

Can historical data predict population responses to climate change experiments?

Andrew R. Kleinhesselink^{*a}, Peter B. Adler^a, Andrew Tredennick^a, Caitlin Andrews^b and John Bradford^b

^aDepartment of Wildland Resources and the Ecology Center, Utah State University, Logan, Utah

^bUS 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.

^{*}Corresponding author. Department of Wildland Resources and the Ecology Center, Utah State University, Logan, Utah Email: arklein@aggiemail.usu.edu

¹ Abstract

² Climate is an important driver of population ecology however there have been few tests of whether
³ observational data correlating population performance with climate variation can be used to predict
⁴ how populations respond to climate change experiments. We tested whether longterm observational
⁵ data relating the effects of climate variation to the survival, growth and recruitment of four common
⁶ plants in a sagebrush steppe could be produce accurate predictions for how each species would
⁷ respond to a drought and irrigation experiment.

⁸ We used rainout shelters and automatic sprinklers to manipulate the soil moisture in sixteen
⁹ new plots established at the same field site as the longterm observational data. After running the
¹⁰ experiment for five years, we analyzed how the growth, survival and recruitment of each species
¹¹ responded to the experimental drought and irrigation treatments. We made predictions using two
¹² types of models fit to the observational data collected prior to the experiment: a no climate model
¹³ that only included the effects of plant size and competition on plant performance in each plot in each
¹⁴ year; and a climate model that also included the effects of three seasonal soil moisture variables.
¹⁵ We then used these models to predict the performance of each species during the course of the
¹⁶ experiment. We also generated one step a head predictions of population size in each experimental
¹⁷ plot using an individual based population model. We compared predictions made by the the no
¹⁸ climate and climate models to the observed experimental responses.

¹⁹ Over the course of the experiment, average cover of the perennial grasses, *Hesperostipa*
²⁰ *comata* and *Pseudoroegneria spicata* declined significantly in the drought treatment. At the level
²¹ of individual vital rates, experimental drought reduced the survival of *Hesperostipa comata* and *P.*
²² *spicata* and the growth of the grass *Poa secunda*. In contrast, drought increased the growth of the
²³ shrub *A. tripartita*. The climate model better predicted the response to the experiment than the
²⁴ no climate model in six out of twelve cases. Across all species and vital rates, there was a strong
²⁵ positive correlation between the observed effects of the experiment and the effects predicted by
²⁶ the climate model fit to the observational data. At the population-level, the climate model made
²⁷ better predictions of the changes in species cover than the no climate model for *P. secunda* and *P.*
²⁸ *spicata*.

29 Observational climate data held valuable information for predicting species' responses to
30 a climate change experiment in this ecosystem. Treatment responses often matched the direction
31 of predicted responses even when the effects were not significant. We were better able to predict
32 species' responses to the drought treatment than to the control and irrigation treatments. This
33 suggests that soil moisture is an important factor in predicting the population dynamics of these
34 species but only when water is truly limiting.

35 **Introduction**

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

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

57 come with artifacts that may make them less than ideal models for understanding and predicting
58 the effects of future variation in climate (Wolkovich et al., 2012).

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

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

87 There is abundant evidence that plant performance shows high year to year variation, both
88 at the level of individual growth, survival and reproduction and total population abundance. Pre-
89 cipitation is often especially important in driving interannual variation in net primary productivity
90 among terrestrial plants (Knapp and Smith, 2001; Hsu and Adler, 2014), the annual growth rings
91 in trees (Yang et al., 2014) and smaller plants (Srur and Villalba, 2009; Franklin, 2013), as well as
92 the germination and reproductive output of many annuals (Venable, 2007).

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

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

115 In this study, we report how the four dominant plant species at the USSES respond to a
116 five year drought and irrigation experiment and use the results to address two research questions:

117 first, how much do the growth, recruitment and survival of our target species differ between the
118 precipitation manipulation treatments? If our experiment does affect species vital rates we interpret
119 that as strong evidence that changes in precipitation should have an effect on populations in the
120 future. Second, can we predict each species' response to the experimental conditions based on
121 how they respond to natural climate variation in the observational data? If models based on
122 observational data can predict the response of species to this experiment this should give us more
123 confidence in longterm population monitoring data to predict species response to climate change.

124 Methods

125 Study site and data set description

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

132 Scientists at the USSES established 26 1-m² quadrats between 1926 and 1932. Eighteen
133 quadrats were distributed among four ungrazed exclosures, and eight were distributed in two pad-
134 docks grazed at medium intensity spring through fall. All quadrats were located on similar topog-
135 raphy and soils. In most years until 1957, all individual plants in each quadrat were mapped using
136 a pantograph (Blaisdell, 1958). The historical data set is public and available online (Zachmann
137 et al., 2010). In 2007, we located 14 of the original quadrats, all of which are inside permanent live-
138 stock exclosures, and resumed annual mapped censusing using the traditional pantograph method.
139 Daily temperature and precipitation has been monitored throughout this period at a climate sta-
140 tion located at the USSES headquarters (station id: GHCND:USC00102707) which located within
141 2 km of the research plots. We downloaded daily and monthly tmin, tmax, and precipitation data
142 from the National Climate Data Centers online database.

143 We extracted data on survival, growth, and recruitment from the mapped quadrats based
144 on plants' spatial locations. Our approach tracks genets representing individual plants. For the

145 shrub, each genet is associated with the basal position of a stem. For the bunchgrasses, each genet
146 represents a spatially distinct polygon in the mapped quadrat. These genets may fragment and/or
147 coalesce over time. Each mapped polygon is classified as a surviving genet or a new recruit based
148 on its spatial location relative to genets present in previous years (Lauenroth and Adler, 2008). We
149 modeled vital rates using data from 21 year-to-year transitions between 1929 and 1957, and four
150 year-to-year transitions from 2007 to 2011.

151 **Precipitation experiment**

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

167 We monitored soil moisture and air temperature in four of the precipitation experiment plot
168 pairs using Decagon Devices (Pulman Washington) 5TM and EC-5 soil moisture sensors and 5TE
169 temperature sensors. We installed two soil moisture sensors in each monitored plot at 5 cm and
170 two at 25 cm deep in the soil. Air temperature was measured underneath the roofing of the shelter
171 at 30 cm above ground. For each pair of manipulated plots, we also installed sensors in a nearby
172 area to measure ambient rainfall and temperature. Data were logged automatically every four
173 hours. We augmented automatic monitoring of the climate in these plots with direct measurements

174 of soil moisture with a handheld EC-5 soil moisture sensor at six points around all 16 plots on
175 6/6/2012, 4/29/2015, 5/7/2015, 6/9/2015 and 5/10/2016. We analyzed these spot measurements
176 for significant treatment effects on soil moisture using a linear mixed effects model with the *lmer*
177 package in *R*, with plot, plot group, and date as random effects in the model ((Bates et al., 2015)).

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

184 **Soil moisture modeling**

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

198 We averaged daily soil moisture predictions from SOILWAT from upper 40 cm of soil and
199 then averaged these seasonally to serve as the covariates in the vital rate regressions for each species.
200 Because we did not monitor soil moisture directly in all control, drought and irrigation plots, we
201 used a model to describe the average treatment effects on soil moisture during the course of the
202 experiment. To do this we first averaged observed soil moisture data by date and plot and then

203 standardized these by the mean and standard deviation of the control soil moisture conditions
204 observed within each plot group. We then found the difference between the soil moisture in the
205 treated plots and the ambient conditions. We then modeled these treatment effects as a function of
206 season and whether a day was rainy or dry. We expected that our drought and irrigation treatments
207 might be more effective during rainy weather than during dry weather. Rainy days were defined
208 as any days when any precipitation was recorded and average temperatures were above 3 degrees
209 C. The day immediately following rainfall was also classified as rainy. We fit this model using
210 the *lmer* package in *R* with random effects for plot group and date (Bates et al., 2015). We then
211 used this model to predict the treatment effects on soil moisture for the entire study period from
212 the ambient soil moisture values predicted from the SOILWAT model described above. These
213 adjusted soil moisture values reflected the average season and rainfall dependent effects of the
214 experimental treatments on soil moisture and could be used as covariates for predicting the effects
215 of our manipulation on each species demographic rates.

216 Overview of the analysis

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

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

230 Next in order to test how well we could predict the responses in the experimental plots, we fit
231 two classes of models to the observational dataset only. Thus when we used these models to predict

232 the response of each species to the five years of experimental data we are generating predictions
 233 for out of sample data. We refer to the first type of model as the "no climate" models. These
 234 models include parameters for the effects of competition on each vital rate and the size dependence
 235 of survival and growth but they do not include climate or treatment effects. Lastly, we fit the
 236 "climate" models which include the effects of annual variation in soil moisture on each vital rate.
 237 The no climate model provides us a baseline by which to measure the accuracy of the predictions
 238 from the climate model. Because much of the variation in growth, survival and recruitment in this
 239 system can be explained by plant size and competition, we expect that these two models will make
 240 similar predictions for individual plant performance in the experiment. However, if the climate
 241 model makes more accurate predictions than the no climate model, this indicates that the climate
 242 parameters it includes contain useful information for prediction.

243 Statistical models of vital rates

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

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

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

253 where φ is the spatial group dependent intercept, γ is a year-effect, β is year-dependent coefficient
 254 that represents the effect of log genet size, u , on survival in year t . $\boldsymbol{\omega}$ is a vector of interaction
 255 coefficients which determine the impact of crowding, \mathbf{W} , by each species on the focal species. The
 256 vector \mathbf{W} includes crowding from the four dominant species, *A. tripartita*, *P. spicata*, *H. comata*,
 257 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
 258 R. This model is the no climate model for survival.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

301 where λ is the mean intensity and θ is the size parameter. In turn, λ depends on the composition
 302 of the quadrat in the previous year:

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

303 where the superscript R refers to Recruitment, $C'_{jq,t}$ is the ‘effective cover’ (cm^2) of species j in
 304 quadrat q at time t , φ is a group dependent intercept, γ is a random year effect, $\boldsymbol{\omega}$ is a vector of

305 coefficients that determine the strength of intra- and interspecific density-dependence, and \mathbf{C}' is
306 the vector of “effective” cover of each species in the community. Following previous work (Adler
307 et al., 2010), we treated year as a random factor allowing intercepts to vary among years.

308 Because plants outside the mapped quadrat could contribute recruits to the focal quadrat or
309 interact with plants in the focal quadrat, we estimated effective cover as a mixture of the observed
310 cover, C , in the focal quadrat, q , and the mean cover, \bar{C} , across the spatial location, g , in which
311 the quadrat is located: $C'_{jq,t} = p_j C_{jq,t} + (1 - p_j) \bar{C}_{jg,t}$, where p is a mixing fraction between 0 and
312 1 that was estimated as part of fitting the model.

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

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

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

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

332 **Selecting soil moisture covariates**

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

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

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

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

357 To simulate changes in cover in each quadrat from year t to year $t + 1$, we initialized the
358 IBM with the observed genet sizes and locations of the four focal species observed in year t in each
359 quadrat. For every individual genet in a quadrat, we projected its size and survival probability in
360 the next year using the growth and survival models and the appropriate crowding and soil moisture

361 or treatment covariates for that year and quadrat. Likewise we projected the number of new recruits
 362 in the quadrat in the next year using the recruitment model. We calculated the expected cover in
 363 year $t + 1$ as the total area of new recruits, plus the sum of the predicted area of each existing plant
 364 at time $t + 1$ multiplied by each plant's expected survival probability from time t to $t + 1$.

365 We generated predictions using 1000 samples from the posterior distributions of each model
 366 parameter which allowed us to carry forward all of the uncertainty of the fitted vital rate models into
 367 our cover predictions. Because we were interested in comparing model predictions to observations,
 368 and were not interested in the effects of demographic stochasticity, we used a deterministic version
 369 of the models (e.g., recruitment is the λ of (6), rather than a random draw from a negative binomial
 370 distribution with a mean of λ).

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

373 Quantifying predictive accuracy

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

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

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

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

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

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

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

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

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

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

402 Results

403 Effects on soil moisture

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

408 The continuously recorded soil moisture data also showed treatment effects, but these were
 409 weaker on average than the spot measurements and depended on season and recent rainfall (table

410 2; fig 2). We saw weaker effects during the spring than during the fall and summer: the drought
411 plots were about 20-30% drier than ambient in the fall and summer but only 7 to 14% drier during
412 the spring, while the irrigated plots were 30% wetter during the fall and summer but only 20-25%
413 wetter during the spring. Treatment differences were slightly larger during rainy periods, especially
414 in the spring (rainfall effect in table 2). We did not find evidence that the drought shelters and the
415 irrigation treatments consistently affected air temperature at 30 cm above the plots.

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

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

430 Effects on cover and vital rates

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

435 Our treatment models fit to the experimental and observational data indicated a variety of
436 treatment effects on the vital rates of each species. Based on the WAIC scores with and without the
437 size by treatment effects, we retained size by treatment effects in the growth models for *A. tripartita*
438 and *P. secunda*, and the survival model for *P. secunda*. For *A. tripartita* we found significant size

439 by treatment effects of drought: drought had positive effects on plants of average size and smaller
440 (fig 5), but plants larger than the mean size by more than 1.5 standard deviations grew slightly less
441 in the drought treatment than in the controls. *A. tripartita* showed the opposite response in the
442 irrigated plots, (although the irrigation parameters were not significant at the 95% confidence level):
443 irrigation reduced growth for small plants while irrigation increased growth of plants more than
444 1.5 standard deviations larger than the mean size. Drought led to a strong (but not significant)
445 decrease in *H. comata* growth, while irrigation had no effect on growth. Like *A. tripartita*, we
446 saw size by treatment effects on *P. secunda* growth, with the negative effects of drought becoming
447 greater for larger plants. *P. secunda* showed the opposite response in the irrigation plots with larger
448 plants showing the largest increase in growth in response to irrigation (although not significant).
449 *P. spicata* growth was relatively unaffected by the drought and irrigation treatments.

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

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

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

463 Effects of soil moisture on vital rates

464 We choose three seasonal soil moisture variables for each species' based on their correlation with the
465 random year effects in the no climate model (table 7). We included size by soil moisture variables
466 for *A. tripartita* and *P. secunda* based on the treatment response we observed in the experiment.
467 All three time lags and all three seasons show up in the selected variables. After fitting the vital rate

468 models with the selected soil moisture variable we observed a trend towards positive soil moisture
469 effects on growth of all three grasses (fig 8). For *H.comata* the soil moisture of the most recent
470 summer (*summer*₁) had a significantly positive effect while the soil moisture during *summer*₀ and
471 *fall_{lag}* were also positive but not significant. For *A. tripartita* *fall*₀ and *summer*₀ had strong
472 negative effects on growth. There were also strong positive size by climate interaction effects for
473 these variables: soil moisture had a stronger negative effect on small plants and a positive effect
474 only on the largest plants (fig A-1).

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

482 There were only two significant effects of soil moisture on recruitment: *fall_{lag}* soil moisture
483 had a positive effect on *P. secunda*, and *summer_{lag}* soil moisture had a negative effect on *P.*
484 *spicata* recruitment (fig 10). Soil moisture of *summer*₀ also had a strong negative effect on *P.*
485 *spicata* recruitment.

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

488 Evaluating the predictions

489 For most models adding climate covariates did not improve our ability to predict species responses
490 in the experiment (table 8). However, the climate models did improve overall prediction MSE for
491 growth of *A. tripartita* and growth and survival of *P. secunda* (table 8). In terms of lppd, the
492 climate model outperformed the no climate model in six out of twelve models: for *A. tripartita*
493 growth, *H. comata* recruitment, *P. secunda* growth and survival and *P. spicata* recruitment (table
494 8).

495 When we look at the predictions for each treatment separately we see that climate covariates
496 improved model predictions more often in the drought treatments than in the control or irrigation

497 treatments (table A-1). For all four species, the climate model outperformed the no climate model
498 for predicting the response of growth to drought in terms of lppd (table A-1). The climate model
499 also outperformed the no climate model for predicting irrigation effects on growth for all species
500 except *H. comata*.

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

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

527 model only for *P. spicata* and *P. secunda* in the control treatment and *P. secunda* and *H. comata*
528 in the drought treatment.

529 Discussion

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

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

552 In contrast to the effects on the grasses, the effects on *A. tripartita* are more complicated.
553 The effects of drought and irrigation on *A. tripartita* growth were highly dependent on plant size
554 (figs 5): drought had a positive effect except possibly for the largest plants (fig A-1). This result
555 matches the negative effects of precipitation on this species estimated by (Adler et al., 2012) and

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

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

585 A similar result was recently reported from another precipitation experiment in a warm
586 desert grassland: the effect of increased soil moisture on grass productivity in that system showed a
587 concave down pattern of response to increased soil moisture (Gherardi and Sala, 2015). Increasing
588 precipitation had a weaker effect on annual grass productivity than decreasing precipitation by
589 an equal amount. A pattern like this is qualitatively similar to our finding significant decreases in
590 either the cover or performance of the three grasses in drought plots, but not seeing any statistically
591 significant increase in performance in the irrigated plots.

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

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

610 Using the IBM, we demonstrated that the climate model predictions of the vital rates could
611 be used to generate population-level predictions that were also better than the no climate models in
612 some cases: the climate model improved one step ahead predictions of cover changes of *P. spicata*
613 and *P. secunda* (table 9). In the drought treatment our one step ahead cover predictions for *H.*
614 *comata* and *P. secunda* were also better than the no climate model. In this analysis, we either

615 used the climate model, or the no climate model to generate predictions for the IBM. In theory
616 we could have refined our predictions by climate model to predict all three vital rates in the IBM.
617 Improvements of our cover predictions may have been possible if we only included climate models
618 in the IBM when those models improved vital rate predictions—for instance, the *A. tripartita* IBM
619 would only include the climate model for growth but not for survival and recruitment.

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

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

642 Our results give us more confidence that historical observational data can be used to de-
643 tect and predict the demographic effects of climate change among sagebrush steppe plants. This
644 should encourage more researchers to try and use observational data to predict the future in both

645 experimental and natural settings (Houlahan et al., 2016). Nevertheless, our success at predicting
646 the short-term response of two species to a small-scale climate manipulation is not likely to be very
647 reassuring to applied ecologists and resource managers wishing to make accurate predictions about
648 the effects of climate change in large complex systems. Clearly more work is needed to learn how
649 best to meet the challenge of prediction in ecological systems. Towards that goal, perhaps the best
650 way forward is to conduct more tests like this one.

651 **Acknowledgements**

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

657 **References**

- 658 Adler, P. B., K. M. Byrne, and J. Leiker. 2013. Can the past predict the future? Experimental
659 tests of historically based population models. *Global Change Biology*, **19**:1793–1803.
- 660 Adler, P. B., H. J. Dagleish, and S. P. Ellner. 2012. Forecasting plant community impacts of
661 climate variability and change: when do competitive interactions matter? *Journal of Ecology*,
662 **100**:478–487.
- 663 Adler, P. B., S. P. Ellner, and J. M. Levine. 2010. Coexistence of perennial plants: an embarrassment
664 of niches. *Ecology Letters*, **13**:1019–1029.
- 665 Adler, P. B., J. HilleRisLambers, and J. M. Levine. 2009. Weak effect of climate variability on
666 coexistence in a sagebrush steppe community. *Ecology*, **90**:3303–3312.
- 667 Adler, P. B., A. R. Kleinhesselink, G. Hooker, B. Taylor, B. J. Teller, and S. P. Ellner. 2016. Weak
668 interspecific interactions in a sagebrush steppe: Evidence from observational data, population
669 models, and a removal experiment.
- 670 Aubry, L. M., R. F. Rockwell, E. G. Cooch, R. W. Brook, C. P. H. Mulder, and D. N. Koons.
671 2013. Climate change, phenology, and habitat degradation: drivers of gosling body condition
672 and juvenile survival in lesser snow geese. *Global Change Biology*, **19**:149–160.
- 673 Bates, D., M. Mchler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using
674 lme4. *Journal of Statistical Software*, **67**:1–48.
- 675 Bazzaz, F. A. 1979. The Physiological Ecology of Plant Succession. *Annual Review of Ecology and*
676 *Systematics*, **10**:351–371.
- 677 Blaisdell, J. P. 1958. Seasonal development and yield of native plants on the upper Snake River
678 Plains and their relation to certain climatic factors. U.S. Dept. of Agriculture.
- 679 Bradford, J. B., D. R. Schlaepfer, and W. K. Lauenroth. 2014. Ecohydrology of Adjacent Sage-
680 brush and Lodgepole Pine Ecosystems: The Consequences of Climate Change and Disturbance.
681 *Ecosystems*, **17**:590–605.

- 682 Caldwell, M. M., J. H. Richards, J. H. Manwaring, and D. M. Eissenstat. 1987. Rapid shifts in
683 phosphate acquisition show direct competition between neighbouring plants. *Nature*, **327**:615–
684 616.
- 685 Chu, C. and P. B. Adler. 2015. Large niche differences emerge at the recruitment stage to stabilize
686 grassland coexistence. *Ecological Monographs*.
- 687 Chu, C., A. R. Kleinhesselink, K. M. Havstad, M. P. McClaran, D. P. Peters, L. T. Vermeire,
688 H. Wei, and P. B. Adler. 2016. Direct effects dominate responses to climate perturbations in
689 grassland plant communities. *Nature communications*, **7**.
- 690 Compagnoni, A. and P. B. Adler. 2014. Warming, soil moisture, and loss of snow increase *Bromus*
691 *tectorum*s population growth rate. *Elementa: Science of the Anthropocene*, **2**:000020.
- 692 Dalgleish, H. J., D. N. Koons, M. B. Hooten, C. A. Moffet, and P. B. Adler. 2010. Climate influences
693 the demography of three dominant sagebrush steppe plants. *Ecology*, **92**:75–85.
- 694 Davis, M. B. and R. G. Shaw. 2001. Range Shifts and Adaptive Responses to Quaternary Climate
695 Change. *Science*, **292**:673–679.
- 696 Doak, D. F. and W. F. Morris. 2010. Demographic compensation and tipping points in climate-
697 induced range shifts. *Nature*, **467**:959–962.
- 698 Ehrlén, J., W. F. Morris, T. von Euler, and J. P. Dahlgren. 2016. Advancing environmentally explicit
699 structured population models of plants. *Journal of Ecology*, **104**:292–305.
- 700 Elmendorf, S. C., G. H. R. Henry, R. D. Hollister, A. M. Fosaa, W. A. Gould, L. Hermanutz,
701 A. Hofgaard, I. I. Jnsdttir, J. C. Jorgenson, E. Lvesque, B. Magnusson, U. Molau, I. H. Myers-
702 Smith, S. F. Oberbauer, C. Rixen, C. E. Tweedie, and M. Walker. 2015. Experiment, monitoring,
703 and gradient methods used to infer climate change effects on plant communities yield consistent
704 patterns. *Proceedings of the National Academy of Sciences*, **112**:448–452.
- 705 Franklin, R. S. 2013. Growth response of the alpine shrub, *Linanthus pungens*, to snowpack and
706 temperature at a rock glacier site in the eastern Sierra Nevada of California, USA. *Quaternary*
707 *International*, **310**:20–33.

- 708 Franks, S. J. 2011. Plasticity and evolution in drought avoidance and escape in the annual plant
709 *Brassica rapa*. *New Phytologist*, **190**:249–257.
- 710 Gelman, A., J. Hwang, and A. Vehtari. 2014. Understanding predictive information criteria for
711 Bayesian models. *Statistics and Computing*, **24**:997–1016.
- 712 Gerber, B. D., W. L. Kendall, M. B. Hooten, J. A. Dubovsky, and R. C. Drewien. 2015. Optimal
713 population prediction of sandhill crane recruitment based on climate-mediated habitat limita-
714 tions. *Journal of Animal Ecology*, **84**:1299–1310.
- 715 Germino, M. J. and K. Reinhardt. 2014. Desert shrub responses to experimental modification of
716 precipitation seasonality and soil depth: relationship to the two-layer hypothesis and ecohydro-
717 logical niche. *Journal of Ecology*, **102**:989–997.
- 718 Gherardi, L. A. and O. E. Sala. 2013. Automated rainfall manipulation system: a reliable and
719 inexpensive tool for ecologists. *Ecosphere*, **4**:1–10.
- 720 Gherardi, L. A. and O. E. Sala. 2015. Enhanced precipitation variability decreases grass- and
721 increases shrub-productivity. *Proceedings of the National Academy of Sciences*, **112**:12735–
722 12740.
- 723 Hilborn, R. 2016. Correlation and Causation in Fisheries and Watershed Management. *Fisheries*,
724 **41**:18–25.
- 725 Houlahan, J. E., S. T. McKinney, T. M. Anderson, and B. J. McGill. 2016. The priority of prediction
726 in ecological understanding. *Oikos*, pages n/a–n/a.
- 727 Hsu, J. S. and P. B. Adler. 2014. Anticipating changes in variability of grassland production due
728 to increases in interannual precipitation variability. *Ecosphere*, **5**:1–15.
- 729 Jenouvrier, S., H. Caswell, C. Barbraud, M. Holland, J. Strve, and H. Weimerskirch. 2009. De-
730 mographic models and IPCC climate projections predict the decline of an emperor penguin
731 population. *Proceedings of the National Academy of Sciences*, **106**:1844–1847.
- 732 Jentsch, A., J. Kreyling, and C. Beierkuhnlein. 2007. A new generation of climate-change experi-
733 ments: events, not trends. *Frontiers in Ecology and the Environment*, **5**:365–374.

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

- 761 Post, E. and M. C. Forchhammer. 2002. Synchronization of animal population dynamics by large-
762 scale climate. *Nature*, **420**:168–171.
- 763 Sala, O. E., W. K. Lauenroth, and W. J. Parton. 1992. Long-Term Soil Water Dynamics in the
764 Shortgrass Steppe. *Ecology*, **73**:1175–1181.
- 765 Srur, A. M. and R. Villalba. 2009. Annual growth rings of the shrub *Anarthrophyllum rigidum*
766 across Patagonia: Interannual variations and relationships with climate. *Journal of Arid Envi-*
767 *ronments*, **73**:1074–1083.
- 768 Sturges, D. L. 1989. Response of Mountain Big Sagebrush to Induced Snow Accumulation. *Journal*
769 *of Applied Ecology*, **26**:1035–1041. ArticleType: research-article / Full publication date: Dec.,
770 1989 / Copyright 1989 British Ecological Society.
- 771 Teller, B. J., P. B. Adler, C. B. Edwards, G. Hooker, and S. P. Ellner. 2016. Linking demography
772 with drivers: climate and competition. *Methods in Ecology and Evolution*, **7**:171–183.
- 773 Tredennick, A. T., M. B. Hooten, and P. B. Adler. 2016. Do we need demographic data to forecast
774 plant population dynamics? *Methods in Ecology and Evolution*, pages n/a–n/a.
- 775 Tylianakis, J. M., R. K. Didham, J. Bascompte, and D. A. Wardle. 2008. Global change and species
776 interactions in terrestrial ecosystems. *Ecology Letters*, **11**:1351–1363.
- 777 Venable, D. L. 2007. Bet Hedging in a Guild of Desert Annuals. *Ecology*, **88**:1086–1090.
- 778 Walther, G.-R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J.-M. Fromentin,
779 O. Hoegh-Guldberg, and F. Bairlein. 2002. Ecological responses to recent climate change. *Nature*,
780 **416**:389–395.
- 781 Wang, J., P. M. Rich, and K. P. Price. 2003. Temporal responses of NDVI to precipitation and
782 temperature in the central Great Plains, USA. *International Journal of Remote Sensing*, **24**:2345–
783 2364.
- 784 Whittaker, R. H. 1975. *Communities and ecosystems*. Macmillan, New York, second edition edition.
- 785 Williams, J. W. and S. T. Jackson. 2007. Novel climates, no-analog communities, and ecological
786 surprises. *Frontiers in Ecology and the Environment*, **5**:475–482.

- 787 Wolkovich, E. M., B. I. Cook, J. M. Allen, T. M. Crimmins, J. L. Betancourt, S. E. Travers, S. Pau,
788 J. Regetz, T. J. Davies, N. J. B. Kraft, T. R. Ault, K. Bolmgren, S. J. Mazer, G. J. McCabe,
789 B. J. McGill, C. Parmesan, N. Salamin, M. D. Schwartz, and E. E. Cleland. 2012. Warming
790 experiments underpredict plant phenological responses to climate change. *Nature*, **485**:494–497.
- 791 Yang, B., C. Qin, J. Wang, M. He, T. M. Melvin, T. J. Osborn, and K. R. Briffa. 2014. A 3,500-year
792 tree-ring record of annual precipitation on the northeastern Tibetan Plateau. *Proceedings of the*
793 *National Academy of Sciences*, **111**:2903–2908.
- 794 Zachmann, L., C. Moffet, and P. Adler. 2010. Mapped quadrats in sagebrush steppe: long-term
795 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

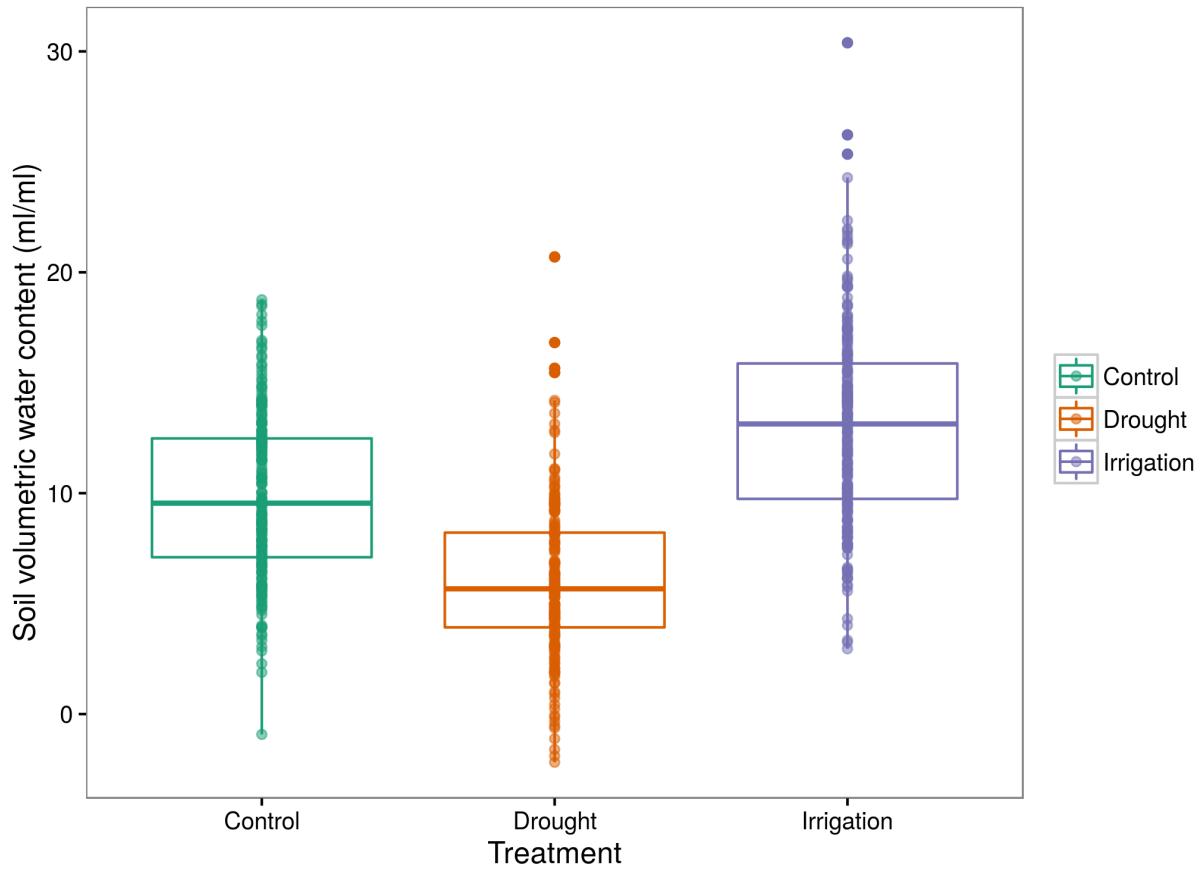


Figure 1: Soil moisture in the upper 5 cm of drought and irrigated plots compared to ambient controls. Soil moisture was measured at six locations around each plot at five different dates during the spring. Control plots were nearby areas of experiencing ambient soil moisture. Box plots show the median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

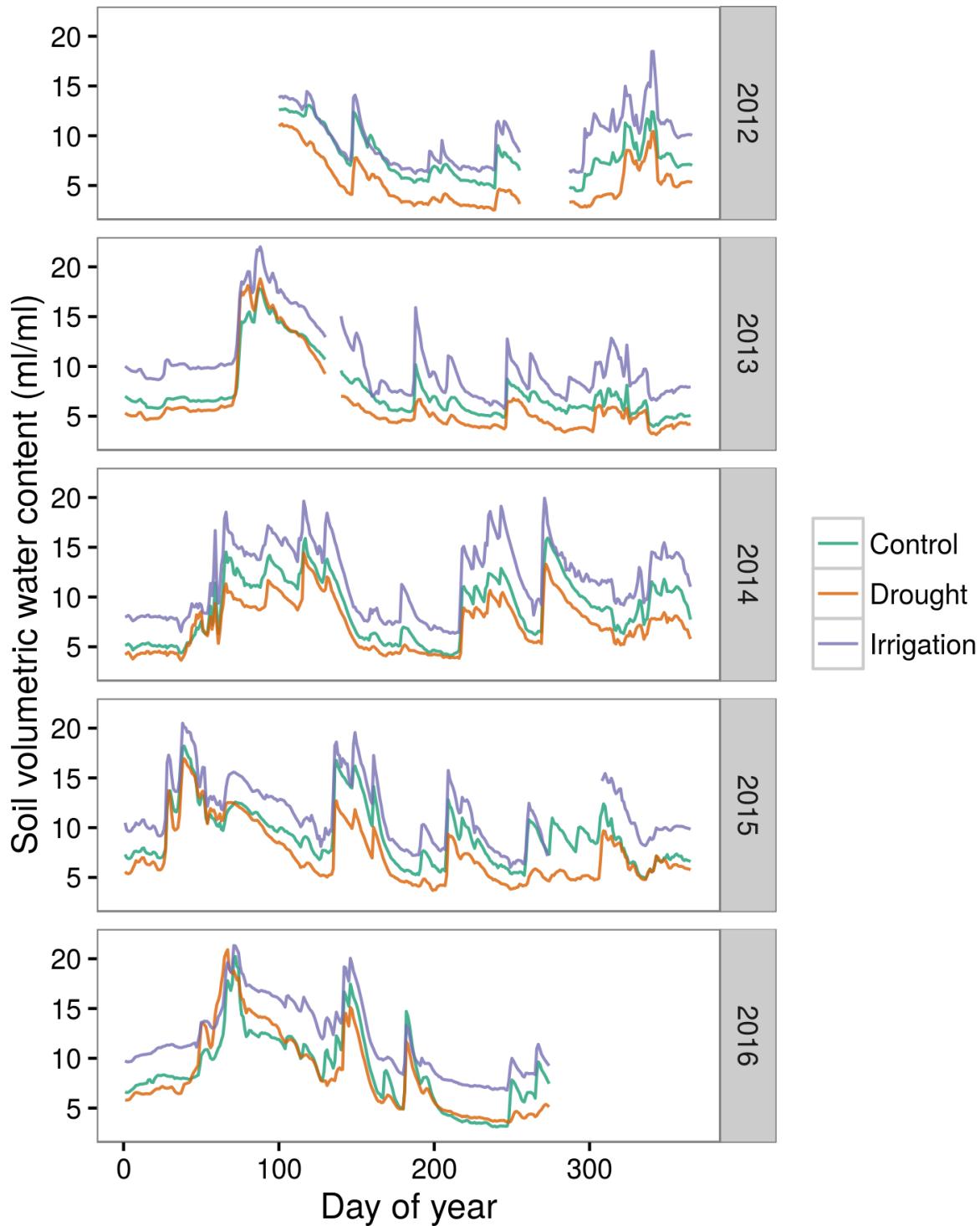


Figure 2: Average soil moisture in the control, drought, and irrigation treatments during each year of the experiment. Soil moisture was monitored in four drought plots, four irrigated plots and four ambient control plots. Two sensors were installed at 5 cm depth at each plot and two at 25 cm and data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

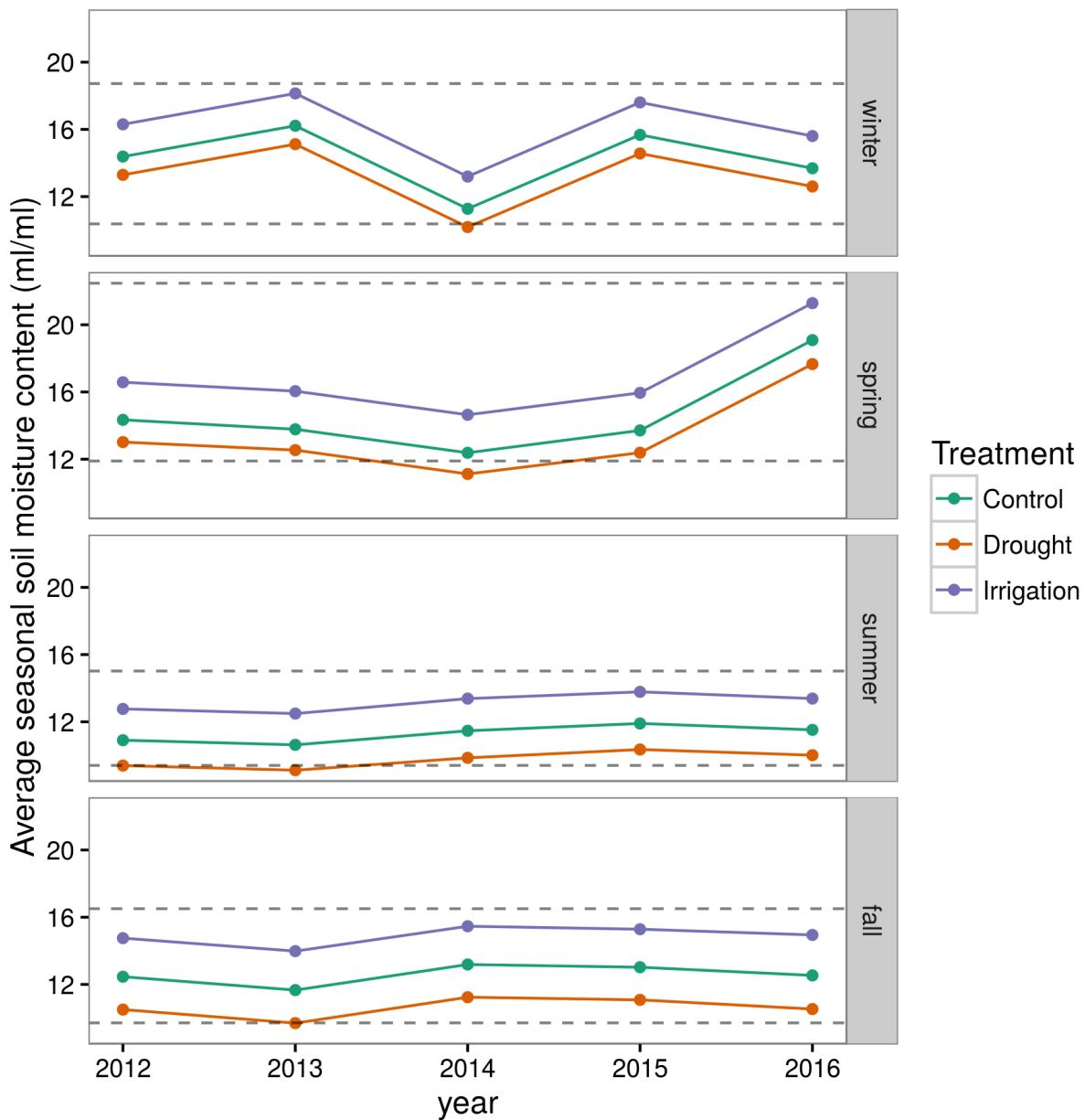


Figure 3: Average seasonal soil moisture in the control, drought, and irrigation treatments during each year of the experiment. The dashed gray lines give the 5th percentile and 95th percentile limits for seasonal soil moisture in the historical record (1929 to 2010).

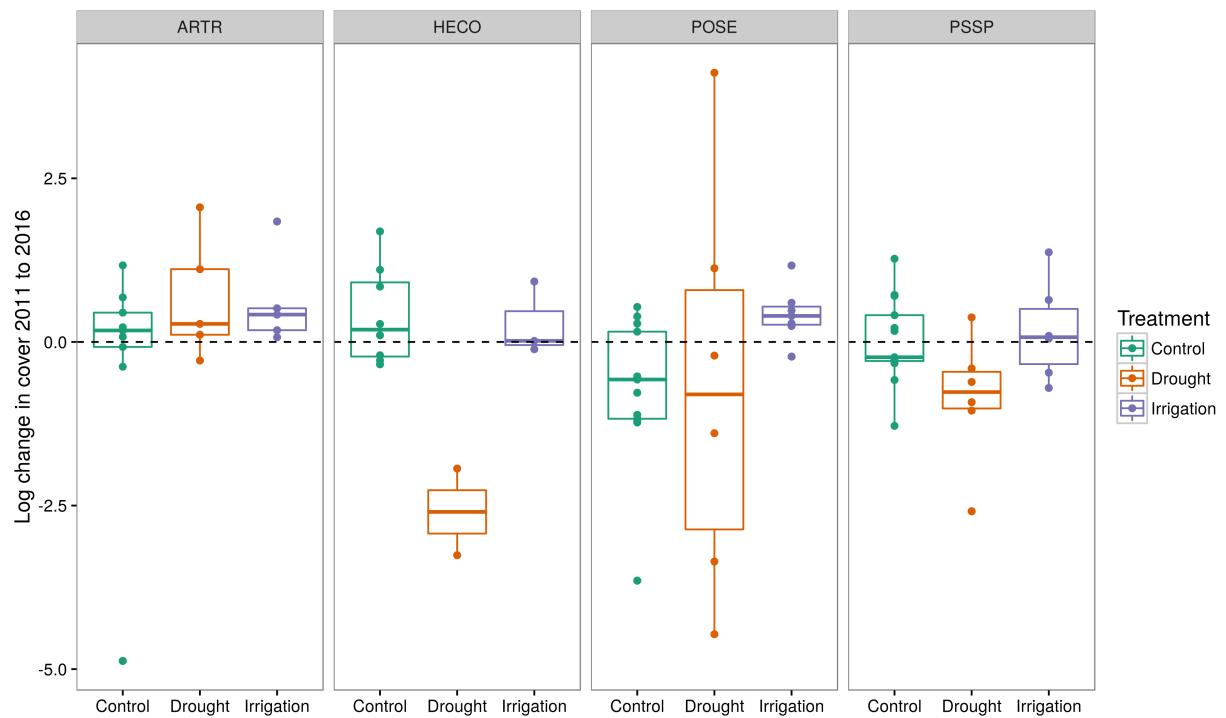


Figure 4: Log change in cover in each of the experimental plots from the pre-treatment monitoring in 2011 to the last year of the experiment in 2016. Box plots show the median cover change and the interquartile range. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

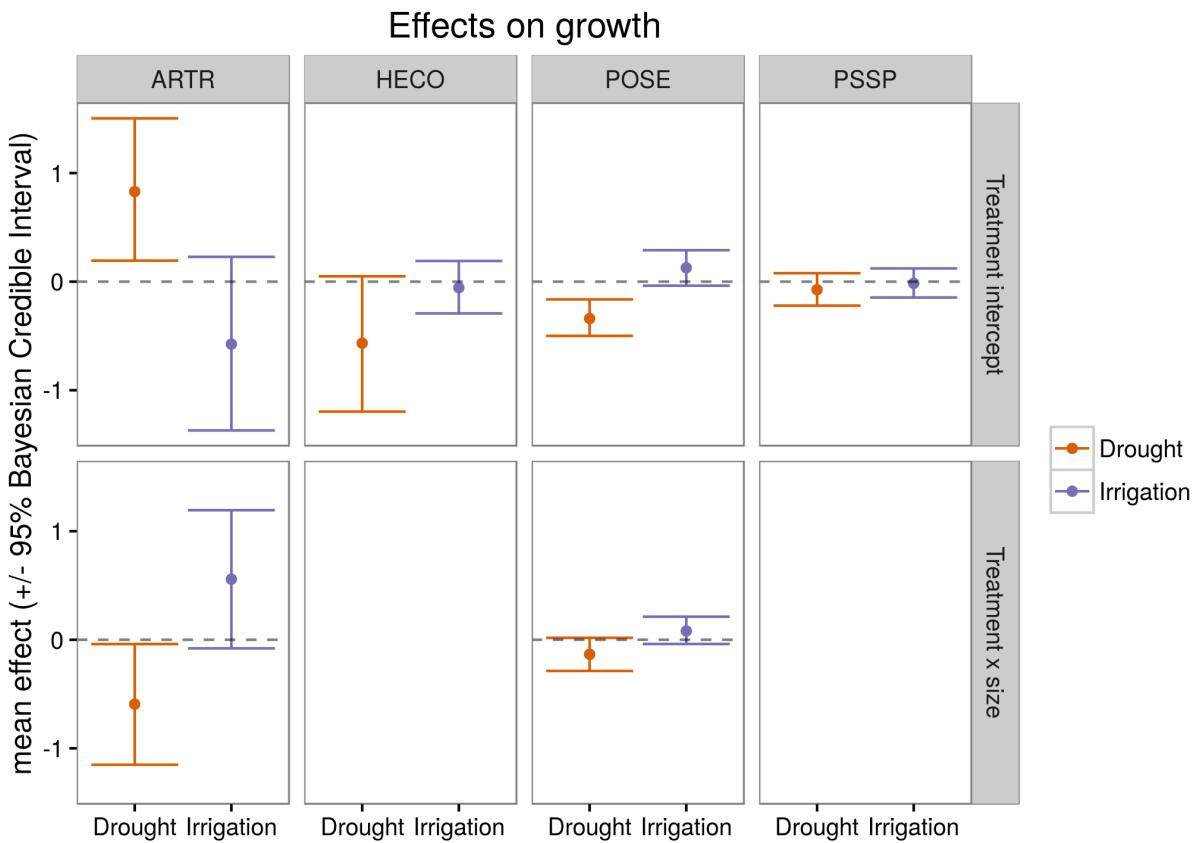


Figure 5: Parameter estimates for the effects of treatment on growth of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. Size by treatment interactions were only fit for ARTR, and POSE. Plant size was centered on mean size and scaled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

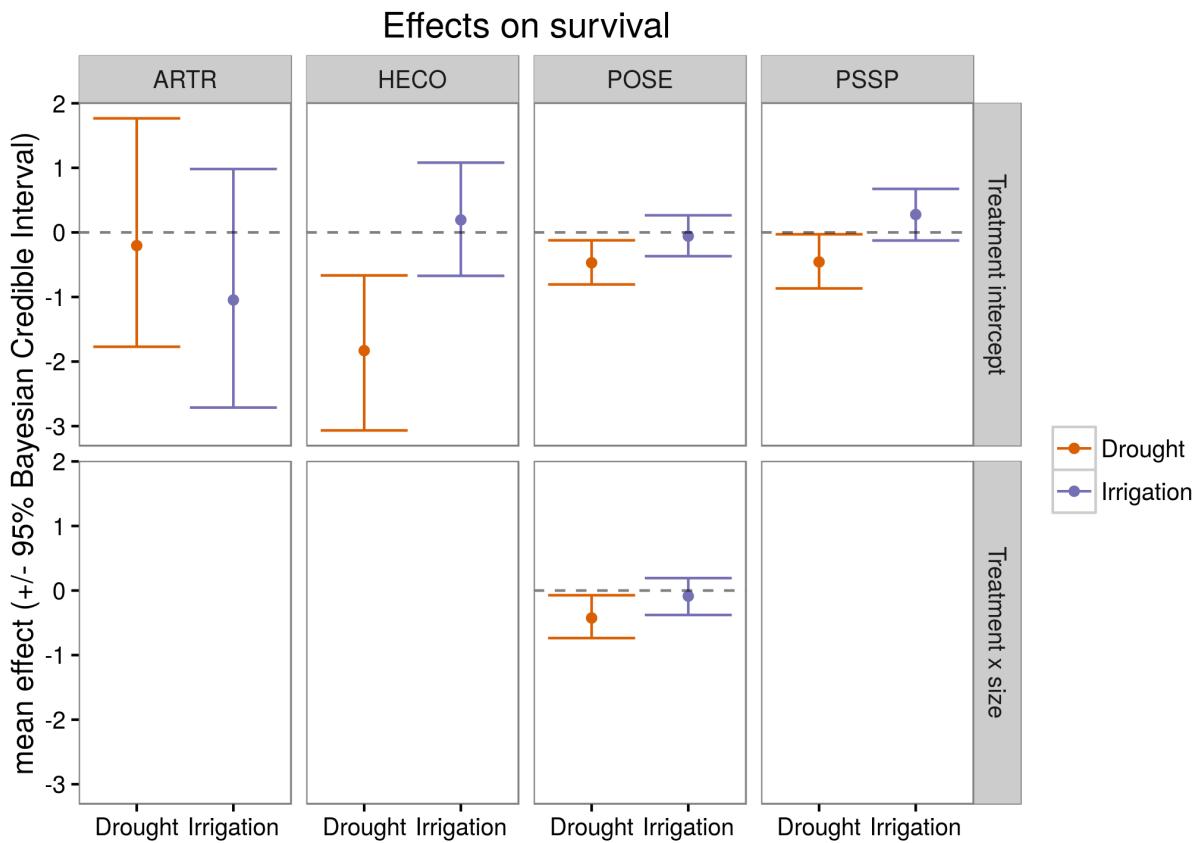


Figure 6: Parameter estimates for the effects of treatment on survival of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. Size by treatment interactions were only fit for POSE. Plant size was centered on mean size and scaled by its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

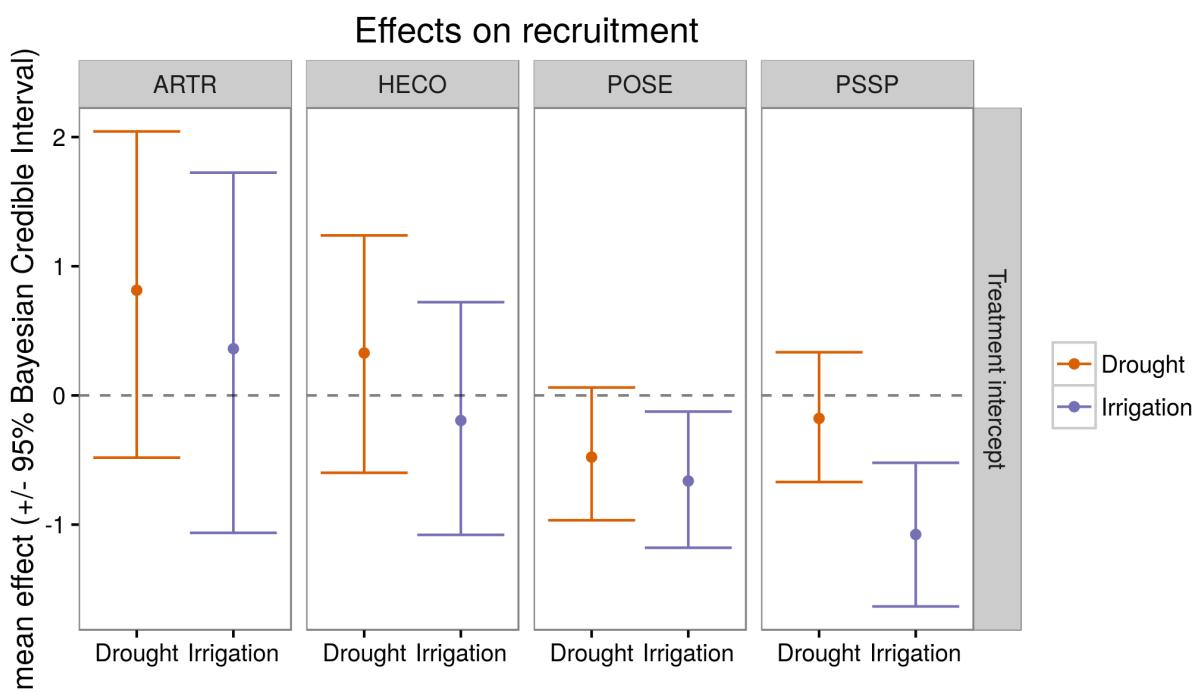


Figure 7: Parameter estimates for the effects of treatment on recruitment of all four species. We assessed a parameter as significant when the 95% Bayesian credible intervals did not overlap zero. ARTR = *A. tripartita*, HECO = *H. comata*, POSE = *P. secunda*, PSSP = *P. spicata*.

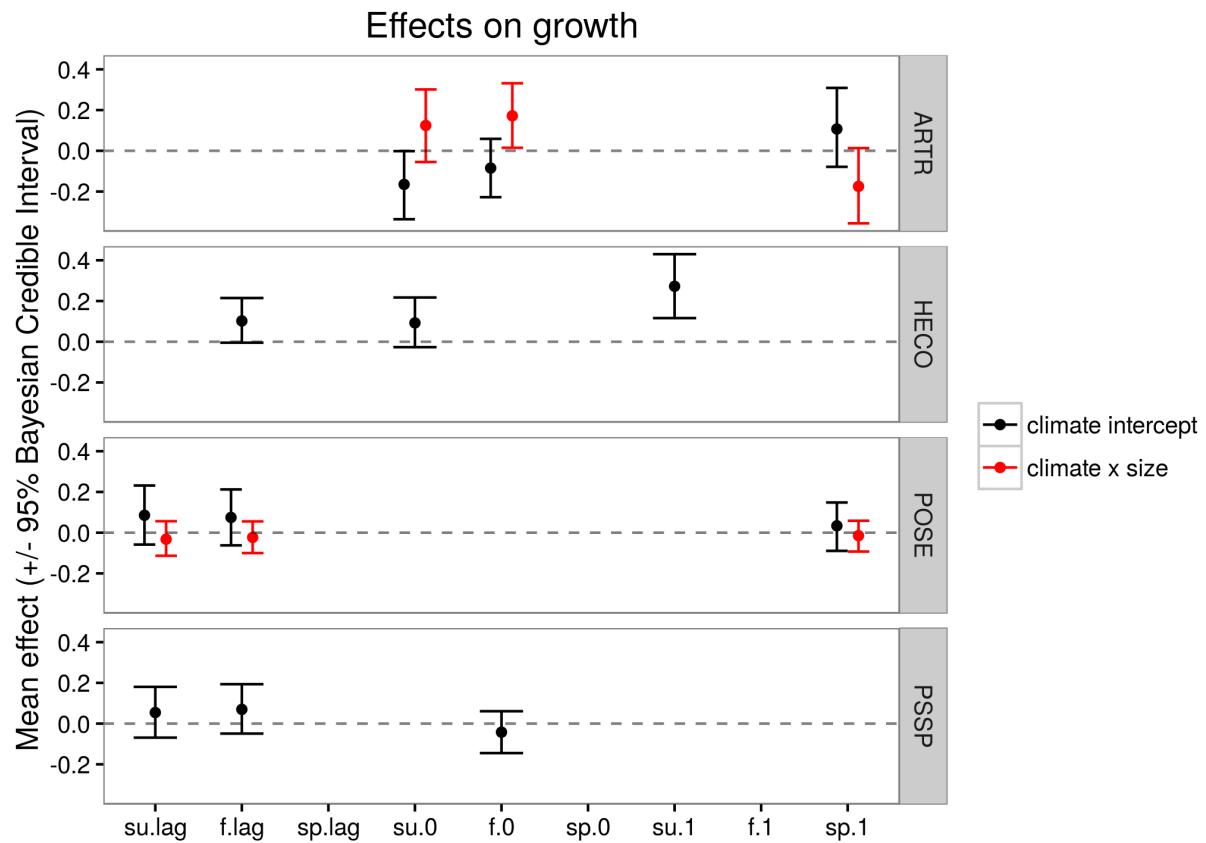


Figure 8: Parameter estimates for the selected seasonal soil moisture covariates on the growth of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. Red parameters show size x climate interaction effects. We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

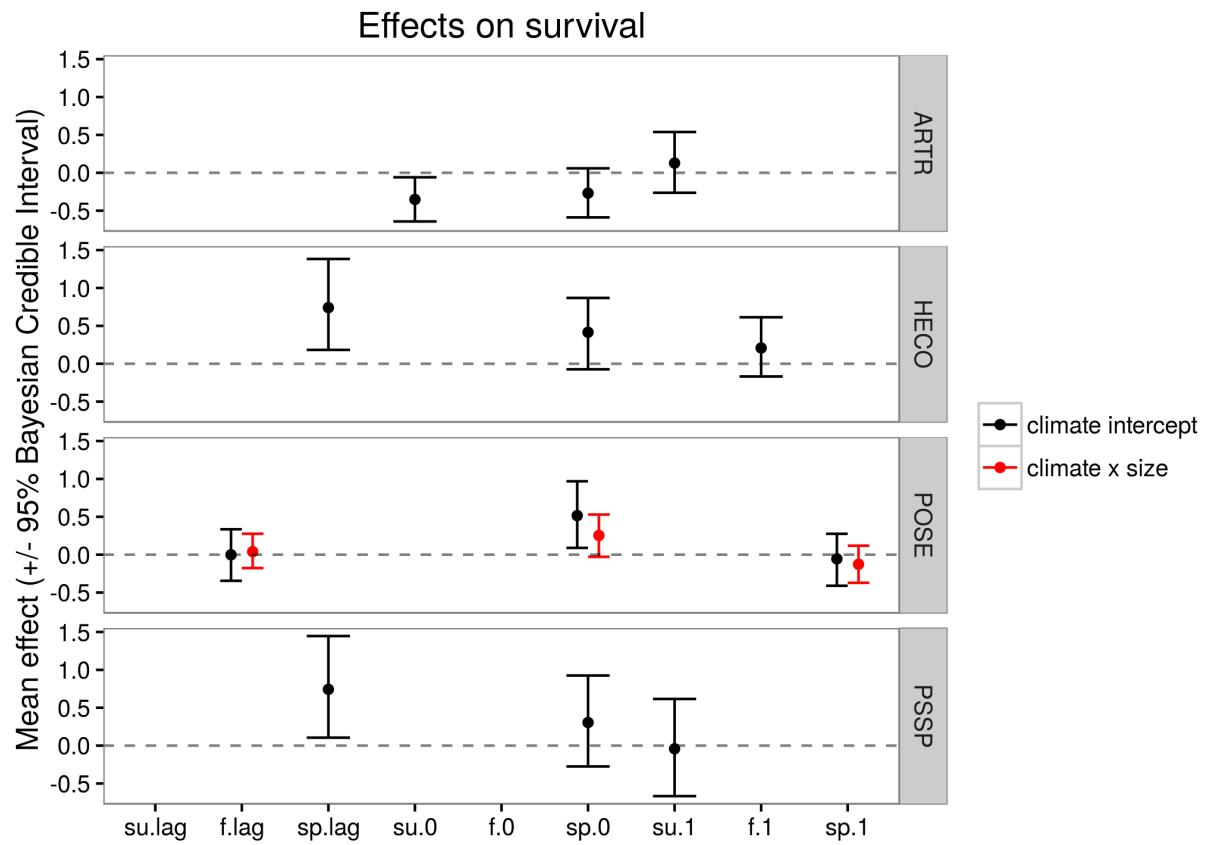


Figure 9: Parameter estimates for the selected seasonal soil moisture covariates on the survival of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. Red parameters show size x climate interaction effects.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

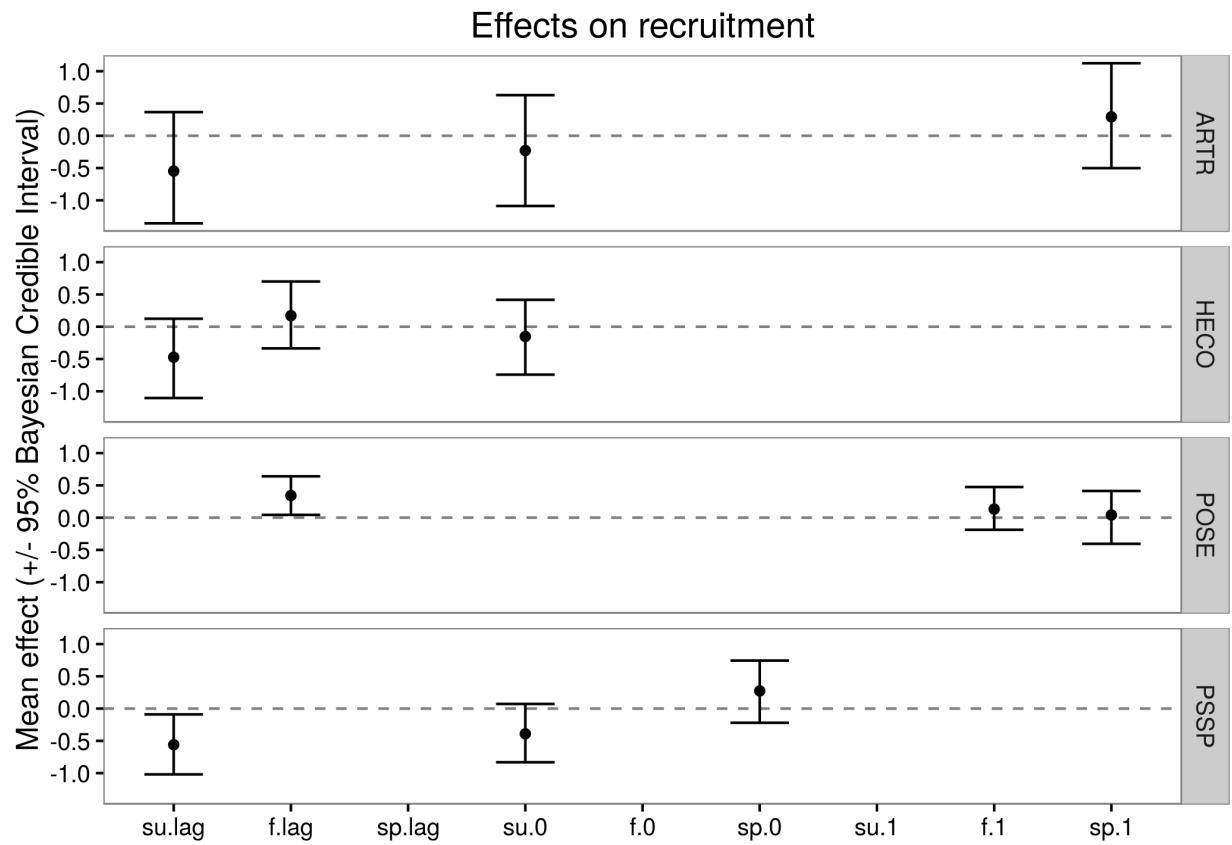


Figure 10: Parameter estimates for the selected seasonal soil moisture covariates on the recruitment of all four species. Parameters are ordered chronologically from most recent to the current growing season on the right to most distant on the left. We assessed a parameter as significant when the 95%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

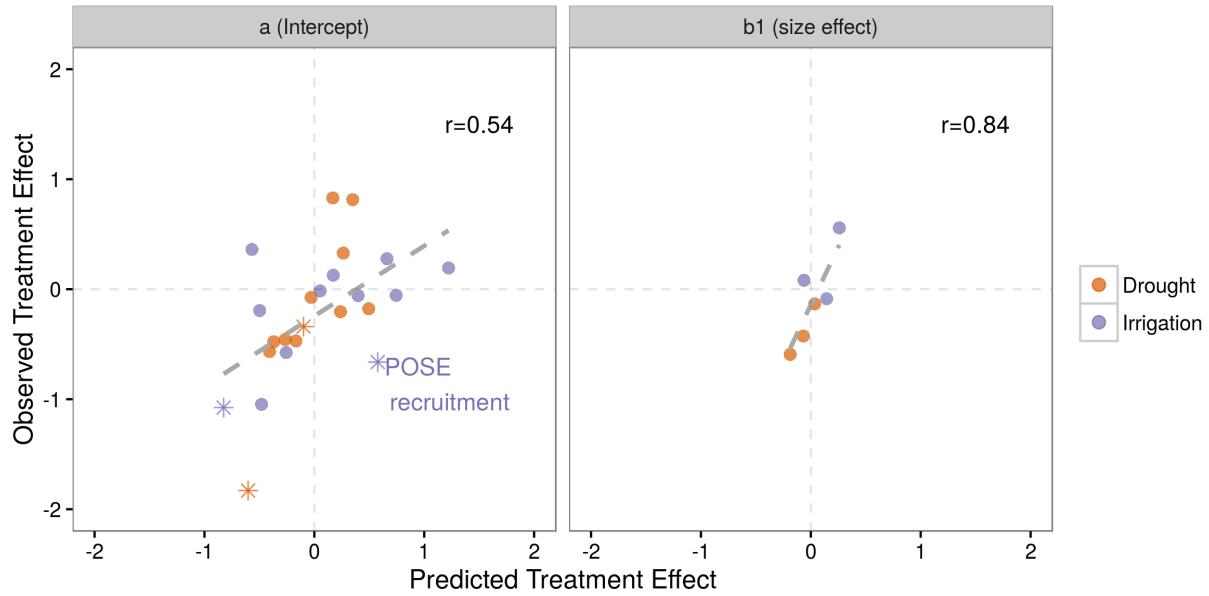


Figure 11: The treatment effects predicted by the climate model compared to the treatment effects observed for the intercept parameters (left side) and size by climate/treatment effects (right side). Parameters from all species and vital rates are shown together. The observed treatment effects come from 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.

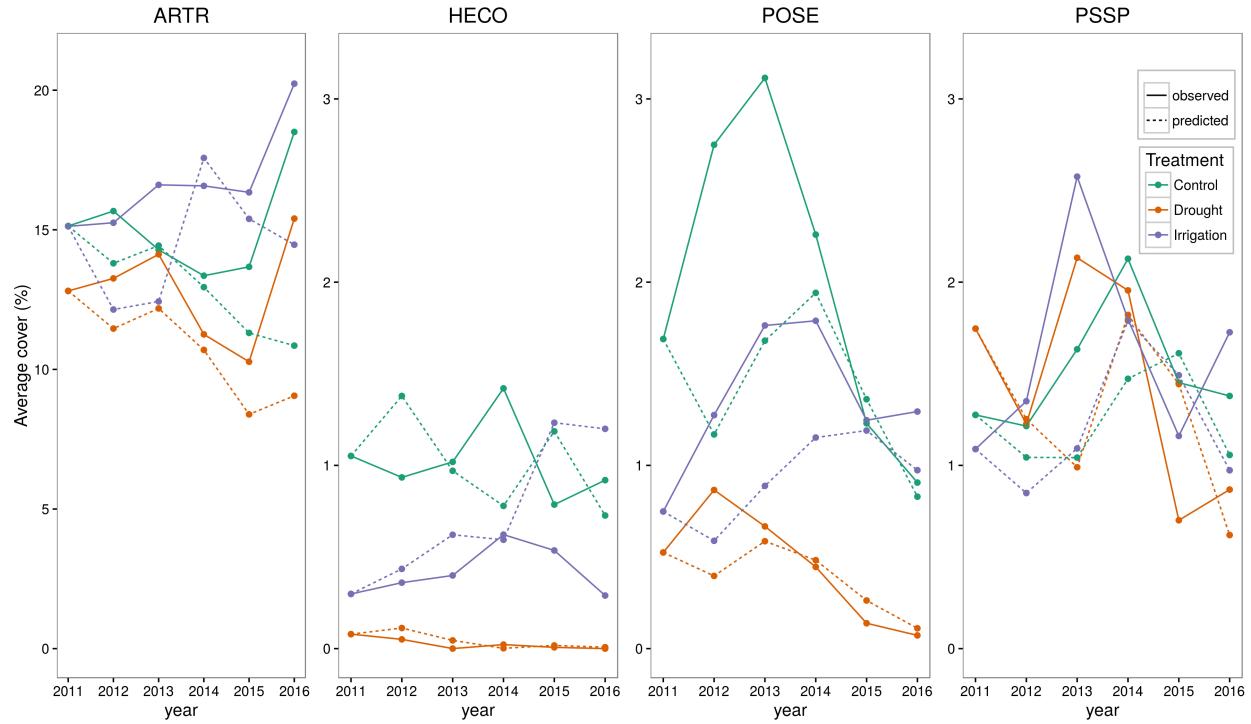


Figure 12: Observed average cover per quadrat in the experimental and control plots and one step ahead cover predictions from the climate model. Cover predictions for each year are generated from the IBM based on the observed distribution of plants in each quadrat in the current year. Quadrat cover was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

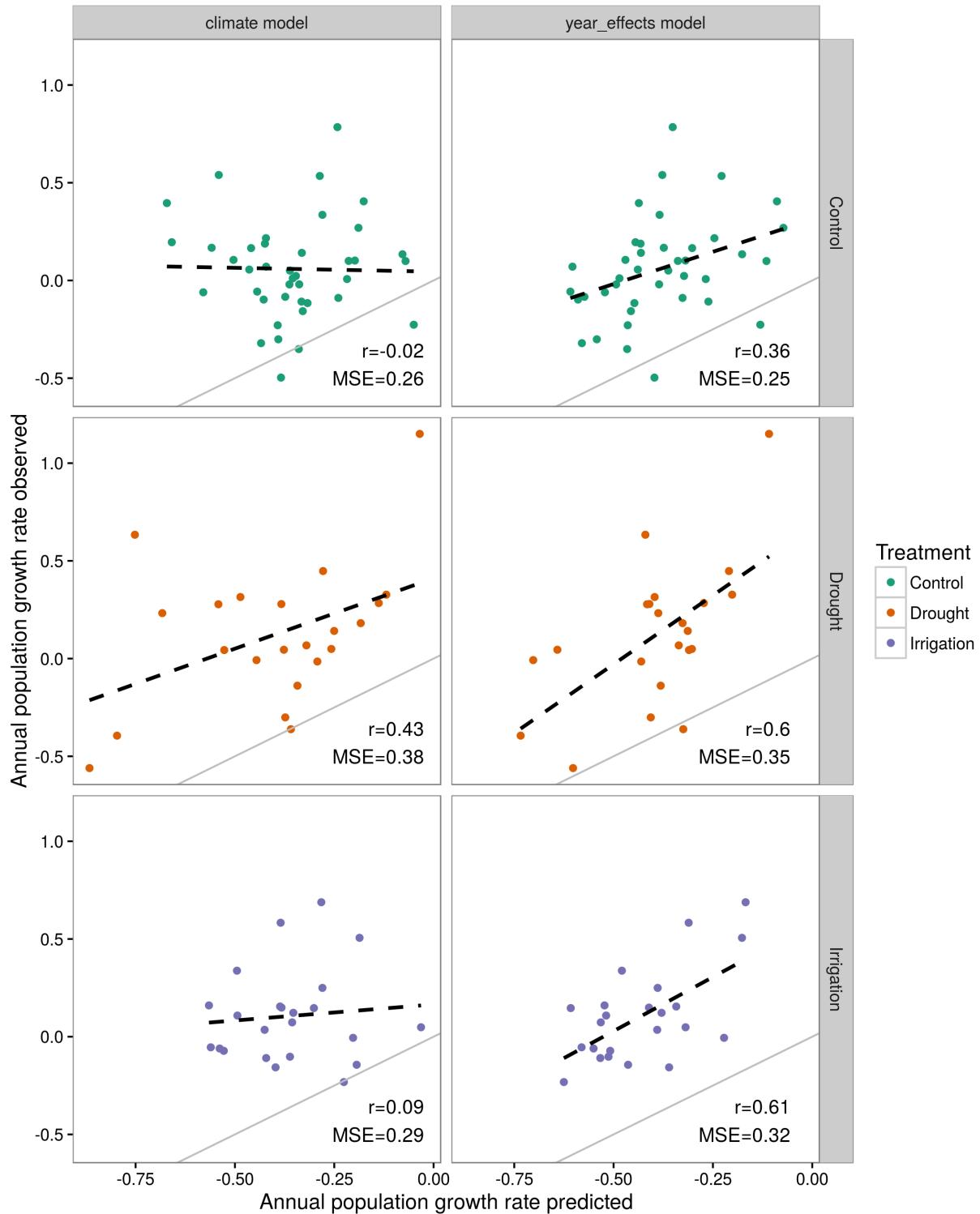


Figure 13: Observed and predicted one step ahead log change in *A. tripartita* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations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

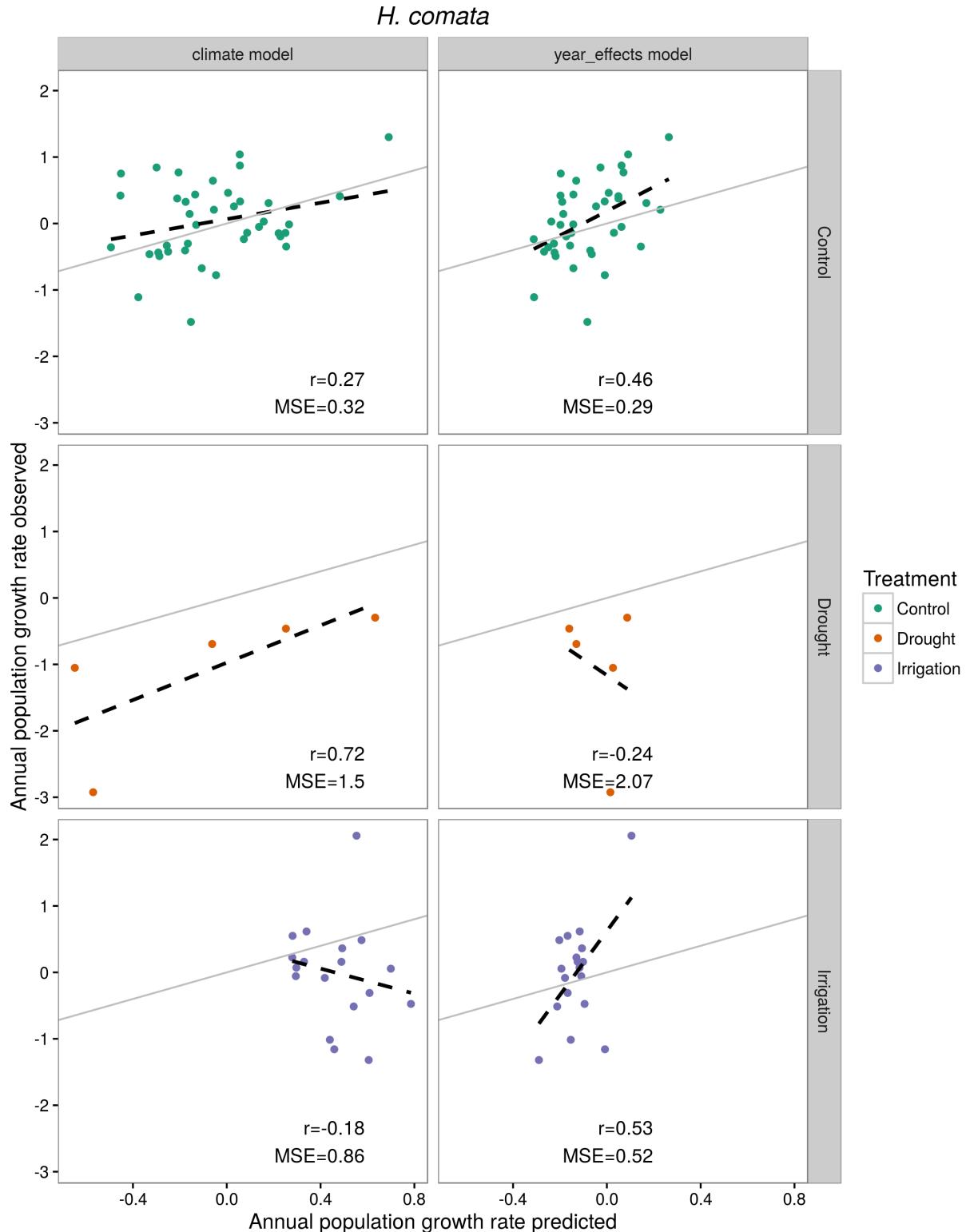


Figure 14: Observed and predicted one step ahead log change in *H. comata* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations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

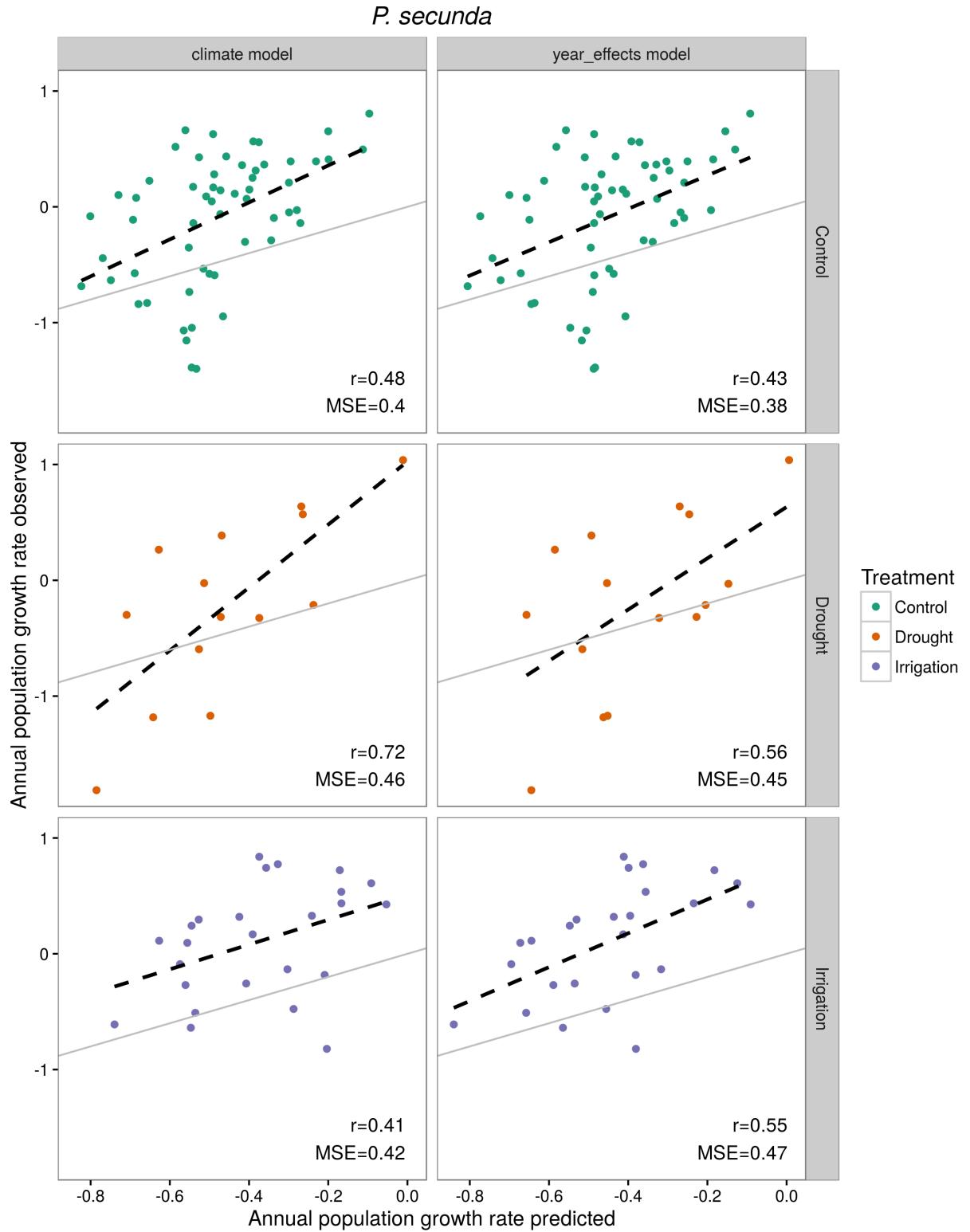


Figure 15: Observed and predicted one step ahead log change in *P. secunda* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations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

P. spicata

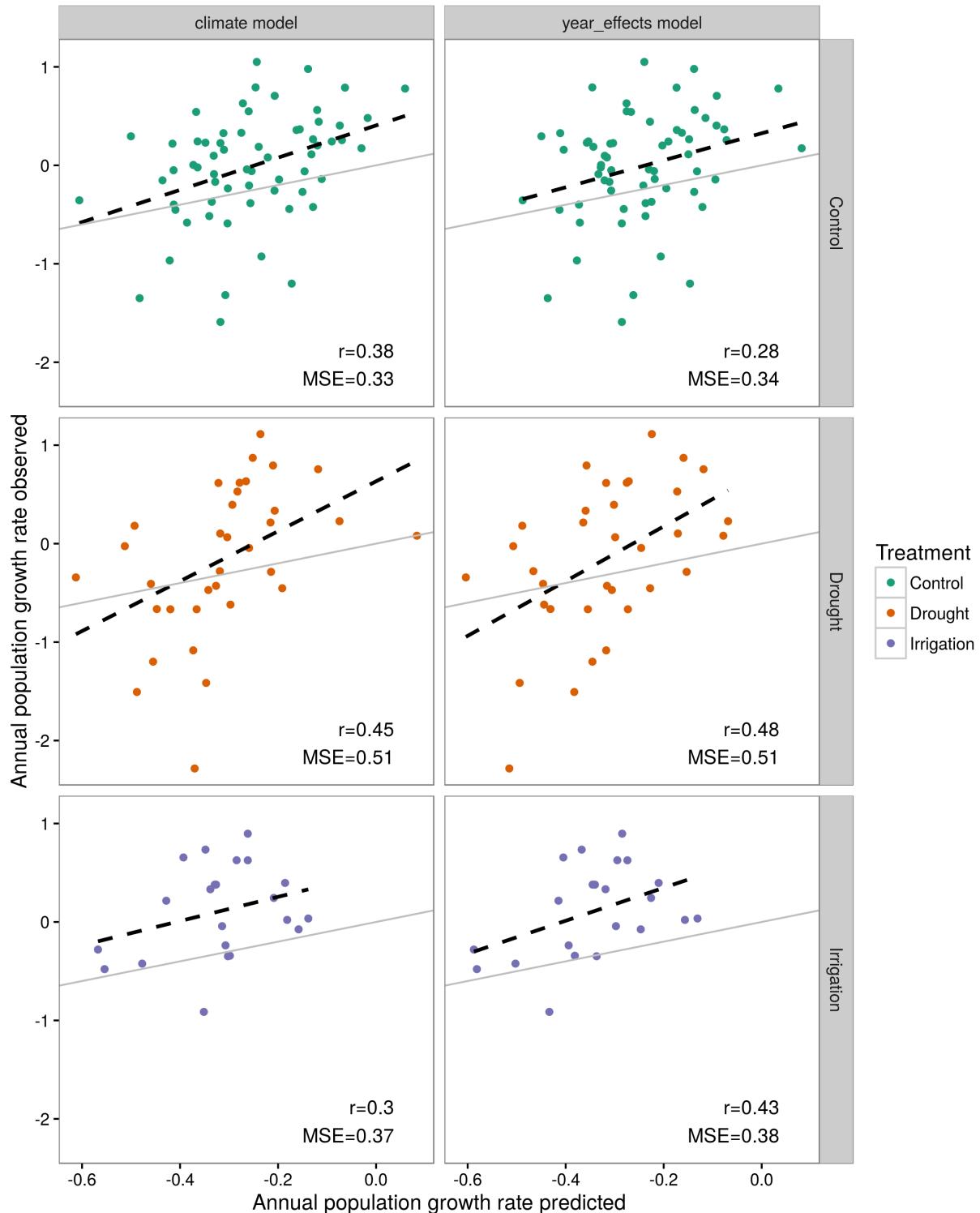


Figure 16: Observed and predicted one step ahead log change in *P. spicata* cover in the experiment. Changes in cover predicted by the climate model are shown on the left and those predicted by the no climate model are shown on the right. Correlations coefficients between predictions and observations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

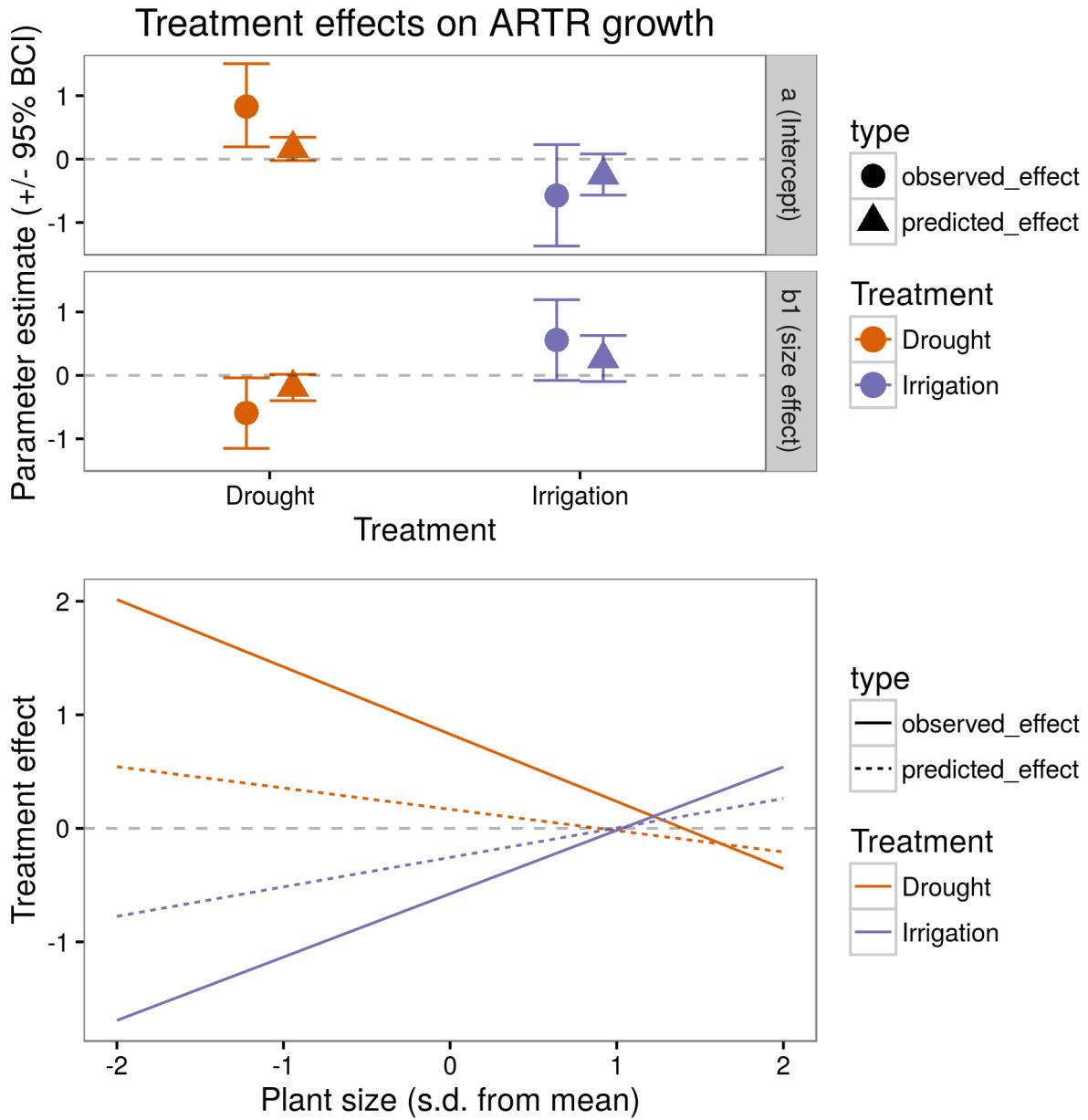


Figure A-1: Comparison of treatment effects predicted and observed for *A. tripartita* growth. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to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

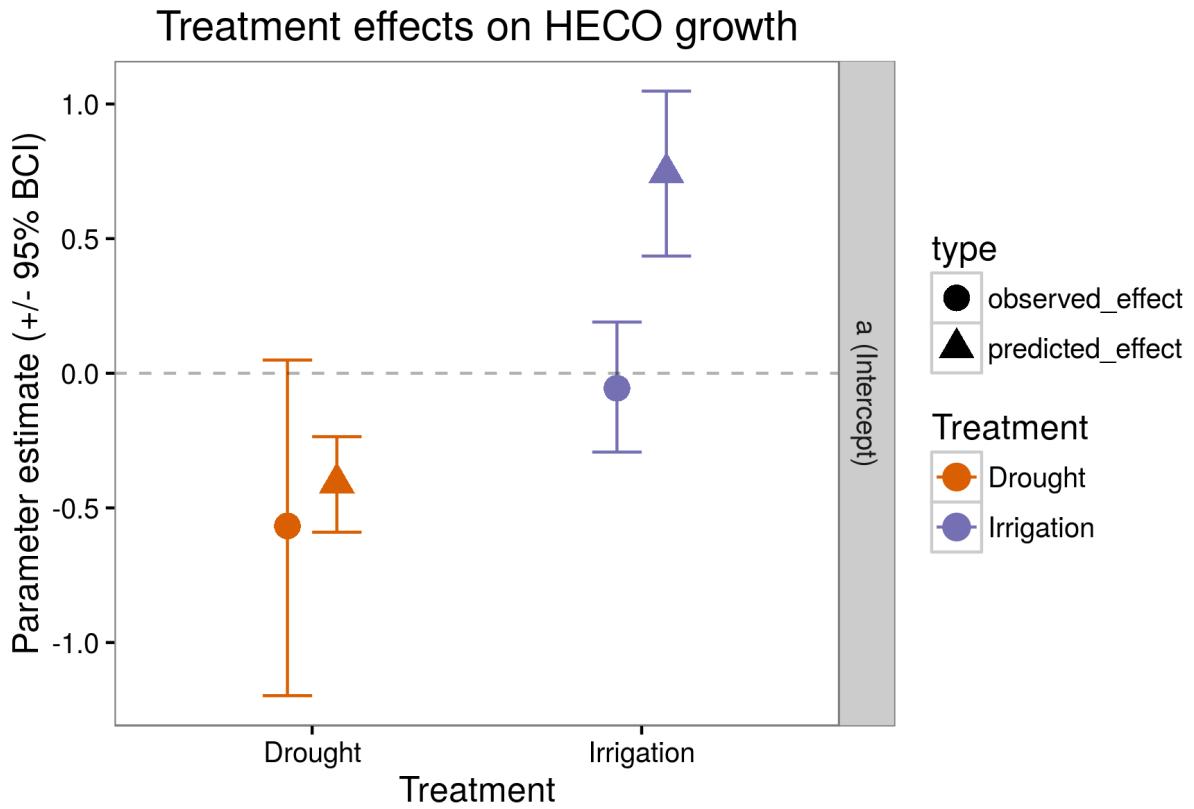


Figure A-2: Comparison of treatment effects predicted and observed for *H. comata* growth. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational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

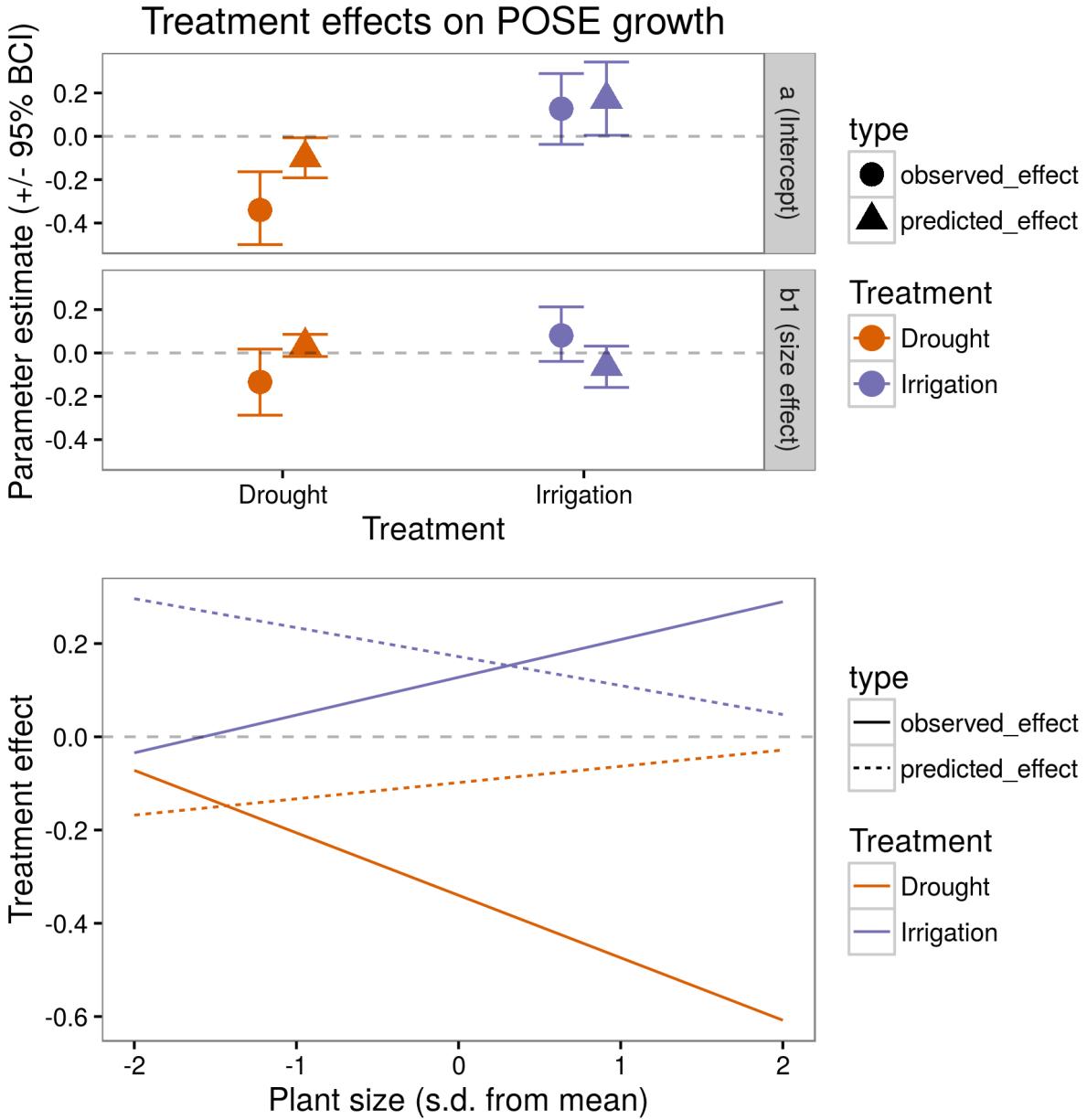


Figure A-3: Comparison of treatment effects predicted and observed for *P. secunda* growth. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to all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

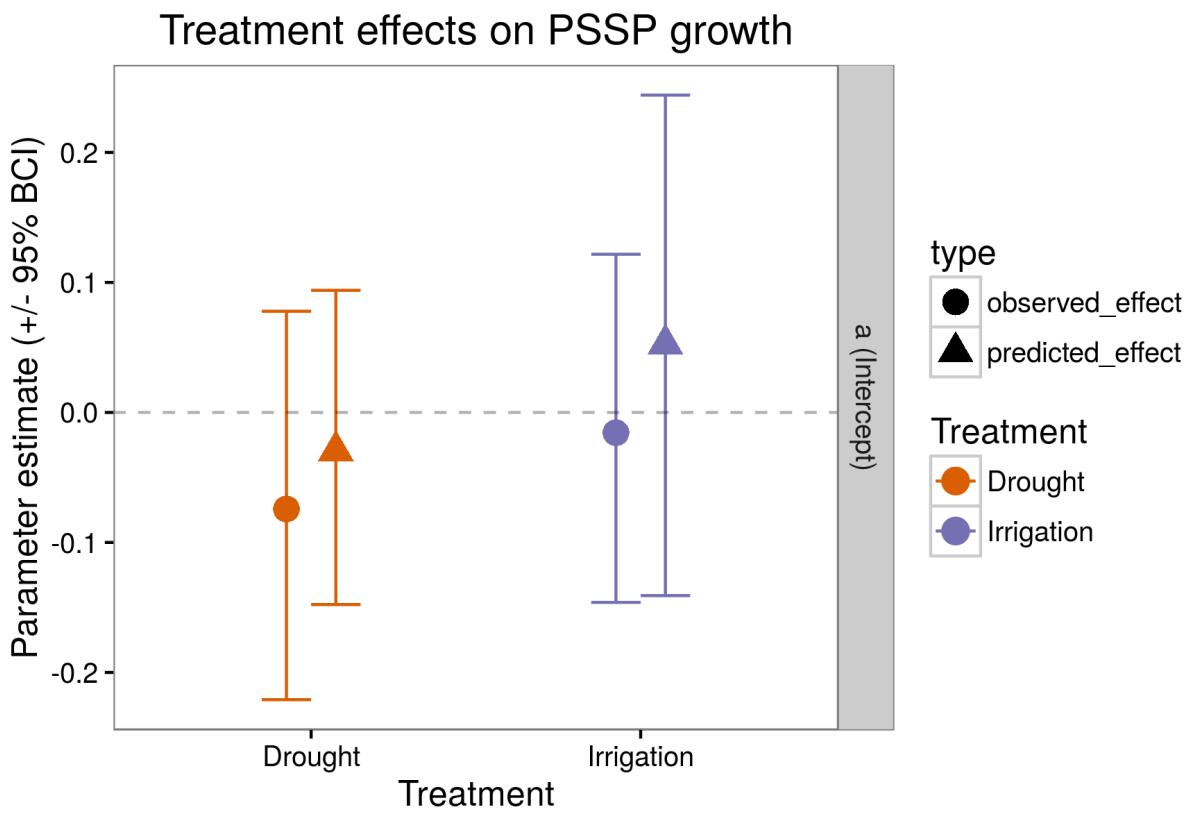


Figure A-4: Comparison of treatment effects predicted and observed for *P. spicata* growth. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational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

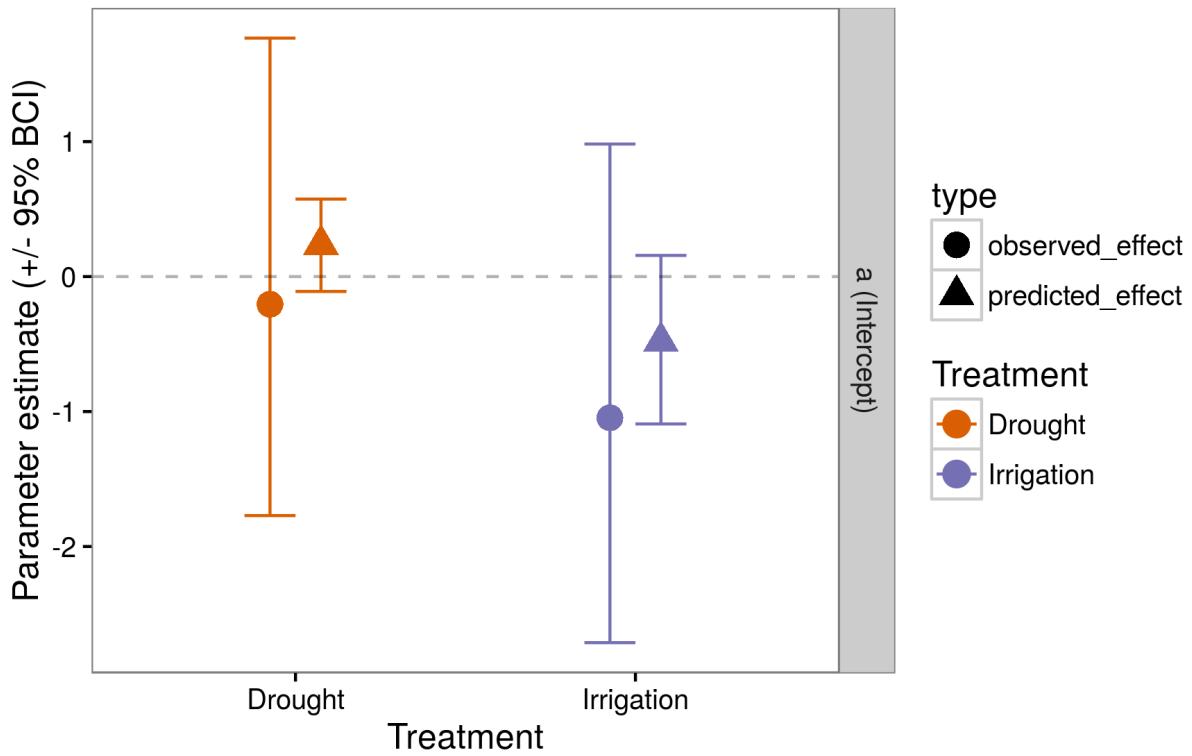


Figure A-5: Comparison of treatment effects predicted and observed for *A. tripartita* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of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

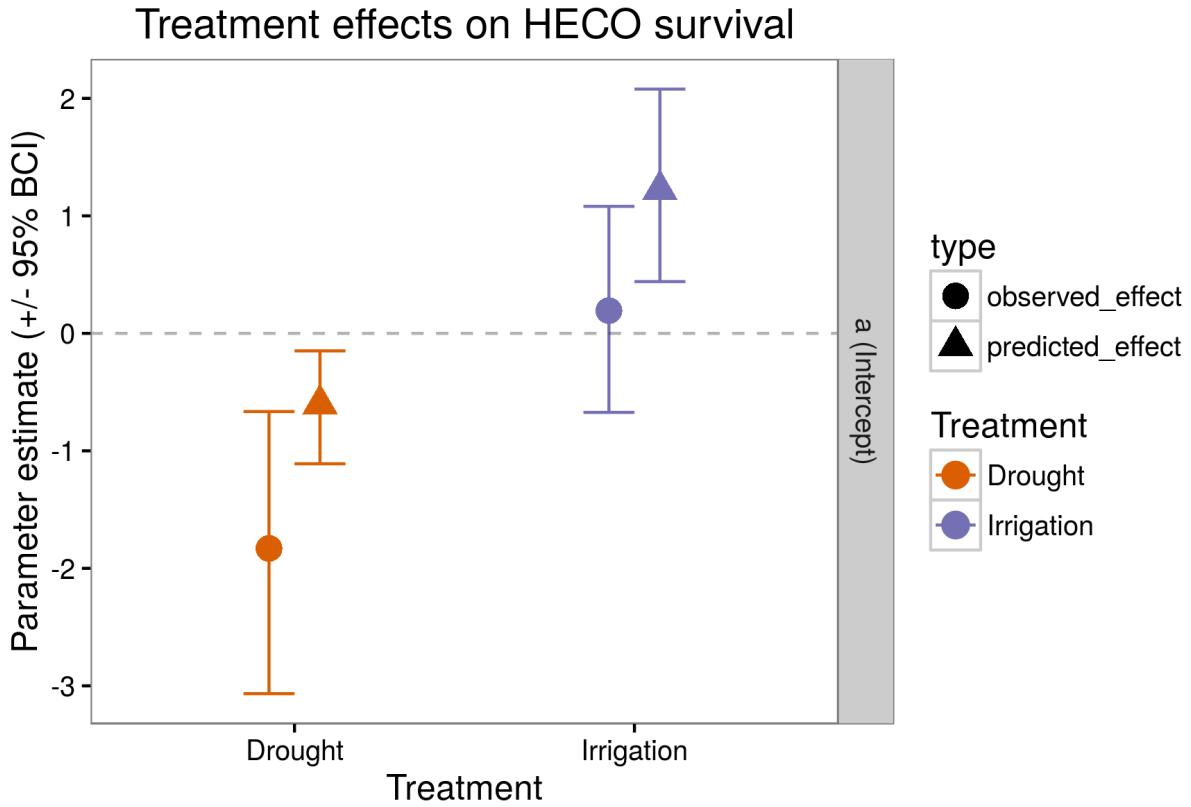


Figure A-6: Comparison of treatment effects predicted and observed for *H. comata* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational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

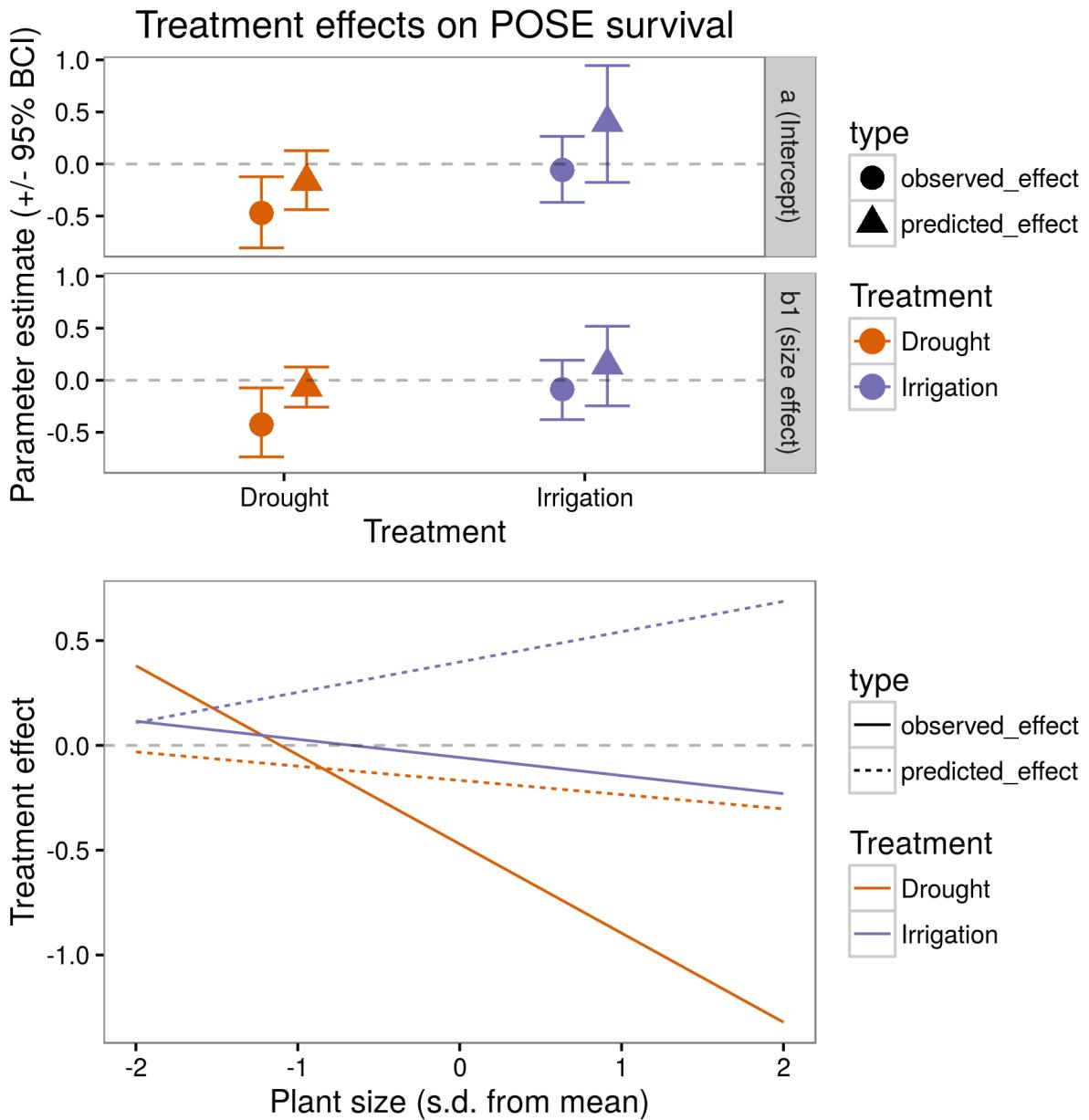


Figure A-7: Comparison of treatment effects predicted and observed for *P. secunda* survival. Upper figure shows each parameter estimate separately while the lower figure shows the effect of treatment as a function of plant size. The observed treatment effects come from the treatment model fitted to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

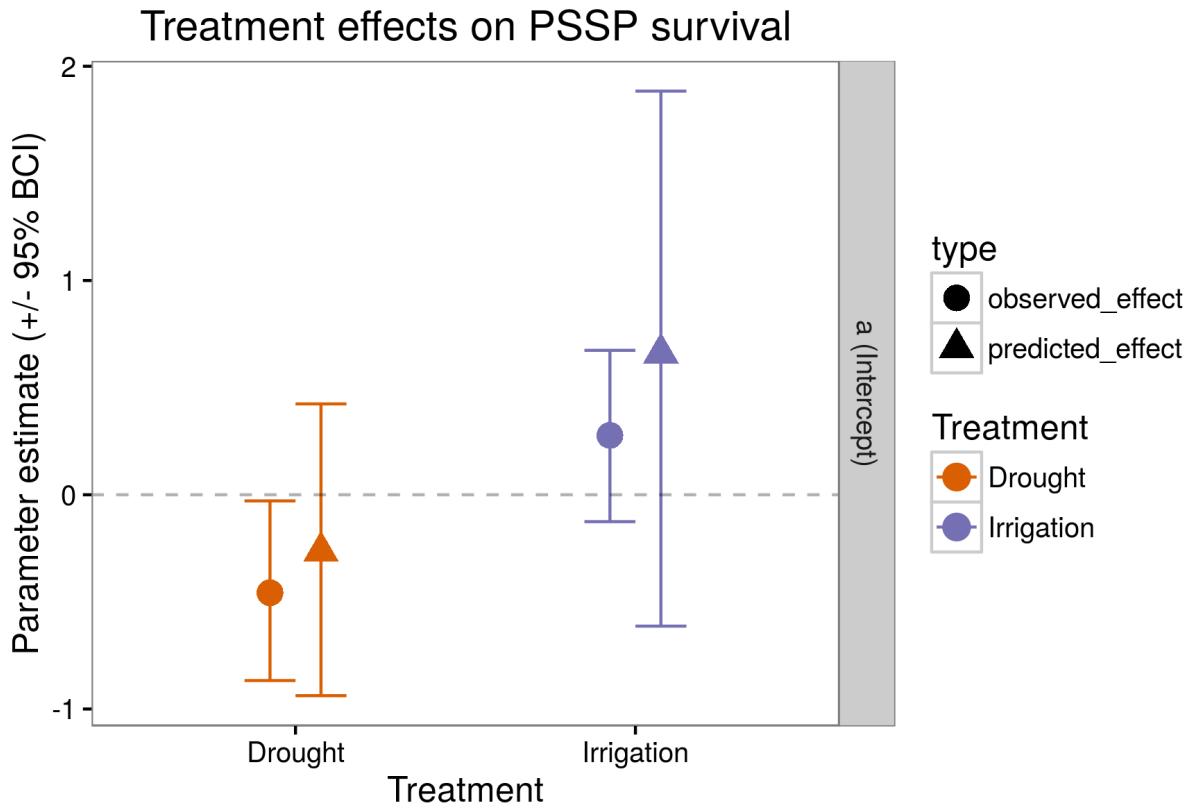


Figure A-8: Comparison of treatment effects predicted and observed for *P. spicata* survival. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of observational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

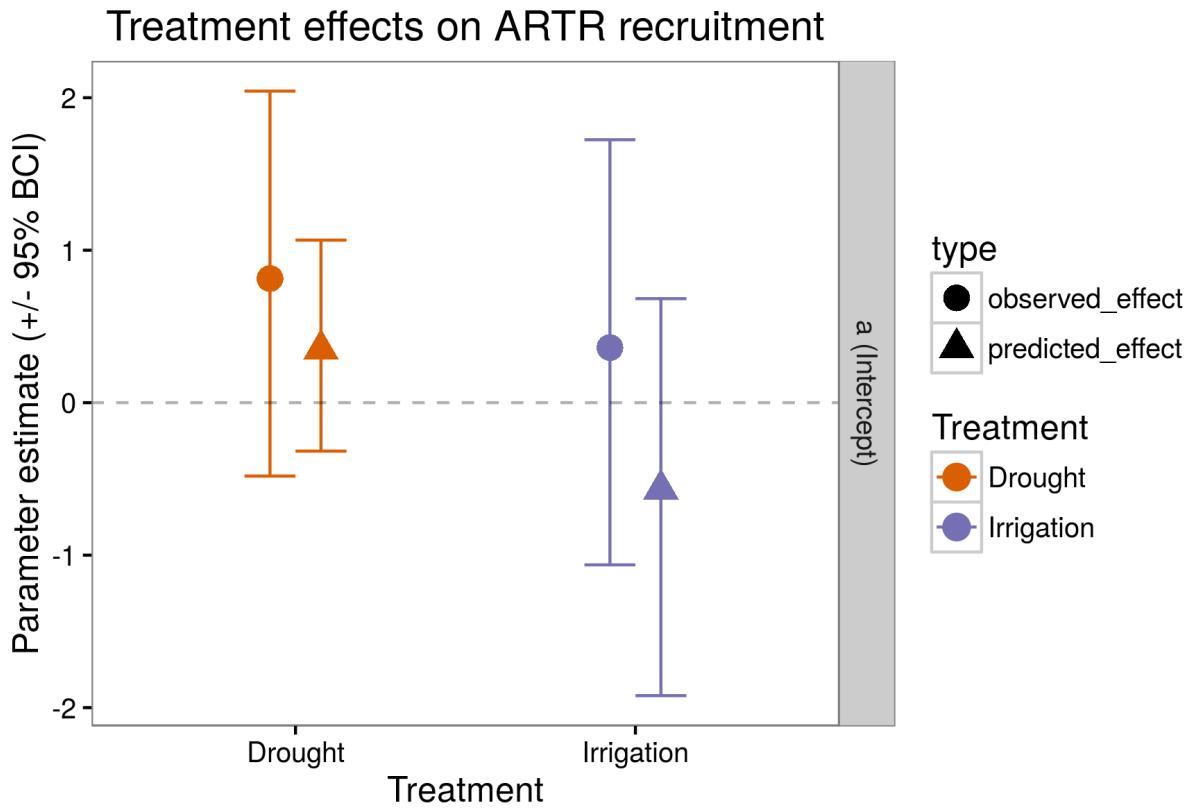


Figure A-9: Comparison of treatment effects predicted and observed for *A. tripartita* recruitment. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years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

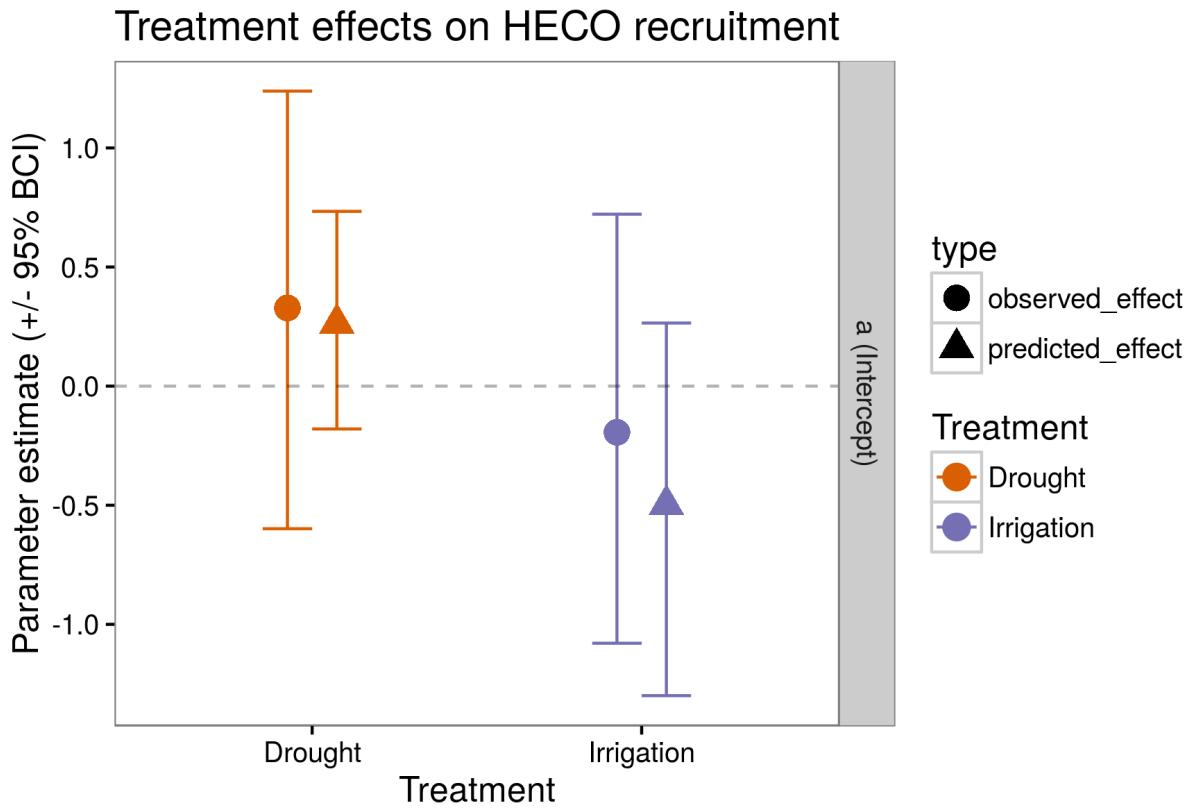


Figure A-10: Comparison of treatment effects predicted and observed for *H. comata* recruitment. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of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

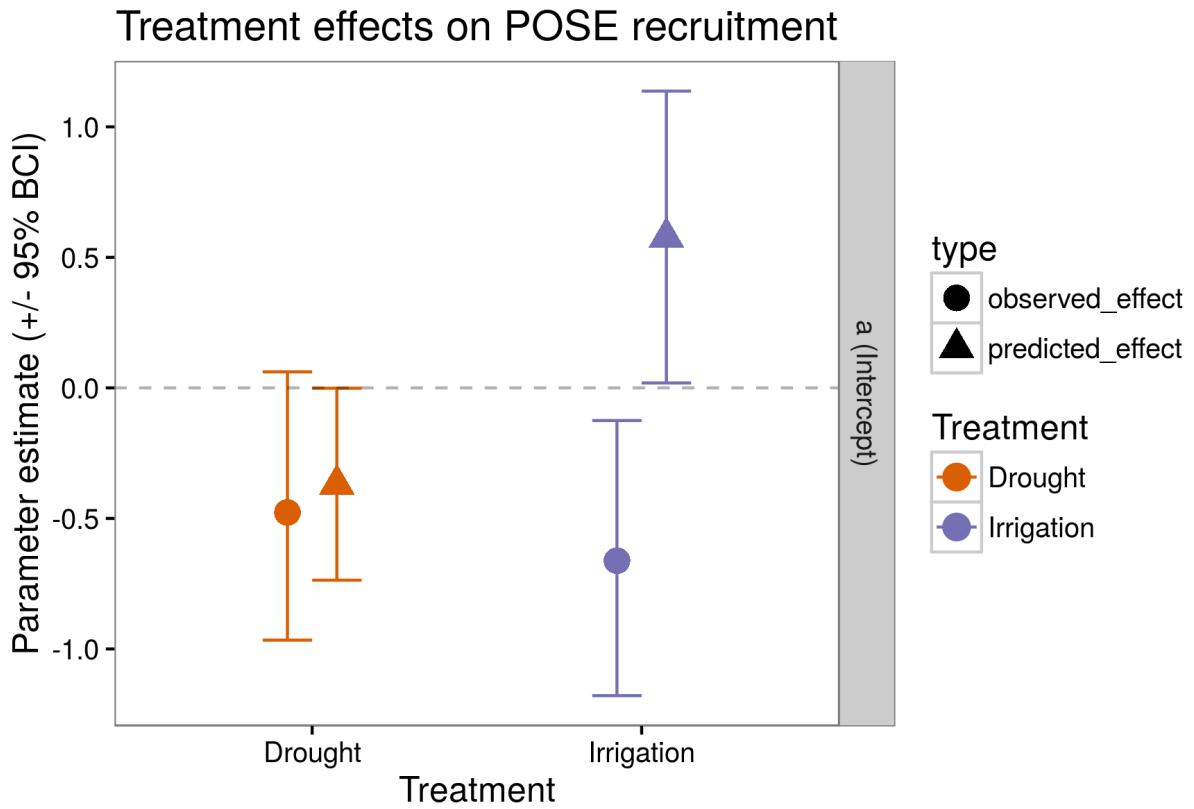


Figure A-11: Comparison of treatment effects predicted and observed for *P. secunda* recruitment. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of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

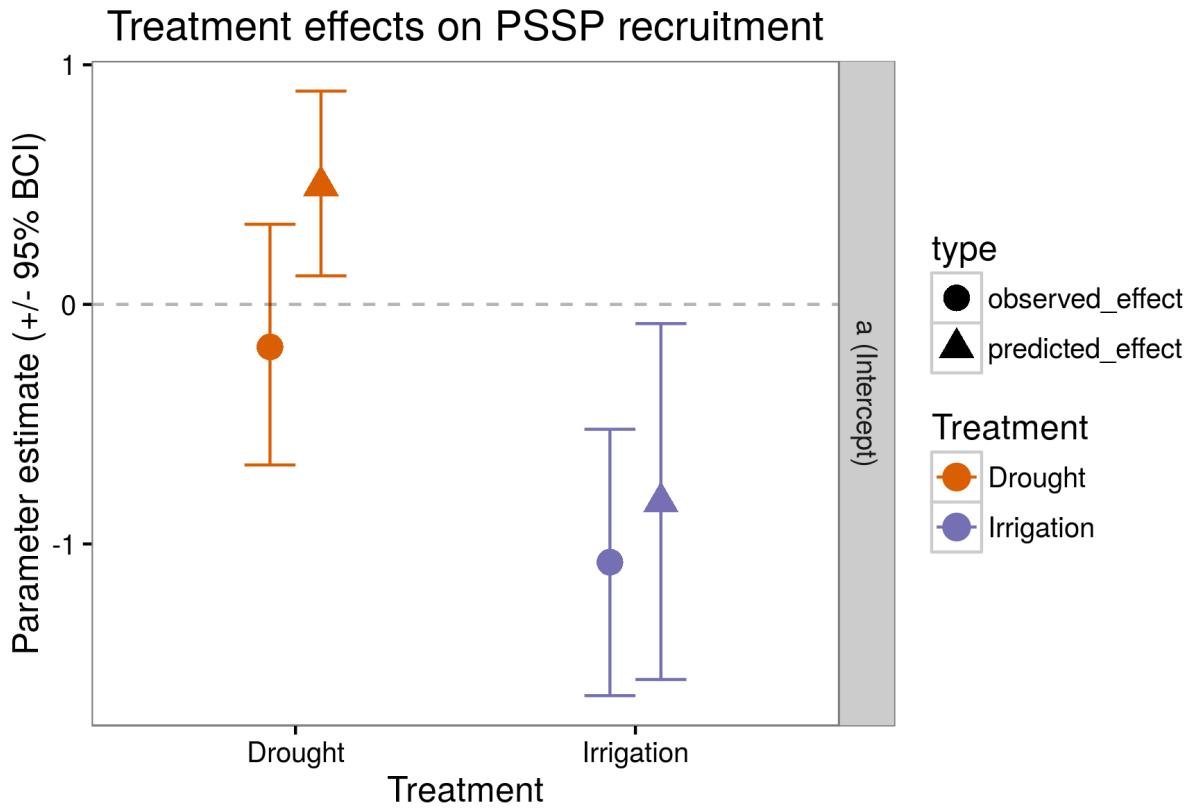


Figure A-12: Comparison of treatment effects predicted and observed for *P. spicata* recruitment. The observed treatment effects come from the treatment model fitted to all the data including the five years of the experiment. The predicted parameters come from the climate model fitted to all years of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