# Understanding Unintentional Tree Mortality in Selectively Logged Areas of Tapajós National Forest

Chainey Boroski, Israel Golden, Caitlin Markus

11/16/2020


## ABSTRACT

Our team used data gathered by the Oak Ridge National Laboratory and NASA that described the effect of logging on non-target trees in a selectively logged stand in Tapajós National Forest. The data describe the type and severity of damage done to non-target trees as well as several measures of biomass, both directly measured and calculated via allometric equations. We examined the effects of these measurements on tree mortality using a poisson regression model and two binary logistic regression models. Bole snap fraction, DBH, crown damage and the interaction between crown damage and bole snap fraction were found to have significant effects on tree mortality in the poisson regression. In the second model, concerned specifically with proportional damage, only bole snap fraction was found to have a significant effect. In the third model, bole snap fraction and DBH had a significant effect on tree survival, with the effects of damage code noticeable on survival outcomes. Understanding the relationship between tree mortality and these variables can help logging operation managers to better protect non-target trees critical to stand regeneration. This analysis suggests that reducing bole snap fraction damage, especially among trees with high DBH, can reduce mortality in non-target trees and aid foresters in preserving the ecological integrity of selectively logged stands in critical habitat.


## INTRODUCTION

Sustainable timber harvesting practices are crucial for the health of ecological communities, as well as the maintenance of a steady source of forest resource products for human consumption (Asner, 2006). Forest Stewardship Council (FSC) certification requires that timber harvesters abide by strict regulations regarding maximum timber harvests per acre and preservation of mature trees for stand regeneration (Ribeiro 2020). However, even in sustainably logged stands, collateral damage to neighboring, non-target trees may still occur (Figueira et al., 2008). Thus, the mortality of neighboring trees and its causes must be considered when sustainably harvesting timber.

In our data analysis, we analyzed unintentional logging damage data compiled by NASA and Oak Ridge National Laboratory (ORNL) in the Tapajós National Forest to assess the causes of mortality of neighboring trees in a sustainably harvested stand. Located in Pará, Brazil, it is a part of the Amazon Rainforest that is designated for "sustainable exploitation" (Ribeiro 2020). Logging may alter habitat through deforestation and the creation of edges. Deforestation in the Amazon has implications not only for its immediate surroundings, but also for the world as its carbon storage capacity is critical to the effort to curb global climate change (Lovejoy & Nobre, 2019). It is for these reasons that a closer look at mortality (and what effects it) in neighboring trees associated with sustainable harvesting is warranted.

Our research investigated the question of which factors contribute significantly to tree mortality in selectively logged stands, specifically in relation to 1) morphological traits (e.g. DBH, biomass measurements) and 2) type of damage (damage code). We also investigated the factors contributing to tree mortality at the plot level, and whether number of tree fatalities could be predicted. We hypothesized that 1) trees with greater stature and biomass would have a higher likelihood of surviving unintentional damage, and 2) that the relationship between variables contained in the dataset would allow for the prediction of tree mortality at the individual and plot level. If the second hypothesis were supported, this would have the potential to inform the industry on how to improve the practice of sustainable logging and reduce unintentional tree damage.

## METHODOLOGY

The data for this analysis were collected as part of the Large-Scale Biosphere-Atmosphere Experiment (LBA-ECO) in Amazonia, a joint investigation between NASA and Oak Ridge National Laboratory (ORNL) of the impacts of land use in the tropical rainforests of Amazonia ("Large Scale Biosphere," n.d.). This data set catalogs damage done to 982 non-target trees damaged in a 2001 reduced-impact logging operation in an 18 hectare plot in Tapajós National Forest. Prior to logging, the plot was surveyed in the year 2000 and a grid of transects 25 meters apart was established. The plot was then logged throughout September and December of 2001 and then re-surveyed for damage done to non-target trees ("LBA-ECO", 2011). For every damaged tree, its location within the plot and type of damage were recorded along with measurements of diameter at breast height (DBH), bole snap height, leaf area lost and whether the tree was living or dead. Pre-damage estimates of crown, bole, and total biomass as well as bole height and total leaf area were calculated from allometric equations (Chambers et al., 2001). Canopy mass lost, leaf area lost, bole snap mass, bole snap fraction, and crown loss fraction were calculated with estimates of original values and results from the post-logging survey. Based on this information it is assumed that the data are independent and identically distributed random variables, a key assumption for statistical analysis.

The dependent variable in each model ~~analyzed~~ was tree mortality ("Live_dead"). Though variables were numerous, many were derivative of a core set of data, and care needed to be taken to avoid autocorrelation. The independent variables used in the models to follow were DBH (cm; mean of 18.17 used for trees without a DBH measurement), crown damaged (0-1: ratio of leaf area lost to total leaf area), bole snap fraction (0-1, where 0 is no snap and 1 represents a snap at ground level, calculated as the snap height to bole height ratio), and damage code (bole snap (BS), crown loss (CL), severe lean (SL), standing (ST) and entire tree down (DN)). The dataset's biomass measurements were calculated from DBH – therefore, DBH was used as a proxy for most biomass variables in Models 1 and 3. No data were missing from observed variables.

We used three models to explore our research questions. The first, a poisson regression mixed model with DBH, bole snap fraction, and crown damage as predictors for tree mortality. We analyzed this data without transformation (as count data need not be normally distributed) and checked assumptions of lack of multicollinearity and overdispersion, as well as independence of residuals, via a pairs plot, an overdispersion test, and analysis of a residuals vs fitted plot. Thereafter, we remedied deviations from these assumptions. The second and third models were binary logistic regression models. Model 2 used bole snap fraction, crown damage, and damage code as predictors for tree mortality, coded 1 for Dead, 0 for Living. Crown damage was partitioned into low (< 0.8) and high (>= 0.8) categories to account for left skew, but data were otherwise not transformed. We used the VIF function to check that the data had little to no multicollinearity. Model 3 used damage code, DBH, and bole snap fraction as predictors of tree survival, coded 1 for living, 0 for dead. We used a pairs plot to check for multicollinearity. Despite the partial skew in DBH, we used this data without transformation, as transformation worsened model fit and residual spread. For the logistic regressions, the response was dichotomous and the assumption of a logit-linear predictor-response relationship was tested via the model. For all three models, sample size was large (~ 1000), meeting assumptions of each model. We selected minimum adequate models as those with the lowest AIC, where maximum number of predictors were significant and the model was supported by R2 values.

## RESULTS

We made a pairs plot for variables from each model to ensure correlation values were low (Figure 1). Due to this, we modified model 1 from its original form (which included biomass and leaf loss measurements) to avoid autocorrelation; we removed two of the 982 observations due to excessively high residuals, and separated the data into gap identification plots as outlined in the dataset (A-K); there were varying numbers of recorded observations per gap ID. Numbers of observations within each damage code, used in Models 2 and 3, varied: 213 for Tree Down, 304 for Crown Loss, 15 for Standing, 447 for Bole Snap, and 3 for Severe Lean. Both Models 2 and 3 used all 982 observations. This raised concern for accuracy of predictions for the two smallest observation groups but was encouraging for those of the other three groups. DBH and Crown Damage were not normally distributed; however, we remedied/accounted for these as aforementioned.

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

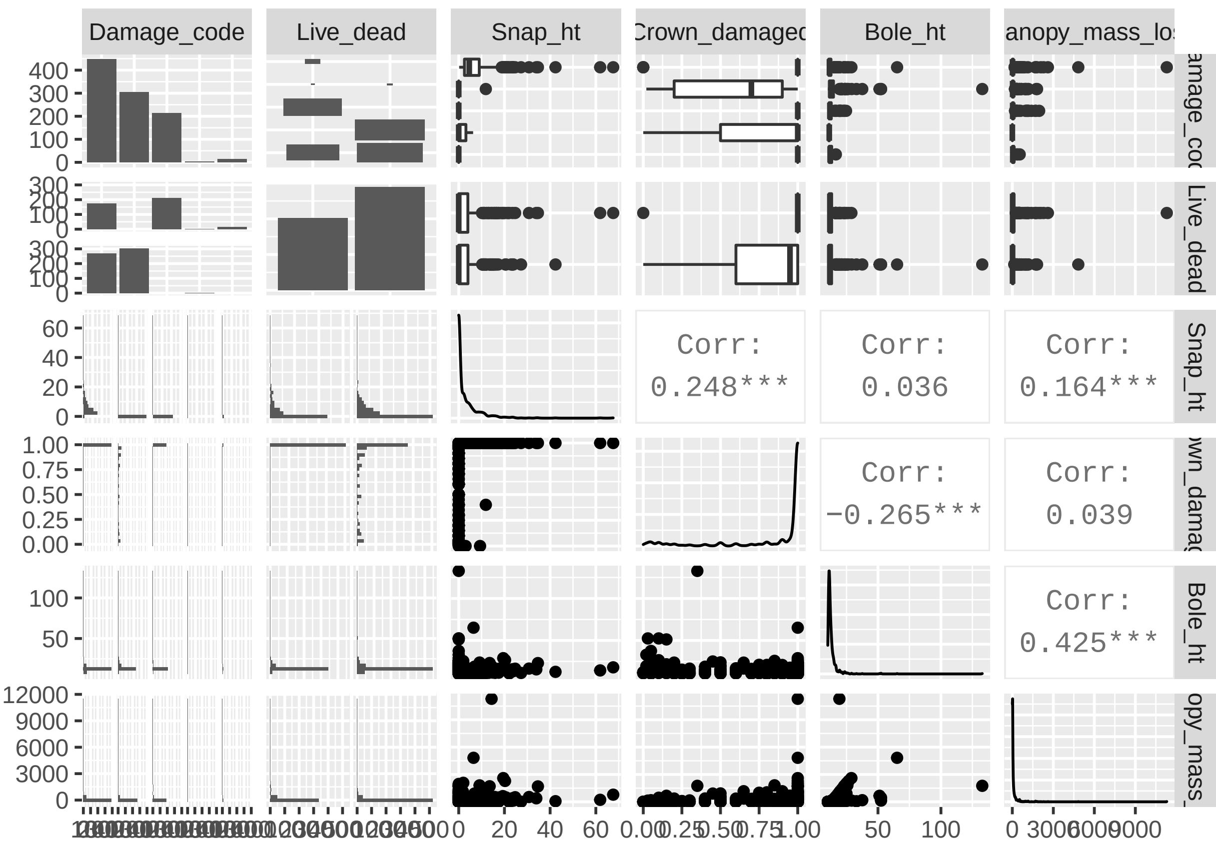


Figure 1: GGPairs Plot of All Variables Used in the Three Models

Statistical Analyses

Model 1's $H_o$ was that all included independent variables and interactions would be equal and have no effect on the number of dead trees found in a plot ($H_a$: at least one predictor would have a significant effect/non-zero coefficient). The minimum adequate model (lowest AIC) included the main effects and interactions of DBH, bole snap fraction, and crown damage and gap identification as a random effect. However, this model was overdispersed, and we ran a negative binomial model in its place to correct this. Back-transformed results were as follows. The intercept (dead tree with bole snap fraction 0, 0 DBH, and 0 crown damage) had a coefficient of 4.21e+01 (z = 1.085, df = 973, p = 0.278). DBH had a significant effect on the dead tree count in a plot, with a multiplicative effect of 1.01 on count for every unit increase of DBH (z = 2.284, df = 973, p = 0.0224). Crown damage had a significant influence on bole snap fraction's effect (and vice versa) on dead tree count as well, with a multiplicative effect of 3.45e+12 (z = 4.555, df = 973, p = 5.23e-06). The interaction between crown damaged and bole snap fraction had a very strong effect on the number of dead trees found within a plot (Figure 2, where curves represent crown damage's influence on different levels of bole snap fraction). The higher the crown damage and bole snap fraction, the higher the count of dead trees within a plot.

```
## boundary (singular) fit: see ?isSingular
```

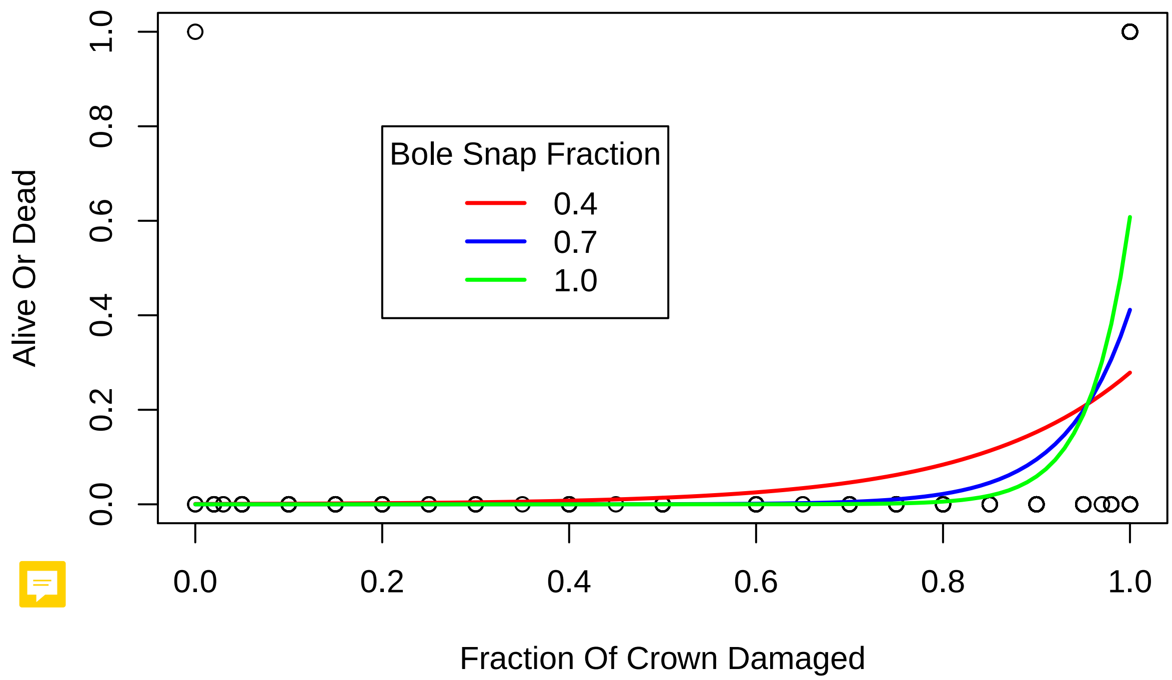


Figure 2: Effect of Interaction between Crown Damage and Bole Snap Fraction on the Count of Dead Trees within a Plot

Model 2's $H_o$ was that bole snap fraction, proportion of crown damage, and recorded damage code had coefficients equal to zero – essentially, that tree mortality had no relationship with any of the independent variables ($H_a$: at least one predictor had a significant effect/non-zero coefficient). The model with all main effects and no interactions was best fit to the data (Chi-squared ANOVA, p=0.60). This model also had the lowest AIC of its nested counterparts. Though this was the best model of all nested models, not all terms were significant. The only significant terms in the model were its intercept, representing a tree with a bole snap fraction of 0, high crown damage, and a damage code of bole snap (z = -3.57, df = 974, p $<$ 0.01), and bole snap fraction. Though damage code and crown damage had insignificant effects (p>0.05) the model was better fit when they were included.

Inverse logit transformation of log-odds indicated that the intercept tree had a 28% probability of mortality. Exponentiated log-odds indicated that for every one-unit change in Bole Snap fraction, the odds of mortality increased by a factor of 2.15 (z = 2.12, df = 974, p = 0.03). Low crown damage was associated with decrease in mortality relative to high crown damage with a 0.671 decrease in log odds of mortality (z = 1.21 , df = 974, p = 0.58) (Figure 3). This indicates that trees are more likely to suffer mortality at high bole snap fractions than at low bole snap fractions. The decreased probability of mortality associated with low crown damage is demonstrated through the effects of bole snap and severe lean damage codes (p $>$ 0.05). The gradual slopes suggest poor explanatory power for predicting mortality. All other damage codes do not exhibit this effect;

all downed and standing trees died and all trees with crown loss survived.

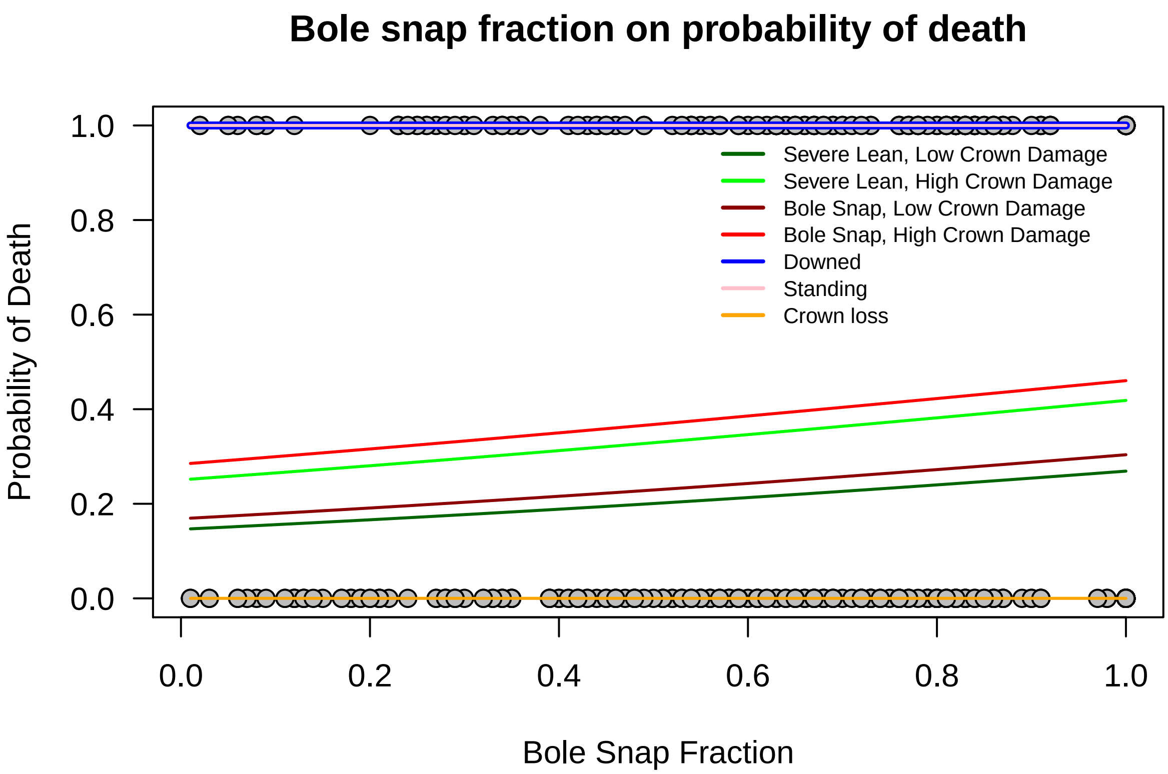


Figure 3: Effect of Bole Snap Fraction on Probability of Death

The third model's null hypothesis was that bole snap fraction, DBH, and damage code had coefficients equal to zero - in essence, that they did not explain the variation in tree survival ($H_a$: at least one predictor had a significant effect/non-zero coefficient). The minimum adequate model included the three independent variables without interaction (null deviance: 1331.06 on 981 DF; residual deviance: 572.45 on 975 DF; AIC: 586.45). The intercept represented a DBH of 0, a bole snap fraction of 0, and the damage code "Bole Snap," and was highly significant ($z = 5.197$, $p < 0.0001$); it should be noted that this intercept is nonsensical, given that a tree cannot have a bole snap fraction of 0 and also be coded as "Bole Snap" in its damage code. DBH was also highly significant ($z = -4.651$, $p < 0.0001$), which reduces the odds of survival by a factor of .96 per unit increase with a probability of 48.9%. Damage codes were non-significant ($p > .92$), however, there was a clear correlation between the damage codes and tree survival. CL (canopy loss) never led to tree mortality and TD (tree down) and ST (standing) produced mortality 100% of the time. SL (severe lean) and BS (bole snap) were the only damage codes for which survival probability was variable - however, as stated previously, there were only three observations under the SL damage code. Bole snap fraction was significant ($z = -2.054$, $p = 0.040$), with a unit increase in bole snap reducing the odds of survival by a factor of .453 with a probability of 31.2%. Figure 4 shows survival probability as separated by damage code over the range of DBH values, clearly indicating the negative relationship between DBH and tree survival in those trees which had a severe lean or whose boles were snapped; Figure 5 indicates a similar (though less visually dramatic) trend in the relationship between bole snap fraction and probability of tree survival.

Adherence to Model Assumptions

We checked Model 1's assumptions of little to no multicollinearity, no overdispersion, and independence of the datapoints via pairs plots, a dispersion test, and residual plots (Figure 5). We removed overly correlated

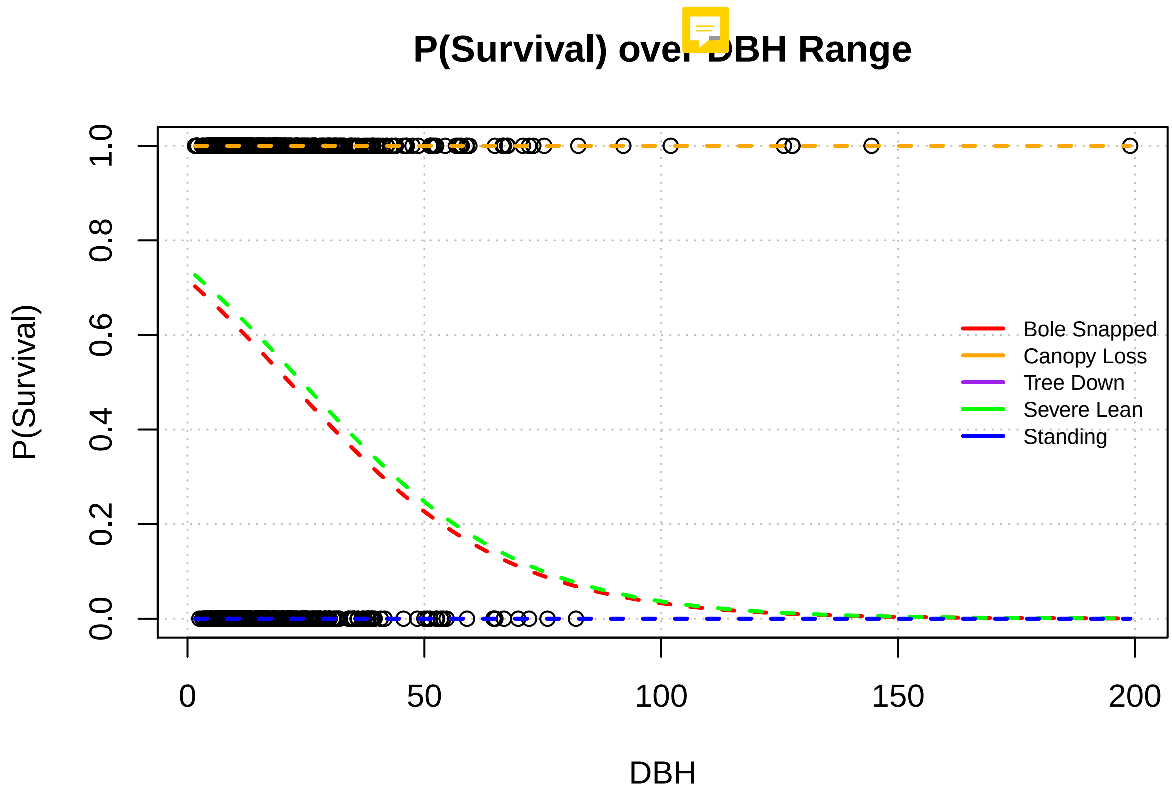


Figure 4: Probability of Tree Survival over the Range of DBH Values, as Predicted by the mix5 Model

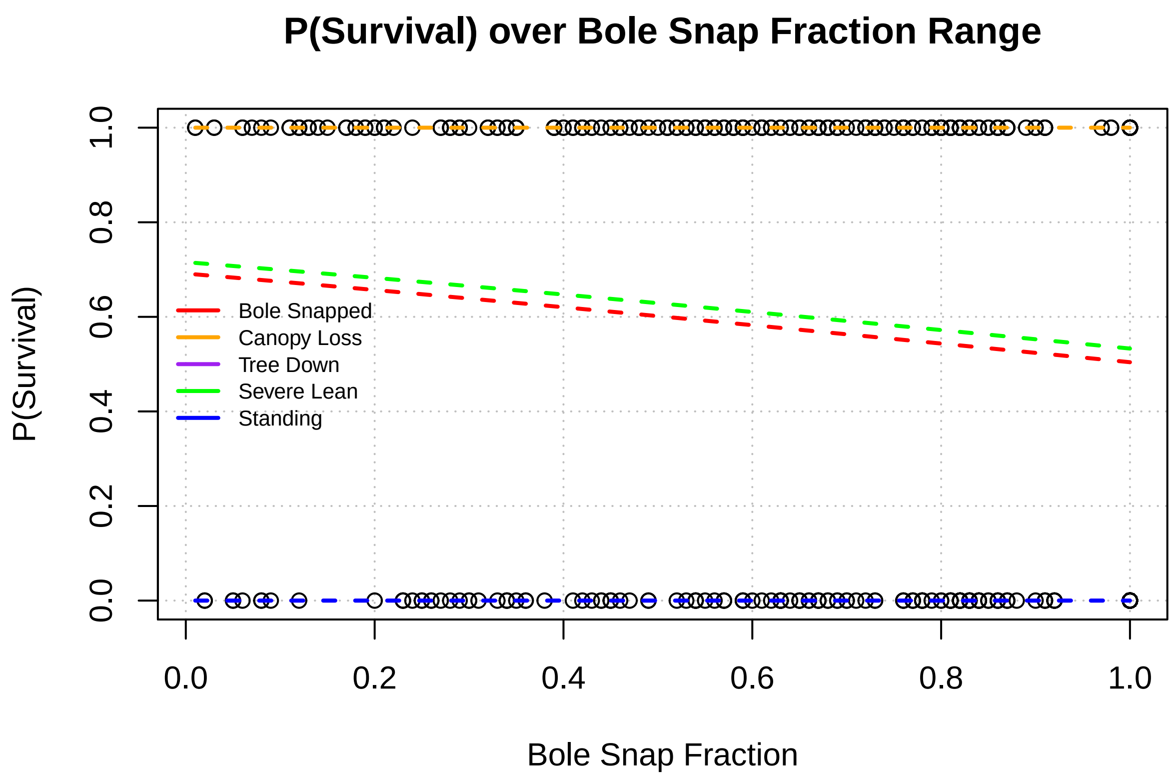


Figure 5: Probability of Tree Survival over the Range of Bole Snap Fraction Values, as Predicted by the mix5 Model

data prior to model formation and remedied overdispersion via a negative binomial model. Large sample size assumptions were met (54 observations in the smallest plot), and residuals were patterned, but less than 2 on the whole. R2 values for this model were .922, indicating that it accounted for most of the variance. The random variable did not account for any additional amount of variance, but was retained due to its relevance to the question being asked. Model 2 and 3 met assumptions of a dichotomous response variable, a logit-linear predictor-response relationship, and were individually IID. VIF analyses for the two models produced values below the 5 threshold, indicating a lack of multicollinearity. Chi-square goodness of fit tests for Models 2 and 3 produced very low values (6.04e-156; 1.345135e-160, respectively), but McFadden's R2 values were similarly high (~55%; ~57%). While neither Model 2 nor 3 was an exceptionally good fit, both captured general trends within the data.

```
## boundary (singular) fit: see ?isSingular
```

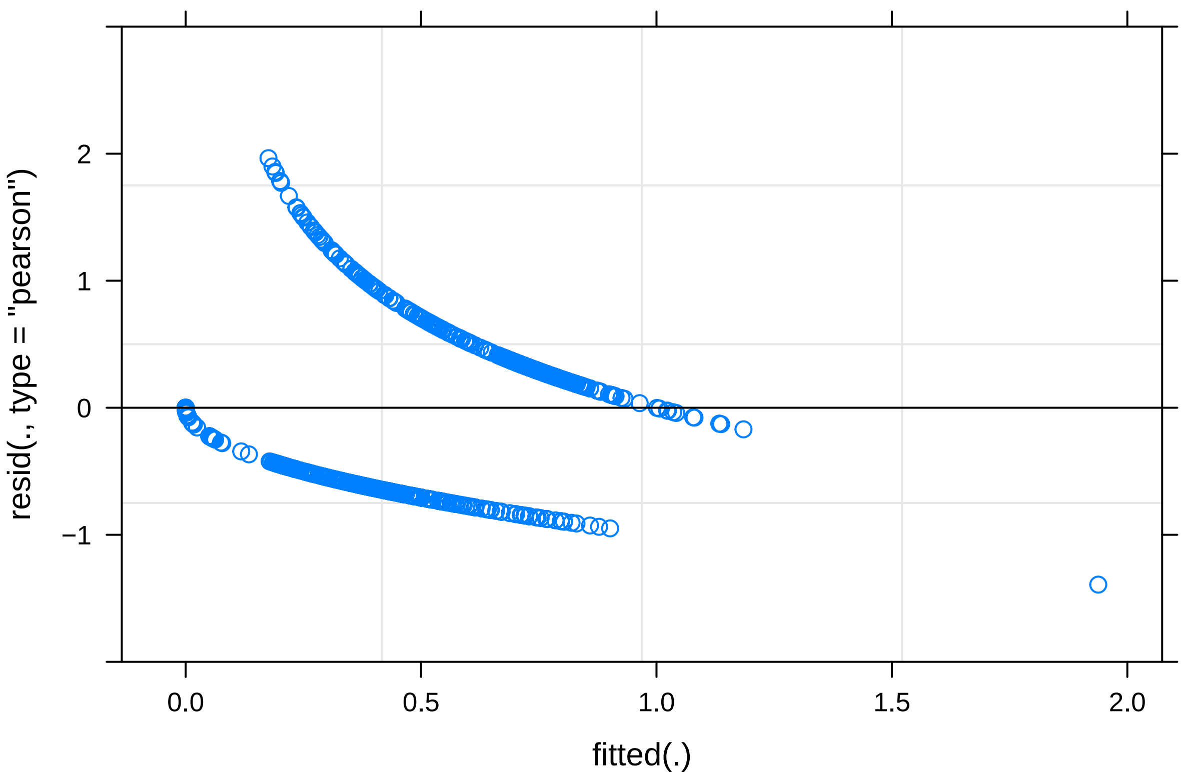


Figure 6: Residuals VS Fitted Diagnostic Plot for Model 1

## DISCUSSION and CONCLUSION

On the whole, effects of the independent variables tended to be small, and fits generally poor (for Models 2 and 3). Intercepts for Models 2 and 3 were nonsensical, as a tree coded "Bole Snap" with a snap fraction of 0 could not occur. This stated, results were consistent among models, and the sample size of data was large enough to inspire confidence in the response of the dependent variable to the predictors. Additionally, outside of a purely statistical interpretation, significant predictors were generally commonsensical, and adhered to reasonable expectations (with some exception, to follow).

Our initial hypothesis that the variables included within these data would be able to produce a predictive model for tree mortality was loosely supported. Crown damage (Model 1) and bole snap fraction (Models 1, 2, 3) had significant effects on the probability of mortality, with crown damage being of note to describing the number of dead trees in each plot. However, we anticipated that the interaction between bole snap fraction and crown damage would be significant, but this was not the case. Further, proportional variables alone were not enough to accurately predict tree mortality (as can be seen when comparing Model 2 and Model 3). Model 3, which included the biomass proxy variable DBH (as did Model 1), did explain a reasonable proportion of variation, but was on the whole a poor fit.

The hypothesis that greater stature and biomass would increase probability of survival was not supported, and rather was directly contradicted by Models 1 and 3. DBH negatively correlated with probability of survival, especially in trees for which the bole had snapped. This could speak to the force of impact (as it would take greater force to snap a thicker trunk), or potentially the age of the trees, as older trees of the same species would have a tendency to have a higher DBH than younger trees, and could theoretically be less resilient to damage. Without the data to support this, however, interpretations of possible reasons for this trend are hypothetical.

It should be further stated that we did not anticipate the relationship of damage code to tree mortality/survival to be as minimally variable as analyses revealed. As previously stated, three of the five damage codes produced fixed mortality probabilities. Of the two which were not fixed, only one − bole snap − was of a sample size that could be statistically analyzed. Given this, it is unsurprising that bole snap fraction was a significant predictor, as it was one of the only predictors to which there was a response via mortality.

When considering the poor fit of the Chi-squared tests, we feel it is reasonable to conclude that there are additional variables responsible for the variation observed in this dataset. For one example, in a richer dataset with more variation within each damage code subcategory, more predictive elements could have potentially been revealed. Further, more detailed tree demographic information (species, age, direct biomass measurements) as well as environmental data concerning elements such as air quality, drought conditions, and pest outbreaks, could more completely represent a tree's natural proclivity to mortality, allowing the effects of direct damage to be teased apart from pre-conditions. Relevant land use history, presence of additional stressors, and proximity of the land to other land types could further enrich these data.

Despite concerns regarding model fit and a lack of more information, our results indicate that for types of damage for which mortality is variable, bole snap fraction and crown damage are significant predictors of tree mortality, and DBH is negatively correlated with tree survival when damage to the bole is incurred. The addition of more independently derived predictor variables could further bolster this result, as well as reveal additional relationships. While these data come specifically from Tapajós National Forest, and are specific to this population, the results of our analyses have applicability beyond the Tapajós to locales where forests of similar species of trees are being selectively logged. These results, and those of similar studies, could be used to inform more sustainable logging practices in selectively logged areas around the world.

## BIBLIOGRAPHY

Asner, Gregory P., et al. "Condition and Fate of Logged Forests in the Brazilian Amazon." Proceedings of the National Academy of Sciences of the United States of America, vol. 103, no. 34, 2006, pp. 12947–12950. JSTOR, www.jstor.org/stable/30051114. Accessed 4 Oct. 2020.

Chambers, J. Q., Santos, J. D., Ribeiro, R. J., & Higuchi, N. (2001). Tree damage, allometric relationships, and above-ground net primary production in central Amazon forest. Forest Ecology and Management, 152(1-3), 73-84. doi:10.1016/s0378-1127(00)00591-0

Figueira, A.M.E.S., S.D. Miller, C.A.D. de Sousa, M.C. Menton, A.R. Maia, H.R. da Rocha, and M.L. Goulden. 2008. Effects of selective logging on tropical forest tree growth. Journal of Geophysical Research-Biogeosciences 113. G00B05

Large Scale Biosphere-Atmosphere Experiment (LBA-ECO). (n.d.). Retrieved November 11, 2020, from https://daac.ornl.gov/cgi-bin/dataset_lister.pl?p=11

LBA-ECO CD-04 Logging Damage, km 83 Tower Site, Tapajos National Forest, Brazil. (2011, September 7). Retrieved November 11, 2020, from https://daac.ornl.gov/LBA/guides/CD04_Logging_Damage.html

Lovejoy, T., & Nobre, C. (2019, December 01). Amazon tipping point: Last chance for action. Retrieved October 05, 2020, from https://advances.sciencemag.org/content/5/12/eaba2949

TRANCOSO, R., FILHO, A., TOMASELLA, J., SCHIETTI, J., FORSBERG, B., & MILLER, R. (2009). Deforestation and conservation in major watersheds of the Brazilian Amazon. Environmental Conservation, 36(4), 277-288. Retrieved October 4, 2020, from http://www.jstor.org/stable/44520025

Ribeiro, M. F. (2020, July 22). A Brazilian forest community shows certified timber really does work. Retrieved October 4, 2020, from Mongabay.

Schwartz, G., Lopes, J., Kanashiro, M., Mohren, G., & Peña-Claros, M. (2014). Disturbance Level Determines the Regeneration of Commercial Tree Species in the Eastern Amazon. Biotropica, 46(2), 148-156. Retrieved October 4, 2020, from http://www.jstor.org/stable/24030114