Using AUC to Assess the Utility of Land Surface Temperature in Defining Forest Regeneration Limits

Robin Rank

November 2020

1 Introduction

Forests play a central role in global water, energy, and biogeochemical cycles, serve as important sources of biodiversity, are the focus of climate change mitigation strategies, and provide essential ecosystem services to communities around the globe [4, 7, 11, 33]. Forests also face accelerating threats due to increases in the severity and frequency of drought and heat stress associated with climate change [1] which manifest as changes in their extent and composition [3]. These permanent changes to forested landscapes occur through both mature die-off events [3, 12] and through reductions in seedling establishment, or forest regeneration after disturbance [5, 38, 17].

Forest regeneration is increasingly garnering the attention of research and management efforts given that this vital rate is considered uniquely sensitive to climate trends and weather extremes. Viewing forest distribution through the lens of regeneration is particularly useful when forests are in disequilibrium with their climate, as is often the case with such long-lived woody vegetation [10]. Mature forest stands, with deeper access to groundwater resources and greater tolerance to high temperature, may persist at a site long after local conditions have passed thresholds suitable for the growth of seedlings with smaller thermal and hydrologic niches [9, 10, 16]. This disequilibrium will persist until a disturbance, such as a stand-replacing fire, removes the mature overstory [23]. An explicit focus on forest regeneration will thus be important in predicting changing forest distribution in semi-arid landscapes globally, where patterns of recurring wildfire and recovery are important in establishing landscape dynamics [23], and which are expected to experience more frequent and intense fires in the coming decades [44, 24].

Tree seedling survival has been traditionally understood as an interaction between hydraulic stress, temperature stress, and biotic agent attack [30, 29]. The proximal cause of seedling mortality, or whether there even is a single identifiable cause of mortality, under specific environmental conditions is a subject of debate [2, 30, 29, 37]. When investigating seedling survival and subsequent forest regeneration, there are many reasonable choices of predictive variables, each with its own advantages and disadvantages. Here we choose land surface temperature (LST) as a predictor of seedling success, as it provides an integrative measure of both the climatologic and hydrologic processes experienced by tree seedlings.

LST is a radiometric measure of the energy balance at the Earth's surface: it is governed by net radiation and soil moisture [32, 22] and has strong effects on seedling physiology [25]. It can differ substantially from air temperatures measured just centimeters above the ground [40, 43, 22], and high surface temperatures have been recognized as a direct cause of seedling mortality since the early 20th century [20]. The heat conducted from soil surrounding the seedling causes irreversible damage to enzymes and proteins in the protoplasm and cell membranes, leading to stem damage and death [19]. Importantly, satellite-based radiometric measures of LST provide continuous spatial coverage, enabling analyses of forest regeneration potential at multiple spatial scales. In a single, easily retrievable measurement, LST incorporates information about the energy and water balance at a site, providing useful information about seedling viability.

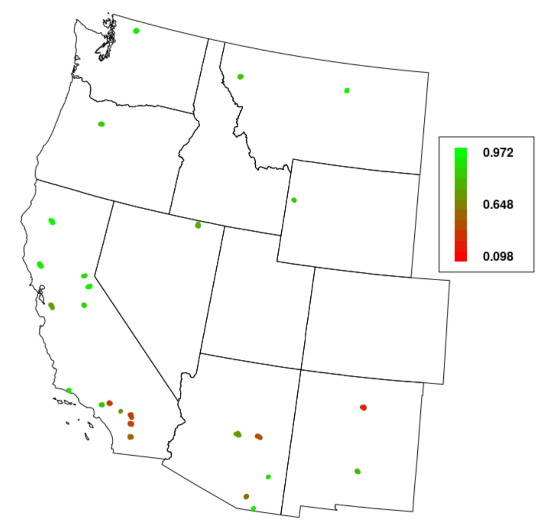

Figure 1: Locations of study basins used in this analysis. Basins are colored by their AUC statistic when using a VCF forest cover threshold of 5%.

For the very same reasons LST serves as an indicator of seedling survival (it integrates information from the surface energy and water balance) it is challenging to model across landscapes, where complex interactions of topography, vegetation, energy, and hydrology must be represented. Despite this complexity, a mechanistic approach to modelling LST has many advantages to using simple correlative models or current satellite observations. Satellite-derived measurements of surface temperature in forested areas capture temperature at the canopy surface, not the soil surface, missing the climatologic and hydrologic drivers of seedling survival [32]. Correlative models may overcome this to predict surface temperatures under current forest cover, but do not provide any insight into process. Mechanistic models of LST can incorporate information about vegetative cover, climate, and hydrology to capture surface temperatures under current forest canopies and generalize to novel temporal or spatial conditions, providing a powerful tool to investigate forest regeneration dynamics.

Here, we use LST estimates from one such mechanistic model, Ech2o-SPAC, to investigate surface temperature thresholds that best predict forest cover across a set of study basins distributed across the western United States. We compare these LST thresholds to experimentally-derived survival curves of four common conifers in response to LST intensity and duration to assess the ...

2 Methods

2.1 Ecohydrologic model

Ech2o-SPAC is a dynamic, spatially explicit coupled ecohydrologic and plant hydraulics model that combines a vertical energy balance scheme, a hydrologic model with lateral and vertical water redistribution, and a dynamic forest growth component. [28], [27], [26], and [41] provide comprehensive descriptions of Ech2o-SPAC, so we will not describe its processes in detail here. We independently calibrated Ech2o-SPAC in 29 basins distributed across the western U.S (Fig. 1). These study basins were selected from the GAGE II dataset, which provides geospatial data for stream gages maintained by the U.S. Geological Survey (USGS) which have had 20 or more complete years of discharge record since 1950 or are currently active as of water year 2009 [18]. Ech2o-SPAC was calibrated to match USGS streamflow [18] and mean-field soil moisture from the Soil Moisture Active Passive (SMAP) satellite [36] using Latin Hypercube Sampling (LHS) (Leonardo Calle, unpublished data).

After calibration, we ran Ech2o-SPAC at a 3-hourly timestep for the growing season (April 1 - August 31) of years 2003 through 2017, using a cell size of 240 meters. Daily gridded temperature, relative humidity and solar radiation inputs to drive the model were extracted from [21]. We retrieved daily precipitation data from 4 km PRISM data which was then resampled to model resolution with bilinear interpolation [13]. We extracted daily mean wind speed grids from the North American regional reanalysis data [31]. This mean daily wind speed was used at each 3-hour time step, while total daily precipitation was distributed evenly across each 3-hour period. We used model outputs to create maps of seasonal maximum LST for each year.

2.2 ROC and AUC

We then evaluated the skill of these seasonal maximum LST estimates at predicting forest cover as measured by MODIS Vegetation Continuous Fields (VCF)

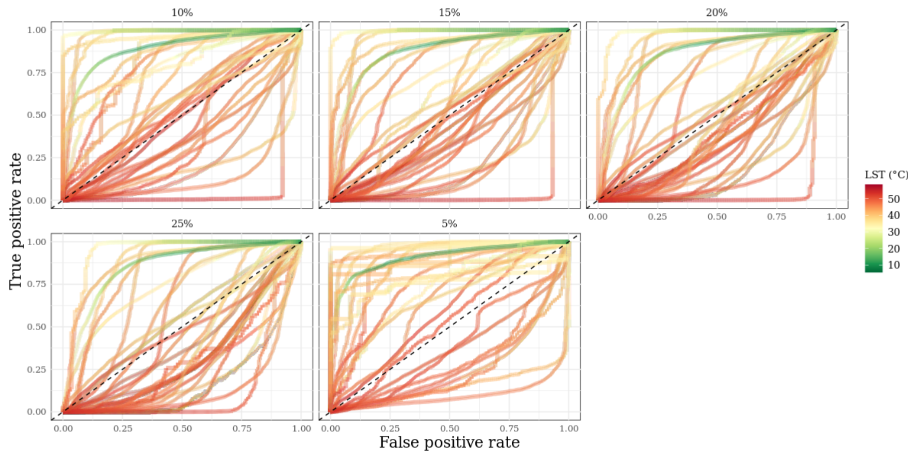

Figure 2: ROC curves for each of the 29 study basins at five different classification thresholds for VCF forest cover (5%, 10%, 15%, 20%, and 25%), colored by LST cutoff value. Dashed lines are the 1:1 line that represents an AUC of 0.5, or a model that only predicts cover correctly by chance. As the VCF forest cover threshold becomes less permissive (i.e. fewer pixels are classified as forest), LST estimates become less predictive of forest cover. Lower LST cutoff values generally result in higher true positive and false positive rates, with true positive rates generally increasing faster than false positives as the cutoff value decreases.

[15]. We converted these VCF measurements to binary cover/no cover estimates using five different threshold values - 5%, 10%, 15%, 20%, and 25% - and then created receiver operating characteristic (ROC) curves for each threshold with LST as the predictor variable and forest cover as the binary response variable. This allowed us to avoid the problems associated with arbitrary classification thresholds and evaluate the skill of our LST estimates at predicting forest cover over a range of LST thresholds. An ROC curve is obtained by varying this probability threshold and plotting the resulting true positive rate against the false positive rate, with the area under the curve (AUC) representing the accuracy of the model. An AUC of 0.5 denotes a model that is correct by pure chance, while a model that always correctly predicts the response variable would have an AUC of 1.

We then used these ROC curves to determine which threshold values of LST best predicted observed forest cover. The optimal cutoff value (OCV) of LST on the ROC curves was determined by calculating the Youden index, C [45]:

$$C = \frac{TPR}{TPR + FNR} + \frac{TNR}{TNR + FPR} - 1 \tag{1}$$

where TPR is the true positive rate, TNR is the true negative rate, FPR is the false positive rate, and FNR is the false negative rate. The OCV is obtained by searching the ROC curve for the point that maximizes C. All ROC and AUC calculations were performed in R [34] using the package ROCR [42].

2.3 Conifer survival model

We also compared the OCV obtained from Ech2o-SPAC with experimentally-derived survival curves of four common conifer species in response to LST intensity and duration (Robin Rank, unpublished data). These survival curves were based on previously published data in which populations of conifer seedlings were exposed to constant elevated surface temperatures for known time intervals [6, 14, 40, 39, 8]. A multilevel Bayesian proportional hazards model was used to quantify conifer response, with methods more closely described in [35].

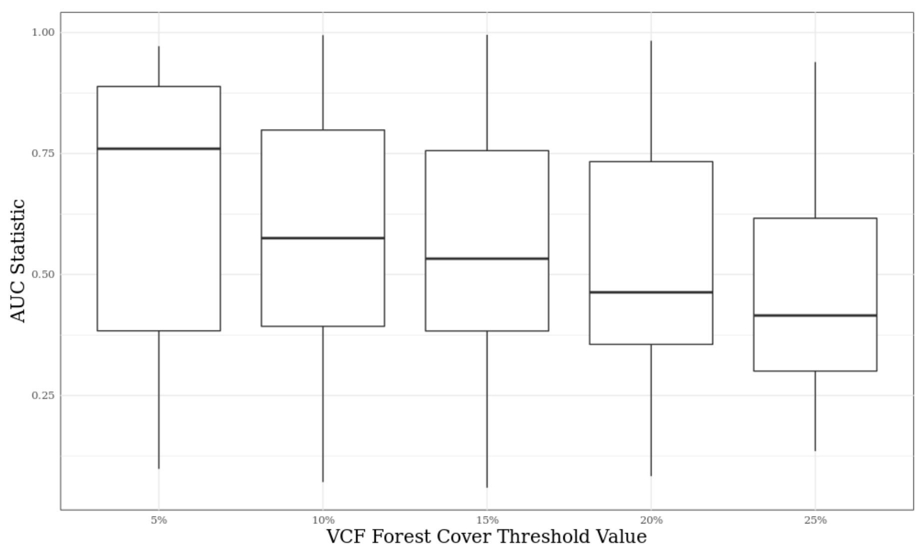

Figure 3: The distribution of AUC statistics across study basins for different VCF forest cover thresholds. Bold lines are median values, while the edges of the box correspond to the first and third quartiles, and the whiskers extend to the minimum and maximum. There is a steep decrease in model quality after the 5% VCF threshold, with estimated LST predicting forest cover worse than pure chance after the VCF threshold increases past 15%.

3 Results

The median accuracy of Ech2o-SPAC surface temperature estimates in predicting forest cover, as measured by the AUC statistic, ranged from 0.76 to 0.42, depending on the VCF threshold used to define "no cover" (Fig. 3). At a VCF classification of 5%, our model does a fair job (median AUC = 0.76) of predicting forest cover across our study basins. The majority of ROC curves for this VCF classification (75% of them) lie above the 1:1 line that represents a "right by chance" model with an AUC of 0.5 (Fig. 2). In other words, as the cutoff value of LST decreases, true positive rates tend to increase faster than false positive rates (Fig. 2).

The median OCV for surface temperature also depends on the VCF threshold value used, and ranges from 42.5°C to 40.4°C (Fig. 4). The VCF classification of 5% produces an OCV that is closest to the surface temperature thresholds resulting in at least 50% seedling mortality as determined by posterior predictive draws from the conifer survival model (Fig. 5). The median surface temperature OCV when using a 5% VCF threshold, which results in the most accurate predictions of forest cover, is 5.1°C to 7.1°C lower than the 50% mortality thresholds obtained from experimental survival curves (Fig. 5).

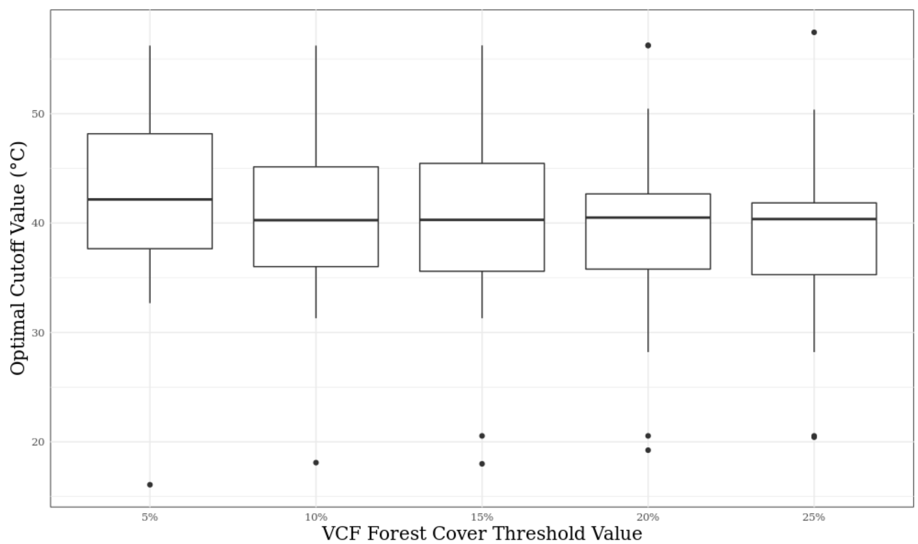

Figure 4: The distribution of OCV across study basins for different VCF forest cover thresholds. Bold lines are median values, while the edges of the box correspond to the first and third quartiles, and the whiskers extend no further than 1.5 times the interquartile range. Individual dots are outliers that lie past 1.5 times the interquartile range.

4 Discussion

With the complex feedbacks involved in mechanistically modelling LST, as well as the uncertainties inherent in both model inputs and observed forest cover response, it is impressive that LST estimates generated by Ech2o-SPAC are able to achieve of a median AUC of 0.76, especially since model parameters were calibrated only to water balance metrics, not to temperature. This is consistent with previous research demonstrating the importance of soil moisture in determining LST [32]. It is also not surprising that the LST value which best predicts forest cover is 5 to 7 degrees off from experimentally-derived mortality thresholds, especially since seasonal maximum values do not address the transient changes in LST duration to which first-year germinants are responding (Fig. 5).

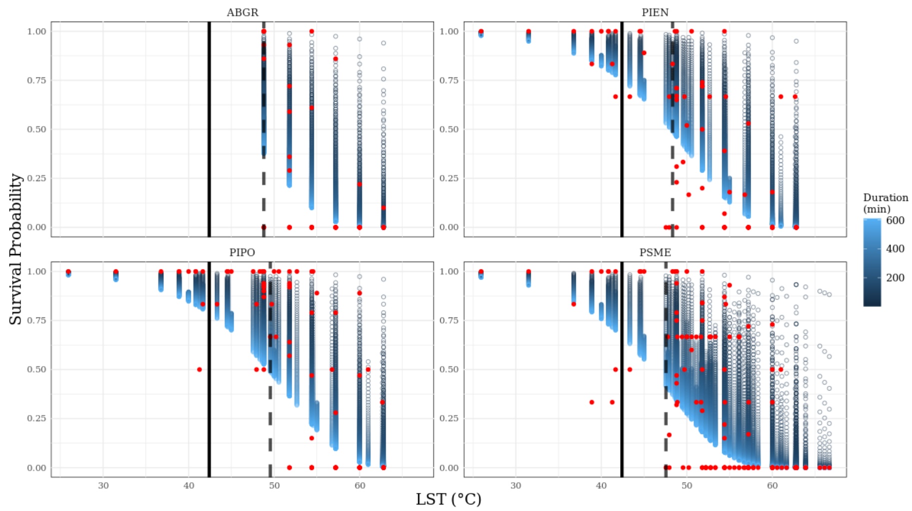

Figure 5: Modelled and observed survival of common western U.S. conifer species (*Abies grandis* (ABGR), *Picea engelmannii* (PIEN), *Pinus ponderosa* (PIPO), and *Pseudotsuga menziesii* (PSME)) plotted against LST intensity. Observed values are in red, while modelled data from posterior predictive draws are in blue, colored by experiment duration. Solid vertical lines represent the median OCV when using a 5% threshold for VCF forest cover. Dashed vertical lines represent species-specific LST thresholds past which seedlings experience 50% mortality.

Previous research has shown that LST measurements signifying the onset of seedling mortality range from 52°C to 66°C [20]. 63°C is an oft-cited threshold for the complete mortality of Ponderosa pine seedlings, even for exposures of less than a minute [25]. The optimum LST threshold determined here, 42.5°C, is well below these values. This could be explained by the wide range of topographic, climatic, and hydrologic conditions represented in this study. The OCV may be so low in order to accommodate the variance in modelled LST estimates and forest cover while maximizing true positives and minimizing false positives.

Interestingly, there seems to be a spatial pattern to these AUC scores (Fig. 1), with LST estimates serving as poor predictors of forest cover in the Southwest, but achieving high AUC scores in the Pacific Northwest and the Rocky Mountains. It could be the case that surface temperatures do not control seedling survival as strongly in the more arid Southwest. Ultimately, we will need to investigate these spatial patterns further.

This work represents an important step towards identifying climatic boundaries to forest regeneration based on LST, but much remains unknown. This is a relatively small sample of western ecosystems, and without more data it will be difficult to investigate the observed spatial bias in model accuracy. It would be helpful to understand how well LST predicts forest cover in different hydrologic and climatologic scenarios, as well. This is a promising start, however, that suggests LST has great potential as an intuitive metric of climate risk with well-understood consequences for seedling survival that can be mapped across landscapes.

References

[1] John T Abatzoglou and A Park Williams. "Impact of anthropogenic climate change on wildfire across western US forests". In: *Proceedings of the National Academy of Sciences* 113.42 (2016), pp. 11770–11775.

[2] Henry D Adams et al. "A multi-species synthesis of physiological mechanisms in drought-induced tree mortality". In: *Nature ecology & evolution* 1.9 (2017), pp. 1285–1291.

[3] Craig D Allen et al. "A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests". In: *Forest ecology and management* 259.4 (2010), pp. 660–684.

[4] William RL Anderegg et al. "Tree mortality predicted from drought-induced vascular damage". In: *Nature Geoscience* 8.5 (2015), pp. 367–371.

[5] Kristina J Anderson-Teixeira et al. "Altered dynamics of forest recovery under a changing climate". In: *Global change biology* 19.7 (2013), pp. 2001–2021.

[6] Frederick S Baker. "Effect of excessively high temperatures on coniferous reproduction". In: *Journal of Forestry* 27.8 (1929), pp. 949–975.

[7] Jean-Francois Bastin et al. "The global tree restoration potential". In: *Science* 365.6448 (2019), pp. 76–79.

[8] Carlos G Bates and Jacob Roeser. *Relative resistance of tree seedlings to excessive heat*. 1263. US Department of Agriculture, 1924.

[9] David M Bell, John B Bradford, and William K Lauenroth. "Early indicators of change: divergent climate envelopes between tree life stages imply range shifts in the western U nited S tates". In: *Global Ecology and Biogeography* 23.2 (2014), pp. 168–180.

[10] Romain Bertrand et al. "Ecological constraints increase the climatic debt in forests". In: *Nature communications* 7 (2016), p. 12643.

[11] Gordon B Bonan. "Forests and climate change: forcings, feedbacks, and the climate benefits of forests". In: *science* 320.5882 (2008), pp. 1444–1449.

[12] Jonathan D Coop et al. "Influences of prior wildfires on vegetation response to subsequent fire in a reburned Southwestern landscape". In: *Ecological Applications* 26.2 (2016), pp. 346–354.

[13] Christopher Daly et al. "Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States". In: *International Journal of Climatology: a Journal of the Royal Meteorological Society* 28.15 (2008), pp. 2031–2064.

[14] RF Daubenmire. "Soil temperature versus drought as a factor determining lower altitudinal limits of trees in the Rocky Mountains". In: *Botanical Gazette* 105.1 (1943), pp. 1–13.

[15] C Dimiceli et al. "MOD44B MODIS/Terra vegetation continuous fields yearly L3 global 250 m SIN grid V006". In: *NASA EOSDIS Land Processes Distributed Active Archive Center* (2015).

[16] Solomon Z Dobrowski et al. "Forest structure and species traits mediate projected recruitment declines in western US tree species". In: *Global Ecology and Biogeography* 24.8 (2015), pp. 917–927.

[17] Daniel C Donato, Brian J Harvey, and Monica G Turner. "Regeneration of montane forests 24 years after the 1988 Yellowstone fires: A fire-catalyzed shift in lower treelines?" In: *Ecosphere* 7.8 (2016), e01410.

[18] James A Falcone. *GAGES-II: Geospatial attributes of gages for evaluating streamflow*. Tech. rep. US Geological Survey, 2011.

[19] David M Gates. *Biophysical ecology*. Courier Corporation, 2012.

[20] Ole T Helgerson. "Heat damage in tree seedlings and its prevention". In: *New Forests* 3.4 (1989), pp. 333–358.

[21] Zachary A Holden et al. "Decreasing fire season precipitation increased recent western US forest wildfire activity". In: *Proceedings of the National Academy of Sciences* 115.36 (2018), E8349–E8357.

[22] Menglin Jin and Robert E Dickinson. "Land surface skin temperature climatology: Benefitting from the strengths of satellite observations". In: *Environmental Research Letters* 5.4 (2010), p. 044004.

[23] Jill F Johnstone et al. "Changing disturbance regimes, ecological memory, and forest resilience". In: *Frontiers in Ecology and the Environment* 14.7 (2016), pp. 369–378.

[24] W Matt Jolly et al. "Climate-induced variations in global wildfire danger from 1979 to 2013". In: *Nature communications* 6.1 (2015), pp. 1–11.

[25] Peter F Kolb and Ronald Robberecht. "High temperature and drought stress effects on survival of Pinus ponderosa seedlings". In: *Tree Physiology* 16.8 (1996), pp. 665–672.

[26] Sylvain Kuppel et al. "What can we learn from multi-data calibration of a process-based ecohydrological model?" In: *Environmental Modelling & Software* 101 (2018), pp. 301–316.

[27] Francisco Javier Lozano Parra, Marco P Maneta, and Susanne Schnabel. "Climate and topographic controls on simulated pasture production in a semiarid Mediterranean watershed with scattered tree cover". In: (2014).

[28] MP Maneta and NL Silverman. "A spatially distributed model to simulate water, energy, and vegetation dynamics using information from regional climate models". In: *Earth Interactions* 17.11 (2013), pp. 1–44.

[29] Nate McDowell et al. “Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought?” In: *New phytologist* 178.4 (2008), pp. 719–739.

[30] Nate G McDowell et al. “The interdependence of mechanisms underlying climate-driven vegetation mortality”. In: *Trends in ecology & evolution* 26.10 (2011), pp. 523–532.

[31] Fedor Mesinger et al. “North American regional reanalysis”. In: *Bulletin of the American Meteorological Society* 87.3 (2006), pp. 343–360.

[32] David J Mildrexler, Maosheng Zhao, and Steven W Running. “A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests”. In: *Journal of Geophysical Research: Biogeosciences* 116.G3 (2011).

[33] Yude Pan et al. “The structure, distribution, and biomass of the world’s forests”. In: *Annual Review of Ecology, Evolution, and Systematics* 44 (2013), pp. 593–622.

[34] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2020. URL: `https://www.R-project.org/`.

[35] Robin Rank and Solomon Dobrowski. “Conifer Seedling Survival in Response to Surface Temperature Intensity and Duration”. In: *NASA EOSDIS Land Processes Distributed Active Archive Center* (2015).

[36] Rolf H Reichle et al. “Version 4 of the SMAP Level-4 Soil Moisture Algorithm and Data Product”. In: *Journal of Advances in Modeling Earth Systems* 11.10 (2019), pp. 3106–3130.

[37] Anna Sala, Frida Piper, and Günter Hoch. “Physiological mechanisms of drought-induced tree mortality are far from being resolved”. In: *New Phytologist* 186.2 (2010), pp. 274–281.

[38] Melissa Savage, Joy Nystrom Mast, and Johannes J Feddema. “Double whammy: high-severity fire and drought in ponderosa pine forests of the Southwest”. In: *Canadian Journal of Forest Research* 43.6 (2013), pp. 570–583.

[39] KW Seidel. “Tolerance of seedlings of ponderosa pine. Douglas-fir, grand fir, and engelmann spruce for high temperatures”. In: (1986).

[40] Roy R Silen. “Lethal surface temperatures and their interpretation for Douglas-fir”. In: (1960).

[41] Caelan Simeone et al. “Coupled ecohydrology and plant hydraulics modeling predicts ponderosa pine seedling mortality and lower treeline in the US Northern Rocky Mountains”. In: *New Phytologist* 221.4 (2019), pp. 1814–1830.

[42] Tobias Sing et al. “ROCR: visualizing classifier performance in R”. In: *Bioinformatics* 21.20 (2005), pp. 3940–3941.

[43] Olli Vaartaja. "High Surface Soil Temperatures on Methods of Investigation, and Thermocouple Observations on a Wooded Heath in the South of Finland". In: *Oikos* 1.1 (1949), pp. 6–28.

[44] Anthony LeRoy Westerling. "Increasing western US forest wildfire activity: sensitivity to changes in the timing of spring". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 371.1696 (2016), p. 20150178.

[45] William J Youden. "Index for rating diagnostic tests". In: *Cancer* 3.1 (1950), pp. 32–35.