**Methods**

*Data Sources*

The scope of inference for this study (Figure X) is approximately 788,000 ha (7880 km2) centered primarily on western Plumas county. A digital elevation model was acquired from the US Geological Survey National Elevation Dataset (https://viewer.nationalmap.gov/) at 1/3 arc-second (8 x 10 m) resolution, and converted to corresponding slope and aspect layers using QGIS. Climate data were acquired from the Basin Characterization Model dataset (Flint et al. 2013) for the same geographic footprint at 270 m resolution. For spatial predictive models, all data were scaled and aligned at 270m resolution.

We identified FIA plots that fell within the footprint defined above (n = 294), and conducted a series of filtering exercises to select those plots comparable to our historical data. …leaving this section for Lex to fill in… Ultimately we included 76 FIA plots as a modern comparison for our historical data.

*Analysis*

For both the 1551 lots in the historical dataset and the 76 plots in the modern dataset, we extracted underlying climate and topographic data at 270 m resolution (datasets described above). We calculated the best-fit model to explain three variables – trees per hectare, basal area per hectare, and pine fraction – using two methods: automated model selection using Akaike’s Information Criterion, and a Random Forests approach.

**Results**

The best model according to AIC had five strong predictors of historical TPH (deficit, AET, snowpack, precipitation and slope; Table A1), and several weaker predictors. Four of these five strong predictors were also identified by the Random Forest model as having high importance, with snowpack being replaced by elevation in the top five variables. We proceeded with a model that had the top five variables from the AIC procedure but no other variables, to simplify our regression tree analysis.



**Figure 1**: This is a placeholder for now, but this is an initial regression tree for TPA (not TPH) based on a compromise model that wasn’t the best model (which had more variables) but had all the most important variables (the three shown here, plus slope and snowpack). There is not much variation in TPA which makes it hard. We probably have more work to do on this. Eventually we will want to replicate this for Basal Area and Pine Fraction as well.

Appendices

Appendix A: Table A1 Model results for TPH

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | AIC multimodel selection |  | Random forest model |  |
| Variable rank | T value | Variable rank | Importance value |  |
| 1-cwd | 9.5 | 1-ppt | 48.0% |  |
| 2-aet | 8.3 | 2-elev | 39.3% |  |
| 3-spk | 8.0 | 3-aet | 36.7% |  |
| 4-ppt | -6.8 | 4-slope | 35.9% |  |
| 5-slope | 5.5 | 5-cwd | 31.1% |  |
| 6-aspect | 2.2 | 6-spk | 27.9% |  |

**References**

Flint, L. E., A. L. Flint, J. H. Thorne, and R. Boynton. 2013. Fine-scale hydrologic modeling for regional landscape applications: the California Basin Characterization Model development and performance. Ecological Processes **2**:1-21.