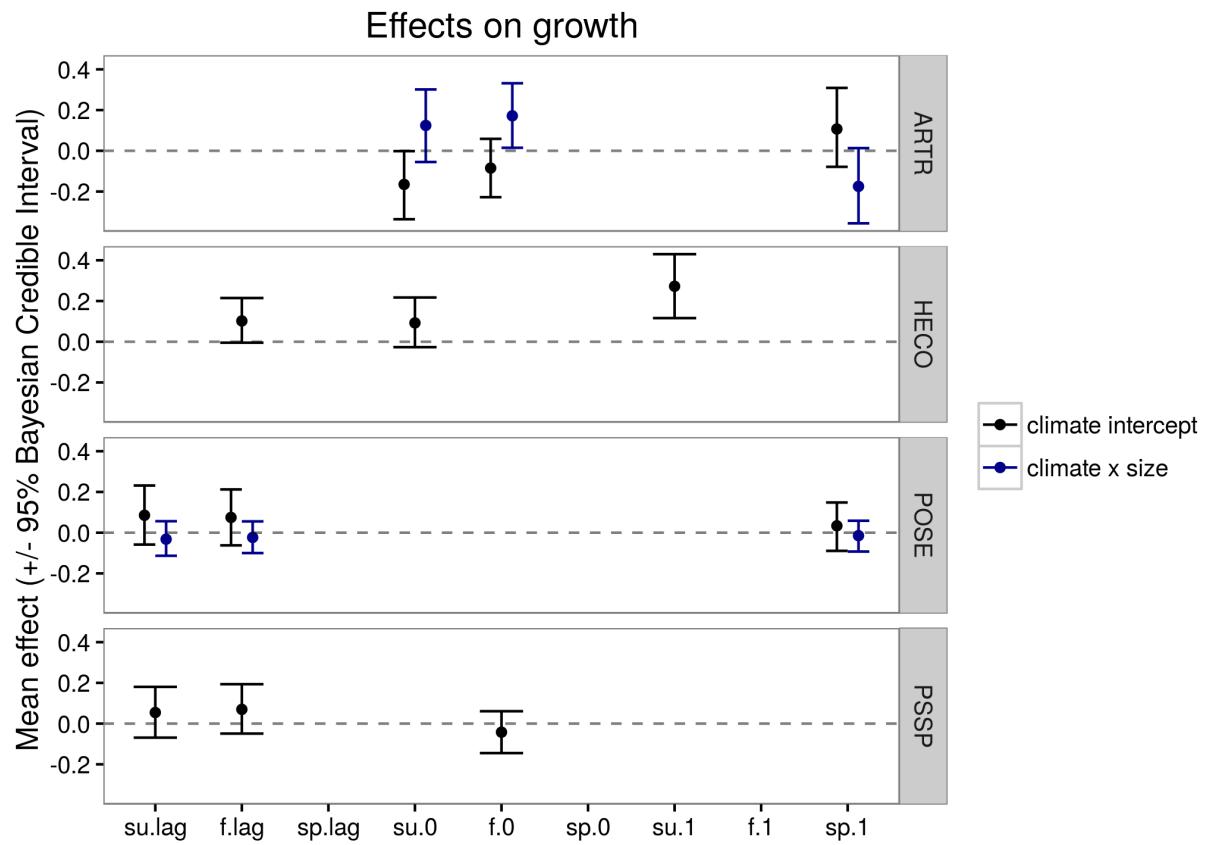


Figure 8: Parameter estimates for the selected seasonal soil moisture covariates on the growth of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. Dark blue parameters show size x climate interaction effects. 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*.

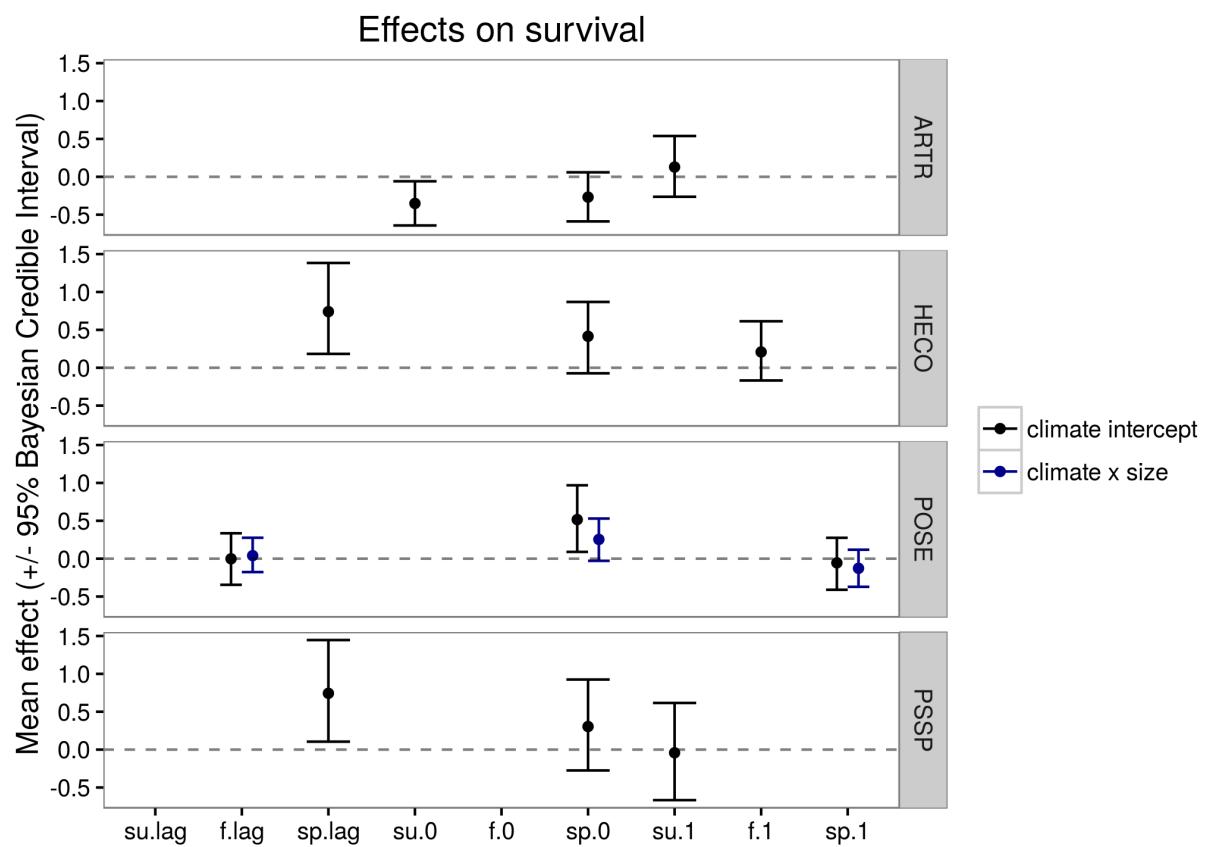


Figure 9: Parameter estimates for the selected seasonal soil moisture covariates on the survival of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. Dark blue parameters show size x climate interaction effects. 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*.

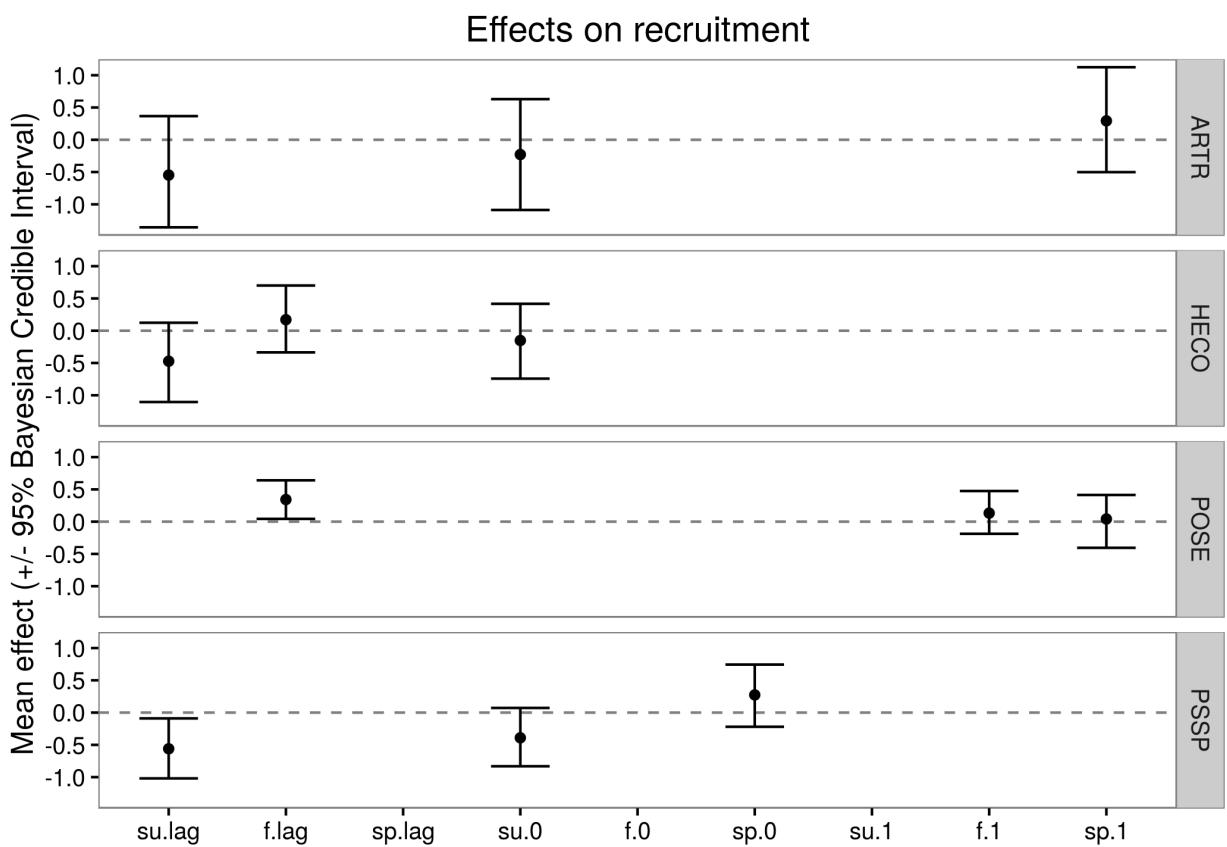


Figure 10: Parameter estimates for the selected seasonal soil moisture covariates on the survival of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. Dark blue parameters show size x climate interaction effects. 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*.

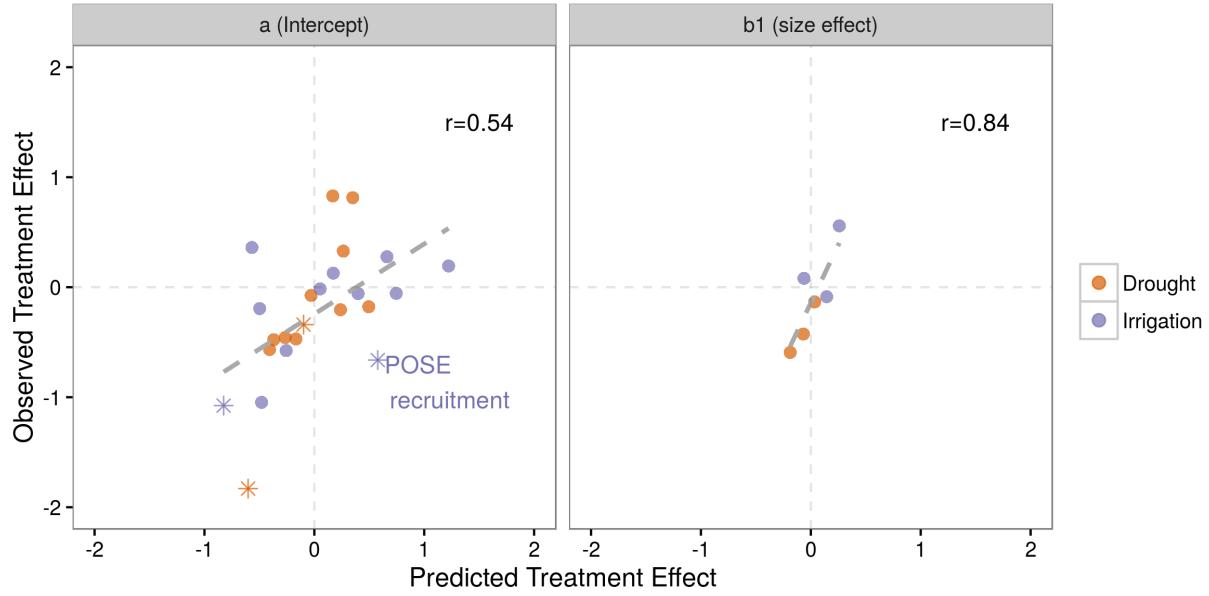


Figure 11: The treatment effects predicted by the climate model compared to the treatment effects observed for the intercept parameters (left side) and size by climate/treatment effects (right side). The observed treatment effects come from the treatment model fitted to all the data including the five 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.

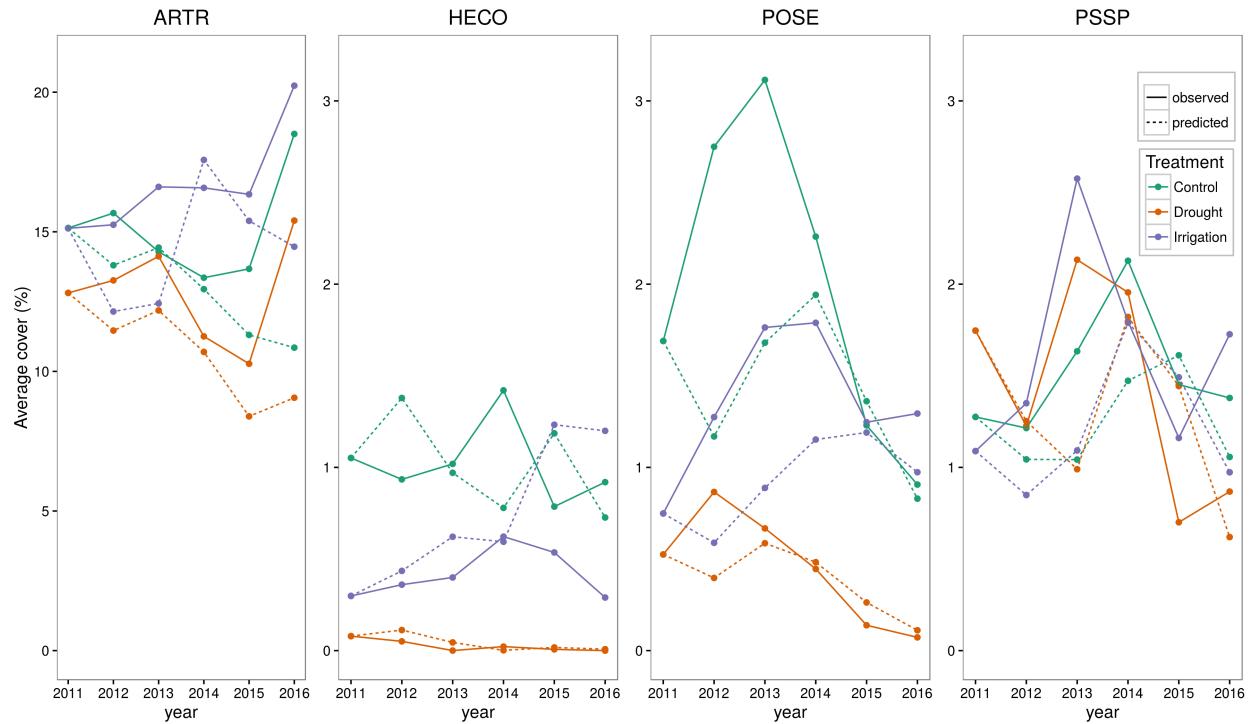


Figure 12: Observed average cover per quadrat in the experimental and control plots and one year ahead cover predictions from the climate model. Cover predictions for each year are generated from the IBM based on the observed distribution of plants in each quadrat in the current year. Quadrat cover was 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*

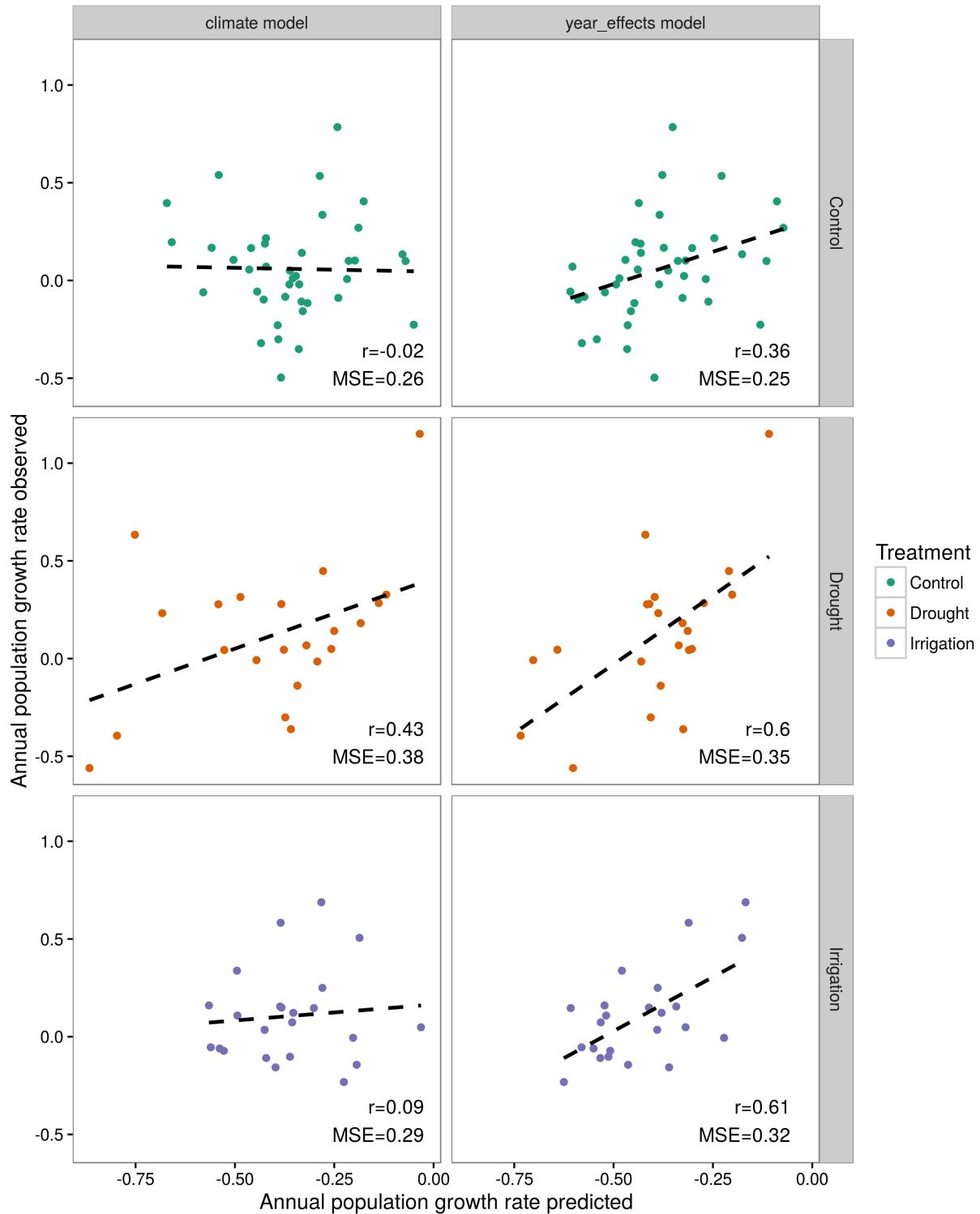


Figure 13: Observed and predicted one year ahead log change in *A. tripartita* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the year effects model are shown on the right. Correlations coefficients between predictions and observations and MSE are shown for each treatment and model.

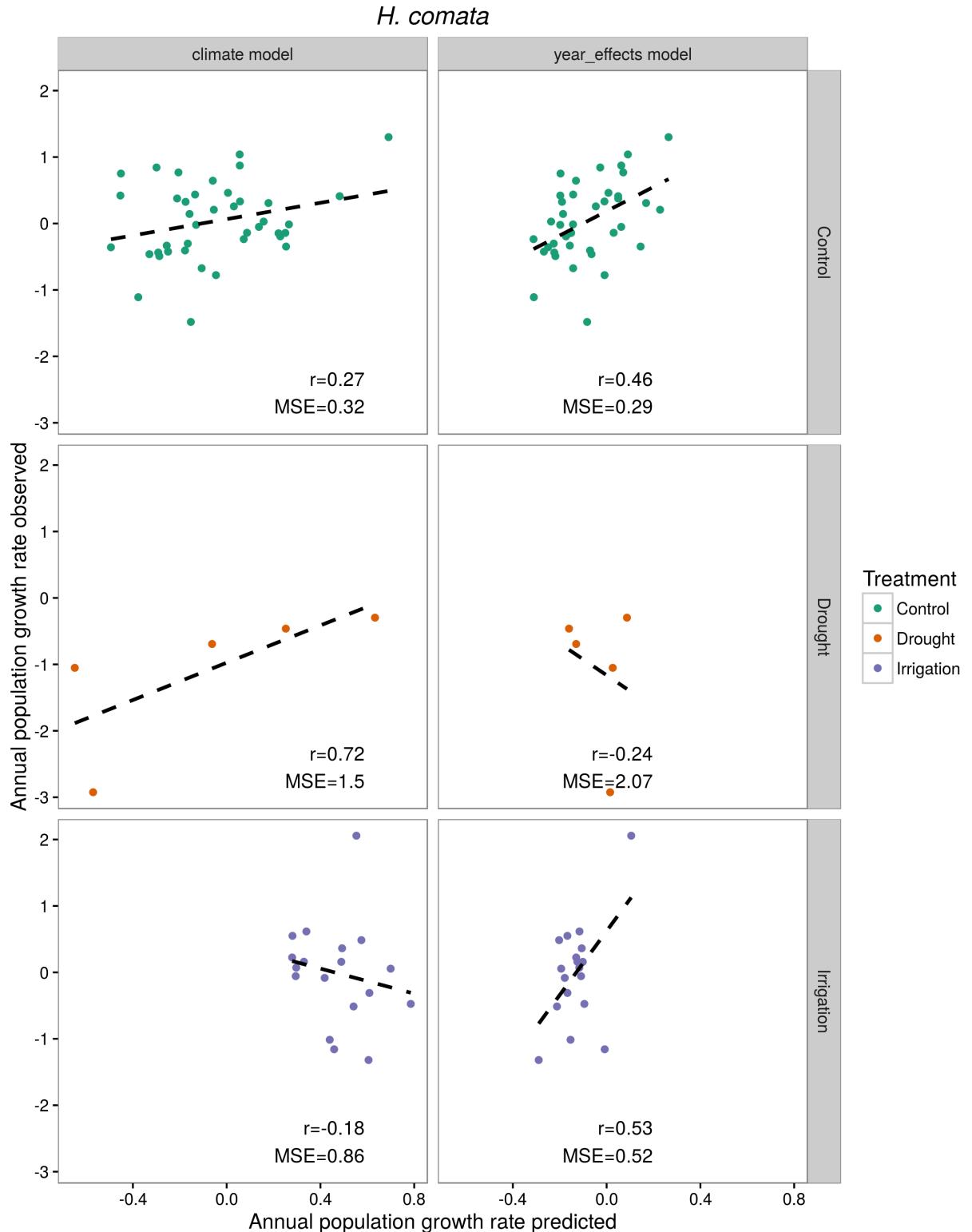


Figure 14: Observed and predicted one year ahead log change in *H. comata* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the year effects model are shown on the right. Correlations coefficients between predictions and observations and MSE are shown for each treatment and model.

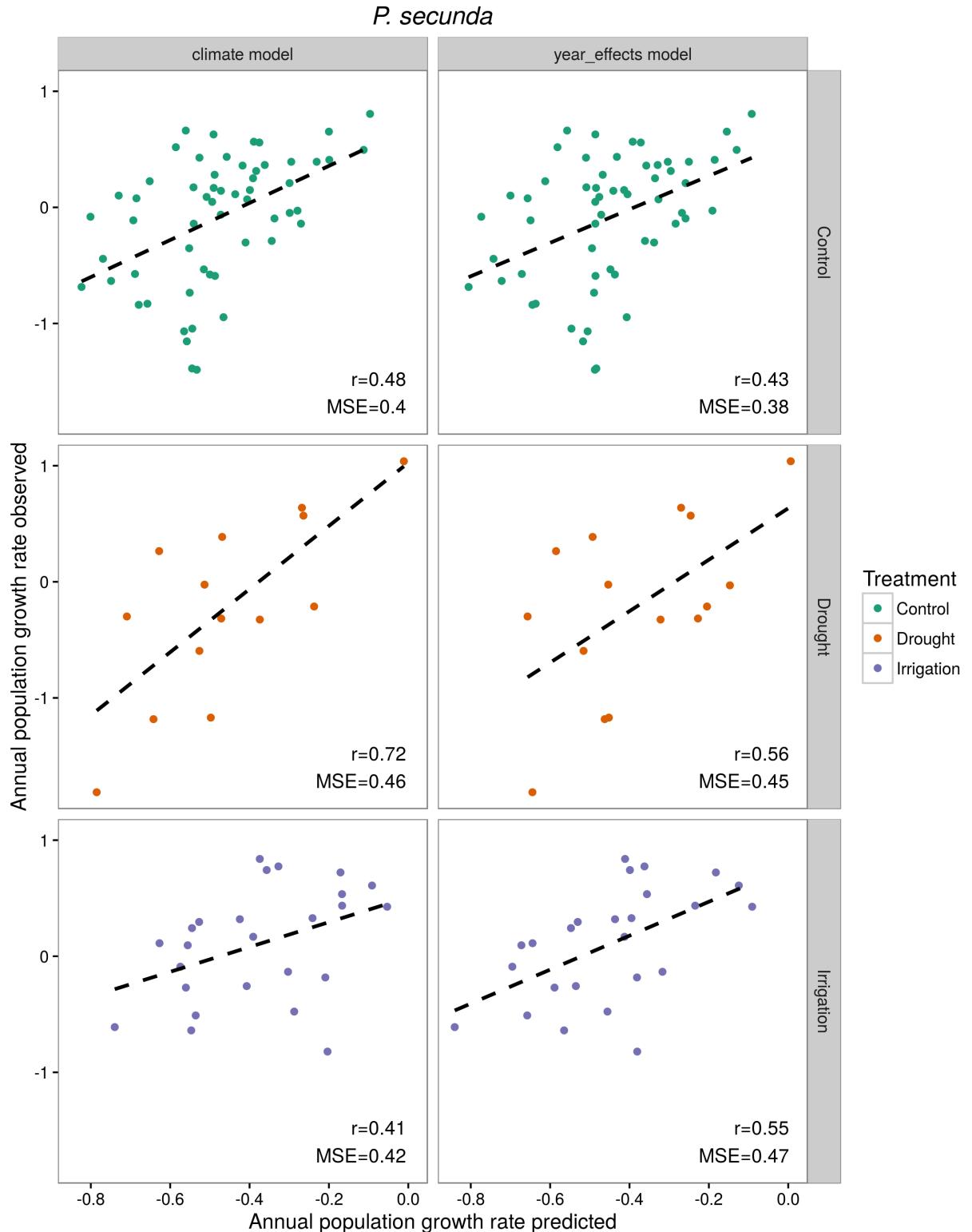


Figure 15: Observed and predicted one year ahead log change in *P. secunda* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the year effects model are shown on the right. Correlations coefficients between predictions and observations and MSE are shown for each treatment and model.

*P. spicata*

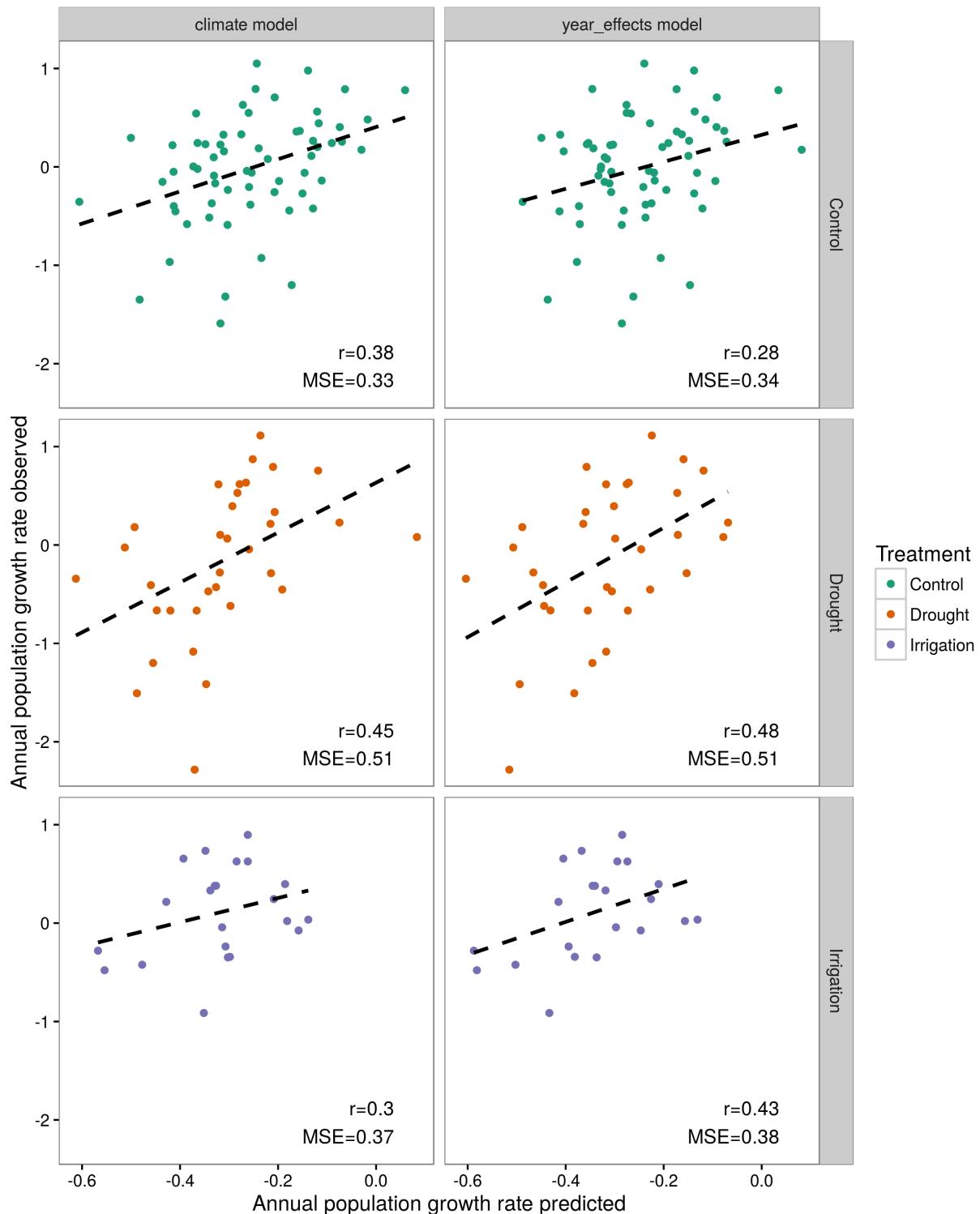


Figure 16: Observed and predicted one year ahead log change in *P. spicata* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the year effects model are shown on the right. Correlations coefficients between predictions and observations and MSE are shown for each treatment and model.

## Supporting Information

Adler et al., “Weak interspecific interactions”

## 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

? 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 ? 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  $x$ .

## Additional Tables

	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	***
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	***

Continued on next page

	species	vital_rate	Treatment	score	climate model	year effects	model	diff	improved
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	
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	

Continued on next page

	species	vital.rate	Treatment	score	climate model	year effects model	diff	improved
71	PSSP	survival	Irrigation	lppd	-100.52	-89.02	-11.50	
72	PSSP	survival	Irrigation	MSE	0.13	0.11	0.01	

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*, HEKO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

## Additional Figures

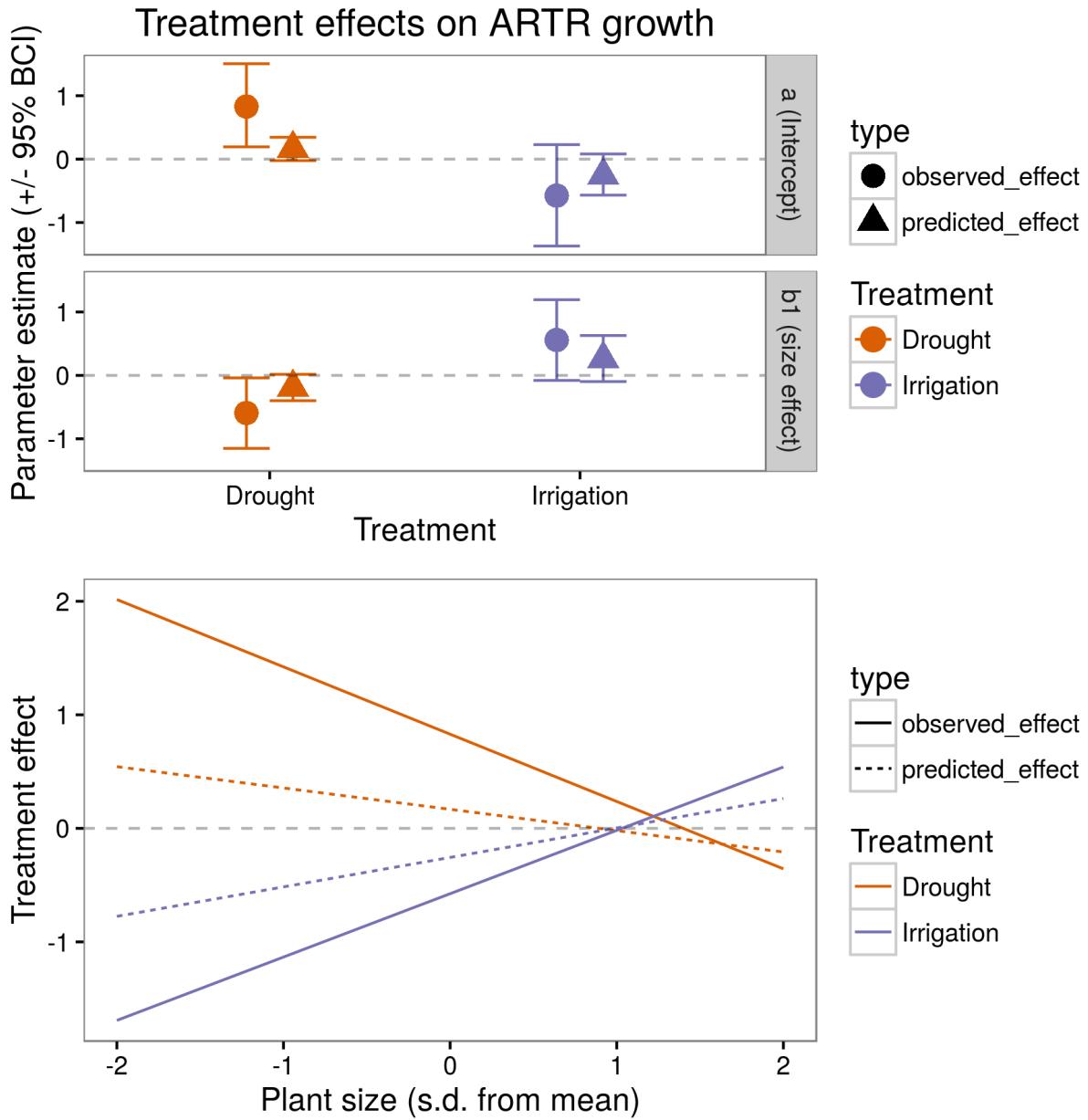


Figure SI-1: Comparison of treatment effects predicted and observed for *A. tripartita* growth. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to 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.

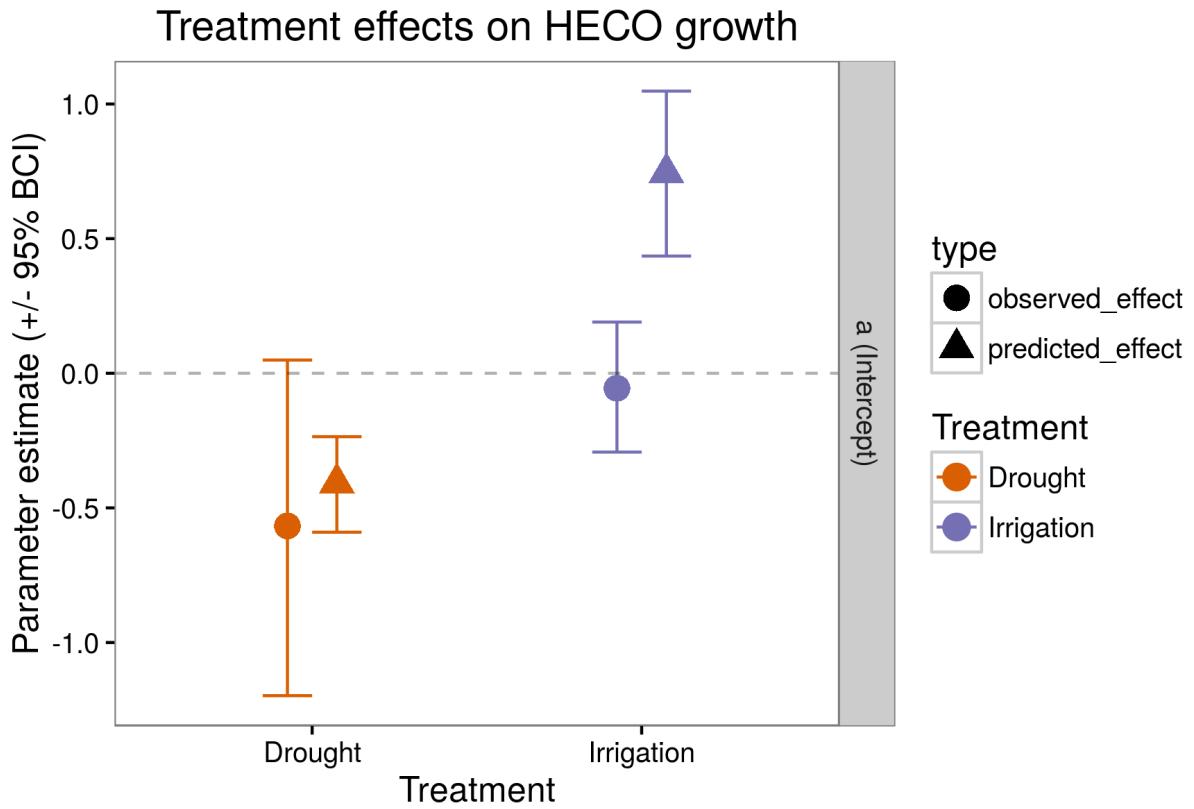


Figure SI-2: Comparison of treatment effects predicted and observed for *H. comata* growth. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational 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.

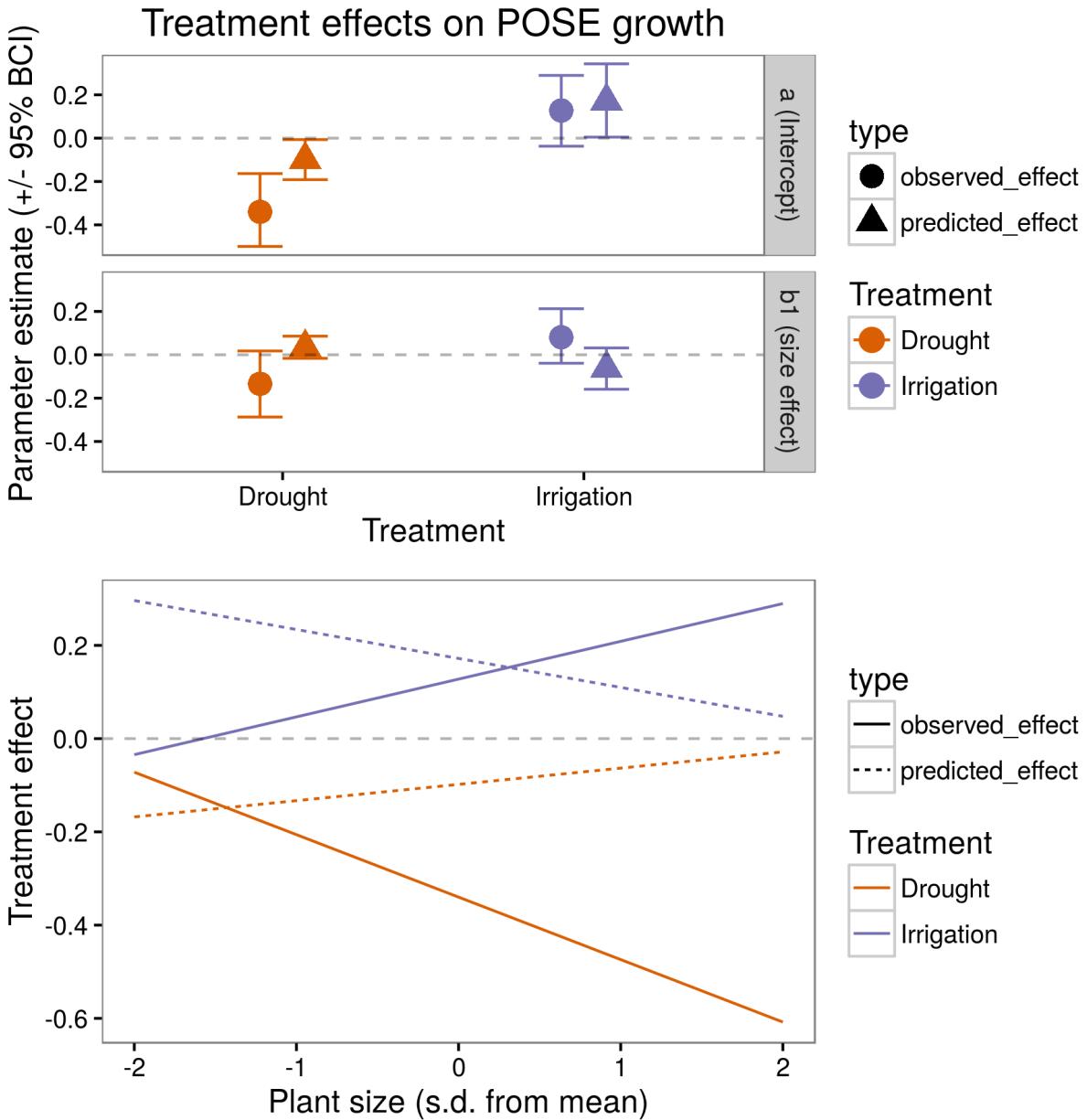


Figure SI-3: Comparison of treatment effects predicted and observed for *P. secunda* growth. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to all 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.

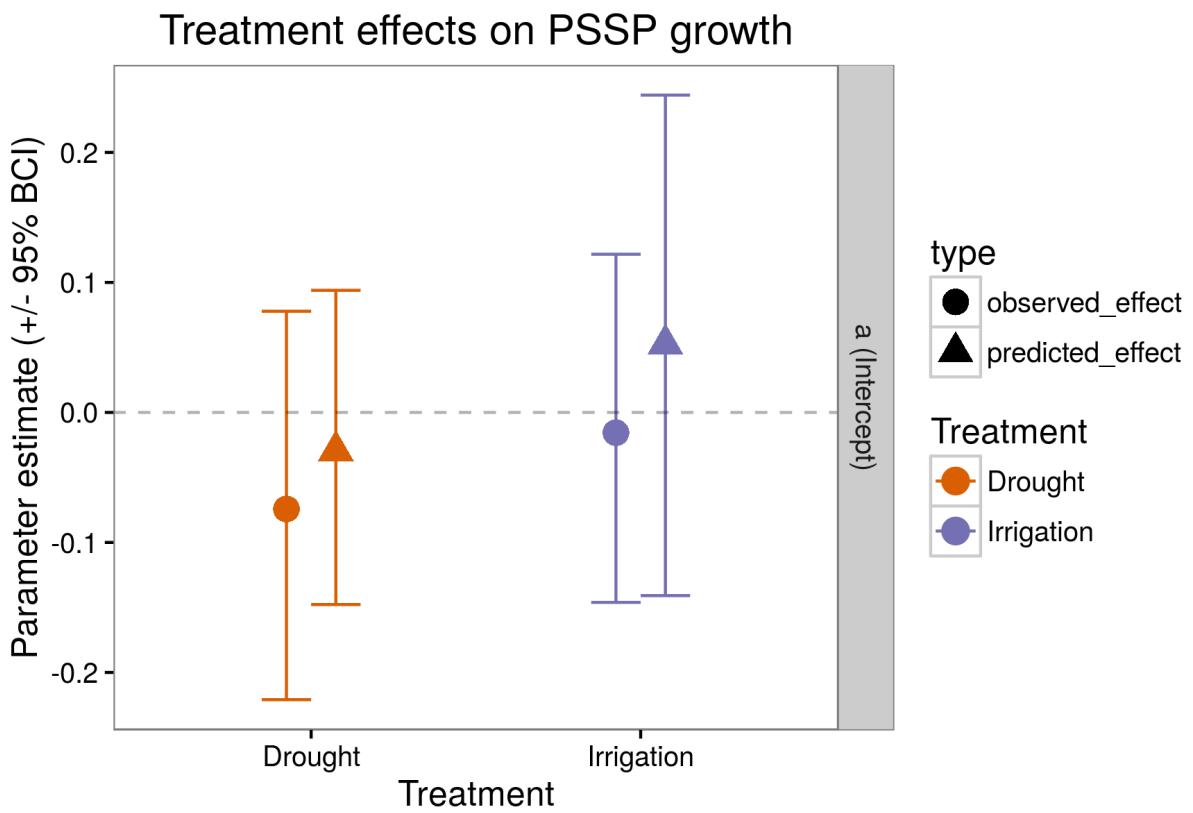


Figure SI-4: Comparison of treatment effects predicted and observed for *P. spicata* growth. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational 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

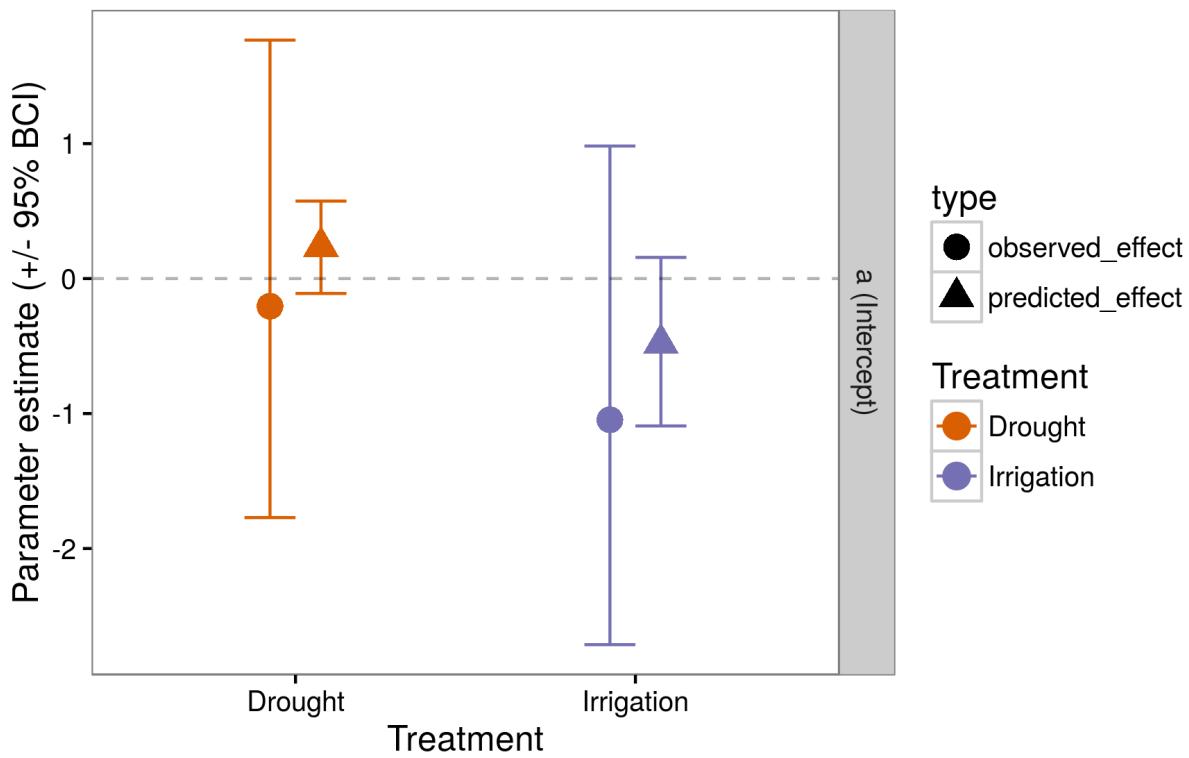


Figure SI-5: Comparison of treatment effects predicted and observed for *A. survival* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational 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.

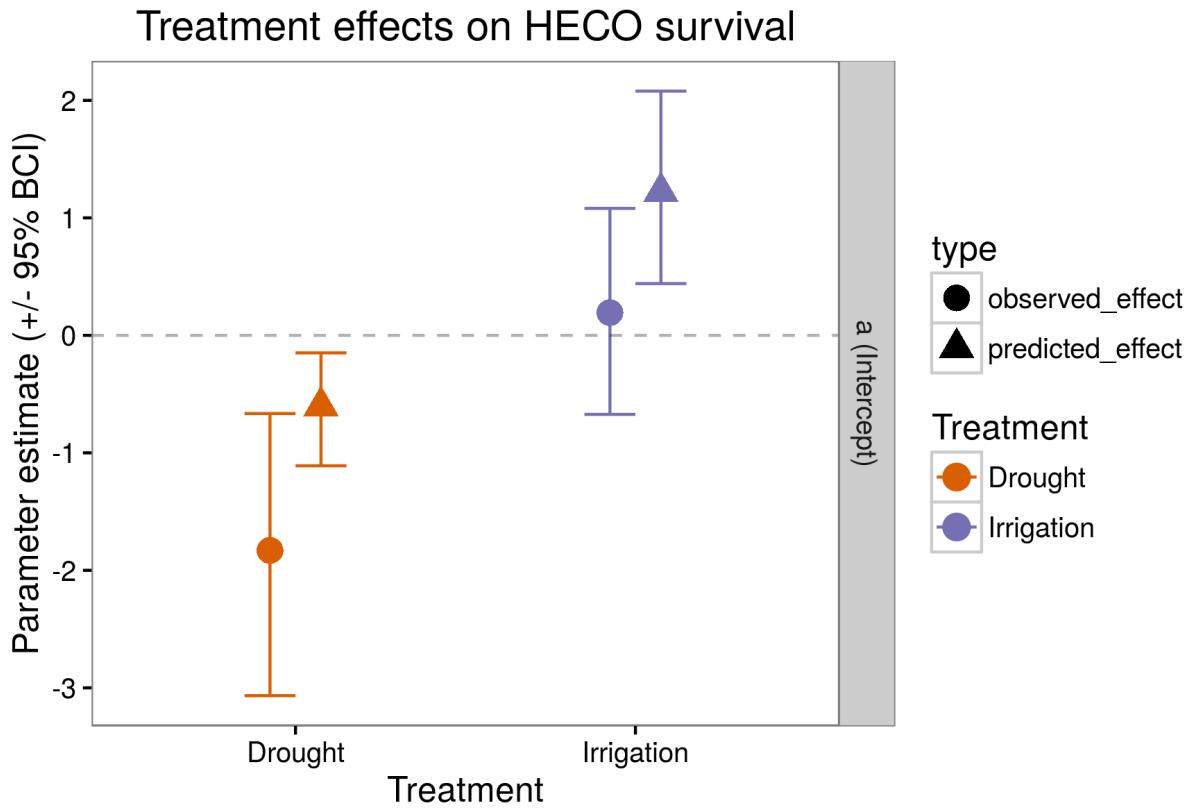


Figure SI-6: Comparison of treatment effects predicted and observed for *H. comata* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational 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.

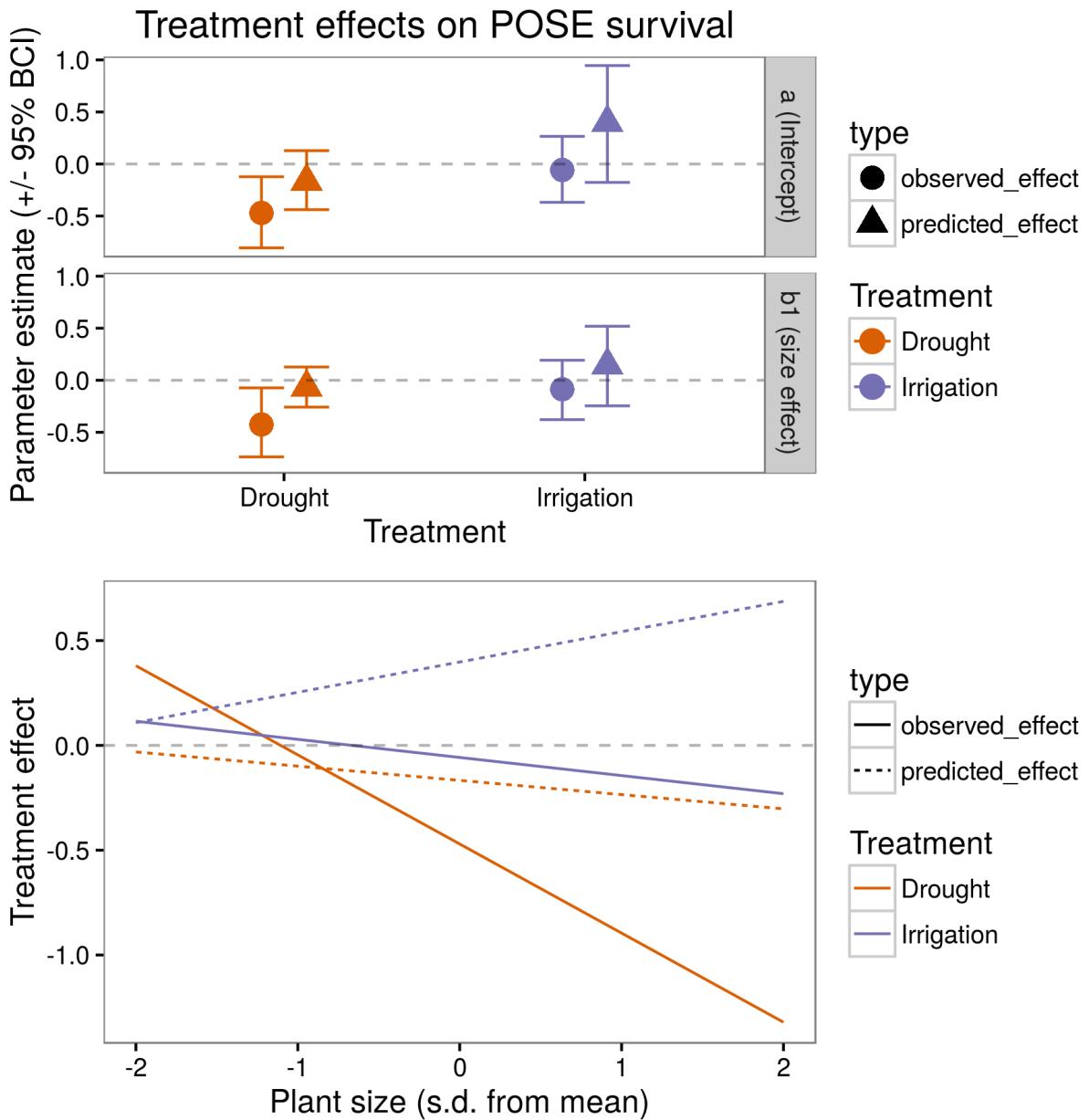


Figure SI-7: Comparison of treatment effects predicted and observed for *P. secunda* survival. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to 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.

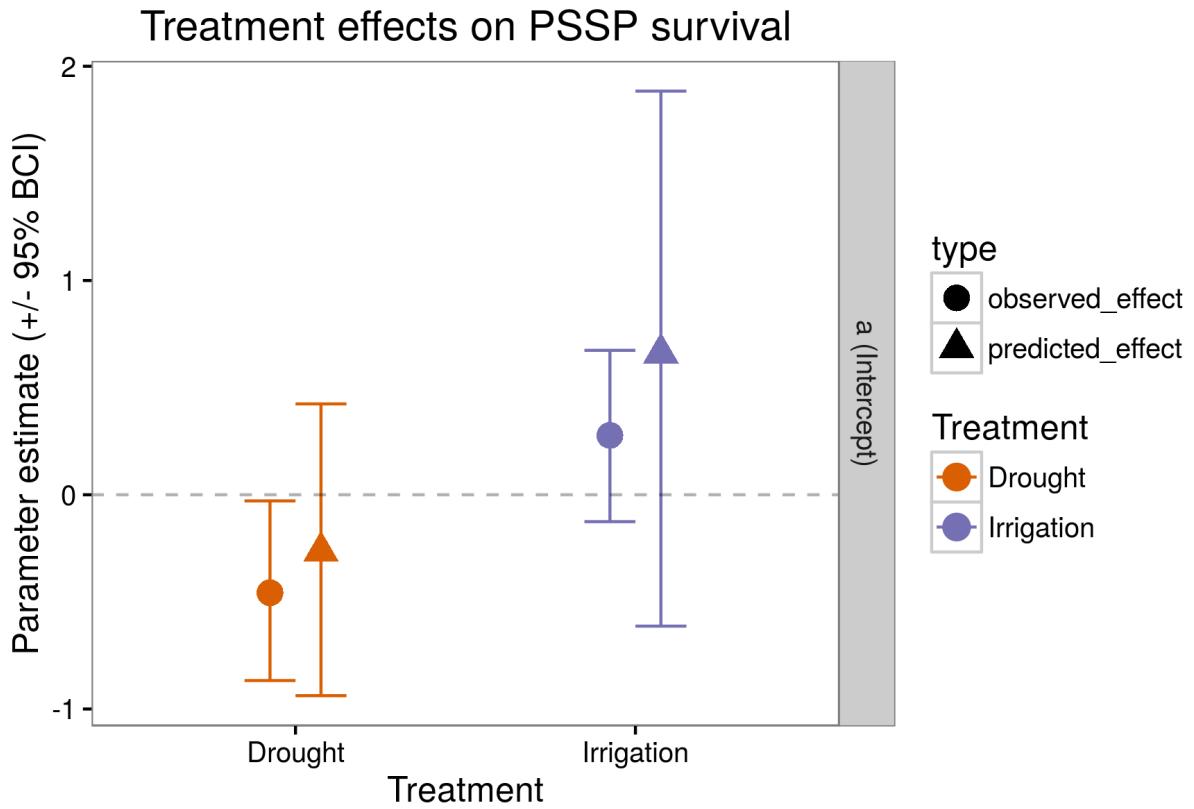


Figure SI-8: Comparison of treatment effects predicted and observed for *P. spicata* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational 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.