

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

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

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

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

Last compile: December 22, 2016

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

**Running title:** Predicting climate response

**Article type:** Letter

**Authorship statement:** ARK and PBA designed the experiment and supervised data collection. AT helped with the data analysis. JB and CA helped with the soil moisture model. ARK wrote the first draft of the manuscript which all authors edited.

---

<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> **Abstract**

<sup>2</sup> Climate is an important driver of population dynamics and annual variation in demographic rates  
<sup>3</sup> often correlate with variation in weather. However, the predictive potential of such correlations is  
<sup>4</sup> largely unknown. We tested how well population models fit using long-term observational could pre-  
<sup>5</sup> dict the responses of four dominant plant species in a sagebrush steppe to a drought and irrigation  
<sup>6</sup> experiment.

<sup>7</sup> We used rainout shelters and automatic sprinklers to manipulate the soil moisture in sixteen  
<sup>8</sup> plots established at the same field site where long-term observational data was collected. We report  
<sup>9</sup> how the growth, survival and recruitment of each species responded to the experimental drought  
<sup>10</sup> and irrigation treatments after five years. We then attempted to predict these treatment responses  
<sup>11</sup> using two models fit to the observational data collected prior to the experiment: a no climate  
<sup>12</sup> model that only included the effects of plant size and competition on plant performance in each  
<sup>13</sup> plot in each year, and a climate model that also included the effects of three seasonal soil moisture  
<sup>14</sup> variables. We compared predictions made by the no climate and climate models to the actual  
<sup>15</sup> experimental responses. We also generated one-step-ahead predictions of population size in each  
<sup>16</sup> experimental plot using individual based population parameterized with the vital rate estimates  
<sup>17</sup> from the climate and no climate models.

<sup>18</sup> Over the course of the experiment, average cover of the perennial grasses, *Hesperostipa*  
<sup>19</sup> *comata* and *Pseudoroegneria spicata* declined significantly in the drought treatment. At the level  
<sup>20</sup> of individual vital rates, experimental drought reduced the survival of *Hesperostipa comata* and  
<sup>21</sup> *P. spicata* and the growth of the grass *Poa secunda*. In contrast, drought increased the growth of  
<sup>22</sup> the shrub *A. tripartita*. The climate model made better predictions of the experimental responses  
<sup>23</sup> than the no climate model in six out of twelve cases. Across all species and vital rates, there was  
<sup>24</sup> a strong positive correlation between the observed responses to the treatments and the responses  
<sup>25</sup> predicted by the climate model. At the population-level, the climate model predicted changes in  
<sup>26</sup> species cover more accurately than the no climate model for *P. secunda* and *P. spicata*.

<sup>27</sup> Observational climate data held valuable information for predicting species' responses to  
<sup>28</sup> a climate change experiment in this ecosystem. Treatment responses often matched the direction  
<sup>29</sup> of predicted responses even when the effects were not significant. We were better able to predict

30 species' responses to the drought treatment than to the control and irrigation treatments, suggesting  
31 that soil moisture is an important factor for predicting the population dynamics of these species  
32 but only when water is truly limiting.

### 33 Introduction

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

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

57 The use of observational data to predict the effects of climate variation on populations  
58 has its own advantages and disadvantages. One important advantage is cost: analyses of already

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

71 The extrapolation of climate-demography correlations presents another potential problem.  
72 In many systems, future precipitation and temperature will fall outside the range of historical  
73 variation. If species performance responds non-linearly to these drivers, fitting linear models for  
74 species responses to climate may produce larger errors when future conditions are outside the range  
75 of observed variation (Doak and Morris, 2010). In addition, climate change will not only alter  
76 average weather, but is also likely to increase the variance in precipitation and the frequency of  
77 extreme events, which will have their own consequences independent of changes in means (Gherardi  
78 and Sala, 2015; Jentsch et al., 2007). Any models based on observations drawn from the historical  
79 range of variation will therefore be extrapolating beyond both the range of observed averages and  
80 variances when used to predict the future (Williams and Jackson, 2007).

81 Here, we combine the strengths of experimental and observational approaches by testing  
82 the ability of models fit to historical data to predict the effects of experimental climate manipu-  
83 lations that generate extreme conditions. A demonstration that the observational approach can  
84 skillfully predict experimental responses would provide strong confirmation that observed climate-  
85 demography correlations are not spurious and will hold even in novel conditions in the future Adler  
86 et al. (2013). (Adler et al., 2013) showed that population models based on observed correlations  
87 between plant population growth rates and precipitation did have some predictive power in de-  
88 scribing species response to a short-term climate manipulation. Three species showed responses

89 to experimentally imposed drought and irrigation that were well predicted by population models  
90 fitted to historical observations. However, the responses of another three species, were not well  
91 predicted by historical observations.

92 Among plant populations, interannual variation in precipitation and or soil moisture often  
93 drives variation in net primary productivity (Knapp and Smith, 2001; Hsu and Adler, 2014), the  
94 annual growth rates of the woody tissue in trees and shrubs ((Yang et al., 2014),(Srur and Villalba,  
95 2009; Franklin, 2013)), and the germination and reproductive output of annuals (Venable, 2007).  
96 Despite clear signs that precipitation shold be important for plant populations, there have been  
97 relatively few studies that clearly link observed variation in precipitation to species performance in  
98 population models (Ehrln et al., 2016).

99 The sagebrush steppe plant community at the US Sheep Experiment Station near Dubois,  
100 Idaho offers an ideal opportunity to test whether the climate effects in plant populations models  
101 derived from observational data can also be used to predict species responses to controlled precip-  
102 itation experiments. The demography of three perennial bunchgrasses and a shrub species at the  
103 USSES have been described in detail in seven different studies since 2010, several of which report  
104 significant effects of seasonal precipitation on the vital rates and overall population growth rates of  
105 these species ((Adler et al., 2010, 2012, 2009, 2016; Chu et al., 2016; Chu and Adler, 2015; Dalgleish  
106 et al., 2010)).

107 In this study, we report how the four dominant plant species at the USSES respond to a  
108 five year drought and irrigation experiment and use the results to address two research questions:  
109 first, how much do the growth, recruitment and survival of our target species differ between the  
110 precipitation manipulation treatments? Significant experimental effects on species vital rates imply  
111 that future changes in precipitation will impact populations. Second, can we predict each species'  
112 response to the experimental conditions based on how they respond to natural climate variation in  
113 the observational data? If models based on observational data can predict the response of species  
114 to this experiment we will gain confidence in using long-term population monitoring data to predict  
115 species responses to future climate change.

<sup>116</sup> **Methods**

<sup>117</sup> **Study site and data set description**

<sup>118</sup> The U.S. Sheep Experiment Station (USSES) is located at Dubois, Idaho (44.2°N, 112.1°W), 1500  
<sup>119</sup> m above sea level. During the period of data collection (1926–2016), mean annual precipitation  
<sup>120</sup> was 270 mm and mean temperatures ranged from -8°C (January) to 21°C (July). The vegetation  
<sup>121</sup> is dominated by a shrub, *Artemisia tripartita*, and three perennial C3 grasses: *Pseudoroegneria*  
<sup>122</sup> *spicata*, *Hesperostipa comata*, and *Poa secunda*. These dominant species account for over 70% of  
<sup>123</sup> basal cover and 60% of canopy cover at this site.

<sup>124</sup> Scientists at the USSES established 26 1-m<sup>2</sup> quadrats between 1926 and 1932. Eighteen  
<sup>125</sup> quadrats were distributed among four ungrazed exclosures, and eight were distributed in two pad-  
<sup>126</sup> docks grazed at medium intensity spring through fall. All quadrats were located on similar topog-  
<sup>127</sup> raphy and soils. In most years until 1957, all individual plants in each quadrat were mapped using  
<sup>128</sup> a pantograph (Blaisdell, 1958). The historical data set is public and available online (Zachmann  
<sup>129</sup> et al., 2010). In 2007, we located 14 of the original quadrats, all of which are inside permanent live-  
<sup>130</sup> stock exclosures, and resumed annual mapped censusing using the traditional pantograph method.  
<sup>131</sup> Daily temperature and precipitation has been monitored throughout this period at a climate sta-  
<sup>132</sup> tion located at the USSES headquarters (station id: GHCND:USC00102707) which located within  
<sup>133</sup> 2 km of the research plots. We downloaded daily and monthly tmin, tmax, and precipitation data  
<sup>134</sup> from the National Climate Data Centers online database.

<sup>135</sup> We extracted data on survival, growth, and recruitment from the mapped quadrats based  
<sup>136</sup> on plants' spatial locations. Our approach tracks genets representing individual plants. For the  
<sup>137</sup> shrub, each genet is associated with the basal position of a stem. For the bunchgrasses, each genet  
<sup>138</sup> represents a spatially distinct polygon in the mapped quadrat. These genets may fragment and/or  
<sup>139</sup> coalesce over time. Each mapped polygon is classified as a surviving genet or a new recruit based  
<sup>140</sup> on its spatial location relative to genets present in previous years (Lauenroth and Adler, 2008). We  
<sup>141</sup> modeled vital rates using data from 21 year-to-year transitions between 1929 and 1957, and four  
<sup>142</sup> year-to-year transitions from 2007 to 2011.

<sup>143</sup> **Precipitation experiment**

<sup>144</sup> In spring 2011, we selected locations for an additional 16 quadrats for the precipitation experiment.  
<sup>145</sup> We located these in a large enclosure containing six of the historical permanent quadrats. We  
<sup>146</sup> avoided plots falling on hill slopes, areas with greater than 20% bare rock, or with over 10% cover  
<sup>147</sup> of the woody shrubs *Purshia tridentata* or *Amelanchier utahensis*. New plots were established  
<sup>148</sup> in pairs, and one plot per pair was randomly assigned to either the precipitation reduction or  
<sup>149</sup> the precipitation addition treatment. We mapped the quadrats in June, 2011 and then built the  
<sup>150</sup> rainfall shelters and set-up the irrigation systems in the fall of 2011. We used a rain-out shelter and  
<sup>151</sup> automatic irrigation design described in (Gherardi and Sala, 2013). Each rain-out shelter covered  
<sup>152</sup> an area of 2.5 by 2 m and consisted of transparent acrylic shingles held up 1.5 to 1 m over the plot  
<sup>153</sup> to channel 50% of incoming rainfall off of the plot and into 75 l reservoirs. The collected water  
<sup>154</sup> was pumped out of reservoirs and sprayed onto paired irrigation treatment plots. Pumping was  
<sup>155</sup> initiated automatically with float switches that were triggered when water levels in the reservoirs  
<sup>156</sup> were approximately 20 l, or equivalently irrigation was triggered once for every 6 mm of rainfall  
<sup>157</sup> collected. We disconnected the irrigation pumps in late fall each year and re-connected them in  
<sup>158</sup> April. The drought shelters remained in place throughout the year.

<sup>159</sup> We monitored soil moisture and air temperature in four of the precipitation experiment plot  
<sup>160</sup> pairs using Decagon Devices (Pulman Washington) 5TM and EC-5 soil moisture sensors and 5TE  
<sup>161</sup> temperature sensors. We installed two soil moisture sensors in each monitored plot at 5 cm and  
<sup>162</sup> two at 25 cm deep in the soil. Air temperature was measured underneath the roofing of the shelter  
<sup>163</sup> at 30 cm above ground. For each pair of manipulated plots, we also installed sensors in a nearby  
<sup>164</sup> area to measure ambient rainfall and temperature. Data were logged automatically every four  
<sup>165</sup> hours. We augmented automatic monitoring of the climate in these plots with direct measurements  
<sup>166</sup> of soil moisture with a handheld EC-5 soil moisture sensor at six points around all 16 plots on  
<sup>167</sup> 6/6/2012, 4/29/2015, 5/7/2015, 6/9/2015 and 5/10/2016. We analyzed these spot measurements  
<sup>168</sup> for significant treatment effects on soil moisture using a linear mixed effects model with the *lmer*  
<sup>169</sup> package in *R*, with plot, plot group, and date as random effects in the model ((Bates et al., 2015)).

<sup>170</sup> We conducted a simple statistical to determine the net effect of the experimental treatments  
<sup>171</sup> on cover in the experiment. First we calculated the log change in cover for each of the four focal

172 species in each quadrat from from the start of the experiment in spring prior to manipulation, to  
173 the last year of the experiment. Log change in cover was defined as ,  $\log(Cover_{2016}/Cover_{2011})$   
174 where  $Cover_{2016}$  is the cover of each species in 2011 and  $Cover_{2015}$  is cover in 2011. We tested for  
175 the effect of precipitation treatment on this measure with a linear model in *R*.

176 **Soil moisture modeling**

177 We expected that our precipitation manipulation experiment would affect plants by altering avail-  
178 able soil moisture during the growing season. Because we do not have direct soil moisture measures  
179 for each year of observed plant cover in the historical record, we used the SOILWAT soil moisture  
180 model to estimate daily soil moisture at the USSES from 1925 to the present (Sala et al., 1992). We  
181 used an enhanced version of SOILWAT that has recently been developed for use in semi-arid shrub-  
182 land ecosystems (Bradford et al., 2014). SOILWAT uses daily weather data, ecosystem specific  
183 vegetation properties and site specific soil properties to estimate water balance processes. SOIL-  
184 WAT specifically estimates rainfall interception by vegetation, evaporation of intercepted water,  
185 snow melt and snow redistribution, infiltration into the soil, percolation through the soil, bare-soil  
186 evaporation, transpiration from each soil layer, and drainage. We parameterized SOILWAT with  
187 the generic sagebrush steppe vegetation parameters and site specific soil texture and bulk density  
188 data. We used daily weather data collected at the USSES from 1925 until the present as weather  
189 forcing data for the SOILWAT predictions.

190 We averaged daily soil moisture predictions from SOILWAT from upper 40 cm of soil and  
191 then averaged these seasonally to serve as the covariates in the vital rate regressions for each species.  
192 Because we did not monitor soil moisture directly in all control, drought and irrigation plots, we  
193 used a model to describe the average treatment effects on soil moisture during the course of the  
194 experiment. To do this we first averaged observed soil moisture data by date and plot and then  
195 standardized these by the mean and standard deviation of the control soil moisture conditions  
196 observed within each plot group. We then found the difference between the soil moisture in the  
197 treated plots and the ambient conditions. We then modeled these treatment effects as a function of  
198 season and whether a day was rainy or dry. We expected that our drought and irrigation treatments  
199 might be more effective during rainy weather than during dry weather. Rainy days were defined  
200 as any days when any precipitation was recorded and average temperatures were above 3 degrees

201 C. The day immediately following rainfall was also classified as rainy. We fit this model using  
202 the *lmer* package in *R* with random effects for plot group and date (Bates et al., 2015). We then  
203 used this model to predict the treatment effects on soil moisture for the entire study period from  
204 the ambient soil moisture values predicted from the SOILWAT model described above. These  
205 adjusted soil moisture values reflected the average season and rainfall dependent effects of the  
206 experimental treatments on soil moisture and could be used as covariates for predicting the effects  
207 of our manipulation on each species demographic rates.

208 **Overview of the analysis**

209 Our analysis consists of two separate datasets and three different categories of vital rate models.  
210 We refer to the first dataset as the observational data. It consists of all the historical data collected  
211 from 1925 to 1957 as well as the contemporary data collected from the same plots from 2007 to  
212 2010. These data record the response of plants in each plot to the ambient climate variation. We  
213 refer to the second dataset as the experimental data. It consists of the data collected from 2011 to  
214 2016 from the 16 new experimentally manipulated plots, as well as from 14 of the original historical  
215 plots which serve as ambient climate controls.

216 In order describe the effects of the experimental treatments on each vital rate, we fit "treat-  
217 ment" models. The treatment models included parameters representing the effects of the drought  
218 and irrigation treatments on each vital rate. We fit these models to using all the experimental and  
219 all the observational data. We combined the datasets because we wanted to focus our predictions  
220 on the effects of the experimental treatments on the vital rates, rather than any differences between  
221 the historical and contemporary periods in the effects of crowding and plant size on the vital rates.

222 Next, in order to test how well we could predict the responses in the experimental plots,  
223 we fit two classes of models to the observational dataset only. The "no climate" models include  
224 parameters for the effects of competition on each vital rate and the size dependence of survival and  
225 growth but they do not include climate or treatment effects. The "climate" models include the  
226 effects of annual variation in soil moisture on each vital rate. The no climate model provides us a  
227 baseline by which to measure the accuracy of predictions from the climate model. Because much of  
228 the variation in growth, survival and recruitment in this system can be explained by plant size and  
229 competition, we expect that these two models will make similar predictions for individual plant

230 performance in the experiment. However, if the climate model makes more accurate predictions  
 231 than the no climate model, this indicates that the climate parameters it includes contain useful  
 232 information for prediction. Note that because these models are fit using only the observation  
 233 dataset, when we use these models to predict experimental responses we are generating true out-  
 234 of-sample predictions.

235 **Statistical models of vital rates**

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

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

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

245 where  $\varphi$  is the spatial group dependent intercept,  $\gamma$  is a year-effect,  $\beta$  is year-dependent coefficient  
 246 that represents the effect of log genet size,  $u$ , on survival in year  $t$ .  $\boldsymbol{\omega}$  is a vector of interaction  
 247 coefficients which determine the impact of crowding,  $\mathbf{W}$ , by each species on the focal species. The  
 248 vector  $\mathbf{W}$  includes crowding from the four dominant species, *A. tripartita*, *P. spicata*, *H. comata*,  
 249 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  
 250 R. This model is the no climate model for survival.

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

256 In the climate model, the above term is replaced with  $\mathbf{C}\xi_j^S$ , where  $\xi$  gives a vector of  
 257 coefficients describing the effects of a set of soil moisture covariates  $\mathbf{C}$  in treatment  $h$  and year  $t$

258 on the survival rate of species  $j$ .  $\mathbf{C}$  can include interaction effects between plant size,  $u$ , and the  
 259 soil moisture covariates allowing the effects of soil moisture to vary with plant size.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

300 Because plants outside the mapped quadrat could contribute recruits to the focal quadrat or  
 301 interact with plants in the focal quadrat, we estimated effective cover as a mixture of the observed  
 302 cover,  $C$ , in the focal quadrat,  $q$ , and the mean cover,  $\bar{C}$ , across the spatial location,  $g$ , in which

303 the quadrat is located:  $C'_{jq,t} = p_j C_{jq,t} + (1 - p_j) \bar{C}_{jq,t}$ , where  $p$  is a mixing fraction between 0 and  
304 1 that was estimated as part of fitting the model.

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

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

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

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

## 324 Selecting soil moisture covariates

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

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

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

### 342 Predicting cover from individual-based models

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

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

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

360 and were not interested in the effects of demographic stochasticity, we used a deterministic version  
361 of the models (e.g., recruitment is the  $\lambda$  of (6), rather than a random draw from a negative binomial  
362 distribution with a mean of  $\lambda$ ).

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

### 365 Quantifying predictive accuracy

366 We assessed the predictive performance of the climate and no climate by calculating the mean  
367 square error (MSE) between the predicted and observed responses in the experimental data as,

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

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

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

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

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

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

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

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

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

## 394 Results

### 395 Effects on soil moisture

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

400 The continuously recorded soil moisture data also showed treatment effects, but these were  
401 weaker on average than the spot measurements and depended on season and recent rainfall (table  
402 2; fig 2). We saw weaker effects during the spring than during the fall and summer: the drought  
403 plots were about 20-30% drier than ambient in the fall and summer but only 7 to 14% drier during  
404 the spring, while the irrigated plots were 30% wetter during the fall and summer but only 20-25%  
405 wetter during the spring. Treatment differences were slightly larger during rainy periods, especially  
406 in the spring (rainfall effect in table 2). We did not find evidence that the drought shelters and the  
407 irrigation treatments consistently affected air temperature at 30 cm above the plots.

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

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

#### 422        Effects on cover and vital rates

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

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

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

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

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

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

## 455 Effects of soil moisture on vital rates

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

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

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

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

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

## 480 Evaluating the predictions

481 For most models adding climate covariates did not improve our ability to predict species responses  
482 in the experiment (table 8). However, the climate models did improve overall prediction MSE for  
483 growth of *A. tripartita* and growth and survival of *P. secunda* (table 8). In terms of lppd, the  
484 climate model outperformed the no climate model in six out of twelve models: for *A. tripartita*  
485 growth, *H. comata* recruitment, *P. secunda* growth and survival and *P. spicata* recruitment (table  
486 8).

487 When we look at the predictions for each treatment separately we see that climate covariates  
488 improved model predictions more often in the drought treatments than in the control or irrigation  
489 treatments (table A-1). For all four species, the climate model outperformed the no climate model  
490 for predicting the response of growth to drought in terms of lppd (table A-1). The climate model  
491 also outperformed the no climate model for predicting irrigation effects on growth for all species  
492 except *H. comata*.

493 Overall our climate models often predicted the correct direction of the drought and irriga-  
494 tion treatments (fig 11). In four cases we both observed and predicted treatment effects significantly  
495 different from zero based on the 95% Bayesian credible interval around the parameter mean: the  
496 drought response of *H. comata* survival (fig A-6), the drought response of *P. secunda* growth (fig  
497 A-3), the irrigation response of *P. spicata* recruitment (fig A-12) and the irrigation response of *P.*  
498 *secunda* recruitment (fig A-11). In only one of these cases, for *P. secunda* recruitment, was the  
499 predicted effect in the opposite direction from the observed treatment effect (fig 11). The overall  
500 correlation between the predicted and observed treatment effects for all treatments, species and  
501 vital rates was  $r = 0.54$ , whereas the correlation for the drought treatment effects,  $r = 0.77$ , was  
502 better than for the irrigation effects,  $r = 0.46$ . Also the correlation between the size by climate and  
503 size by treatment effects for *A. tripartita* growth and *P. secunda* growth and survival was much  
504 stronger than the correlation between the intercept parameter estimates (fig 11).

505 Using the vital rate models for each species we generated one step ahead cover predictions  
506 for each quadrat in each year of the experiment. Average cover predicted by the climate model  
507 tended to be lower than the observed cover each year for *A. tripartita* and *P. secunda* (fig 12).  
508 Comparing the overall population growth rates predicted to those observed in the experiment,  
509 we see that the MSE of the climate model was lower than the MSE of the no climate model for  
510 *P. secunda* and *P. spicata* (table 9). The predictions produced by the climate model for these  
511 species were also slightly more correlated with the observations than the predictions produced by  
512 the no climate model (table 9). Considering each treatment and species separately, the predicted  
513 population growth rates for *A. tripartita*, *P. secunda* and *P. spicata* were all consistently lower  
514 than the observed population growth rates (figs 13, 15, 16). The climate model showed lower MSE  
515 for *A. tripartita*, *P. secunda* and *P. spicata* in the irrigation treatment, *P. spicata* in the control  
516 treatment and *H. comata* in the drought treatment (fig 14). However, the correlations between the  
517 predicted and observed log changes in cover did not always show the same pattern as MSE: the  
518 climate model made more strongly correlated predictions with the observations than the no climate  
519 model only for *P. spicata* and *P. secunda* in the control treatment and *P. secunda* and *H. comata*  
520 in the drought treatment.

521 **Discussion**

522 Our experiment showed that observational data on the response of plant populations to interannual  
523 climate variation can indeed help us predict the direction of species responses to experimental  
524 climate manipulations (fig 11). The historical climate-demography correlations helped predict the  
525 direction of experimental responses even though adding climate parameters to the demographic  
526 models only improved vital rate predictions for half of the models (fig 8). This should give us some  
527 hope that even when climate effects in demographic models fit to observational data are weak or  
528 not significant, they may contain useful qualitative information on the direction of climate effects  
529 in the future.

530 **Comparison of experimental and natural climate effects**

531 While previous studies in this system used the observational data to describe the effects of climate  
532 on demography and survival, this is the first study to demonstrate effects of climate experimentally.  
533 While previous studies of this system fit a variety of climate effects and used different modeling  
534 approaches, we see many points of similarity between these studies in the responses of the four  
535 dominant species to precipitation ((Chu et al., 2016),(Dalgleish et al., 2010), (Adler et al., 2012)).  
536 First, in all three studies the strongest positive effects of precipitation among the four species are  
537 reported for *H. comata*; this matches the negative effects of our drought experiment on this species  
538 (fig 4). This effect is driven by a negative growth and survival response to drought (fig 5, 6).  
539 On the other hand, if we had only conducted an irrigation experiment our results may not have  
540 shown this consistency with previous work as *H. comata* showed no positive response to irrigation.  
541 Previous studies also reported positive effects of precipitation on the other grasses, *P. secunda* and  
542 *P. spicata*. Again our results are consistent with this result: drought led to declines in cover of  
543 *P. spicata*, 4, and in the growth and survival of *P. secunda* (figs 5, 6). As for the *H. comata* the  
544 magnitudes of drought effects were greater than the irrigation effects on these grasses.

545 The effects of precipitation on *A. tripartita* are more complicated. Previous research re-  
546 ported negative direct effects of precipitation on this species ((Adler et al., 2012), (Chu et al.,  
547 2016)). This effect has always seemed odd because it is hard to imagine why precipitation would  
548 have a direct negative effect on a species in this dry ecosystem. But again the largely positive

549 (but size dependent) effects of drought treatments on *A. tripartita* growth should give us more  
550 confidence in the negative effects of precipitation shown in the historical data. These studies also  
551 report relative strong indirect effects of precipitation on *A. tripartita* mediated by its competition  
552 with grasses. It is possible that some of the positive effect on *A. tripartita* growth in our drought  
553 plots is the result of reduced grass cover (fig 12; (Chu et al., 2016)). However, our growth model  
554 includes interspecific crowding and so should take into account any changes in grass abundance  
555 that could be driving a positive response from *A. tripartita*. This leaves us with the question of  
556 why this species would show a positive direct response to drought. Although there is some evidence  
557 that saturated soils in the spring are detrimental for big sagebrush (*A. tridentata*), a closely related  
558 species (Sturges, 1989; Germino and Reinhardt, 2014), soil saturation would conservatively seem to  
559 require soils to be above 30 or 40% volumetric water content for several weeks, something that we  
560 did not observe (fig 2). Another possible explanation is that our drought treatments reduced snow  
561 cover in the winter and early spring, an effect that has been shown to benefit related sagebrush  
562 species in other ecosystems (Perfors et al., 2003).

563 Overall we were somewhat surprised by the weak effects that reducing water availability  
564 by 50% and increasing water availability by 150% had in this arid system. Cross-biome studies  
565 of the relationship between precipitation and ANPP generally show that arid systems are highly  
566 sensitive to water limitation (Huxman et al., 2004). We have two explanations for the seemingly  
567 weak effects of precipitation we observed on demography. First, we measure the size of the pernnial  
568 bunchgrasses in this system by their basal cover, which may not have a strong relationship with  
569 their annual production. It is likely that we would find larger effects of precipitation on these  
570 grasses if we had a more complete measurement of aboveground biomass. Moreover, much of the  
571 growth of these species may be going into roots.

572 Another explanation for the weak effects of precipitation are that perennial species in this  
573 cold desert ecosystems are well adapted to tolerate drought, either through escaping drought by  
574 growing early in the year, or by avoiding drought stress later in the year through high water use  
575 efficiency (Bazzaz, 1979; Franks, 2011). Indeed our soil moisture data generally show a pulse of  
576 soil moisture in the spring when many grasses are actively growing (fig 2). Likewise, *A. tripartita*  
577 is more deeply rooted than the grasses and able to continue its growth throughout the growing  
578 season by drawing from deeper soil water (Germino and Reinhardt, 2014). The adaptations of

579 native perennial plants in cold deserts could make them less sensitive to water availability than  
580 species in a more mesic ecosystem.

581 **Can the past predict the future?**

582 Our second research question was whether we could use long term observational data on species  
583 response to precipitation to predict the response of each species to the experiment. Using the IBM,  
584 we to generate predicted changes in population size for each year, we found that climate model  
585 predictions were indeed better than the no climate models for two species: *P. spicata* and *P. secunda*  
586 (table 9). In the drought treatment our one step ahead cover predictions for *H. comata* and *P.*  
587 *secunda* were also better than the no climate model. Moreover, we also found that climate models  
588 produced better predictions of species vital rates for half of the species/vital rate combinations we  
589 tested (table 8). The rate of overall improvement in predictive ability produced by the climate  
590 models over the no climate models was similar to the results reported by (Adler et al., 2013) who  
591 also reported improved population-level predictions for half of the species predicted. Likewise, in  
592 a within sample cross-validation analysis, (Tredennick et al., 2016) found that including climate  
593 covariates improved population level predictions for two out of four species in a mixed grass prairie  
594 in Montana.

595 We also compared the treatment parameters from the treatment model fit to the experimen-  
596 tal data to the treatment parameters predicted by the climate model fit only to the observational  
597 data 11. Among all the climate effects we predicted and observed, there were only four cases where  
598 vital rate predictions and observations were both significantly different from zero (fig 11) and three  
599 of these cases we successfully predicted the direction of the treatment effects. However, for *P. se-*  
600 *cunda* recruitment we predicted a positive response of irrigation, but observed a negative response  
601 (figs 11). From a statistical standpoint this is our arguably our greatest error in prediction. How-  
602 ever, recruitment decreased in both the drought and irrigation plots for *P. secunda* and also for *P.*  
603 *spicata* (fig 7). So its likely the decrease in *P. secunda* recruitment in the irrigated plots was due  
604 to underlying differences in the set of experimental plots from the historical control plots rather  
605 than the precipitation treatments.

606 The drought effects we observed on the three grasses were often stronger than the effects  
607 we predicted, while the irrigation effects observed were often weaker than predicted (figs A-2, A-6,

608 A-3, A-7, A-4, A-8). In this water limited system, we expected that experimental irrigation would  
609 lead to increases in plant performance, but we saw few cases where irrigation benefited any of the  
610 plants. A pattern qualitatively similar to this shows up in both natural and experimental data  
611 comparing precipitation to ANPP: decreases in grassland ANPP induced by drought are often of  
612 greater magnitude than increases in ANPP induced by experimental irrigation or by above average  
613 precipitation ((Hsu and Adler, 2014; Gherardi and Sala, 2015)). If we had fit our growth and  
614 survival models with a non-linear function for soil moisture, perhaps informed by more mechanistic  
615 understanding of water limitation on the physiology of these plants, we may have made more  
616 accurate predictions of the drought and irrigation effects (Ehrln et al., 2016).

## 617 Conclusion

618 Our results give us more confidence that observational data can be used to detect and predict the  
619 effects of annual soil moisture variation on sagebrush steppe plants. This should encourage more  
620 researchers to try and use observational data to predict population response to climate in both  
621 experimental and natural settings (Houlahan et al., 2016; Ehrln et al., 2016). Nevertheless, our  
622 success at predicting the short-term response of two out of four species to a simple precipitation  
623 manipulation is not likely to impress applied ecologists and policymakers who often need accurate  
624 predictions for the effects of climate change in large complex systems. Clearly more work is needed  
625 to learn how to accurately predict the ecological responses of species to climate change. Towards  
626 that goal, perhaps the best way forward is to conduct more tests like this one.

## 627 Acknowledgements

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

633 **References**

- 634 Adler, P. B., K. M. Byrne, and J. Leiker. 2013. Can the past predict the future? Experimental  
635 tests of historically based population models. *Global Change Biology*, **19**:1793–1803.
- 636 Adler, P. B., H. J. Dagleish, and S. P. Ellner. 2012. Forecasting plant community impacts of  
637 climate variability and change: when do competitive interactions matter? *Journal of Ecology*,  
638 **100**:478–487.
- 639 Adler, P. B., S. P. Ellner, and J. M. Levine. 2010. Coexistence of perennial plants: an embarrassment  
640 of niches. *Ecology Letters*, **13**:1019–1029.
- 641 Adler, P. B., J. HilleRisLambers, and J. M. Levine. 2009. Weak effect of climate variability on  
642 coexistence in a sagebrush steppe community. *Ecology*, **90**:3303–3312.
- 643 Adler, P. B., A. R. Kleinhesselink, G. Hooker, B. Taylor, B. J. Teller, and S. P. Ellner. 2016. Weak  
644 interspecific interactions in a sagebrush steppe: Evidence from observational data, population  
645 models, and a removal experiment.
- 646 Aubry, L. M., R. F. Rockwell, E. G. Cooch, R. W. Brook, C. P. H. Mulder, and D. N. Koons.  
647 2013. Climate change, phenology, and habitat degradation: drivers of gosling body condition  
648 and juvenile survival in lesser snow geese. *Global Change Biology*, **19**:149–160.
- 649 Bates, D., M. Mchler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using  
650 lme4. *Journal of Statistical Software*, **67**:1–48.
- 651 Bazzaz, F. A. 1979. The Physiological Ecology of Plant Succession. *Annual Review of Ecology and*  
652 *Systematics*, **10**:351–371.
- 653 Blaisdell, J. P. 1958. Seasonal development and yield of native plants on the upper Snake River  
654 Plains and their relation to certain climatic factors. U.S. Dept. of Agriculture.
- 655 Bradford, J. B., D. R. Schlaepfer, and W. K. Lauenroth. 2014. Ecohydrology of Adjacent Sage-  
656 brush and Lodgepole Pine Ecosystems: The Consequences of Climate Change and Disturbance.  
657 *Ecosystems*, **17**:590–605.

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

- 683 Gelman, A., J. Hwang, and A. Vehtari. 2014. Understanding predictive information criteria for  
684 Bayesian models. *Statistics and Computing*, **24**:997–1016.
- 685 Gerber, B. D., W. L. Kendall, M. B. Hooten, J. A. Dubovsky, and R. C. Drewien. 2015. Optimal  
686 population prediction of sandhill crane recruitment based on climate-mediated habitat limita-  
687 tions. *Journal of Animal Ecology*, **84**:1299–1310.
- 688 Germino, M. J. and K. Reinhardt. 2014. Desert shrub responses to experimental modification of  
689 precipitation seasonality and soil depth: relationship to the two-layer hypothesis and ecohydro-  
690 logical niche. *Journal of Ecology*, **102**:989–997.
- 691 Gherardi, L. A. and O. E. Sala. 2013. Automated rainfall manipulation system: a reliable and  
692 inexpensive tool for ecologists. *Ecosphere*, **4**:1–10.
- 693 Gherardi, L. A. and O. E. Sala. 2015. Enhanced precipitation variability decreases grass- and  
694 increases shrub-productivity. *Proceedings of the National Academy of Sciences*, **112**:12735–  
695 12740.
- 696 Hilborn, R. 2016. Correlation and Causation in Fisheries and Watershed Management. *Fisheries*,  
697 **41**:18–25.
- 698 Houlahan, J. E., S. T. McKinney, T. M. Anderson, and B. J. McGill. 2016. The priority of prediction  
699 in ecological understanding. *Oikos*, pages n/a–n/a.
- 700 Hsu, J. S. and P. B. Adler. 2014. Anticipating changes in variability of grassland production due  
701 to increases in interannual precipitation variability. *Ecosphere*, **5**:1–15.
- 702 Huxman, T. E., M. D. Smith, P. A. Fay, A. K. Knapp, M. R. Shaw, M. E. Loik, S. D. Smith, D. T.  
703 Tissue, J. C. Zak, J. F. Weltzin, W. T. Pockman, O. E. Sala, B. M. Haddad, J. Harte, G. W.  
704 Koch, S. Schwinning, E. E. Small, and D. G. Williams. 2004. Convergence across biomes to a  
705 common rain-use efficiency. *Nature*, **429**:651–654.
- 706 Jenouvrier, S., H. Caswell, C. Barbraud, M. Holland, J. Strve, and H. Weimerskirch. 2009. De-  
707 mographic models and IPCC climate projections predict the decline of an emperor penguin  
708 population. *Proceedings of the National Academy of Sciences*, **106**:1844–1847.

- 709 Jentsch, A., J. Kreyling, and C. Beierkuhnlein. 2007. A new generation of climate-change experi-  
710       ments: events, not trends. *Frontiers in Ecology and the Environment*, **5**:365–374.
- 711 Knapp, A. K., M. L. Avolio, C. Beier, C. J. W. Carroll, S. L. Collins, J. S. Dukes, L. H. Fraser,  
712       R. J. Griffin-Nolan, D. L. Hoover, A. Jentsch, M. E. Loik, R. P. Phillips, A. K. Post, O. E.  
713       Sala, I. J. Slette, L. Yahdjian, and M. D. Smith. 2016. Pushing precipitation to the extremes  
714       in distributed experiments: recommendations for simulating wet and dry years. *Global Change  
715       Biology*.
- 716 Knapp, A. K. and M. D. Smith. 2001. Variation Among Biomes in Temporal Dynamics of Above-  
717       ground Primary Production. *Science*, **291**:481–484.
- 718 Koons, D. N., P. Terletzky, P. B. Adler, M. L. Wolfe, D. Ranglack, F. P. Howe, K. Hersey, W. Pas-  
719       kett, and J. T. d. Toit. 2012. Climate and density-dependent drivers of recruitment in plains  
720       bison. *Journal of Mammalogy*, **93**:475–481.
- 721 Krner, C. 2012. Alpine Treelines: Functional Ecology of the Global High Elevation Tree Limits.  
722       Springer Science & Business Media.
- 723 Lauenroth, W. K. and P. B. Adler. 2008. Demography of perennial grassland plants: survival, life  
724       expectancy and life span. *Journal of Ecology*, **96**:1023–1032.
- 725 Levine, J. M., A. K. McEachern, and C. Cowan. 2010. Do competitors modulate rare plant response  
726       to precipitation change? *Ecology*, **91**:130–140.
- 727 Lunn, N. J., S. Servany, E. V. Regehr, S. J. Converse, E. Richardson, and I. Stirling. 2016.  
728       Demography of an apex predator at the edge of its range impacts of changing sea ice on polar  
729       bears in Hudson Bay. *Ecological Applications*, pages n/a–n/a.
- 730 Parmesan, C. and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across  
731       natural systems. *Nature*, **421**:37–42.
- 732 Perfors, T., J. Harte, and S. E. Alter. 2003. Enhanced growth of sagebrush (*Artemisia tridentata*)  
733       in response to manipulated ecosystem warming. *Global Change Biology*, **9**:736–742.

- 734 Petchey, O. L., M. Pontarp, T. M. Massie, S. Kfi, A. Ozgul, M. Weilenmann, G. M. Palamara, F. Al-  
735 termatt, B. Matthews, J. M. Levine, D. Z. Childs, B. J. McGill, M. E. Schaeppman, B. Schmid,  
736 P. Spaak, A. P. Beckerman, F. Pennekamp, and I. S. Pearse. 2015. The ecological forecast horizon,  
737 and examples of its uses and determinants. *Ecology Letters*, **18**:597–611.
- 738 Post, E. and M. C. Forchhammer. 2002. Synchronization of animal population dynamics by large-  
739 scale climate. *Nature*, **420**:168–171.
- 740 Sala, O. E., W. K. Lauenroth, and W. J. Parton. 1992. Long-Term Soil Water Dynamics in the  
741 Shortgrass Steppe. *Ecology*, **73**:1175–1181.
- 742 Srur, A. M. and R. Villalba. 2009. Annual growth rings of the shrub *Anarthrophyllum rigidum*  
743 across Patagonia: Interannual variations and relationships with climate. *Journal of Arid Envi-  
744 ronments*, **73**:1074–1083.
- 745 Sturges, D. L. 1989. Response of Mountain Big Sagebrush to Induced Snow Accumulation. *Journal*  
746 of Applied Ecology, **26**:1035–1041. ArticleType: research-article / Full publication date: Dec.,  
747 1989 / Copyright 1989 British Ecological Society.
- 748 Teller, B. J., P. B. Adler, C. B. Edwards, G. Hooker, and S. P. Ellner. 2016. Linking demography  
749 with drivers: climate and competition. *Methods in Ecology and Evolution*, **7**:171–183.
- 750 Tredennick, A. T., M. B. Hooten, and P. B. Adler. 2016. Do we need demographic data to forecast  
751 plant population dynamics? *Methods in Ecology and Evolution*, pages n/a–n/a.
- 752 Tylianakis, J. M., R. K. Didham, J. Bascompte, and D. A. Wardle. 2008. Global change and species  
753 interactions in terrestrial ecosystems. *Ecology Letters*, **11**:1351–1363.
- 754 Venable, D. L. 2007. Bet Hedging in a Guild of Desert Annuals. *Ecology*, **88**:1086–1090.
- 755 Walther, G.-R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J.-M. Fromentin,  
756 O. Hoegh-Guldberg, and F. Bairlein. 2002. Ecological responses to recent climate change. *Nature*,  
757 **416**:389–395.

- 758 Wang, J., P. M. Rich, and K. P. Price. 2003. Temporal responses of NDVI to precipitation and  
759 temperature in the central Great Plains, USA. International Journal of Remote Sensing, **24**:2345–  
760 2364.
- 761 Whittaker, R. H. 1975. Communities and ecosystems. Macmillan, New York, second edition edition.
- 762 Williams, J. W. and S. T. Jackson. 2007. Novel climates, no-analog communities, and ecological  
763 surprises. Frontiers in Ecology and the Environment, **5**:475–482.
- 764 Wolkovich, E. M., B. I. Cook, J. M. Allen, T. M. Crimmins, J. L. Betancourt, S. E. Travers, S. Pau,  
765 J. Regetz, T. J. Davies, N. J. B. Kraft, T. R. Ault, K. Bolmgren, S. J. Mazer, G. J. McCabe,  
766 B. J. McGill, C. Parmesan, N. Salamin, M. D. Schwartz, and E. E. Cleland. 2012. Warming  
767 experiments underpredict plant phenological responses to climate change. Nature, **485**:494–497.
- 768 Yang, B., C. Qin, J. Wang, M. He, T. M. Melvin, T. J. Osborn, and K. R. Briffa. 2014. A 3,500-year  
769 tree-ring record of annual precipitation on the northeastern Tibetan Plateau. Proceedings of the  
770 National Academy of Sciences, **111**:2903–2908.
- 771 Zachmann, L., C. Moffet, and P. Adler. 2010. Mapped quadrats in sagebrush steppe: long-term  
772 data for analyzing demographic rates and plantplant interactions. Ecology, **91**:3427.

## Tables

Table 1: Spring soil moisture

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

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

Table 2: soil moisture model

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

## Figures

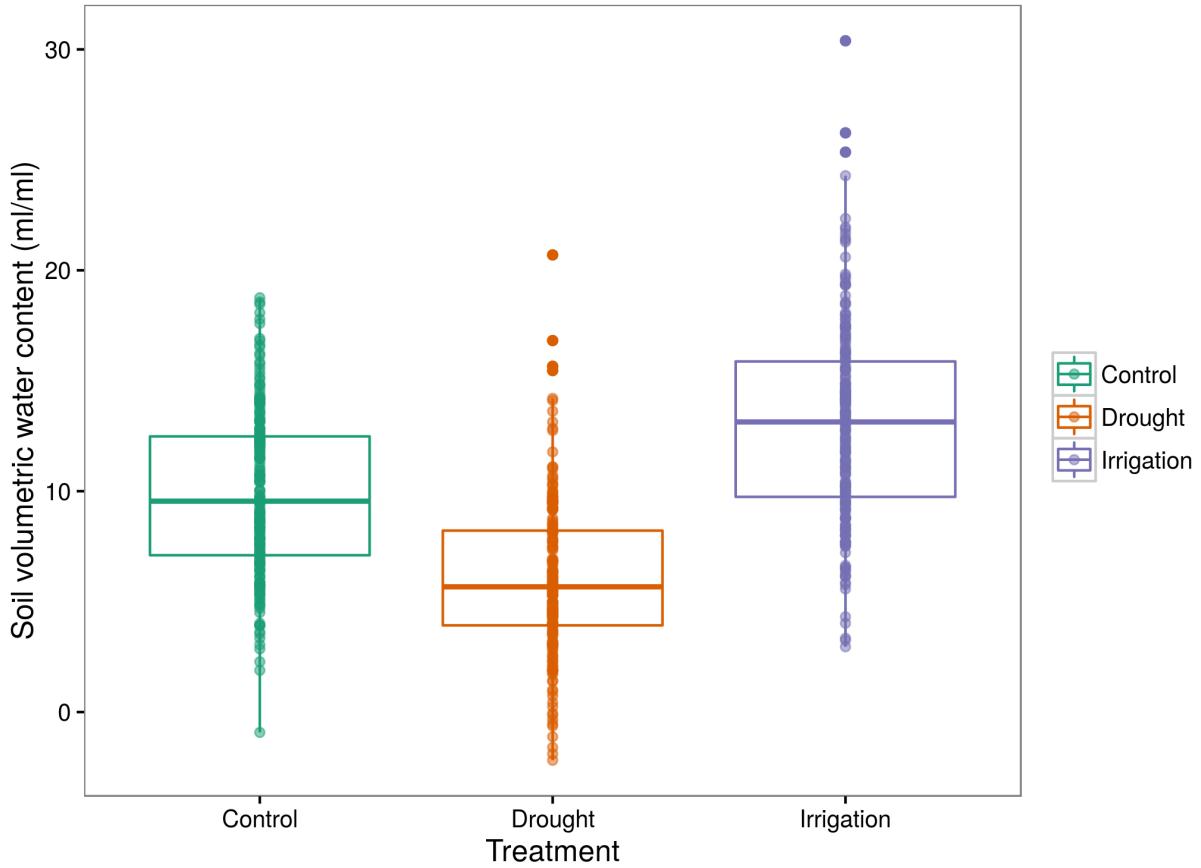


Figure 1: Soil moisture in the upper 5 cm of drought and irrigated plots compared to ambient controls. Soil moisture was measured at six locations around each plot at five different dates during the spring. Control plots were nearby areas of experiencing ambient soil moisture. Box plots show the median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

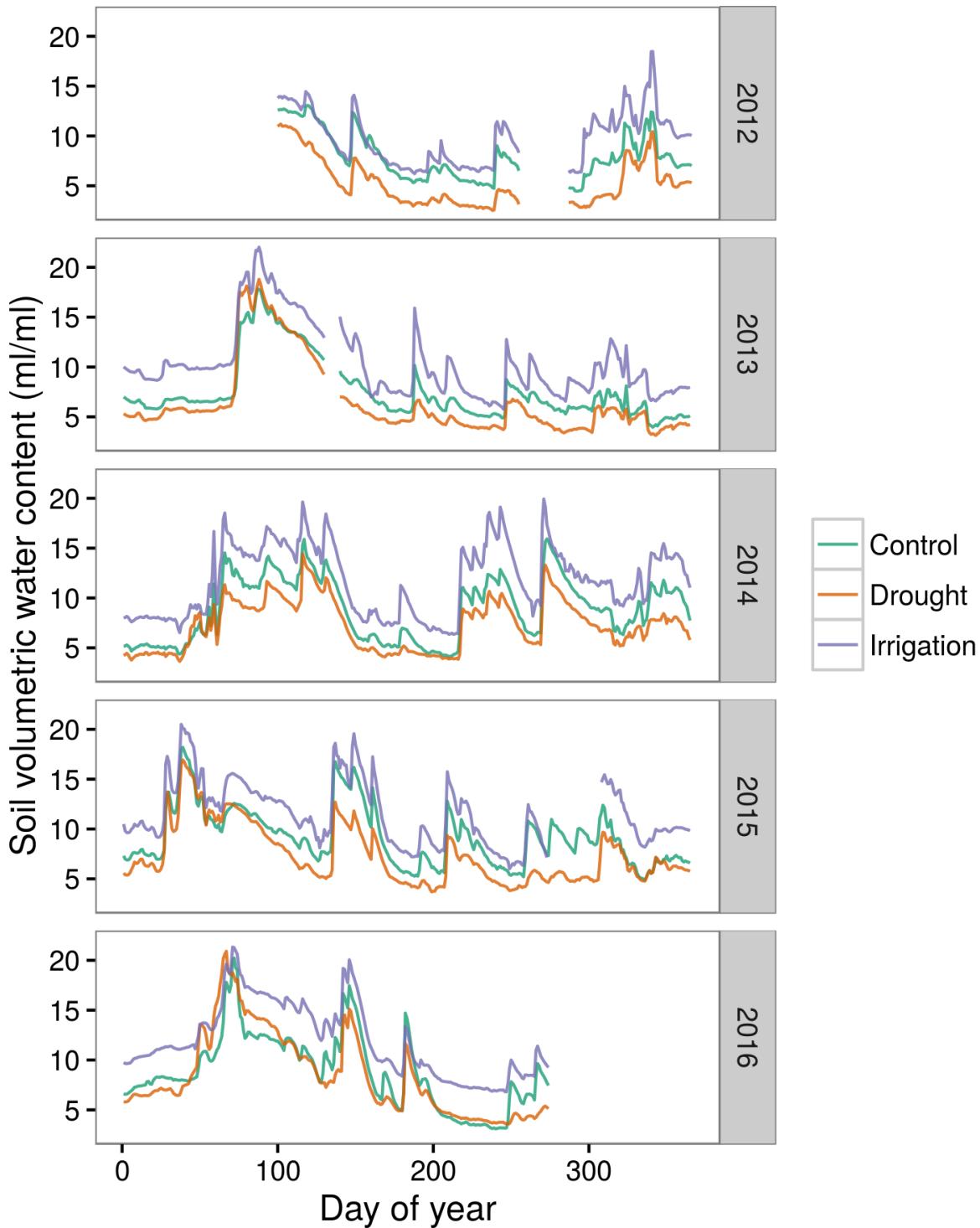


Figure 2: Average soil moisture in the control, drought, and irrigation treatments during each year of the experiment. Soil moisture was monitored in four drought plots, four irrigated plots and four ambient control plots. Two sensors were installed at 5 cm depth at each plot and two at 25 cm and data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

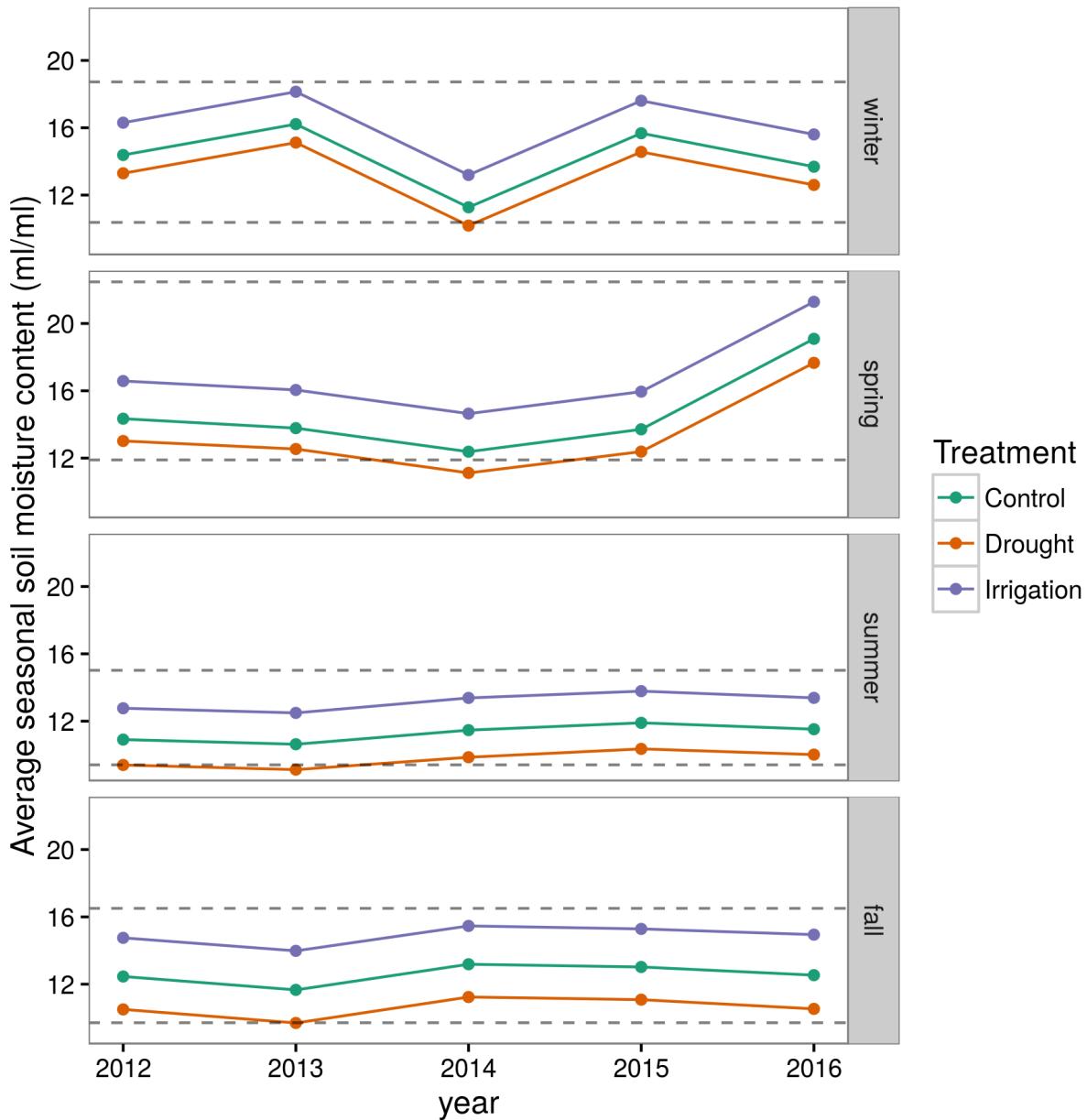


Figure 3: Average seasonal soil moisture in the control, drought, and irrigation treatments during each year of the experiment. The dashed gray lines give the 5th percentile and 95th percentile limits for seasonal soil moisture in the historical record (1929 to 2010).

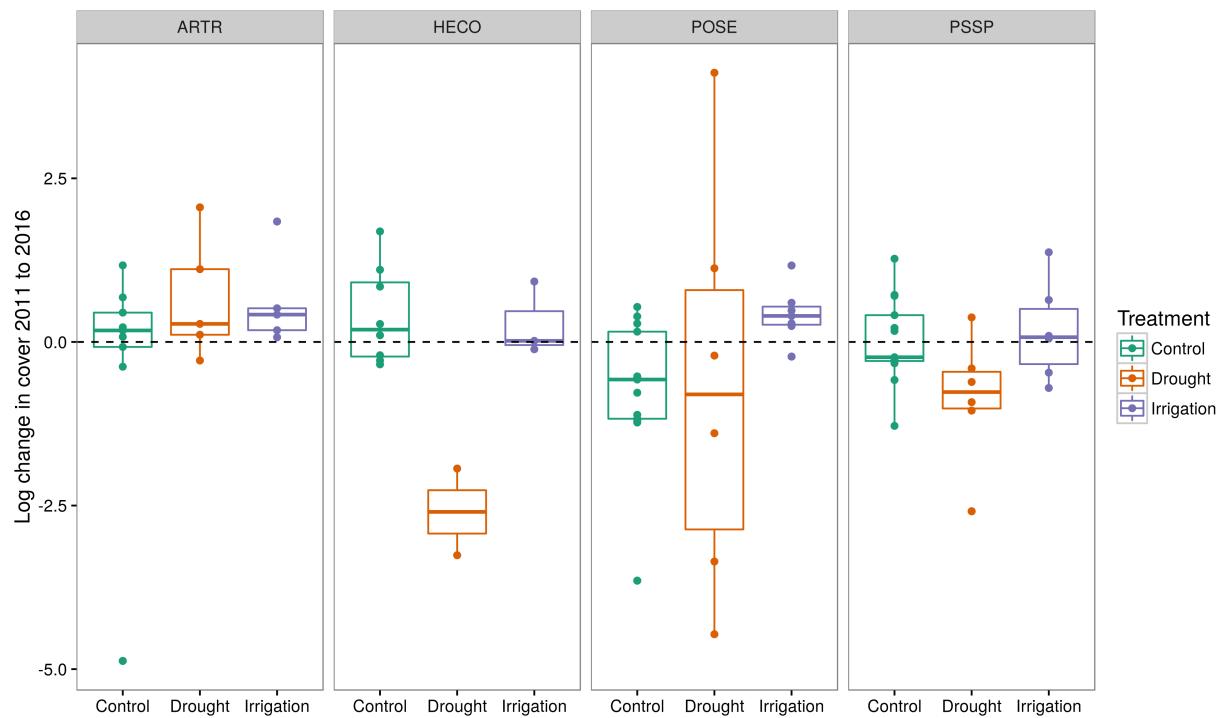


Figure 4: Log change in cover in each of the experimental plots from the pre-treatment monitoring in 2011 to the last year of the experiment in 2016. Box plots show the median cover change and the interquartile range. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

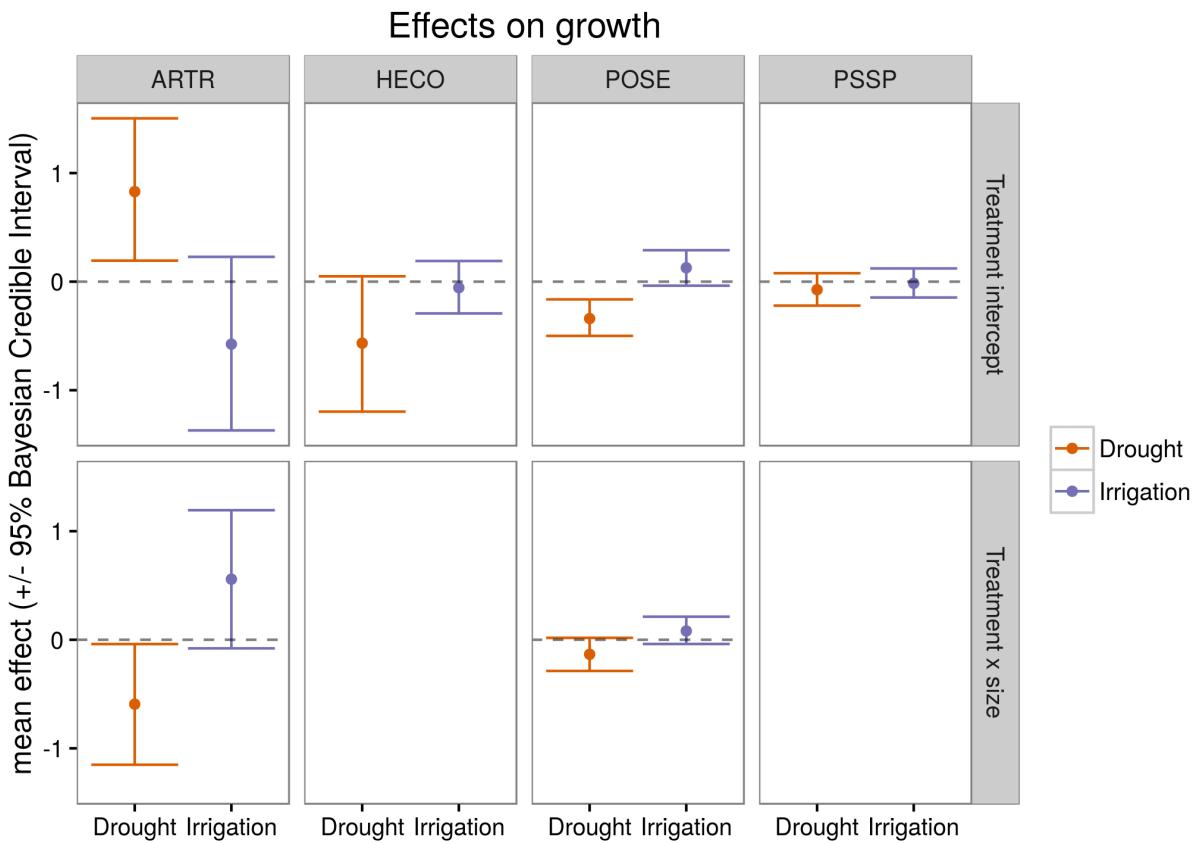


Figure 5: Parameter estimates for the effects of treatment on growth of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. Size by treatment interactions were only fit for ARTR, and POSE. Plant size was centered on mean size and scaled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

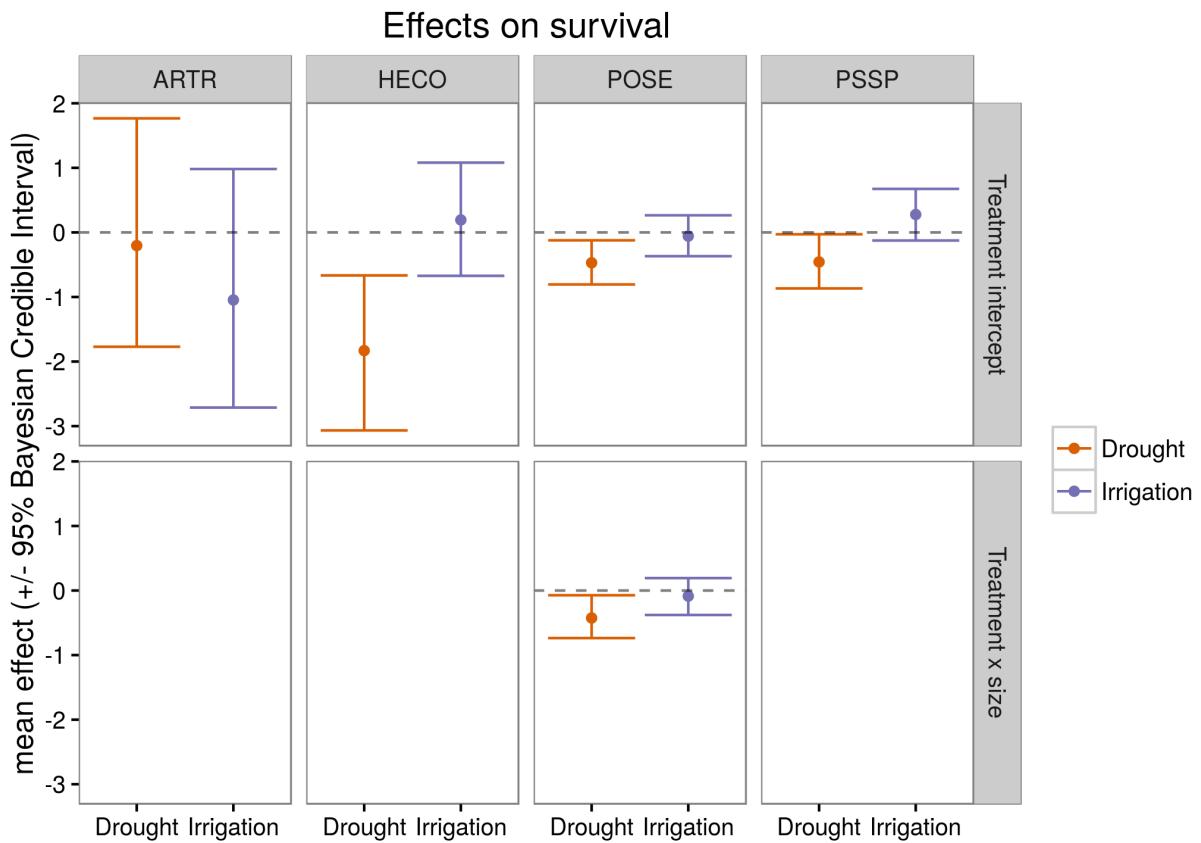


Figure 6: Parameter estimates for the effects of treatment on survival of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. Size by treatment interactions were only fit for POSE. Plant size was centered on mean size and scaled by its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

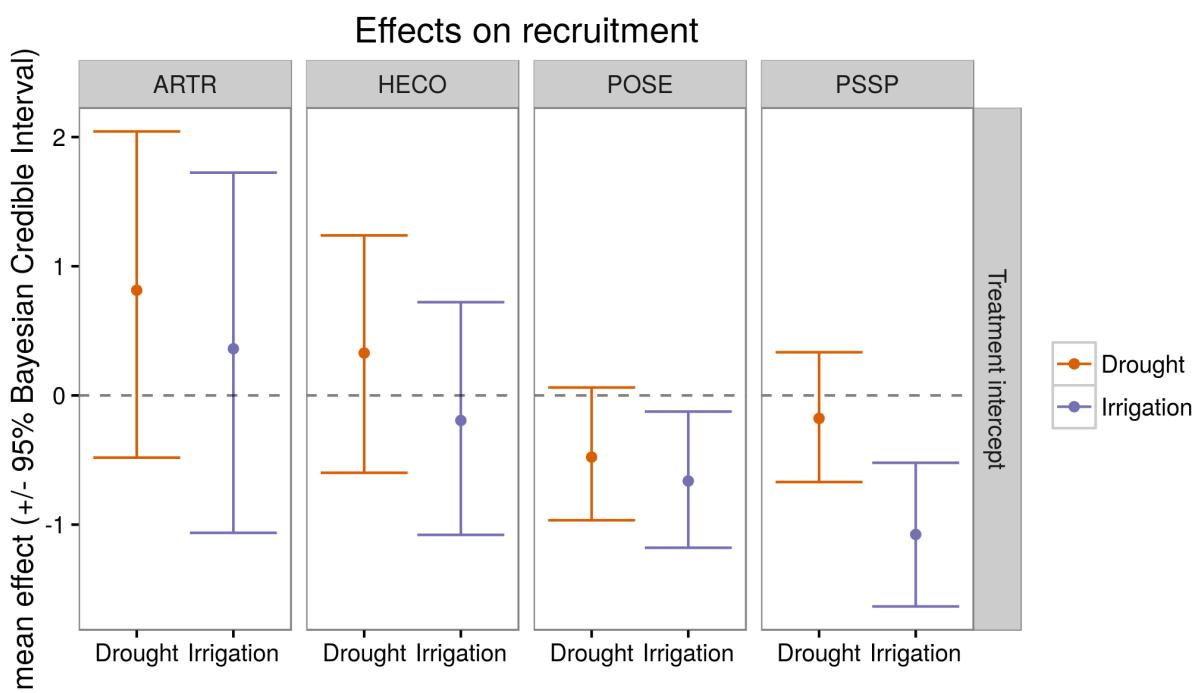


Figure 7: Parameter estimates for the effects of treatment on recruitment of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

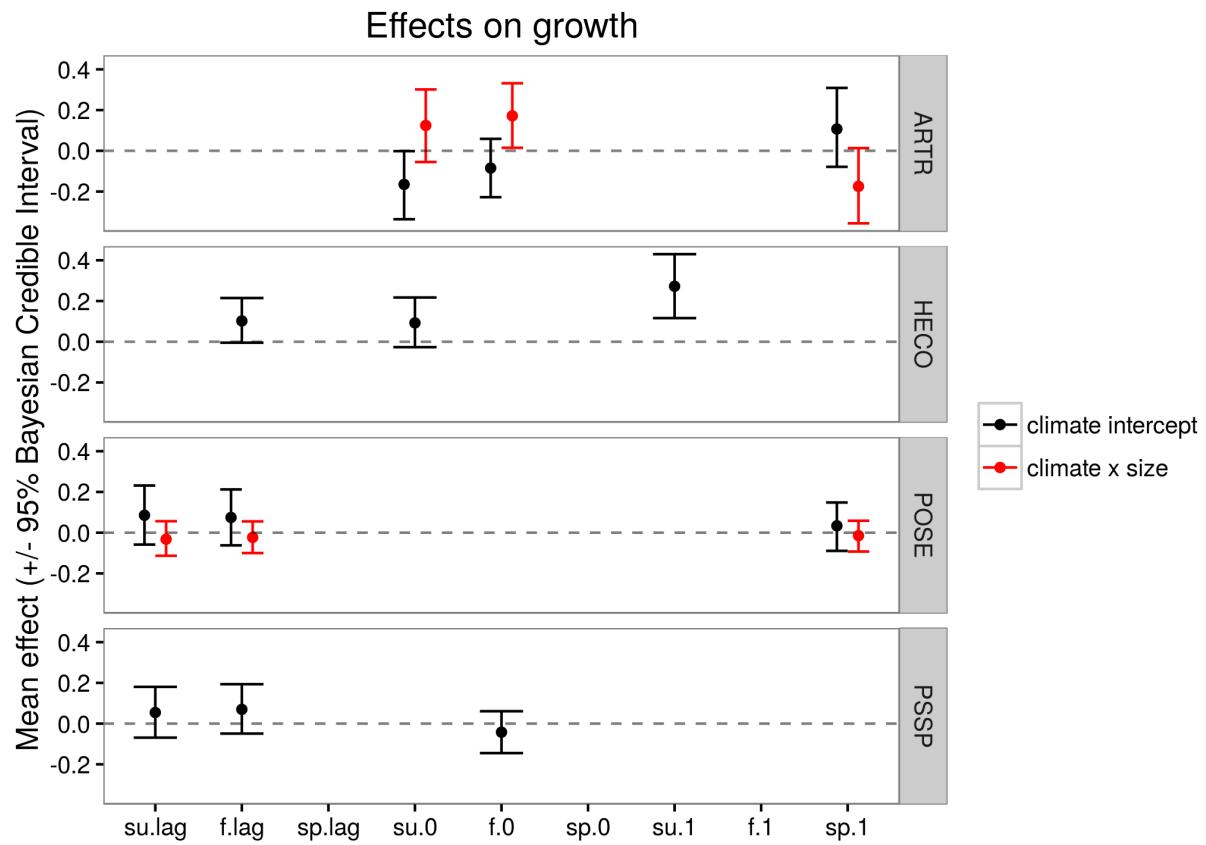


Figure 8: Parameter estimates for the selected seasonal soil moisture covariates on the growth of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. Red parameters show size x climate interaction effects. We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

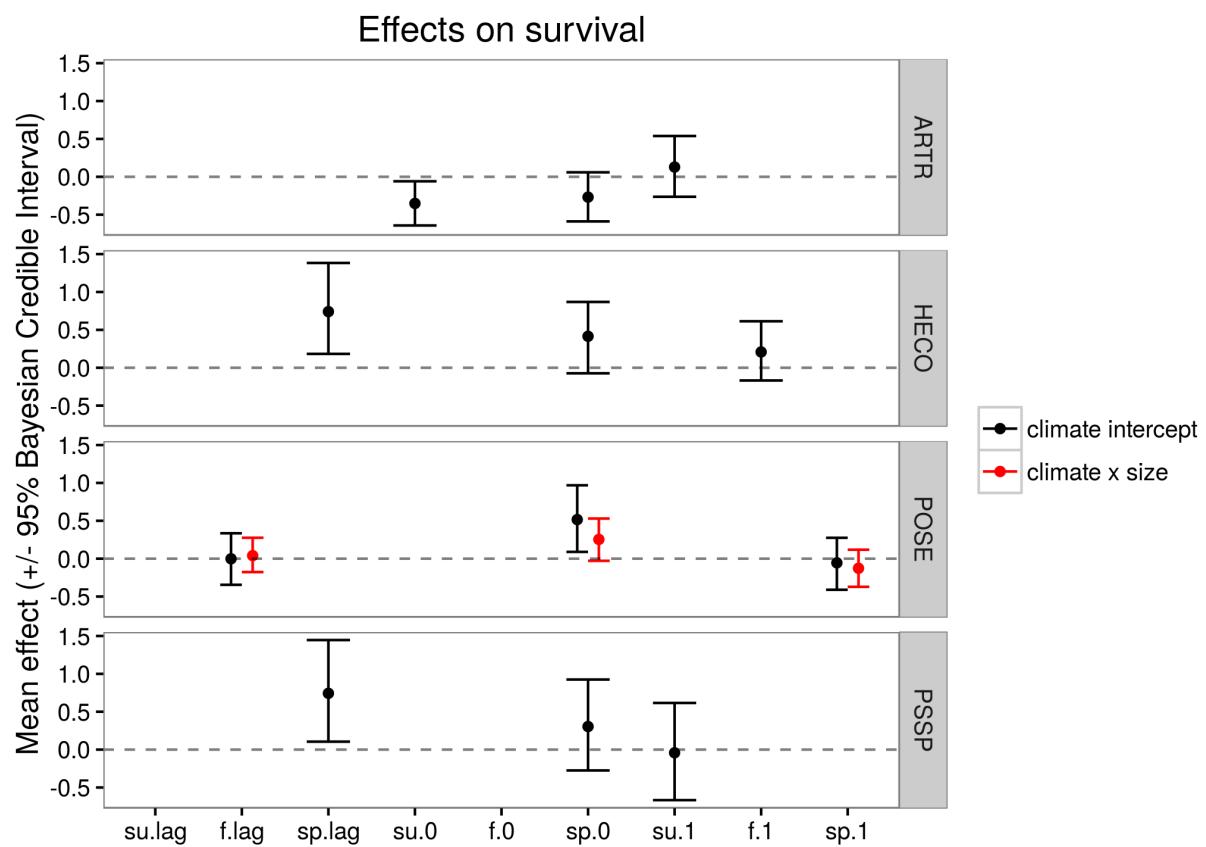


Figure 9: Parameter estimates for the selected seasonal soil moisture covariates on the survival of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. Red parameters show size x climate interaction effects.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

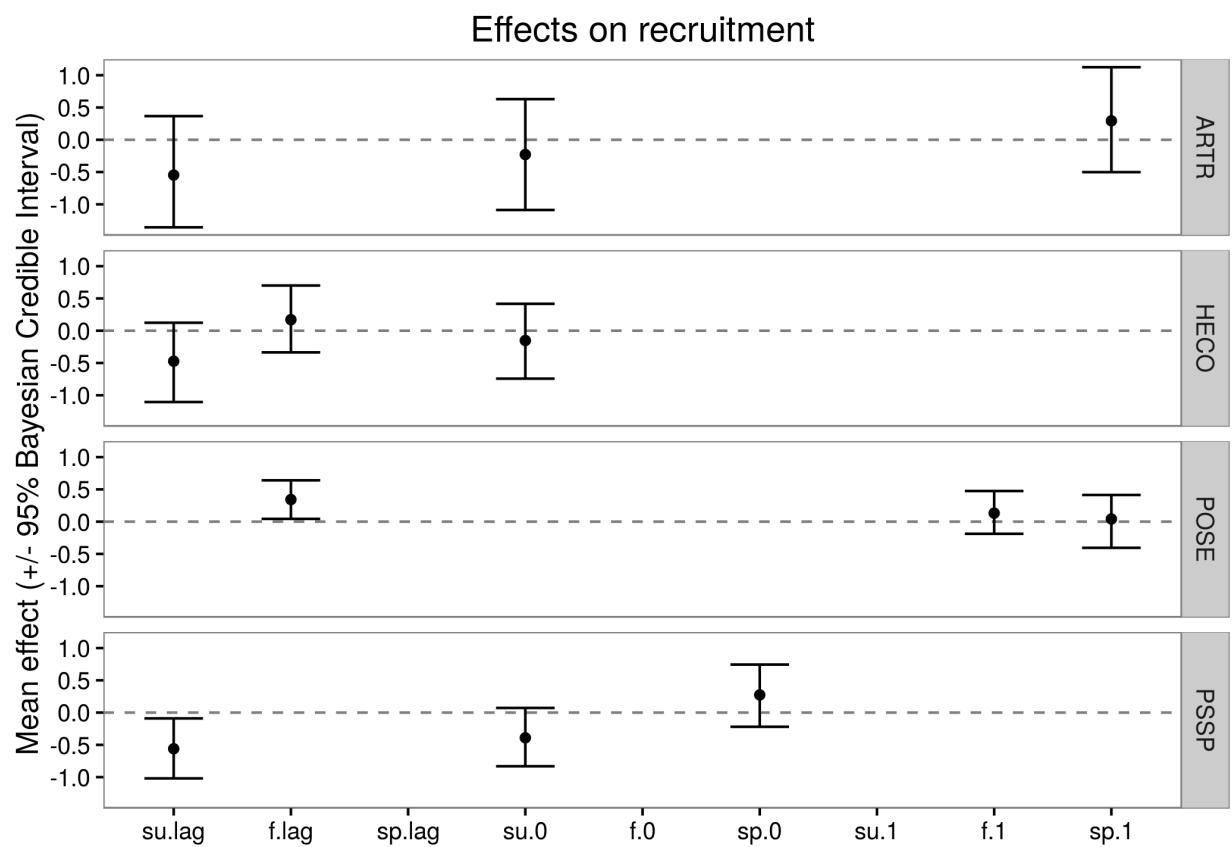


Figure 10: Parameter estimates for the selected seasonal soil moisture covariates on the recruitment of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. We assessed a parameter as significant when the 95%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

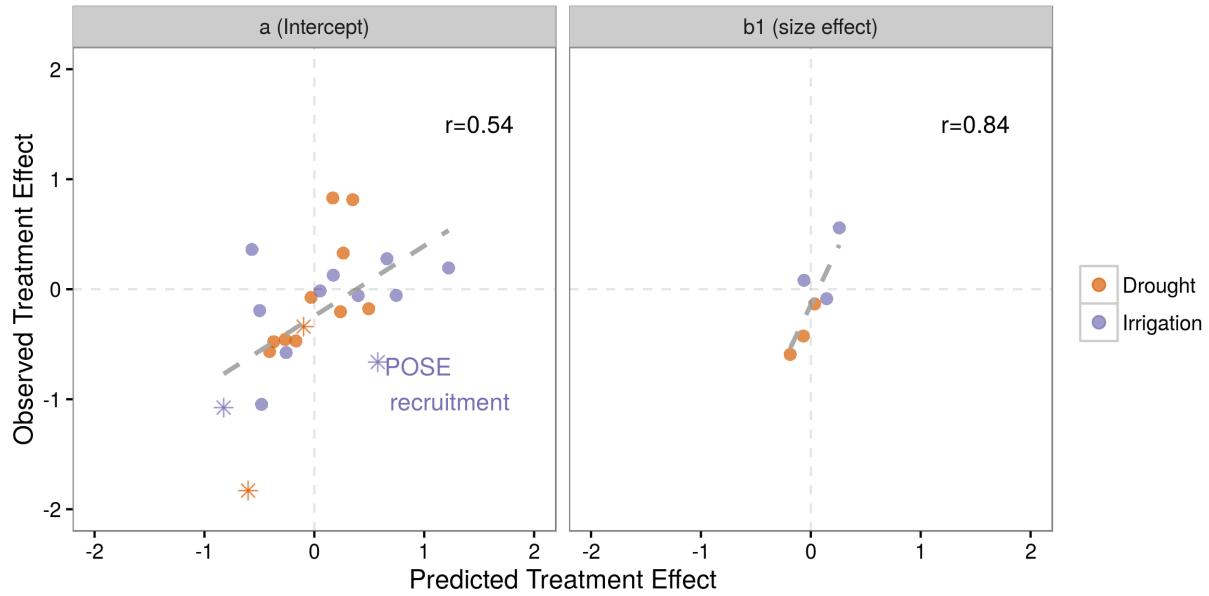


Figure 11: The treatment effects predicted by the climate model compared to the treatment effects observed for the intercept parameters (left side) and size by climate/treatment effects (right side). Parameters from all species and vital rates are shown together. The observed treatment effects come from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

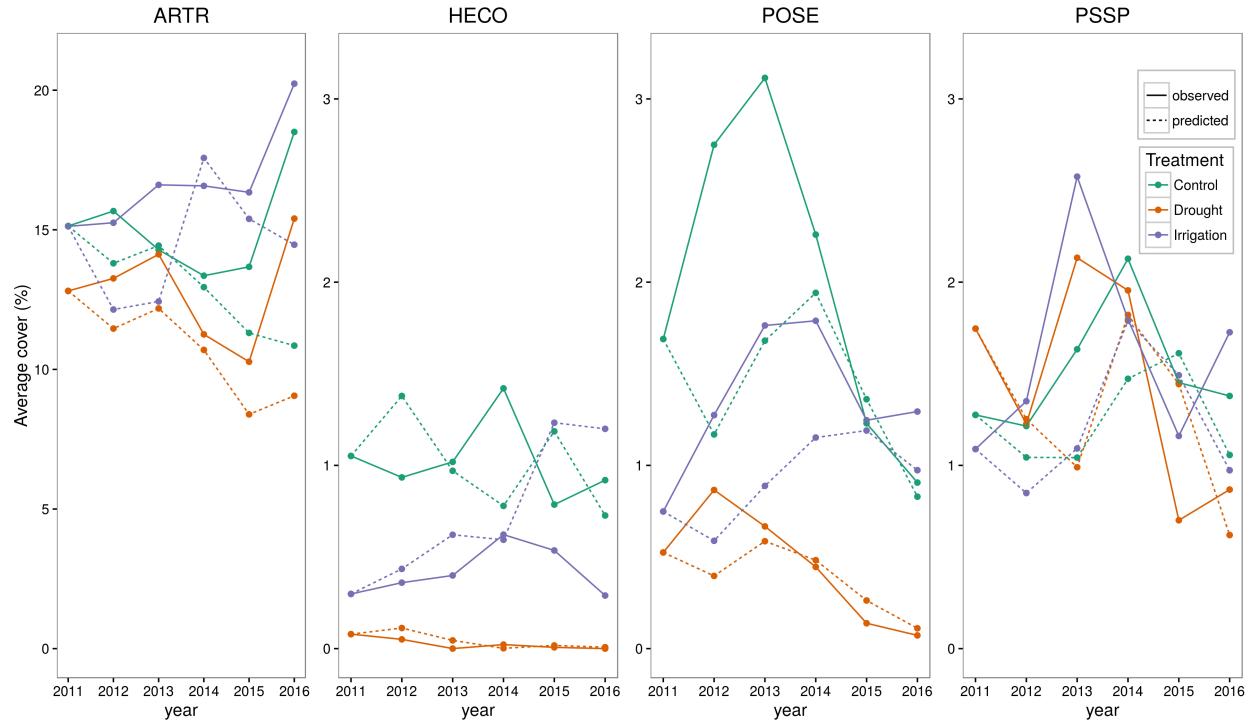


Figure 12: Observed average cover per quadrat in the experimental and control plots and one step ahead cover predictions from the climate model. Cover predictions for each year are generated from the IBM based on the observed distribution of plants in each quadrat in the current year. Quadrat cover was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

### *A. tripartita*

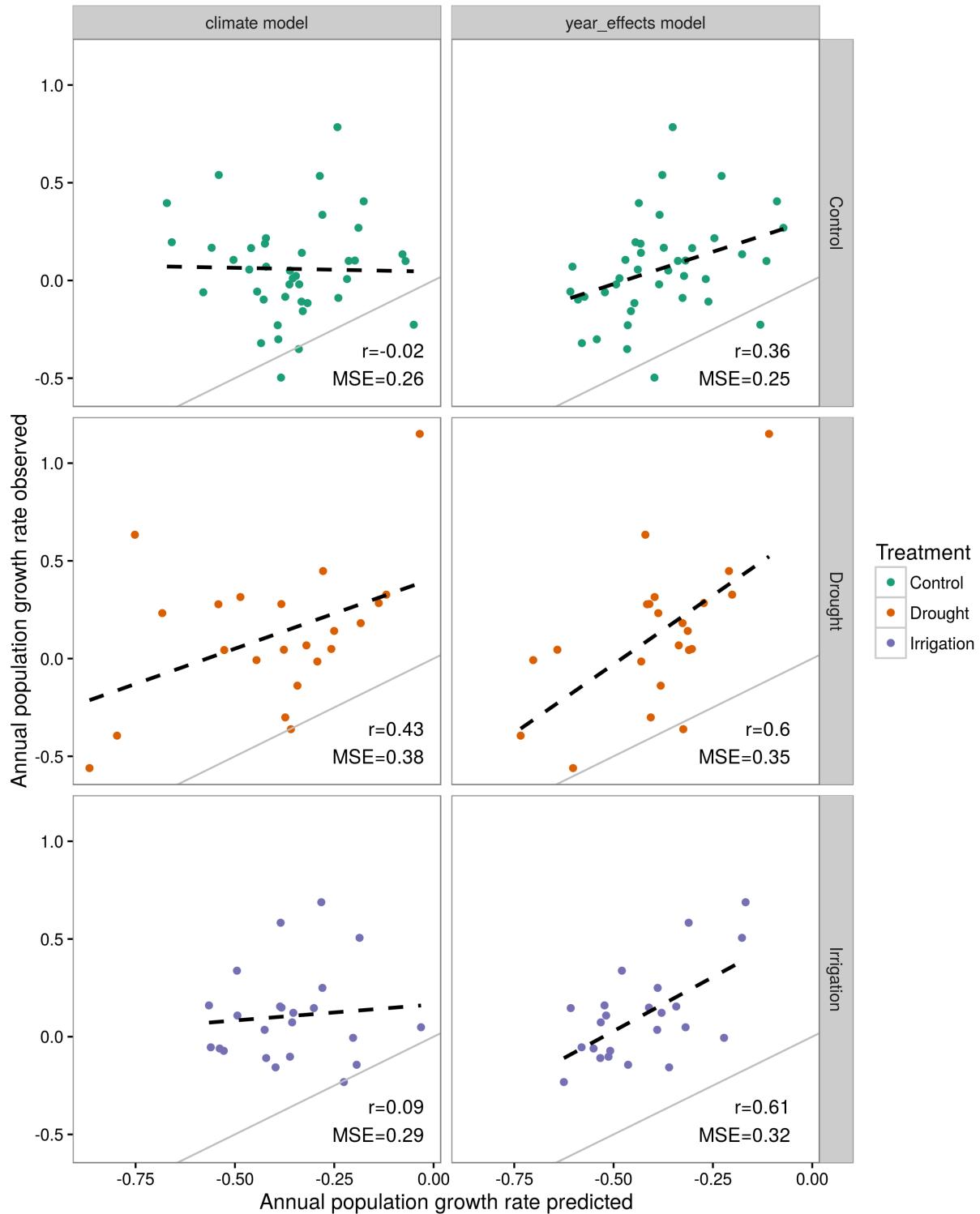


Figure 13: Observed and predicted one step ahead log change in *A. tripartita* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations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

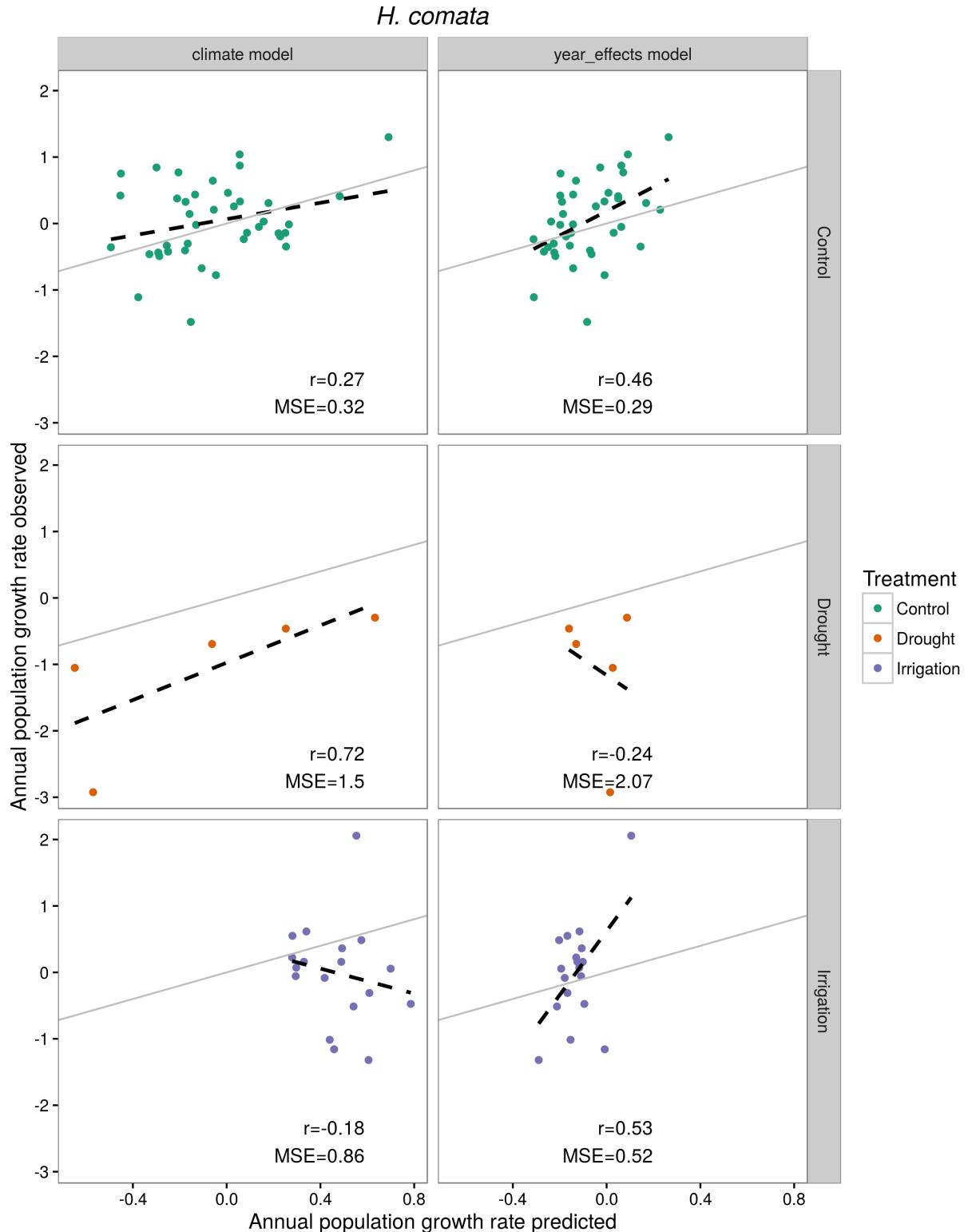


Figure 14: Observed and predicted one step ahead log change in *H. comata* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations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

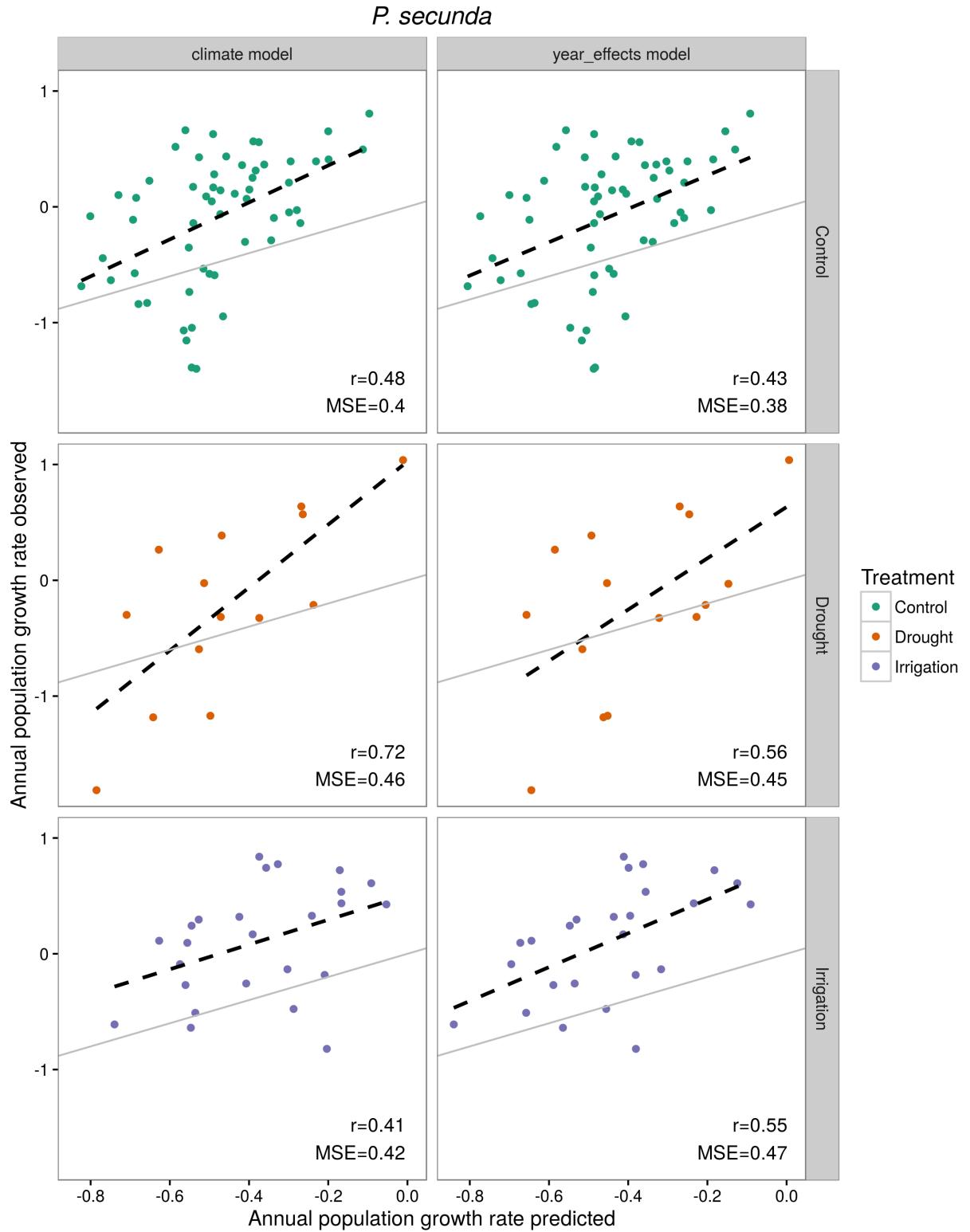


Figure 15: Observed and predicted one step ahead log change in *P. secunda* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations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

*P. spicata*

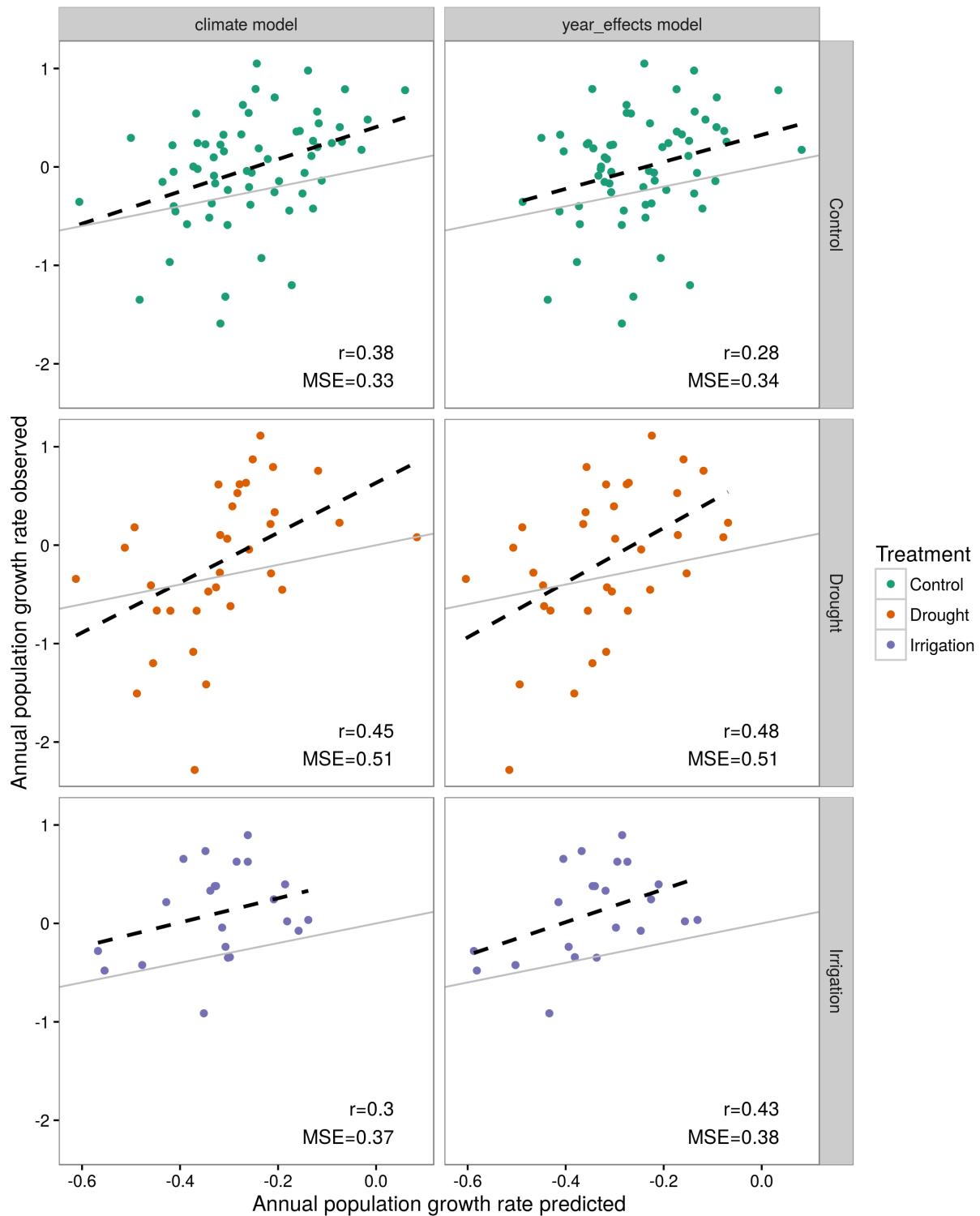


Figure 16: Observed and predicted one step ahead log change in *P. spicata* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations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

## Appendix

Kleinhezelink et al., “Predicting climate response”

## Additional Tables

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

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

Continued on next page

	species	vital_rate	Treatment	score	climate model	year effects	model	diff	improved
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		***
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		***

Continued on next page

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

## Additional Figures

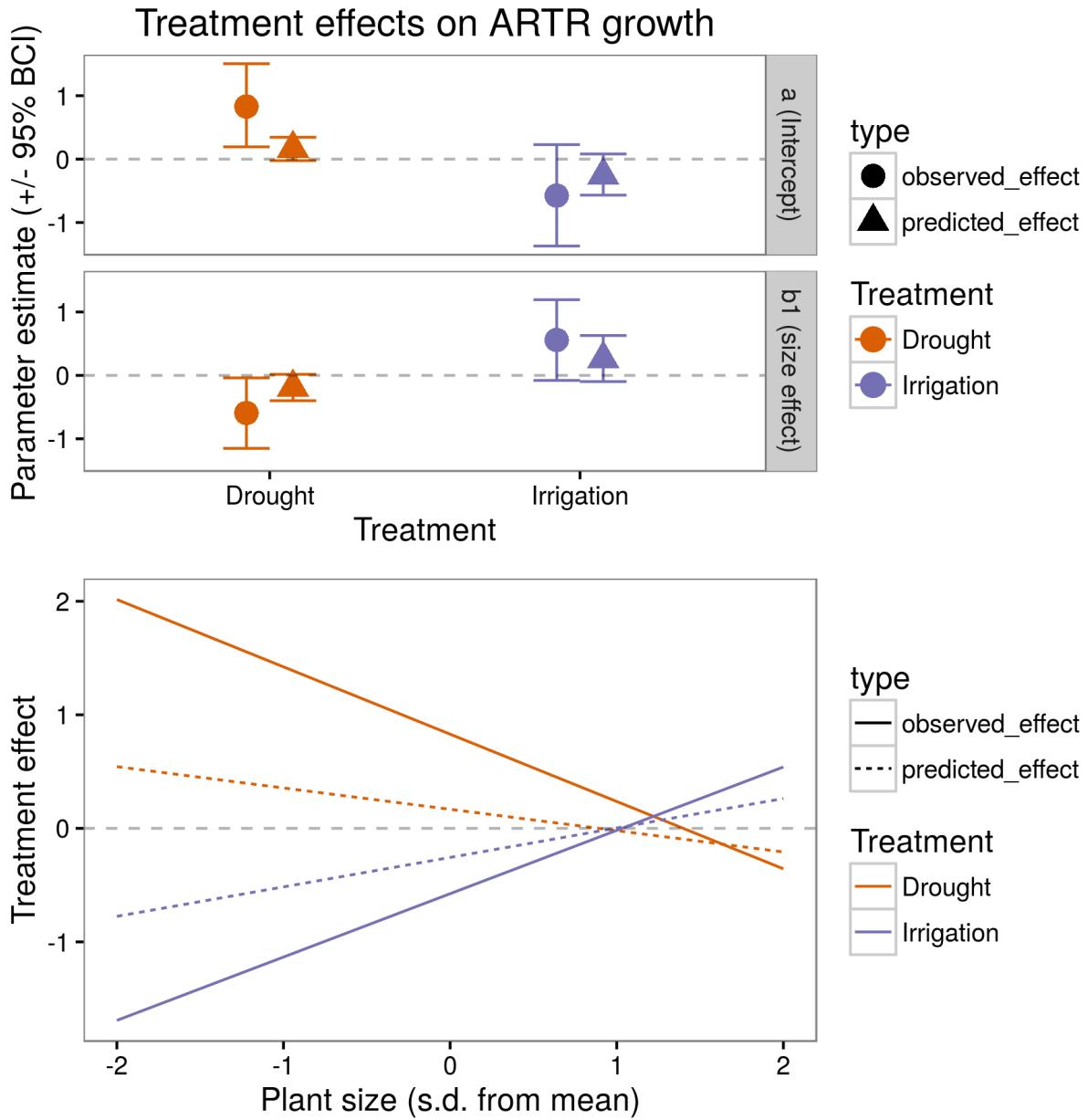


Figure A-1: Comparison of treatment effects predicted and observed for *A. tripartita* growth. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to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

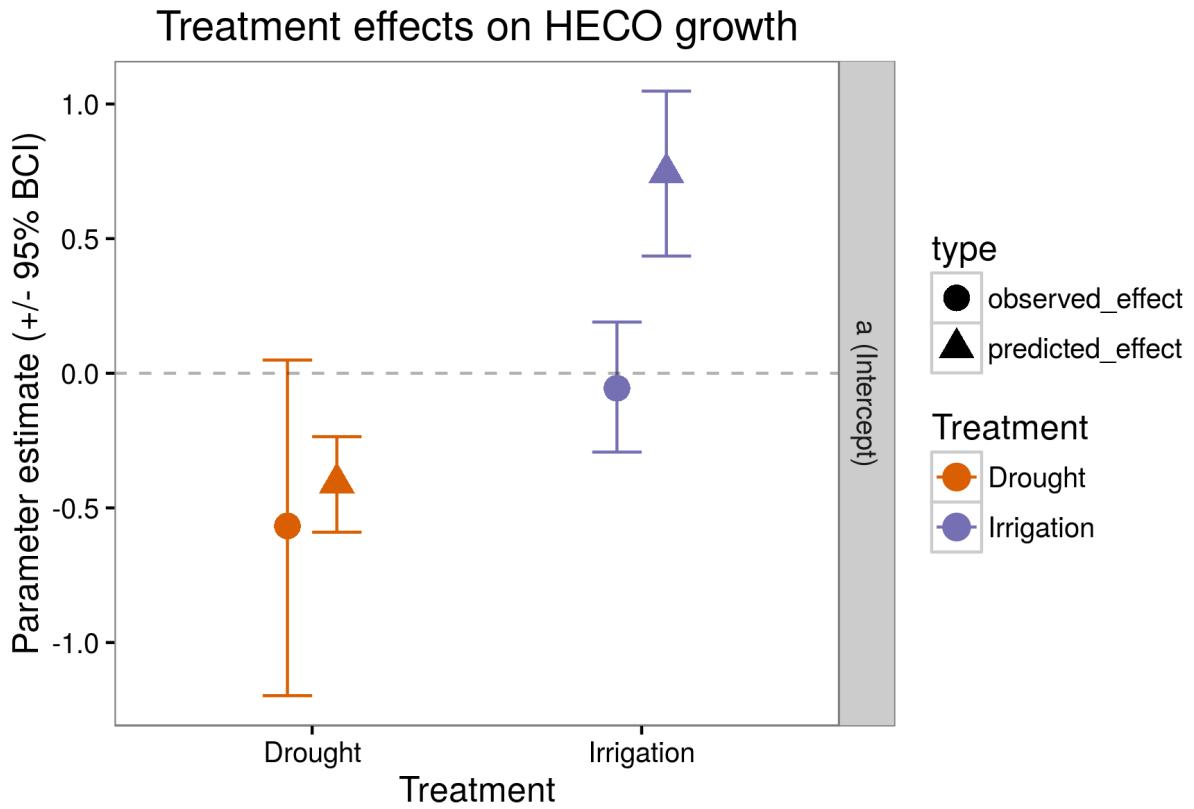


Figure A-2: Comparison of treatment effects predicted and observed for *H. comata* growth. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational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

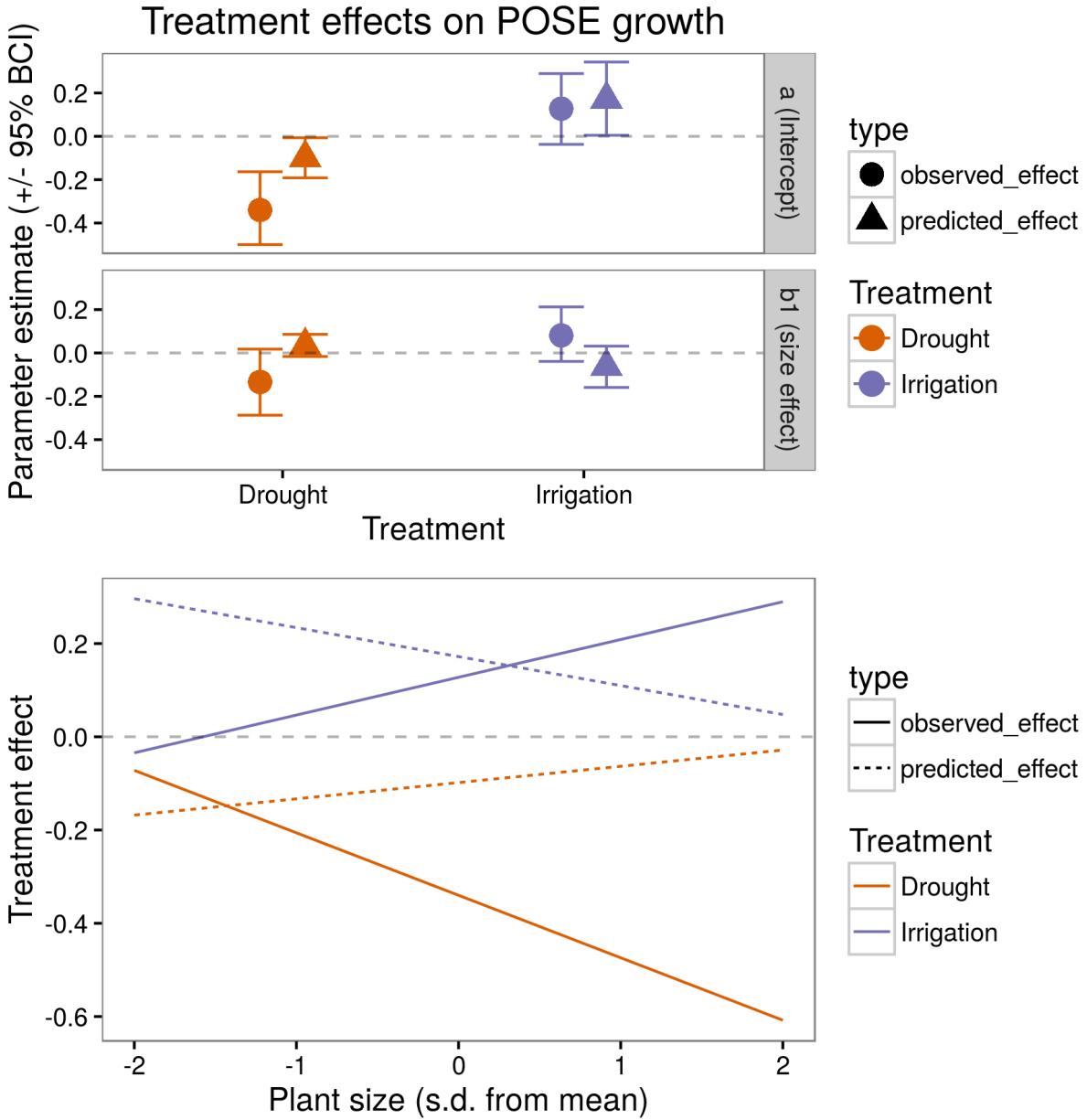


Figure A-3: Comparison of treatment effects predicted and observed for *P. secunda* growth. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to all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

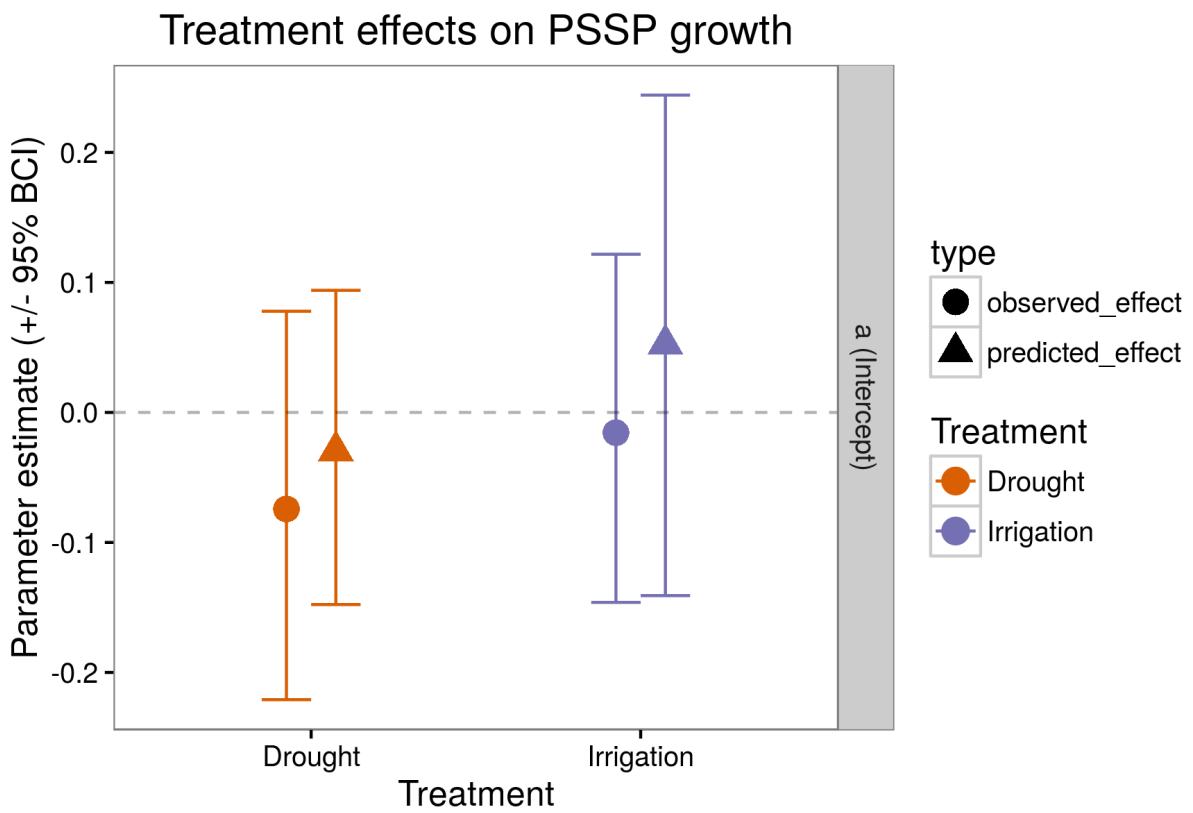


Figure A-4: Comparison of treatment effects predicted and observed for *P. spicata* growth. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational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

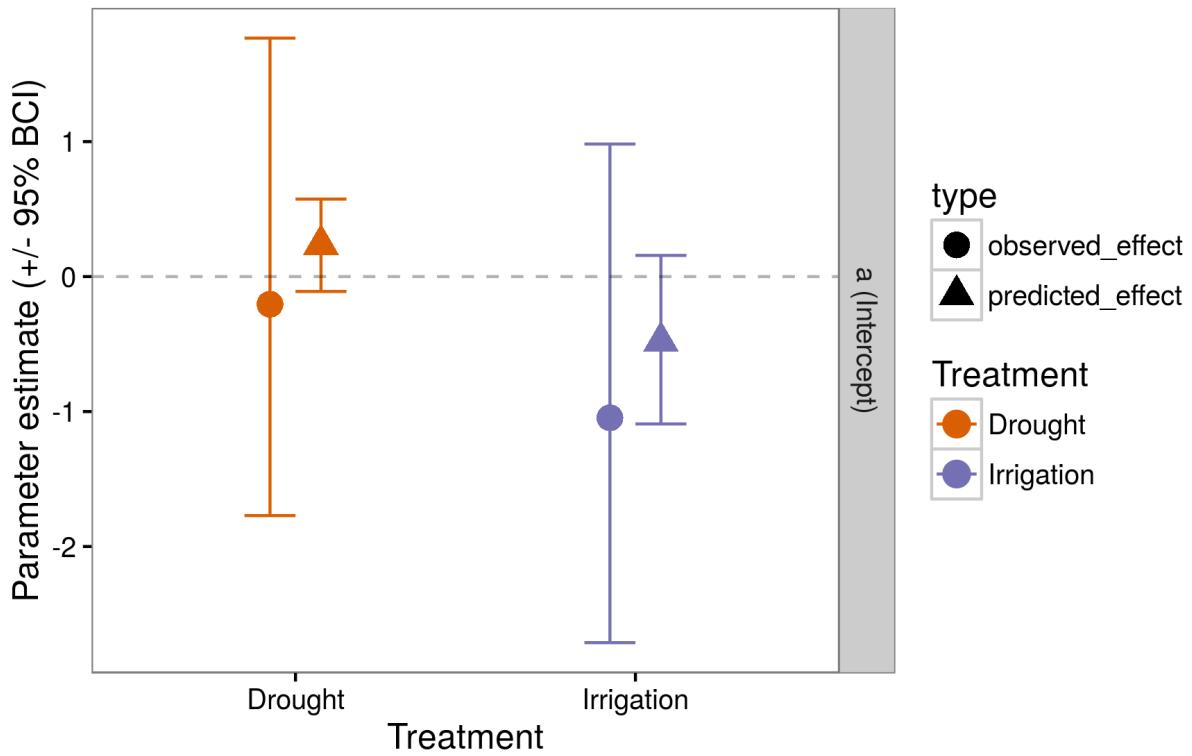


Figure A-5: Comparison of treatment effects predicted and observed for *A. tripartita* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of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

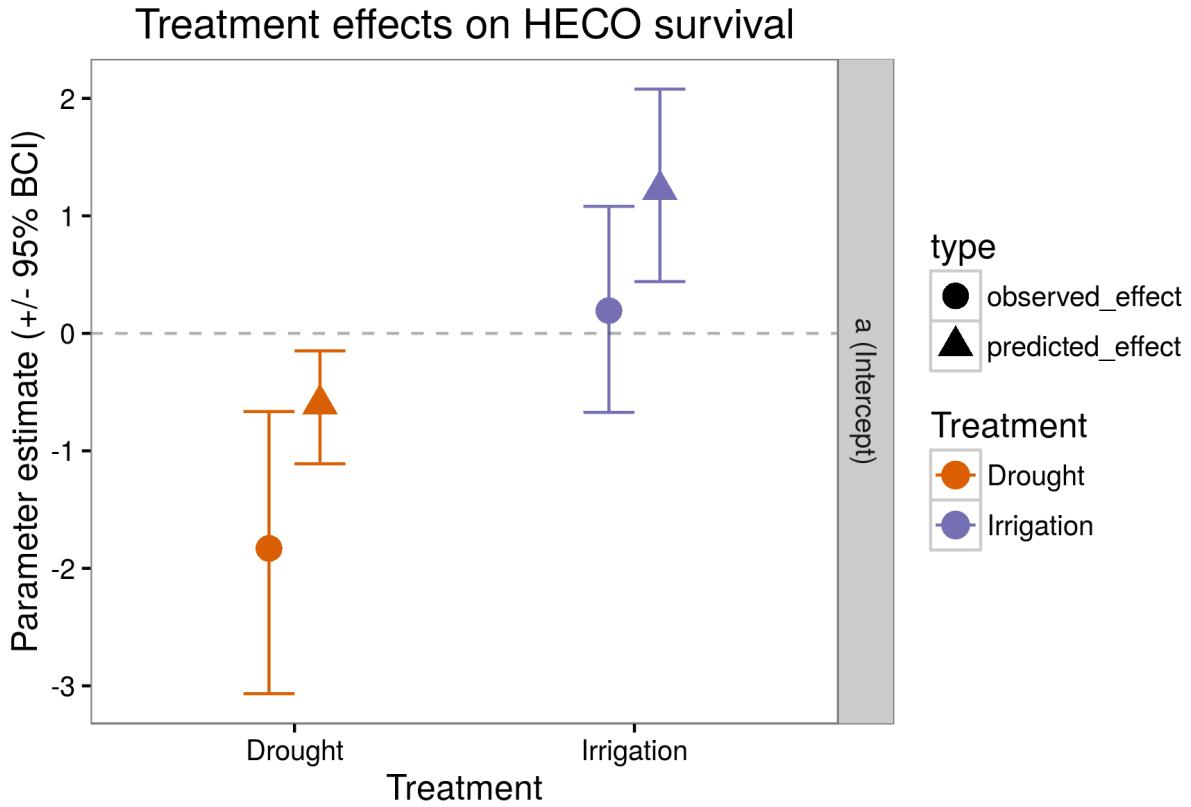


Figure A-6: Comparison of treatment effects predicted and observed for *H. comata* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational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

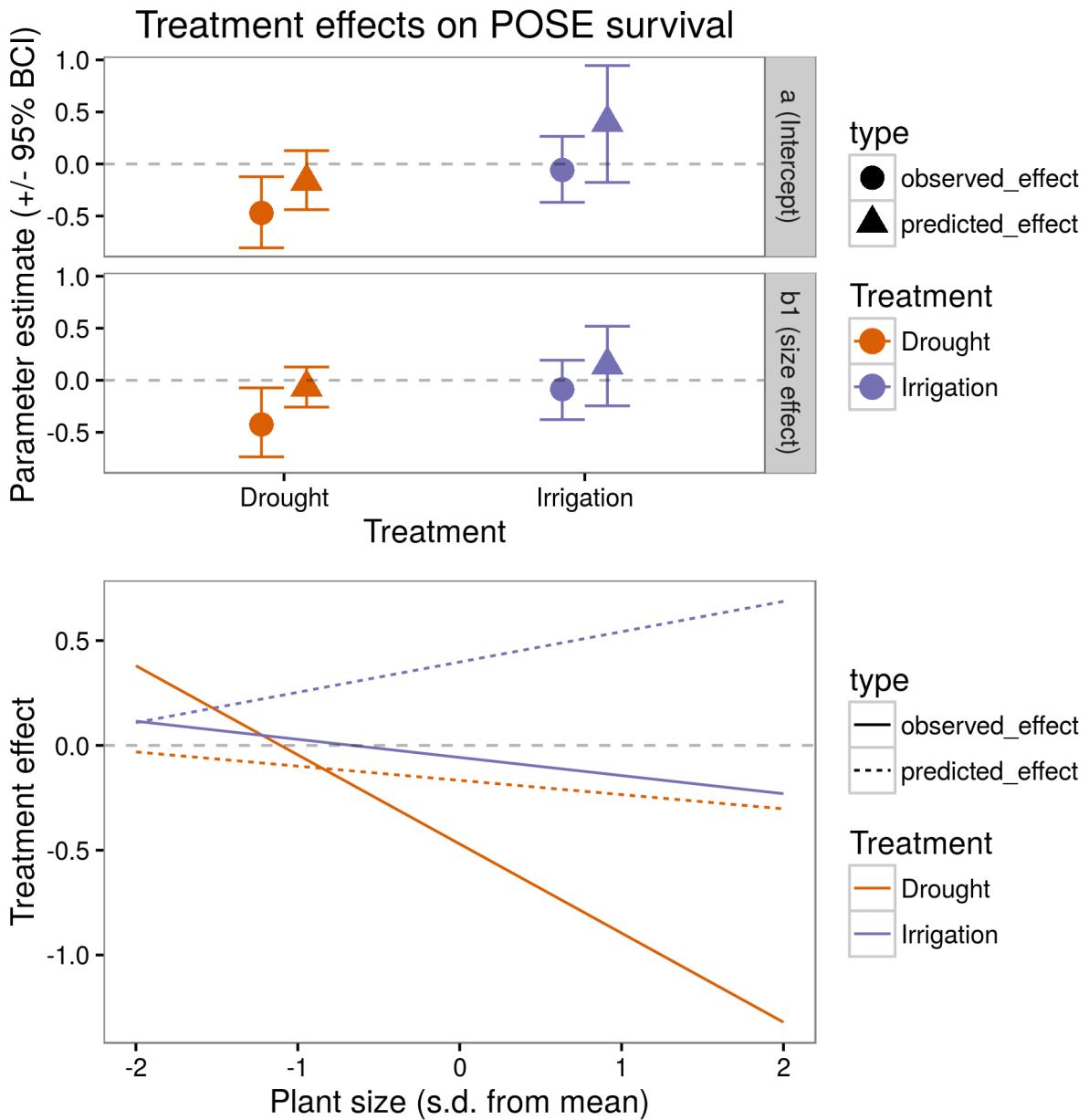


Figure A-7: Comparison of treatment effects predicted and observed for *P. secunda* survival. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to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

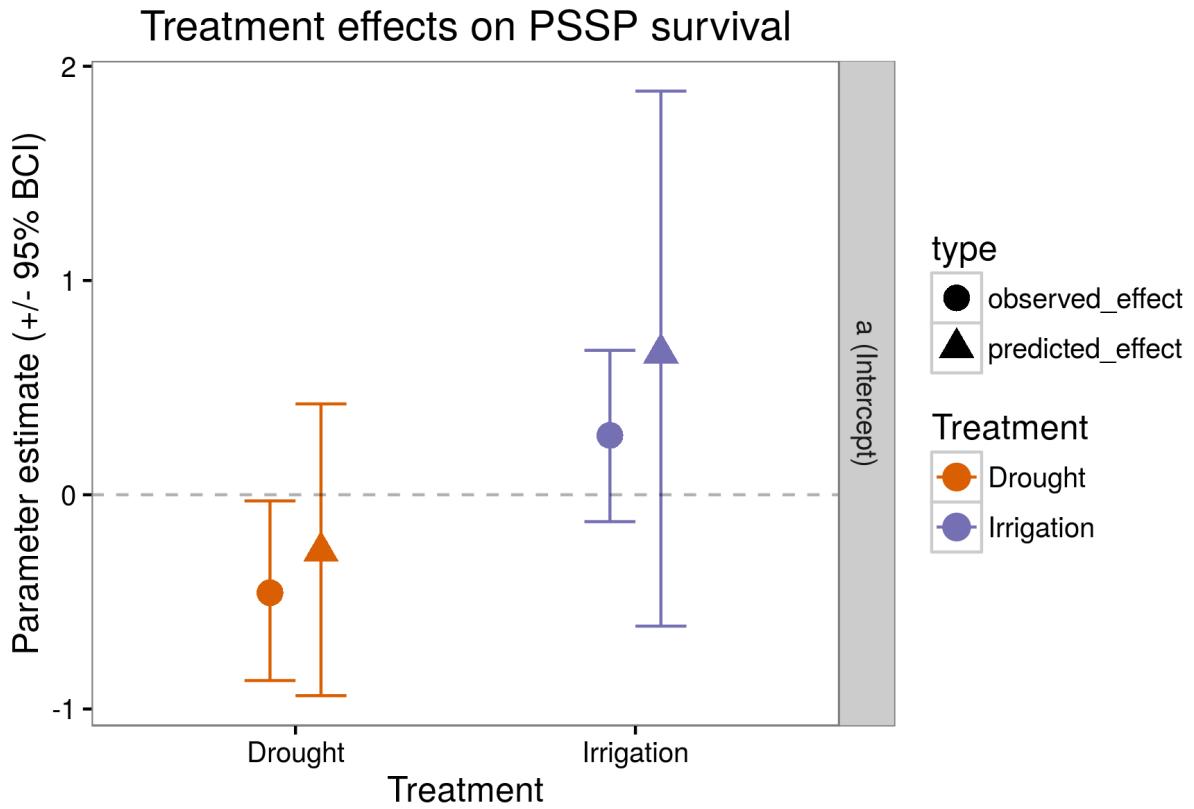


Figure A-8: Comparison of treatment effects predicted and observed for *P. spicata* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational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

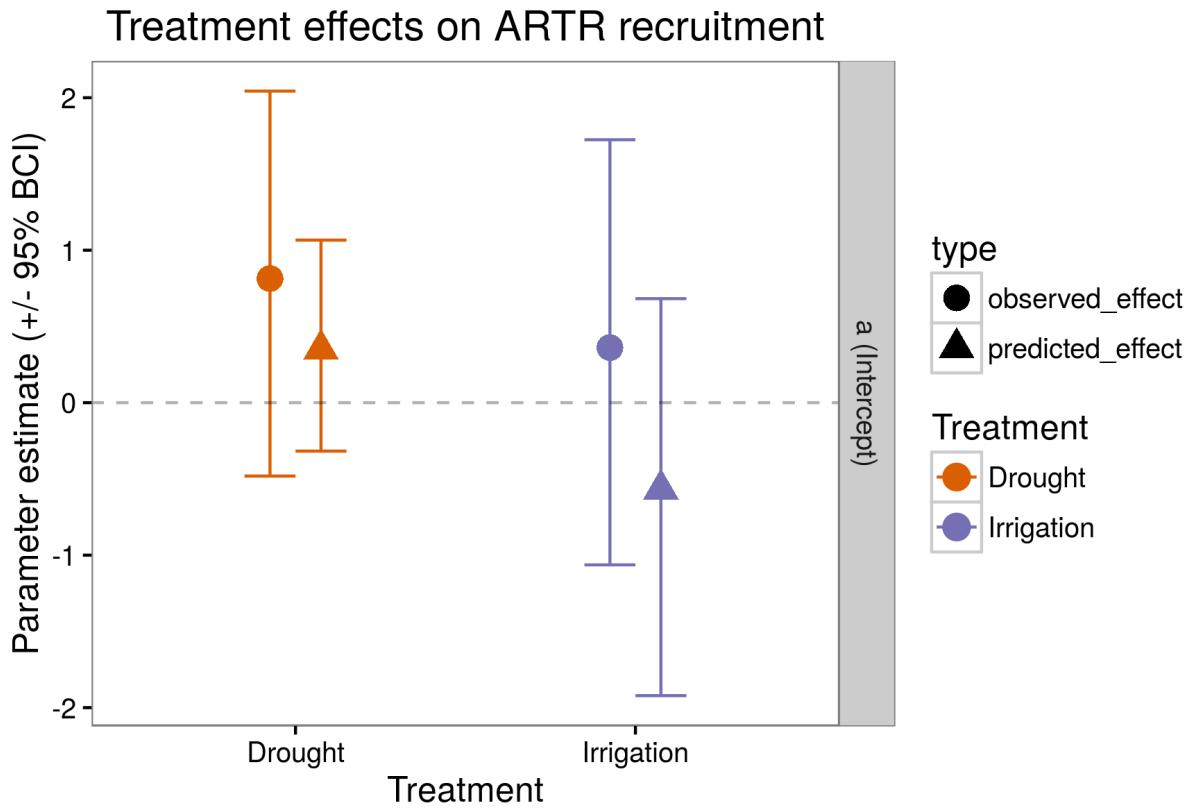


Figure A-9: Comparison of treatment effects predicted and observed for *A. tripartita* recruitment. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years 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.

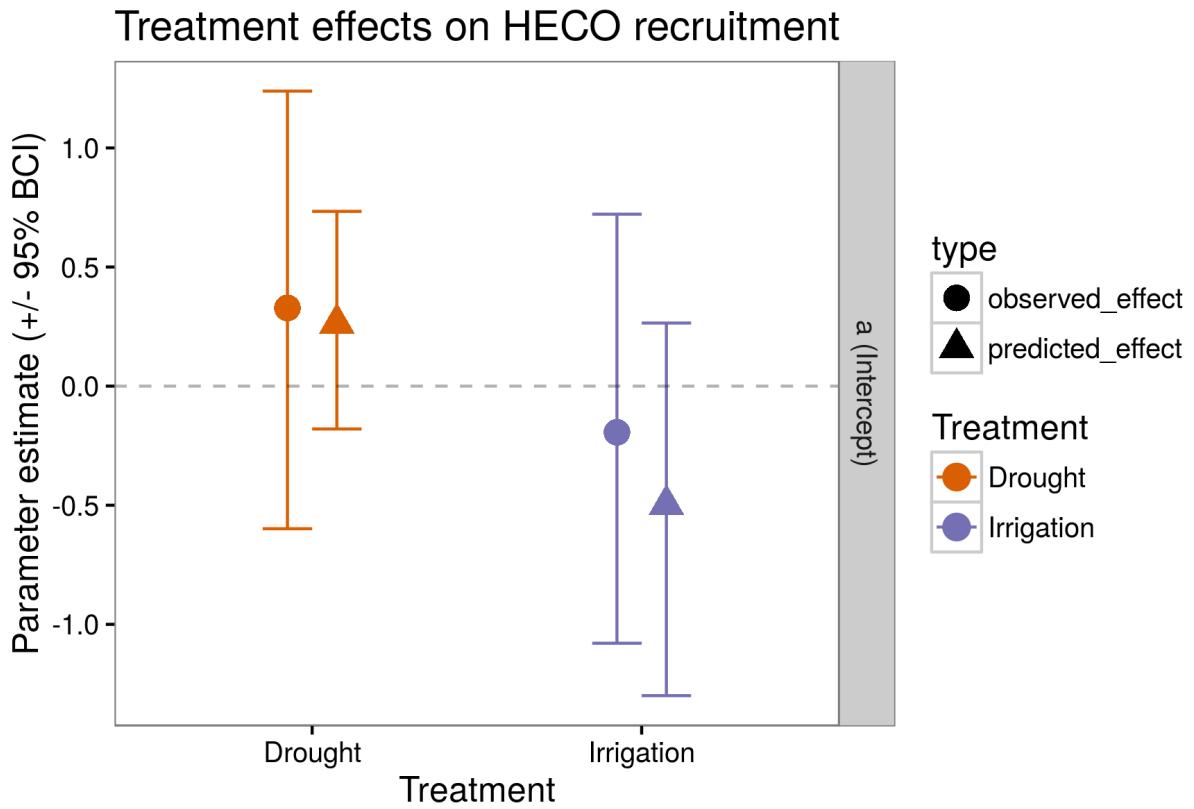


Figure A-10: Comparison of treatment effects predicted and observed for *H. comata* recruitment. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of 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.

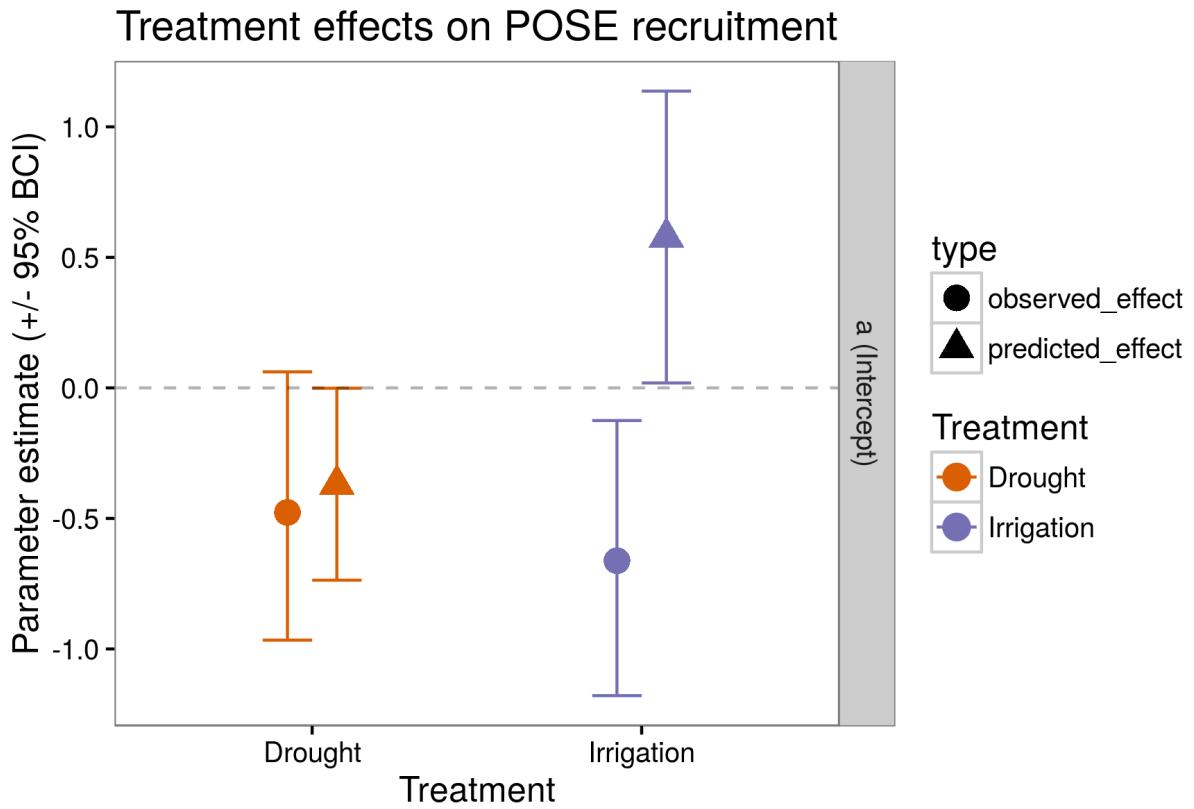


Figure A-11: Comparison of treatment effects predicted and observed for *P. secunda* recruitment. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of 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.

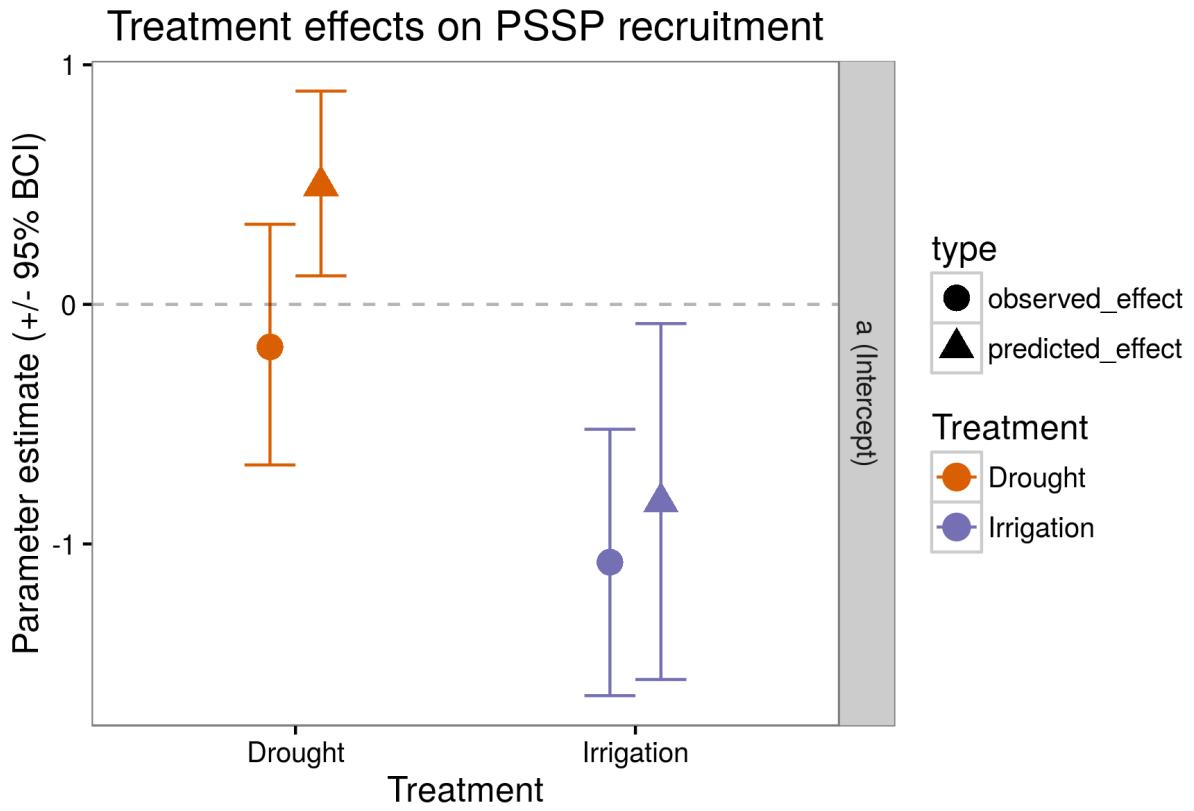


Figure A-12: Comparison of treatment effects predicted and observed for *P. spicata* recruitment. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of 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.