Appendix S1 - Full analytical methods and model results

Supplemental information for Diaz and Ernest, “Maintenance of community function through compensation breaks down over time in a desert rodent community”. In review at Ecology.

Table of Contents

# Data

## Functions

Data are accessed and processed using functions stored in <https://github.com/diazrenata/soar>, archived on Zenodo at <https://doi.org/10.5281/zenodo.5539880>. Install these functions either by running:

remotes::install\_github("diazrenata/soar")

or by downloading the Zenodo archive and installing the package manually.

## Data access

Data can be downloaded directly from the Portal data repository:

plotl <- get\_plot\_totals()  
  
plot\_types <- list\_plot\_types() %>% filter(plot\_type == "EE")

For speed and offline access, data files are also included in this repository in the data directory:

plotl <- read.csv(here::here("data", "plotl.csv"), stringsAsFactors = T)  
plot\_types <- read.csv(here::here("data", "plot\_types.csv"), stringsAsFactors = T)

## Balancing exclosure and control plots

Because there are 5 exclosure plots and 4 control plots in these data, we remove 1 exclosure plot to achieve a balanced design. From the 5 possible exclosures to remove, we randomly select 1 using the seed 1977 (the year the Portal Project was initiated).

plot\_types <- plot\_types %>%   
 filter(plot\_type == "EE")  
  
set.seed(1977)   
remove\_plot <- sample(plot\_types$plot, 1, F) # results in removing plot 19  
  
plotl <- plotl %>%  
 filter(plot != remove\_plot)

## Treatment-level means and quantities of interest

In order to calculate compensation and the total energy ratio, it is necessary to take the treatment-level mean total energy use and energy use by kangaroo rats and small granivores on control plots. For consistency in the main analysis, we take treatment-level means for all quantities.

Because this necessarily elides some degree of variability between plots with treatment types, we also conducted a provisional analysis incorporating between-plot variability for exclosure plots (but not for control plots), with qualitatively the same results (see appendix S4).

To take treatment-level means:

# Treatment-level means:  
treatl <- plots\_to\_treatment\_means(plotl)   
  
# Format column types  
treatl <- treatl %>%  
 mutate(censusdate = as.Date(censusdate),  
 oera = ordered(oera),  
 oplottype = ordered(oplottype))

Calculate proportional energy use of *C. baileyi* on exclosure and control plots. The pb\_nozero dataframe omits the first time period, because during that time *C. baileyi* was essentially absent at the site (and the large number of 0s for an entire treatment-by-factor level combination breaks statistical models).

pb <- get\_pb(treatl)   
  
pb\_nozero <- pb %>%  
 filter(as.numeric(oera) > 1)

Calculate total energy ratio and compensation, comparing exclosure to control plots:

energy\_ratio <- get\_e\_ratio(treatl)  
compensation <- get\_compensation(treatl)

Calculate kangaroo rat (Dipodomys) proportion of total energy use on control plots:

dipo\_c\_dat <- get\_dipo\_c(treatl)

## Variable names for analyses

The variables used in these analyses, and their definitions.

* period: The monthly census period number for each census. Numeric.
* censusdate: The date of the monthly census. Date.
* era: The “time period”, as described in the text. Character, one of a\_pre\_pb (first time period, before *C. baileyi* arrived at the site), b\_pre\_reorg (second time period, after *C. baileyi* established but before the most recent reorganization event), or c\_post\_reorg (third time period, after the last reorganization event).
* oera: era as an ordered factor, for modeling. Ordered factor.
* plot\_type: The treatment, either CC for control or EE for exclosure. Character.
* oplottype: plot\_type as an ordered factor, for modeling. Ordered factor.
* total\_e\_rat, total\_e\_rat\_ma (specific to energy\_ratio): The ratio of total energy use on exclosure plots relative to control plots, and the 6-month moving average. Numeric, unbounded.
* smgran\_comp, smgran\_comp\_ma (specific to compensation): Energetic compensation by small granivores for kangaroo rat removal, and the 6-month moving average. Numeric, unbounded.
* pb\_prop, pb\_prop\_ma (specific to pb and pb\_nozero): The proportion of treatment-level energy use accounted for by *C. baileyi*, and the 6-month moving average. Numeric, proportion bounded 0-1.
* dipo\_prop, dipo\_prop\_ma (specific to dipo\_c\_dat): The proportion of treatment-level energy use accounted for by all kangaroo rats, and the 6-month moving average. Numeric, proportion bounded 0-1.

# Compensation

Fit a generalized least squares accounting for temporal autocorrelation between monthly censuses within each time period using a continuous autoregressive structure of order 1.

comp\_mean\_gls <- gls(smgran\_comp ~ oera, correlation = corCAR1(form = ~ period), data = compensation)

## Model specification and selection

Compared to an intercept-only (null) model, the timeperiod term improves fit. AIC of model with timeperiod term = 69.8; AIC without = 84.7:

comp\_mean\_gls\_notime <- gls(smgran\_comp ~ 1, correlation = corCAR1(form = ~ period), data = compensation)

AIC(comp\_mean\_gls)

## [1] 69.85023

AIC(comp\_mean\_gls\_notime)

## [1] 84.74902

The temporal autocorrelation term improves fit over not including the term. AIC with the autocorrelation term, 69.8; AIC for without, 157.1:

comp\_mean\_gls\_noautoc <- gls(smgran\_comp ~ oera, data = compensation)

AIC(comp\_mean\_gls\_noautoc)

## [1] 157.0973

AIC(comp\_mean\_gls)

## [1] 69.85023

Calculate estimates from the final model:

comp\_mean\_gls\_emmeans <- emmeans(comp\_mean\_gls, specs = ~ oera)

## Results

### Table S1. Coefficients from GLS for compensation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Value | Std.Error | t-value | p-value |
| (Intercept) | 0.3450313 | 0.0294996 | 11.696141 | 0.0000000 |
| oera.L | 0.0647933 | 0.0524103 | 1.236269 | 0.2172146 |
| oera.Q | -0.2833553 | 0.0477359 | -5.935890 | 0.0000000 |

### Table S2. Estimates from GLS for compensation

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| oera | emmean | SE | df | lower.CL | upper.CL |
| a\_pre\_pb | 0.1835362 | 0.0520378 | 44.11081 | 0.0786683 | 0.2884041 |
| b\_pre\_reorg | 0.5763899 | 0.0462641 | 47.37851 | 0.4833383 | 0.6694416 |
| c\_post\_reorg | 0.2751677 | 0.0528010 | 46.75897 | 0.1689314 | 0.3814041 |

### Table S3. Contrasts from GLS for compensation

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| contrast | estimate | SE | df | t.ratio | p.value |
| a\_pre\_pb - b\_pre\_reorg | -0.3928537 | 0.0689413 | 47.89422 | -5.698378 | 0.0000 |
| a\_pre\_pb - c\_post\_reorg | -0.0916315 | 0.0741194 | 45.51740 | -1.236269 | 0.4383 |
| b\_pre\_reorg - c\_post\_reorg | 0.3012222 | 0.0694989 | 49.52957 | 4.334200 | 0.0002 |

# Total energy use

## Model specification and selection

As for compensation, fit a generalized least squares accounting for temporal autocorrelation between monthly censuses within each time period using a continuous autoregressive structure of order 1.

totale\_mean\_gls <- gls(total\_e\_rat ~ oera, correlation = corCAR1(form = ~ period), data = energy\_ratio)

Compared to an intercept-only (null) model, the timeperiod term improves fit. AIC of model with timeperiod term = -133; AIC without = -118.

totale\_mean\_gls\_notime <- gls(total\_e\_rat ~ 1, correlation = corCAR1(form = ~ period), data = energy\_ratio)

AIC(totale\_mean\_gls)

## [1] -132.9214

AIC(totale\_mean\_gls\_notime)

## [1] -118.15

Again, the temporal autocorrelation term improves fit over not including the term. AIC with the autocorrelation term is - 133; without, 13.29.

totale\_mean\_gls\_notautoc <- gls(total\_e\_rat ~ oera, data = energy\_ratio)

AIC(totale\_mean\_gls\_notautoc)

## [1] 13.29396

AIC(totale\_mean\_gls)

## [1] -132.9214

Calculate estimates:

totale\_mean\_gls\_emmeans <- emmeans(totale\_mean\_gls, specs = ~ oera)

## Results

### Table S4. Coefficients from GLS on total energy ratio

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Value | Std.Error | t-value | p-value |
| (Intercept) | 0.5016731 | 0.0271176 | 18.499880 | 0.0000000 |
| oera.L | 0.1413504 | 0.0477646 | 2.959316 | 0.0033001 |
| oera.Q | -0.2503659 | 0.0429312 | -5.831790 | 0.0000000 |

### Table S5. Estimates from GLS on total energy ratio

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| oera | emmean | SE | df | lower.CL | upper.CL |
| a\_pre\_pb | 0.2995118 | 0.0475806 | 36.19943 | 0.2030323 | 0.3959913 |
| b\_pre\_reorg | 0.7060960 | 0.0419773 | 38.51943 | 0.6211550 | 0.7910369 |
| c\_post\_reorg | 0.4994115 | 0.0480066 | 37.62774 | 0.4021956 | 0.5966274 |

### Table S6. Contrasts from GLS on total energy ratio

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| contrast | estimate | SE | df | t.ratio | p.value |
| a\_pre\_pb - b\_pre\_reorg | -0.4065842 | 0.0623398 | 40.51631 | -6.522060 | 0.0000 |
| a\_pre\_pb - c\_post\_reorg | -0.1998997 | 0.0675493 | 37.12310 | -2.959316 | 0.0144 |
| b\_pre\_reorg - c\_post\_reorg | 0.2066845 | 0.0626456 | 41.44768 | 3.299267 | 0.0056 |

# Kangaroo rat proportional energy use

## Model specification and selection

Proportional energy use is bounded 0-1 and cannot be fit with generalized least squares. We therefore use a binomial GLM with no temporal autocorrelation term.

dipo\_glm <- glm(dipo\_prop ~ oera, family = binomial, data= dipo\_c\_dat)

Compared to an intercept-only (null) model, the timeperiod term improves fit. AIC of model with timeperiod term = 258; AIC without = 281.

dipo\_intercept\_glm <- glm(dipo\_prop ~ 1, family = binomial, data = dipo\_c\_dat)

AIC(dipo\_glm)

## [1] 258.3581

AIC(dipo\_intercept\_glm)

## [1] 280.8497

Calculate estimates:

dipoemmeans <- regrid(emmeans(dipo\_glm, specs = ~ oera))

## Results

### Table S7. Coefficients from GLM on Dipodomys energy use.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | z value | Pr(>|z|) |
| (Intercept) | 1.4032480 | 0.1503201 | 9.335068 | 0.0000000 |
| oera.L | -1.1000833 | 0.2871738 | -3.830723 | 0.0001278 |
| oera.Q | 0.5855493 | 0.2304516 | 2.540878 | 0.0110574 |

### Table S8. Estimates from GLM on Dipodomys energy use.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| oera | prob | SE | df | asymp.LCL | asymp.UCL |
| a\_pre\_pb | 0.9183528 | 0.0256462 | Inf | 0.8680872 | 0.9686183 |
| b\_pre\_reorg | 0.7160901 | 0.0398537 | Inf | 0.6379782 | 0.7942020 |
| c\_post\_reorg | 0.7035835 | 0.0456677 | Inf | 0.6140765 | 0.7930905 |

### Table S9. Contrasts from GLM on Dipodomys energy use.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| contrast | estimate | SE | df | z.ratio | p.value |
| a\_pre\_pb - b\_pre\_reorg | 0.2022627 | 0.0473925 | Inf | 4.2678236 | 0.0001 |
| a\_pre\_pb - c\_post\_reorg | 0.2147693 | 0.0523762 | Inf | 4.1005151 | 0.0001 |
| b\_pre\_reorg - c\_post\_reorg | 0.0125066 | 0.0606124 | Inf | 0.2063368 | 0.9768 |

# C. baileyi proportional energy use

## Model specification and selection

As for kangaroo rat energy use, we use a binomial GLM to fit *C. bailyei* proportional energy use. Because *C. baileyi* occurs on both exclosure and control plots, we investigate whether *C. baileyi*’s proportional energy use differs between treatment types. We compare models incorporating separate slopes, separate intercepts, or no terms for treatment modulating the change in *C. baileyi* proportional energy use across time periods. We also test a null (intercept-only) model of no change across time periods.

The best-fitting (lowest AIC; use of stepwise selection via anova yields the same result) model is for an effect of time period and of treatment, but no interaction. We therefore proceed with this model.

pb\_glm\_interaction <- glm(pb\_prop ~ oera \* oplottype, family = binomial, data= pb\_nozero)  
pb\_glm\_nointeraction <- glm(pb\_prop ~ oera + oplottype, family = binomial, data= pb\_nozero)  
pb\_glm\_notreat <- glm(pb\_prop ~ oera, family = binomial, data= pb\_nozero)  
pb\_glm\_notime <- glm(pb\_prop ~ 1, family = binomial, data= pb\_nozero)  
  
AIC(pb\_glm\_interaction)

## [1] 237.7643

AIC(pb\_glm\_nointeraction)

## [1] 231.0963

AIC(pb\_glm\_notreat)

## [1] 460.8477

AIC(pb\_glm\_notime)

## [1] 541.3799

Calculate estimates:

pb\_emmeans <- regrid(emmeans(pb\_glm\_nointeraction, specs = ~ oera))

## Results

### Table S10. Coefficients from GLM on C. baileyi energy use

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | z value | Pr(>|z|) |
| (Intercept) | -1.574028 | 0.1670168 | -9.424368 | 0 |
| oera.L | -1.409273 | 0.2010398 | -7.009921 | 0 |
| oplottype.L | 2.184896 | 0.2267112 | 9.637355 | 0 |

### Table S11. Estimates from GLM on C. baileyi energy use

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| oera | prob | SE | df | asymp.LCL | asymp.UCL |
| b\_pre\_reorg | 0.3595031 | 0.0396644 | Inf | 0.2817622 | 0.4372440 |
| c\_post\_reorg | 0.0710590 | 0.0170265 | Inf | 0.0376876 | 0.1044304 |

### Table S12. Contrasts from GLM on C. baileyi energy use.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| contrast | estimate | SE | df | z.ratio | p.value |
| b\_pre\_reorg - c\_post\_reorg | 0.2884441 | 0.0403673 | Inf | 7.145484 | 0 |