## 2. Methods

#### 2.1 Overview

The effects of biocontrol treatments, time, geography, and climate on *C. arvense* stem counts were evaluated using Poisson generalized linear mixed effects models and ordinary krieging. Due to data sparsity, only main effects were considered in the GLMMs, and dimension reduction was performed on multiple climatic variables using principal component analyses (PCA). All mixed effects models included site ID as a random intercept term.

#### 2.2 Overall analyses

Overall trends from the beginning to the end of the time series were evaluated by modelling final stem counts as a function of initial stem counts, by site. Ordinary krieging was also performed on the percentage change in stem counts from the first to the final survey to test for spatial autocorrelation. These initial results were used to inform the next analyses. In particular, ordinary krieging suggested that terrain ‘flatness’ may have influenced stem counts over the time series.

#### 2.3 Per-timestep analyses

The change in stem counts from time t-1 to time t was also evaluated to account for inconsistent treatment protocols at each site. To do so, we used two different Poisson GLMMs. In both models we included stem count at time t-1 and survey year was covariates, to account for temporal trends and autocorrelation.

#### 2.3.1 Protocol effects

The first model focused on treatment protocol, and included as predictors: treatment type at last treatment (targeted, broadcast, or targeted+broadcast), years since the last treatment, and amount of inoculum used in the last treatment.

#### 2.3.2 Site effects

The second model focused on (site-level) climate and geographic effects, and included as independent variables: annual mean maximum temperature, standard deviation of elevation (within a 2.5 km radius around each site), % cropland or developed landuse (within a 1 km radius around each site), and the first and second principal components extracted from two different PCAs.

#### 2.3.3 Climate PCAs

The first PCA (henceforth ‘treatment-time PCA’) included climate variables from a 2 week window prior to each biocontrol treatment, while the second PCA (henceforth ‘observation-time PCA’) included the same variables but collected from a 2 week window prior to each monitoring survey. Both PCAs included the following predictors (based on daily means averaged over the 2-week window): maximum temperature (C), minimum temperature (C), average temperature (C), total precipitation (mm/day), water vapor pressure (Pa), relative humidity (%), dew point (C), shortwave radiation flux (W/m2), as well as elevation (m).

#### 2.4 Covariate data and software

Analyses were performed with R version 4.2.2 (R Core Team 2023) and RStudio version 2023.03.0 (RStudio Team 2023), using dplyr (Wickham et al. 2021) and sqldf to prepare the data and ggplot2 for visualisation (Wickham 2016). Linear mixed effect models were fitted with lme4 (Bates et al. 2015), ordinary kriging was performed using the kriging package (Olmedo 2022), and PCA was performed using the base R stats library. The car package (citation) was used to obtain Anova tables for each GLMM. The Adaptive Gauss-Hermite Quadrature was used (with nAGQ=0 in lme4) was used to allow model convergence on both GLMMs in the per-timestep analyses, which can cause less accurate but nonetheless acceptable model fits (Stegmann et al. 2018).

Climate data were directly obtained or derived from DAYMET version 4 (Thornton et al. 2022) using the daymetr package (Hufkens et al. 2018). Standard deviation of elevation was derived from the USGS 3DEP model (10.2m) dataset (U.S. Geological Survey 2022) and landuse was derived from the ESA WorldCover (10m) dataset (Zanaga et al. 2022), using Google Earth Engine version 0.1.329 (Gorelick et al. 2017) via the rgee (Aybar 2022) and reticulate (Ushey et al. 2023) packages. The Colorado TIGER/Line shapefile was obtained using the tigris package (Walker 2023). Basemap was obtained from Google Maps (Google n.d.).

## 3. Results

#### 3.1 Overall Stem Counts

At most sites, *C. arvense* stem counts declined from initial measurement to final observation (Table 1), from a mean (± SE) of 87.9 ± 6.5 stems to 44.7 ± 4.2 stems (Figure 1).

Table 1. Poisson GLMM results for final stem counts in *C. arvense* as a function of initial stem counts. Site was included as a random intercept term. Statistically significant results (p < 0.05) highlighted in bold.

| **Parameter** | **𝛘2** | **DF** | ***p*** |
| --- | --- | --- | --- |
| first\_count | 5.53 | 1 | **0.019** |

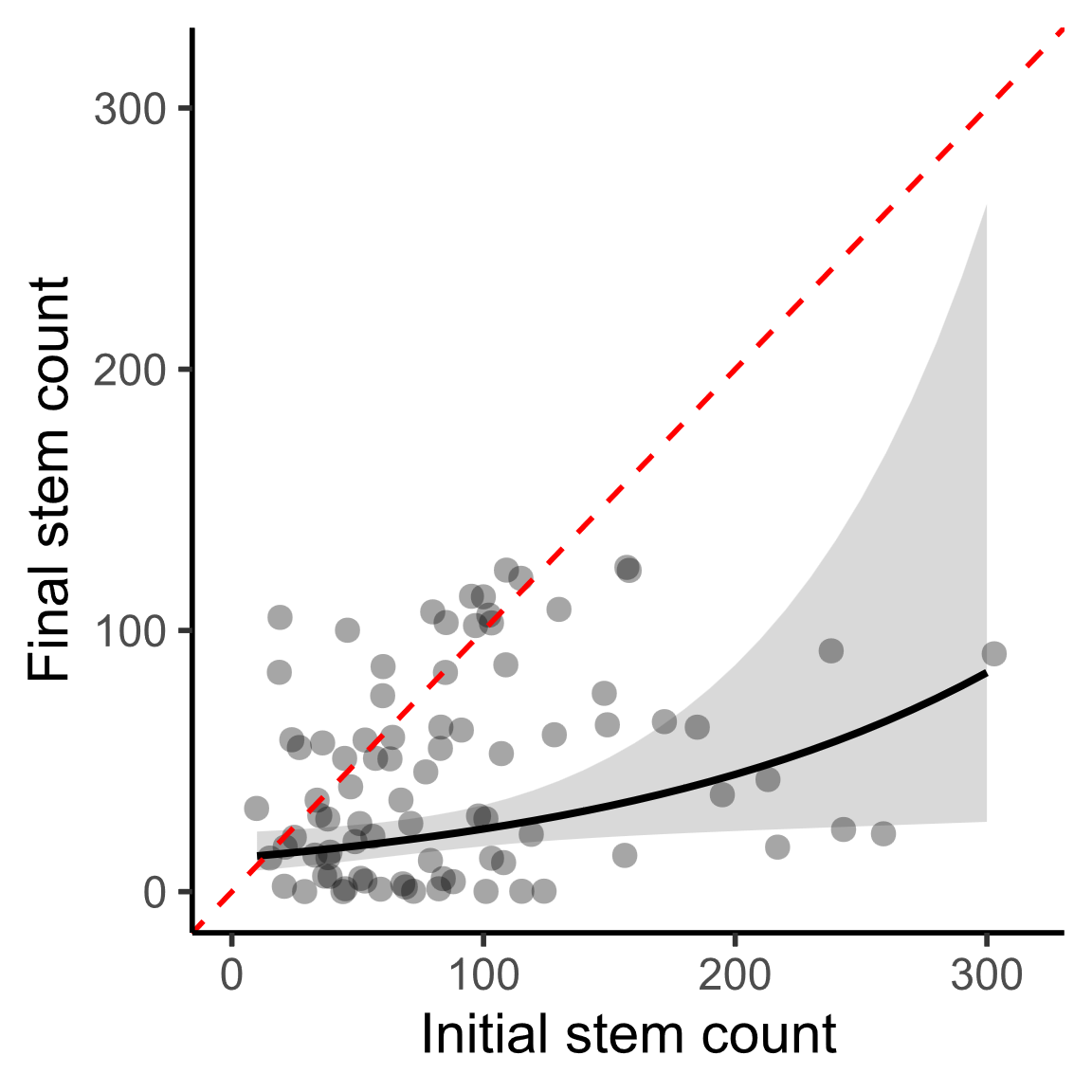
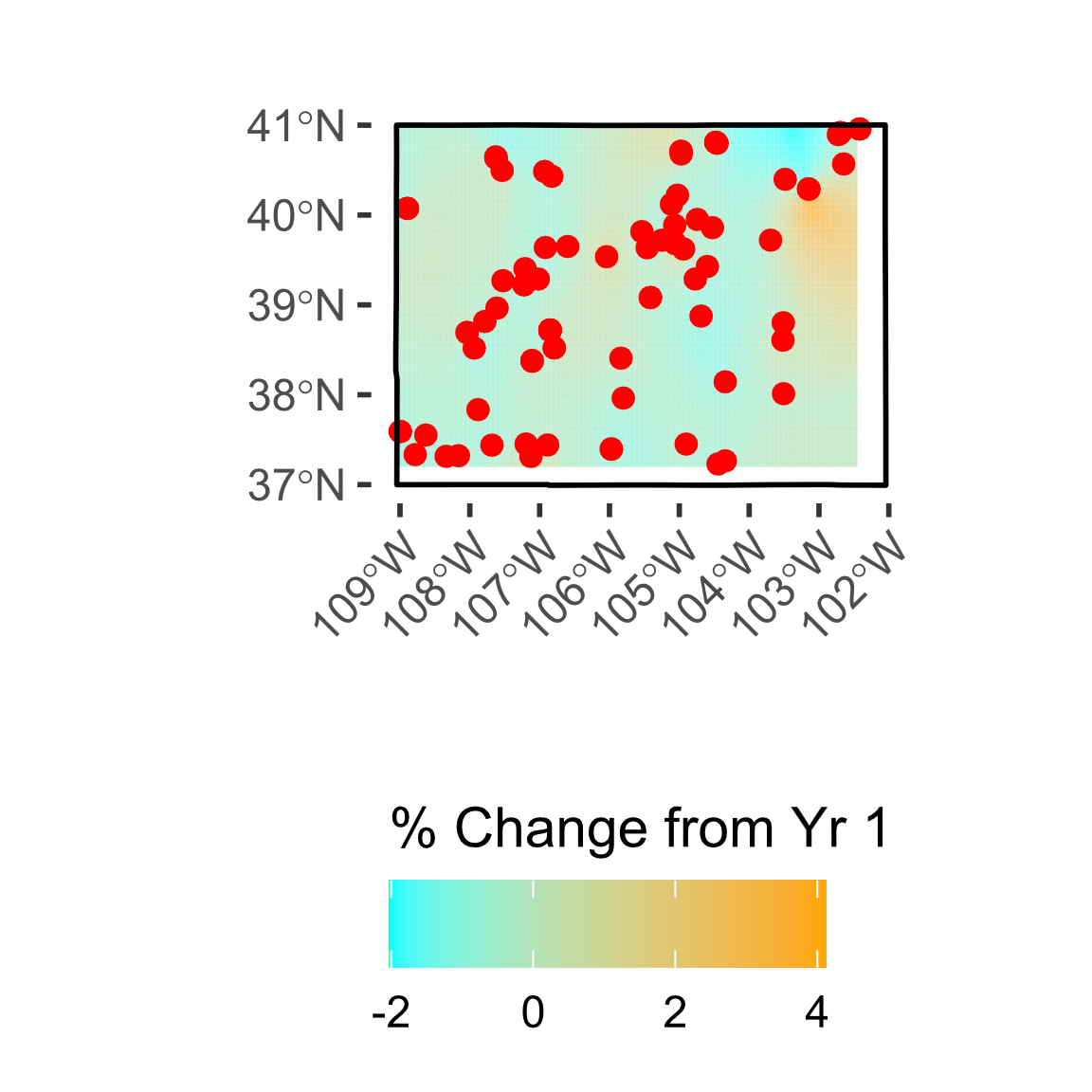


Figure 1. Initial vs final stem counts in *C. arvense*. Points indicate sites. Dashed red line shows y=x: values above this line indicate overall increase in stem counts, and vales below indicate decline. Shaded area is 95% confidence ribbon. N=87.

Percent change from initial to final stem counts was spatially autocorrelated (Moran’s I = 0.186 vs -0.012 expected, p < 0.01). Krieged values are presented in Figure 2.

A picture containing map, atlas, text

Description automatically generated

Figure 2. % change in *C. arvense* stem counts from initial to final survey, interpolated across Colorado with ordinary kriging (left). Red points show site locations on left panel, blue pins show point on right basemap.

#### 3.2 Climate PCAs

In the treatment-time PCA, PC1 explained 52.1% of the variance while PC2 explained 22.8% (Figure 3A). Higher values of PC1 correlated with higher elevation and precipitation, while lower values of PC1 correlated with higher temperature, vapor pressure, and dew point. Higher values of PC2 correlated with higher relative humidity and also correlated with precipitation, while lower values correlated with higher radiance.

In the observation-time PCA, PC1 explained 52.5% of the variance and PC2 explained 30.7% (Figure 3B). Predictor clustering was overall similar to the treatment-time PCA. Higher values of PC1 correlated with elevation, while lower values correlated with temperature, dew point, vapor point, as well as relative humidity and precipitation. Higher values of PC2 correlated with radiance and higher temperatures, while lower values correlated with precipitation and relative humidity.

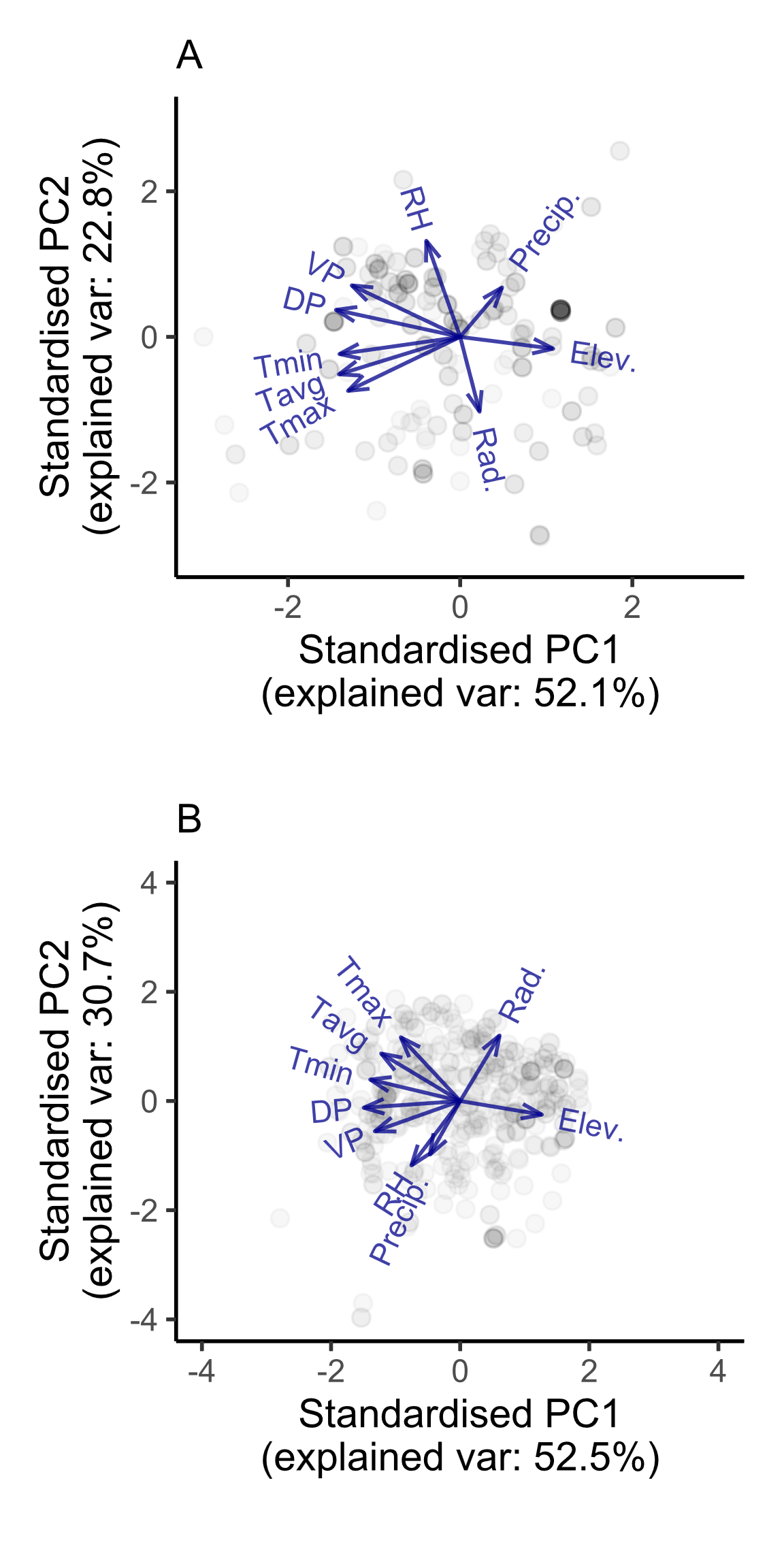


Figure 3. Biplots of A. treatment-time PCA, and B. observation-time PCA. Points represent site-year combinations used to train the PCAs, arrows represent loadings for each predictor. Tmax = daily maximum temperature, Tavg = daily average temperature, Tmin = daily minimum temperature, Rad = radiance, VP = vapor pressure, RH = relative humidity, DP = dew point, Precip = precipitation, Elev = elevation.

#### 3.3 Annual Stem Counts

##### Treatment and protocol effects

Stem count at time t was influenced by stem count at time t-1, monitoring year, last treatment type, last treatment quantity, and years since the last treatment (Table 2). Marginal effects of each predictor are presented in Figure 4.

Table 2. Poisson GLMM results for stem counts in *C. arvense* at time t as a function of: stem counts at time t-1, monitoring year, last treatment type, last treatment quantity (of biocontrol inoculum), and years since last treatment. Site was included as a random intercept term. Statistically significant results (p < 0.05) highlighted in bold.

| **Parameter** | **𝛘2** | **DF** | ***p*** |
| --- | --- | --- | --- |
| last\_stems | 85.54 | 1 | **< 0.001** |
| monitoring\_year | 680.01 | 1 | **< 0.001** |
| yrs\_since\_trt | 25.44 | 1 | **< 0.001** |
| amount\_lasttrt | 16.55 | 1 | **< 0.001** |
| treatment\_lasttrt | 12.50 | 2 | **0.002** |

Stem counts were lower in later monitoring years (Figure 4A), and higher stem counts at time t-1 resulted in lower stem counts at time t, suggesting negative density dependence (Figure 4B). Stem counts were lowest after broadcast treatment (95% CI: 40.2-59.4), highest after targeted treatment (95% CI: 52.4-71.0), and intermediate after broadcast+targeted (95% CI: 46.2-66.2) (Figure 4C). Stem counts also increased with increasing time since the last treatment was applied (Figure 4D), and declined with increasing quantity of innoculum at the last treatment (Figure 4E).

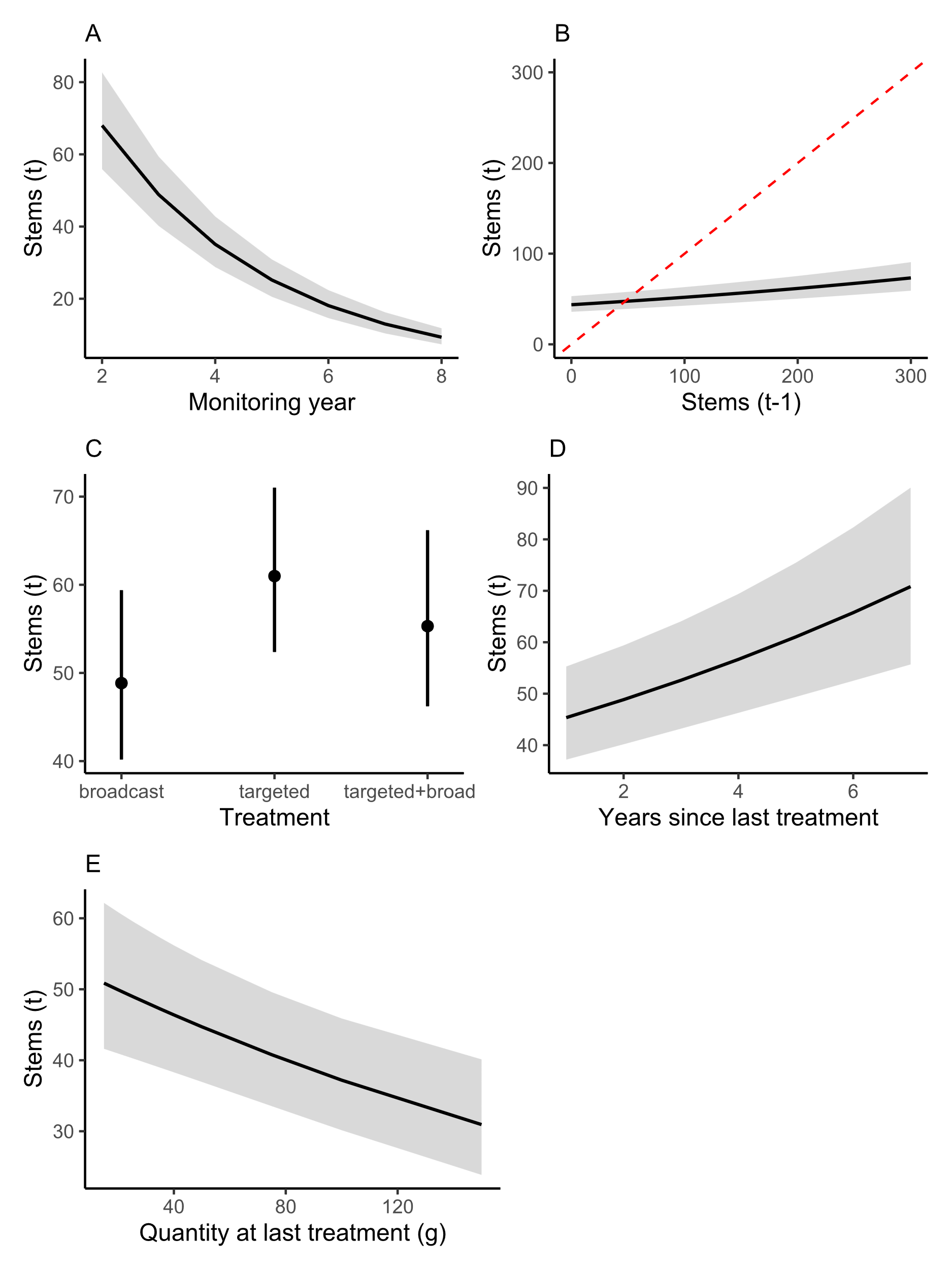


Figure 4. Marginal effects on stem counts at time t of: A. monitoring year, B. stem count in the prior year, C. last treatment type, D. years since last treatment, and E. quantity of *P. punctiformis* inoculum at last treatment. Shaded areas (and error bars in panel C) represent 95% confidence intervals. Dashed red line in panel B shows y=x: values above this line indicate overall increase in stem counts since prior year, and values below indicate decline. N=87.

##### Climate and geographic effects

Controlling for stem counts at t-1 and monitoring year as covariates, stem counts at time t were also affected by mean annual daily maximum temperature and standard deviation of elevation, as well as principal components based on climate data from treatment time and observation time (Table 3).

Table 3. Poisson GLMM results for stem counts in *C. arvense* at time t as a function of: stem counts at time t-1, monitoring year, annual mean maximum temperature, standard deviation of elevation, and PC1 and PC2 from treatment-time PCA and observation-time PCA. Standard deviation of elevation was calculated for a 2.5 radius circle around each site. Site was included as a random intercept term. Statistically significant results (p < 0.05) highlighted in bold.

| **Parameter** | **𝛘2** | **DF** | ***p*** |
| --- | --- | --- | --- |
| last\_stems | 122.29 | 1 | **< 0.001** |
| monitoring\_year | 1325.58 | 1 | **< 0.001** |
| PC1\_obs | 27.92 | 1 | **< 0.001** |
| PC1\_trt | 18.34 | 1 | **< 0.001** |
| PC2\_obs | 4.32 | 1 | **0.038** |
| PC2\_trt | 4.64 | 1 | **0.031** |
| tmax\_ann | 103.15 | 1 | **< 0.001** |
| elev\_stddev\_m | 12.80 | 1 | **< 0.001** |
| pct\_dist | 0.00 | 1 | 0.975 |

Stem counts at time t declined with increasing mean annual daily maximum temperatures (Figure 5A), and with increasing standard deviation of elevation (i.e., less ‘flatness’, Figure 5B).

Stem counts decreased with lower PC1 scores (Figure 5C) and higher PC2 scores (Figure 5D) based on the PCA on climate around treatment time. In other words, stem counts were lower where in the 2 weeks prior to treatment sites had: lower radiance and elevation, and higher temperatures, vapour pressure, dew point, and relative humidity. But the effect of precipitation in the 2 weeks prior to treatment was more ambivalent.

Based on the PCA on climate around observation time, stem counts decreased with higher PC1 (Figure 5E) and PC2 (Figure 5F) scores. In other words stem counts were also lower where, in the 2 weeks prior to monitoring, sites had: higher radiance and elevation, and lower daily average and minimum temperatures, dew point, vapor point, precipitation, and relative humidity. But the effect of daily maximum temperatures was more ambivalent.

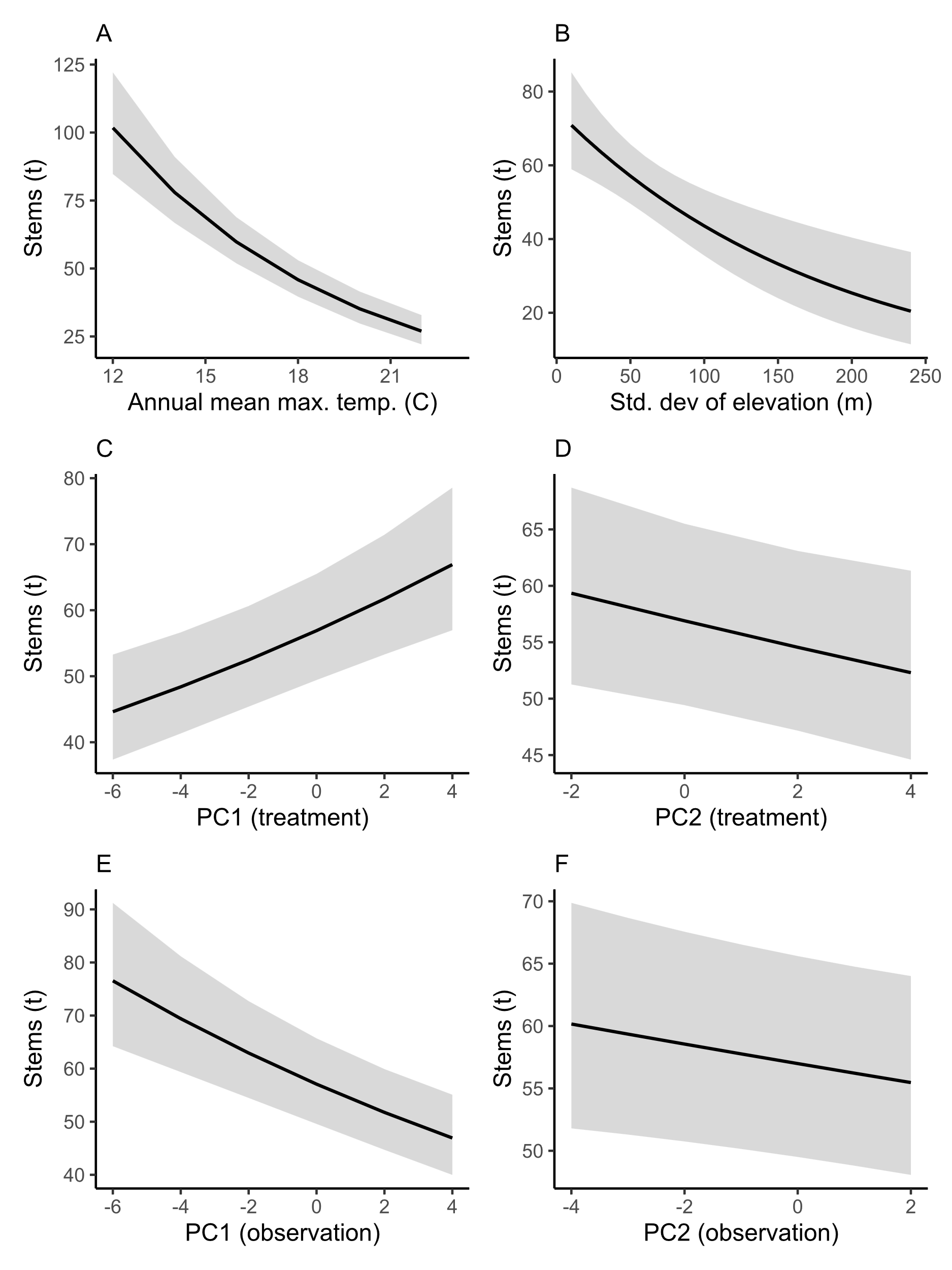


Figure 4. Marginal effects on stem counts at time t of: A. annual mean maximum temperature (in celsius), B. standard deviation of elevation (2.5 km radius around each site), C. PC1 from treatment-time PCA, D. PC2 from treatment-time PCA, E. PC1 from observation-time PCA, F. PC2 from observation-time PCA. Shaded areas represent 95% confidence intervals. N=87.

**Citations**

Ayber, C. (2022) rgee: R bindings for calling the ‘Earth Engine’ API. R package version 1.1.5, https://CRAN.R-project.org/package=rgee

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67, 1–48.

Google. (n.d.) [Google Map of Colorado with sites pinned]. Retrieved June 20, 2023.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18-27.

Hufkins, K., Basler, D., Milliman, T., Melaas, E.K., & Richardson, A.D. (2018) An integrated phenology modelling framework in R: modelling vegetation phenology with phenor. Methods in Ecology & Evolution, 9, 1-10.

Olmedo, O.E. (2022). kriging: Ordinary Kriging. R package version 1.2, https://CRAN.R-project.org/package=kriging

R Core Team. (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

RStudio Team. (2023). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA.

Stegmann, G., Jacobucci, R., Harring, J.R., & Grimm, K.J. (2018). Nonlinear mixed-effects modeling programs in R. Structural Equation Modeling, 25, 160–165. Available at: https://doi.org/10.1080/10705511.2017.1396187.

Thornton, M.M., R. Shrestha, Y. Wei, P.E. Thornton, S-C. Kao, & B.E. Wilson. (2022). Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4 R1. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/2129

U.S. Geological Survey. (2022). 3D Elevation Program 10-Meter Resolution Digital Elevation Model. https://www.usgs.gov/3d-elevation-program

Ushey, K., Allaire, J., Tang, Y. (2023). reticulate: Interface to ‘Python’. R package version 1.29, https://CRAN.R-project.org/package=reticulate

Walker, K. (2023). tigris: Load census TIGER/Line shapefiles. R package version 2.0.3, https://CRAN.R-project.org/package=tigris

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickham, H., François, R., Henry, L. & Müller, K. (2021). dplyr: A Grammar of Data Manipulation. R package version 1.1.0, https://CRAN.R-project.org/package=dplyr

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.E., Xu, P., Ramoino, F., & Arino, O. (2022). ESA WorldCover 10 m 2021 v200. doi:10.5281/zenodo.7254221