Appendix B: Estimating missing sizes

When the field crews encounter a dead tree, they have no way of knowing when it died in the intervening time since the last survey, and the size is recorded in the database as zero. Because we choose to interpolate tree size to create a yearly individual-level variable for our survival models, we need sizes for all trees in all years. When a tree is measured more than twice, the size at potential times of death can be interpolated from those other sizes, but when the tree is measured once and then dies, only one size is available to be interpolated. Therefore, we need a way to estimate that missing size in order to interpolate.

We tried three strategies to address these missing sizes: 1) exclude the trees; 2) use a naive linear model for the trees which do have multiple sizes to estimate the growth as a function of size, averaging over census intervals, and predict the missing size; and 3) use the results from another state-space model for growth, which is designed to estimate missing sizes (Eitzel et al. 2013). We tried strategy three only for white fir.

Note that the ideal solution is to jointly estimate the missing size **and** the missing survival status in each year, but such a joint estimation is beyond the scope of this project. In the case of modeling growth, where size at the previous time predicts size at the next time (Eitzel et al. 2013), interpolating or estimating missing sizes as a separate step would be much more statistically questionable than it is here. In this case, the size is a predictor of mortality, not both a predictor and response variable. Thus the estimation and interpolation introduce minimal error into the models.

Method 1 had disastrous consequences. Excluding these trees biases results (because we eliminate only dead trees and no live trees). Eliminating many dead trees left the annual mortality rates unrealistically low (less than a percent per year, when correct estimates range between 1.21 and 2.3 % (see Table D1). Climate trends emerged for many species and the species-specific stories described in the main text disappeared.

Methods 2 and 3 produced similar final sizes. Examination of posterior sensitivity showed that the methods did not change most parameter estimates at all; the only parameters which were affected were predictably the linear and quadratic size factors. However, no choice for a missing size estimation method changed whether these two parameters were selected during the model selection procedure. Because the two methods gave similar results, and estimating a state-space model for missing sizes for each species is also beyond the scope of this paper, we elected to use Method 2 for the models in this paper.

References

Eitzel, M. V., J. Battles, R. York, J. Knape, and P. de Valpine. 2013. Estimating tree growth from complex forest monitoring data. Ecological Applications **23**:1288–1296.