Tree survival and growth in central Argentina's highlands: impact of wildfires and land management

Daniel Renison, Iván Barberá, Ana M. Cingolani, Juan Manuel Rodriguez, and Isabell Hensen

1. Models specification and assessment

Fire

We modeled the probability of plots being affected by fire in the study period with a GLM, including the management factor as the only predictor:

$$y_j \sim \text{Bernoulli}(p_j)$$

$$\text{logit}(p_j) = \beta_1 \ \text{ranching}_j + \\ \beta_2 \ \text{conservation}_j$$

$$\beta_{1:2} \sim \text{Normal}(0, 10).$$

$y_j \in \{0, 1\}$ is the response variable, indicating whether the jth plot, with $j \in \{1, ..., 139\}$, was affected by fire or not. "ranching" and "conservation" are mutually exclusive binary variables, indicating the management condition for each plot. The Normal distribution is parameterized by mean and standard deviation.
In all models we defined weakly informative priors based on simulations.

Survival

The survival model was defined as follows

$$y_i \sim \text{Bernoulli}(p_i^{15})$$

$$\text{logit}(p_i) = \beta_0 + \\ \beta_1 \ \text{conservation}_{j[i]} + \\ \beta_2 \ \text{burned}_{j[i]} + \\ \beta_3 \ \text{elevation}_{j[i]} + \\ \beta_4 \ \text{treecover}_{j[i]} + \\ \beta_5 \ \text{height}_i + \\ \varepsilon_{j[i]}$$

$$\beta_0 \sim \text{Normal}(0, 5)$$
$$\beta_{1:5} \sim \text{Normal}(0, 3)$$
$$\varepsilon_j \sim \text{Normal}(0, \sigma)$$
$$\sigma \sim \text{Normal}(0, 3) \ T(0, \infty).$$

$y_i \in \{0, 1\}$ is the response variable, indicating whether i^{th} tree, with $i \in \{1, ..., 880\}$, survived or not throughout the study period, with p_i as its corresponding annual survival probability. The probability of surviving 15 years is p_i^{15}. Plots are indexed by $j \in \{1, ..., 139\}$, and $j[i]$ subsets the plot-level variables corresponding to the ith observation. β_0 is the intercept, referring to management = ranching, and fire = unburned; β_1 is the difference between the intercept and the level management = livestock conservation, while β_2 is the difference with the level fire = burned. "conservation" and "burned" are binary variables indicating whether the ith tree corresponds to plots with livestock conservation and a burned plot, respectively. Continuous predictors were standardized before model fitting. ε_j is the plot-level random effect over the intercept. T(,) indicates the lower and upper bounds for truncated distributions.

We included data from trees for which survival was recorded but tags were not found. As these unidentified trees were more likely to be dead than identified trees, excluding them from the model would have biased the survival probability to higher values. But without the id of a tree we could not compute its survival probability, as it depends on the initial tree height. To solve this, the id of each unidentified tree was treated as an extra parameter to estimate. But an id can only belong to one tree, so it is necessary to treat the ids as sets, where each id cannot be repeated. As example, consider a plot where we recorded two tag-less trees, and after discounting the identified trees we have four ids left, with no assigned tree (the remaining two are lost, we did not find them, and do not know whether they lived or died). The unobserved quantity to estimate would be which of all the possible combinations of these four ids taken in sets of two is the correct id. For example, if available ids are $\{13, 14, 15, 16\}$, we should estimate whether our two unidentified trees belong to the id-combinations $c_1 = \{13, 14\}$, $c_2 = \{13, 15\}$, $c_3 = \{13, 16\}$, and so on. But discrete parameters are usually easier to estimate by integrating them out from the likelihood. So, in each plot (j) with unidentified trees we evaluated the joint likelihood of observing the set of unidentified trees ($L_{c,j}$) in each combination (c), and then, summed the likelihood across combinations (C):

$$L_{c,j} = \prod_{i=a}^{i=b} \text{Bernoulli}(y_i \mid p_{i,c,j}^{15})$$

$$L_j = \sum_{c=1}^{C} L_{c,j}.$$

Where a and b denote the first and last recorded but unidentified trees, belonging to the plot j. $p_{i,c,j}$ is the annual survival probability for the ith tree computed from the initial height defined by the id-combination set c. Bernoulli(y | p) stands for the Bernuolli probability mass function.

We assessed model fit by computing randomized cumulative probability residuals (Dunn and Smyth 1996) using the DHARMa R package (Hartig 2022; Fig. S1 and S2).
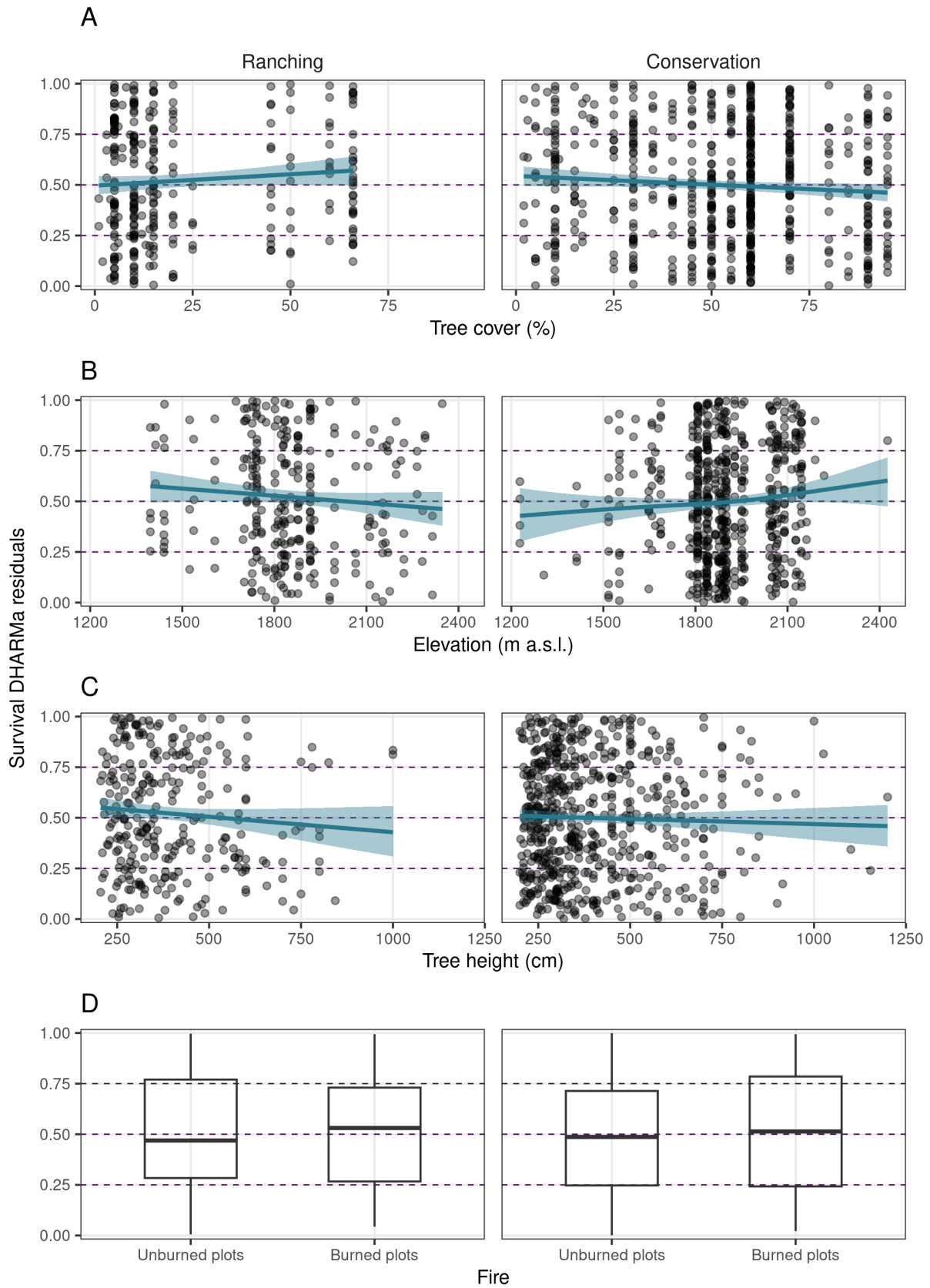
Figure S1. Residuals from the survival model, as a function of explanatory variables. The curves show the prediction from a Beta Generalized Additive Model (GAM) fitted to points. Under good model fit, these residuals show a Uniform distribution in [0, 1], with no patterns as a function of predictors.
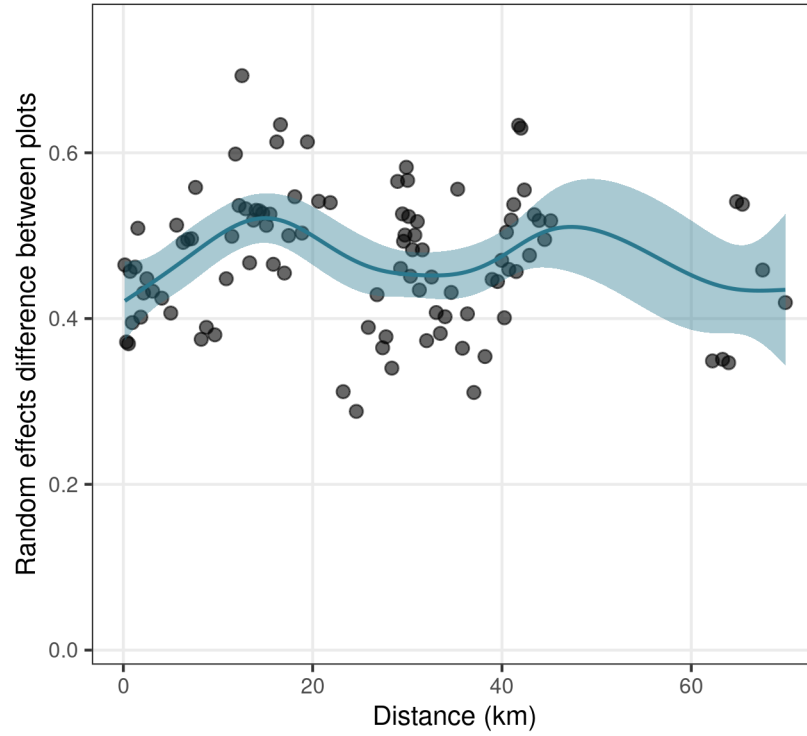
Figure S2. Pairwise absolute difference between plot-level random effects from the survival model (based on posterior means) as a function of the pairwise distance between plots. Data points were averaged in groups of 100 pairs to improve visibility. The line and ribbon show a GAM fitted to points, using nine cubic spline basis functions besides the intercept (maximum likelihood estimate and 95 % confidence interval). An increasing trend at short distance and a plateau afterwards would suggest spatial correlation.

Growth

The annual growth model was defined as follows:

$$y_i \sim \text{Student-t}(\mu_i, \tau_i, \nu)$$

$$\mu_i = \beta_0 +$$
$$\beta_1 \text{ conservation}_{j[i]} +$$
$$\beta_2 \text{ burned}_{j[i]} +$$
$$\beta_3 \text{ conservation}_{j[i]} \text{ unburned}_{j[i]} +$$
$$\beta_4 \text{ elevation}_{j[i]} +$$
$$\beta_5 \text{ treecover}_{j[i]} +$$
$$\beta_6 \text{ treecover}^2_{j[i]} +$$
$$\beta_7 \text{ height}_i +$$
$$\varepsilon_{j[i]}$$

$$\log(\tau_i) = \gamma_1 \text{ ranching}_{j[i]} \text{ unburned}_{j[i]} +$$
$$\gamma_2 \text{ ranching}_{j[i]} \text{ burned}_{j[i]} +$$
$$\gamma_3 \text{ conservation}_{j[i]} \text{ unburned}_{j[i]} +$$
$$\gamma_4 \text{ conservation}_{j[i]} \text{ burned}_{j[i]} +$$
$$\gamma_5 \text{ height}_i$$

$$\beta_{0:7} \sim \text{Normal}(0, 100)$$
$$\varepsilon_j \sim \text{Normal}(0, \sigma)$$
$$\sigma \sim \text{Normal}(0, 100) \text{ T}(0, \infty)$$
$$\nu \sim \text{Gamma}(2, 0.1)$$
$$\gamma_{1:4} \sim \text{Normal}(\log(6.93), 1.5)$$
$$\gamma_5 \sim \text{Normal}(0, 2).$$

Where ν represents the degrees of freedom of the Student-t distribution, and the Gamma distribution is parameterized with shape and rate (usually labeled α and β, respectively). We assumed a Student-t distribution because there were several extreme values that biased the estimation of the mean in the direction of the outliers. We allowed τ to vary with tree height because the residuals analyses suggested higher variance in larger trees, and we also included the combined effects of land management and fire occurrence to account for heterogeneous variance. As we recorded growth only in the trees where we found the tag in 2018, i runs from 1 to 662.
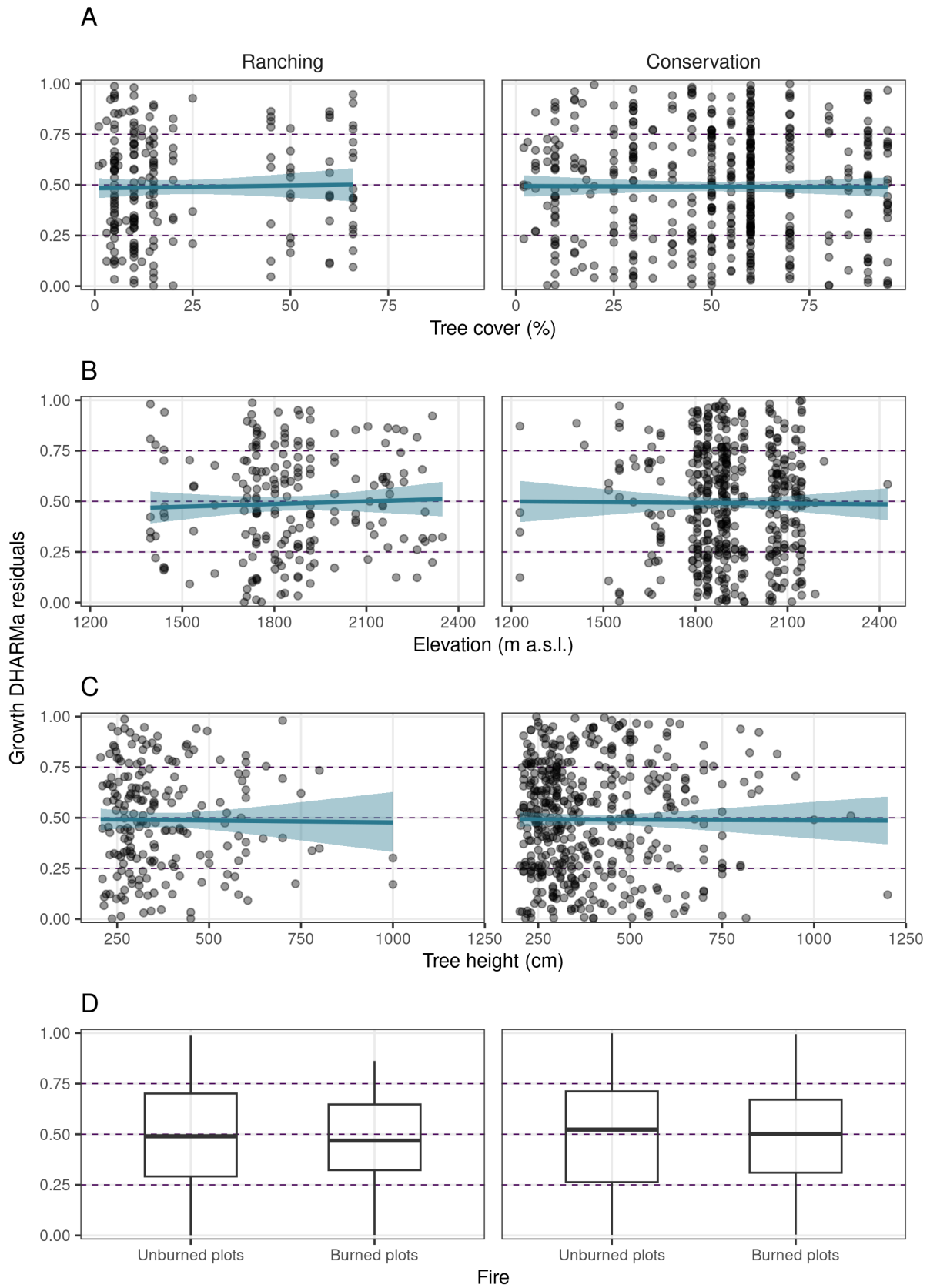
Figure S3. Residuals from the growth model, as a function of explanatory variables. The curves show the prediction from a Beta GAM fitted to points. Under good model fit, these residuals show a Uniform distribution in [0, 1], with no patterns as a function of predictors.
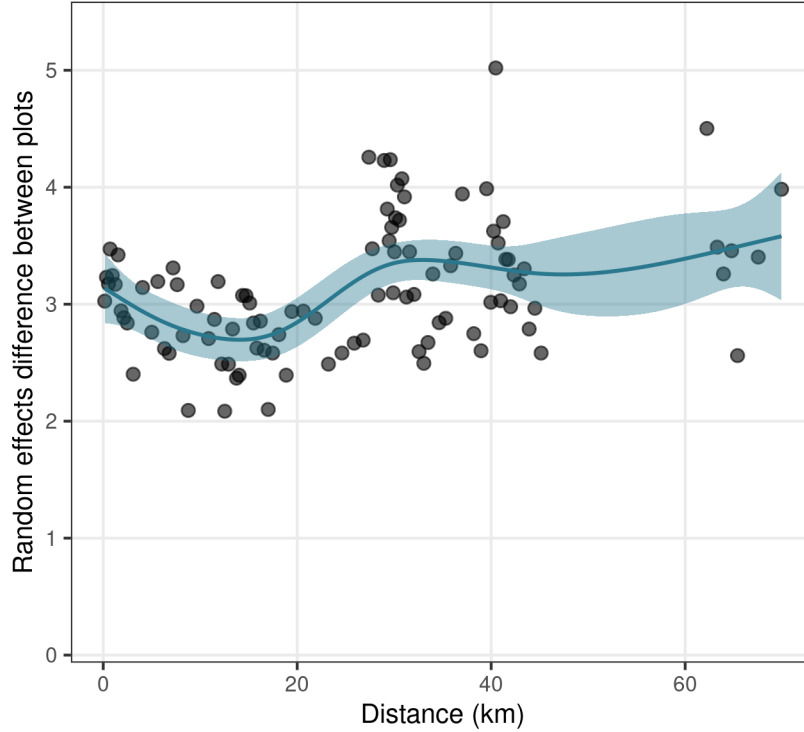
Figure S4. Pairwise absolute difference between plot-level random effects from the survival model (based on posterior means) as a function of the pairwise distance between plots. Data points were averaged in groups of 100 pairs to improve visibility. The line and ribbon show a GAM fitted to points, using nine cubic spline basis functions besides the intercept (maximum likelihood estimate and 95 % confidence interval). An increasing trend at short distance and a plateau afterwards would suggest spatial correlation.

Dieback

Crown dieback, ranging from 0 to 99 %, was transformed to the $(0, 1)$ scale and initially modeled with a Beta distribution. However, residuals suggested that the fitted distribution was not as asymmetric as needed. We first tried to improve this by allowing the dispersion parameter to vary, but this extra complexity did not solve the problem. By transforming the scaled dieback values to the logit scale (replacing zeroes by a small value), we found that the marginal distribution of the data fitted well to a Skew-Normal, with a long tail towards high values (right-skew). Thus we modeled dieback at the logit scale, which implies a Logit-Skew-Normal distribution for the scaled dieback $\in (0, 1)$, with the location and scale as linear functions of predictors. However, if zeroes were converted to very small values (as the logit function is not defined at zero), they were extremely small at the logit scale, and the fitted Skew-Normal lose its skewness to accommodate them, again resulting in a bad fit. Hence, before logit-transforming the data, zeroes were replaced by 0.005 (0.5 % in the original scale, the smallest observed value above zero). In this way, the Skew-Normal fitted a positive skewness and achieved a good fit, even when all predictors were included. Despite the replacement of zeroes to a convenient value may look arbitrary, the proportion of zeroes in the dataset was low. A more principled approach would have been to fit a mixture model, with a Bernoulli component generating zeroes, and a Logit-Skew-Normal generating values above zero. But the results would have been very similar, and the Skew-Normal component would anyway have fitted a right-skewed distribution, due to the small proportion of zeroes. Hence, we decided abandon the endless journey of model refinement at this point. Finally, as the mean of the Logit-Skew-Normal is unknown, we computed it from 500 draws with every posterior sample of the parameters when making predictions.

The model was defined as follows:

$$\text{logit}\left(\frac{y_i}{100}\right) \sim \text{Skew-Normal}(\xi_i, \omega_i, \alpha)$$

$$
\begin{aligned}
\xi_i = \ & \beta_0 \ + \\
& \beta_1 \ \text{conservation}_{j[i]} \ + \\
& \beta_2 \ \text{burned}_{j[i]} \ + \\
& \beta_3 \ \text{conservation}_{j[i]} \ \text{unburned}_{j[i]} \ + \\
& \beta_4 \ \text{elevation}_{j[i]} \ + \\
& \beta_5 \ \text{elevation}^2_{j[i]} \ + \\
& \beta_6 \ \text{treecover}_{j[i]} \ + \\
& \beta_7 \ \text{height}_i \ + \\
& \varepsilon_{j[i]}
\end{aligned}
$$

$$
\begin{aligned}
\omega_i = \ & \gamma_1 \ \text{ranching}_{j[i]} \ \text{unburned}_{j[i]} \ + \\
& \gamma_2 \ \text{ranching}_{j[i]} \ \text{burned}_{j[i]} \ + \\
& \gamma_3 \ \text{conservation}_{j[i]} \ \text{unburned}_{j[i]} \ + \\
& \gamma_4 \ \text{conservation}_{j[i]} \ \text{burned}_{j[i]}
\end{aligned}
$$

$$
\begin{aligned}
\beta_{0:7} &\sim \text{Normal}(0, 10) \\
\varepsilon_j &\sim \text{Normal}(0, \sigma) \\
\sigma &\sim \text{Normal}(0, 10) \ \text{T}(0, \infty) \\
\gamma_{1:4} &\sim \text{Normal}(0, 10) \ \text{T}(0, \infty) \\
\alpha &\sim \text{Normal}(0, 10) \ \text{T}(0, \infty).
\end{aligned}
$$

Where y is the crown dieback, previously replacing zeroes by 0.5 %. ξ is the location parameter, ω is the dispersion parameter, and α is the slant parameter, which determines the skeness. $\alpha = 0$ defines a Normal distribution with $\mu = \xi$ and $\sigma = \omega$. In the model for ω we did not include and intercept, so priors could be defined at the ω scale, rather than at its log, which makes the interpretation harder considering that ω is already at the logit scale. As we recorded growth only in the trees where we found the tag in 2018, i runs from 1 to 662.
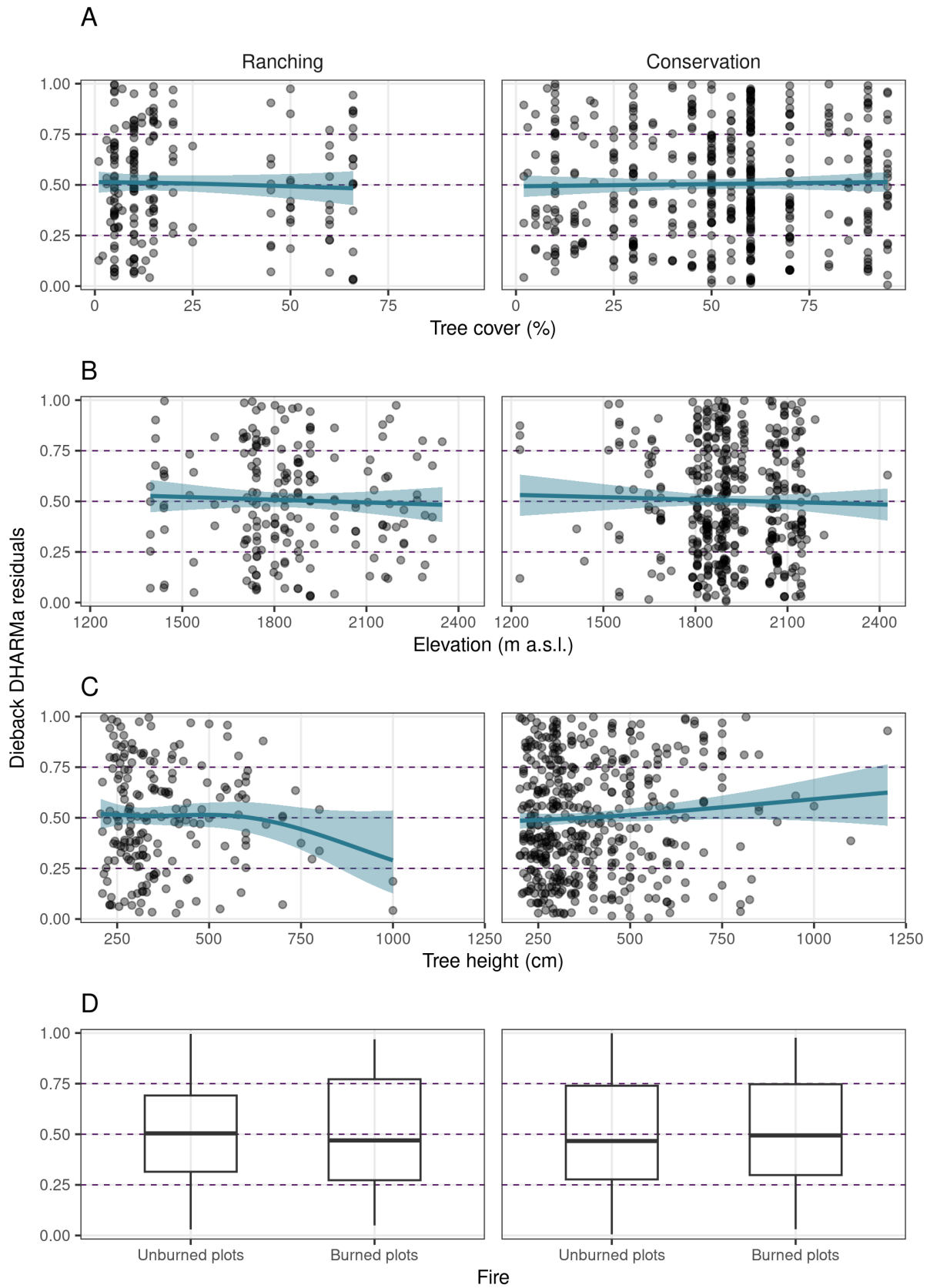
Figure S5. Residuals from the dieback model, as a function of explanatory variables. The curves show the prediction from a Beta GAM fitted to points. Under good model fit, these residuals show a Uniform distribution in [0, 1], with no patterns as a function of predictors.
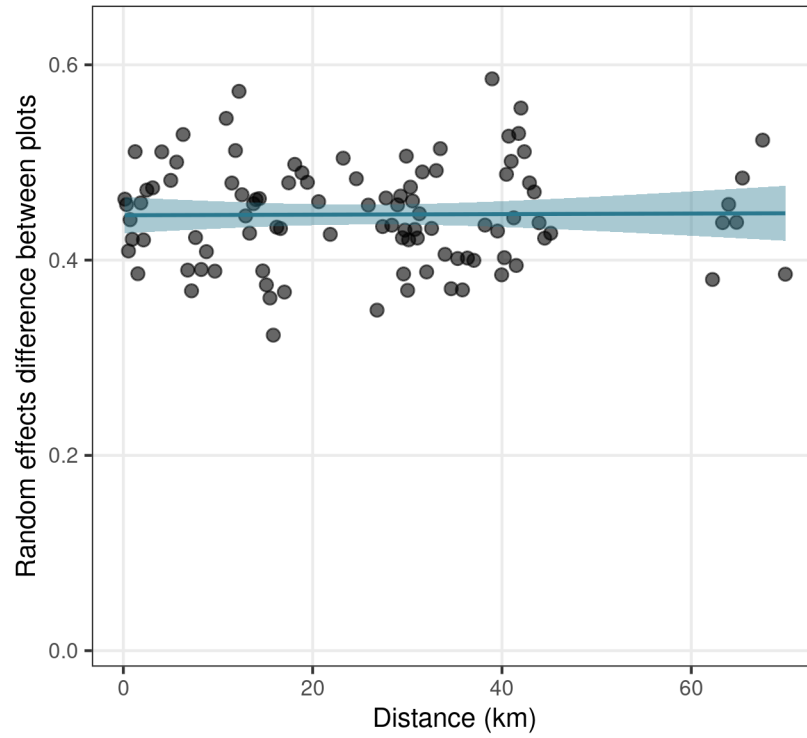
Figure S6. Pairwise absolute difference between plot-level random effects from the survival model (based on posterior means) as a function of the pairwise distance between plots. Data points were averaged in groups of 100 pairs to improve visibility. The line and ribbon show a GAM fitted to points, using nine cubic spline basis functions besides the intercept (maximum likelihood estimate and 95 % confidence interval). An increasing trend at short distance and a plateau afterwards would suggest spatial correlation.

2. Additional figures

Partial predictions from the survival, growth and dieback models for elevation and tree height.
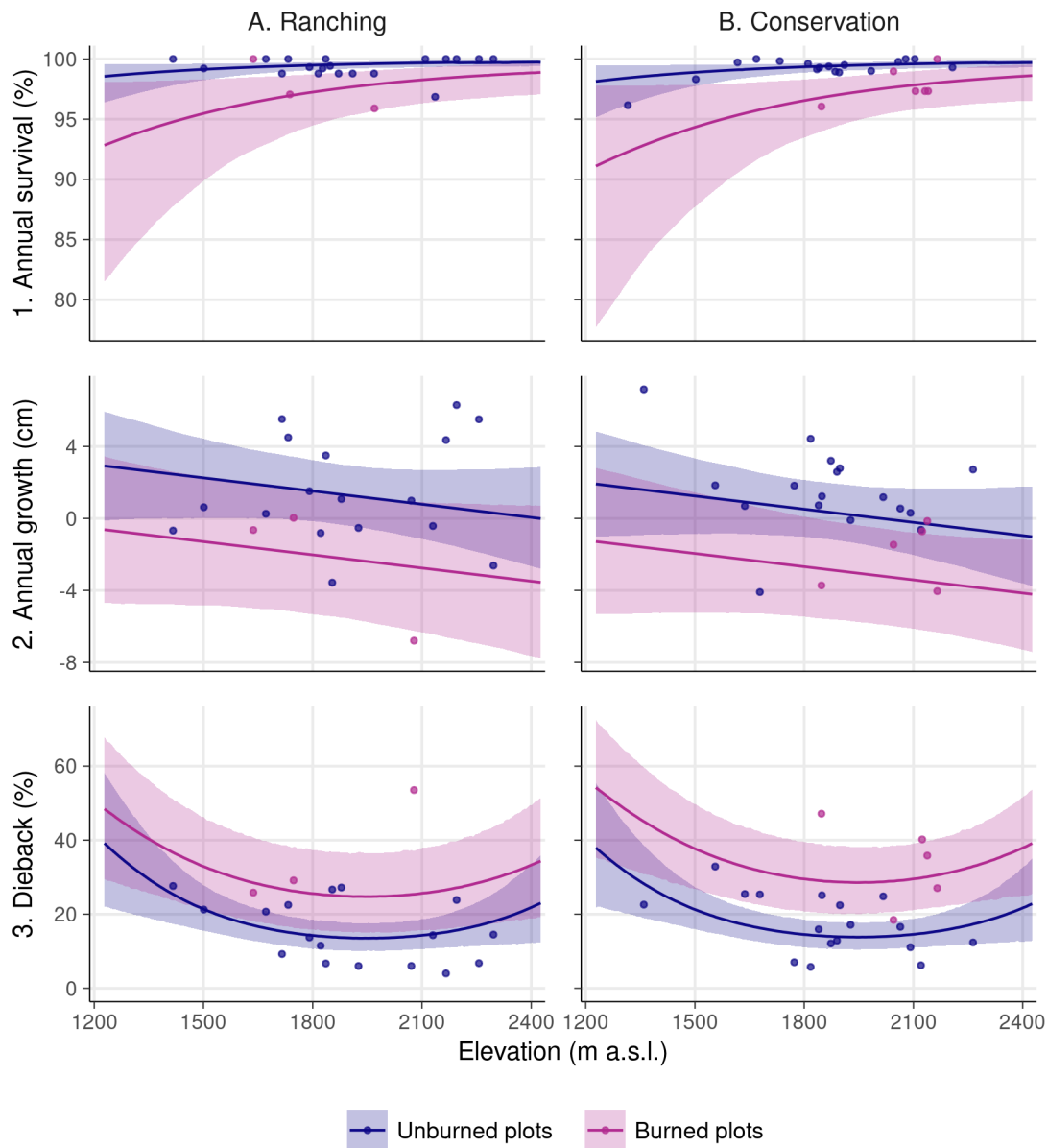

Figure S7. Annual survival probability, annual net growth in height and dieback by land management and fire, as a function of elevation. Lines show the predicted mean for an average plot using the posterior mean, and ribbons show the 95 % equal-tailed credible intervals. Predictions were computed fixing tree cover and tree height at their means across plots (tree cover = 22.8 %; tree height = 385.5 cm). Points show the plot-level means averaged in groups of three plots (for survival proportions, these values were raised to the 1/15th power to approximate annual values).
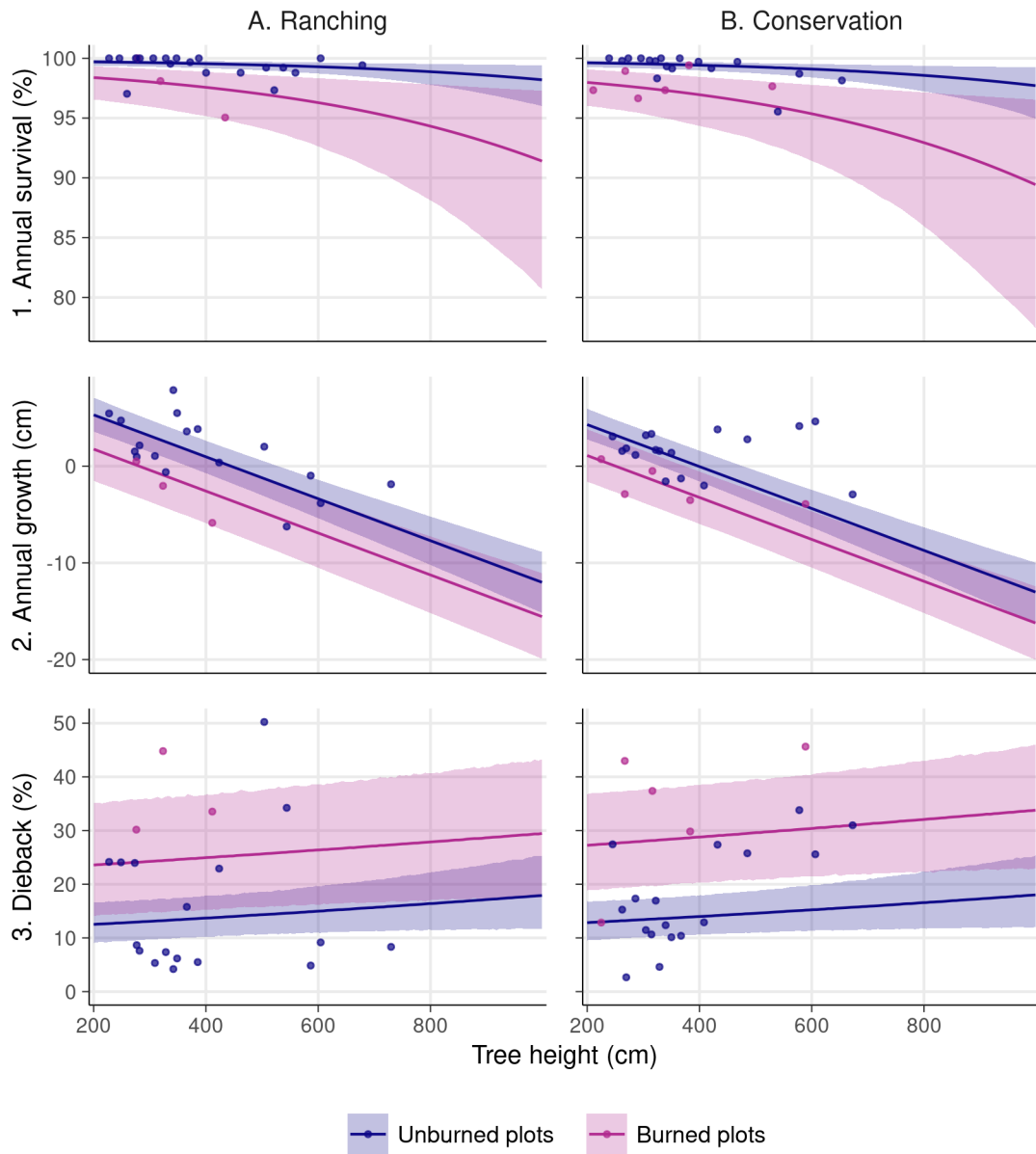
Figure S7. Annual survival probability, annual net growth in height and dieback by land management and fire, as a function of tree height. Lines show the predicted mean for an average plot using the posterior mean, and ribbons show the 95 % equal-tailed credible intervals. Predictions were computed fixing tree cover and elevation at their means across plots (tree cover = 22.8 %; elevation = 1898 m a.s.l.). Points show the plot-level means averaged in groups of three plots (for survival proportions, these values were raised to the 1/15th power to approximate annual values).

3. R^2 at the observation and plot level

Following Gelman et al. (2019), we computed the Bayesian R^2, both at the observation and plot level. In the former, the random effect of the plot was not taken into account, so it was marginalised out through simulation. This avoids getting large R^2 values that are actually related to plot-level unexplained variability.

As the Bayesian R^2 does not depend directly on the observations of the response variable but is based only on the posterior predictive distribution, we used its formulation to compute the R^2 of the mean response at the plot level. By defining a hierarchical model with random effects of the plot, the mean at the plot level is a random variable itself, so we can apply the same computations defined by Gelman et al. (2019), by using the posterior predictive distribution of the mean at the plot level. In particular, it is challenging to define the mean at the plot level if there are predictors at the observation level (like tree height in our models), and even more if predictors at different levels are correlated. To simplify computations we assumed no correlation among tree height and plot-level predictors (elevation, tree cover, land management and fire), as they showed weak to null correlation. Hence, to estimate the conditional mean of the response at the plot level we fixed the tree height at its mean.

To the best of our knowledge, the idea of computing R^2 values at different levels of hierarchical models has not been studied, so this calculation must be considered experimental.

References

- Dunn, P. K., & Smyth, G. K. (1996). Randomized quantile residuals. Journal of Computational and graphical statistics, 5(3), 236-244.

- Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for Bayesian regression models. The American Statistician 73(3), 307–309. doi.org/10.1080/00031305.2018.1549100

- Hartig, F. (2022). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.6, https://CRAN.R-project.org/package=DHARMa.