

UNLOCKING THE CLIMATE RIDDLE IN FORESTED ECOSYSTEMS

Greg C. Liknes, Christopher W. Woodall, Brian F. Walters, and Sara A. Goeking¹

Abstract.—Climate information is often used as a predictor in ecological studies, where temporal averages are typically based on climate normals (30-year means) or seasonal averages. While ensemble projections of future climate forecast a higher global average annual temperature, they also predict increased climate variability. It remains to be seen whether forest ecosystems will respond more to changes in mean climate conditions or changes in climate variability. Our objective was to compare the relative importance of climate mean versus variability metrics as predictors of tree mortality and regeneration. Using the 4-km PRISM and 32-km NARR climate datasets, both mean and variability metrics were derived for Forest Inventory and Analysis (FIA) plot locations across the eastern United States. Tree mortality and seedling abundance data were obtained from FIA plots that were visited twice in the years from 2000 to 2010. A number of statistical approaches (including correlation analysis, and an algorithmic method, Random Forests) were used to examine the relative importance of mean versus variability of climate data in the context of evaluating changes in tree and seedling attributes.

INTRODUCTION

The climate system plays an essential role in forested ecosystems, controlling inputs of energy and moisture and setting constraints that intersect with species' ecological niches to determine forest community composition and structure. As such, climate information is often used in models of species' ecological niches or climate envelopes. In addition, spatial and temporal variability of climate can inform investigations of tree mortality and regeneration dynamics in forested ecosystems. Such studies typically use temporal averages representing climate normals (30-year averages) or seasonal means. Climate change projections predict not only a shift in mean climate regimes in many locations, but also increased variability. If trees are locally adapted to a particular climate, increased variability could lead to

increased mortality or decreased regeneration. Using forest inventory data from the eastern United States, we investigated whether climate variability provides more explanatory power of tree mortality and seedling abundance than climate means by applying classical correlation analysis and the Random Forests algorithm (Breiman 2001a).

METHODS

Forest data from the U.S. Forest Service's Forest Inventory and Analysis (FIA) database (Woudenberg et al. 2010) were analyzed. The analysis was restricted to plots in the eastern United States that had been visited twice during annual inventories between 2000 and 2010, and more than 23,000 plots were available for the analysis. The two continuous response variables, volume mortality ($\text{m}^3 \text{ha}^{-1}$) and change in seedling abundance (stems ha^{-1}), were calculated from differences in adjacent time periods. Other FIA data were used as additional predictor variables, such as stand age, relative density, live/dead biomass, Lorey's height, and disturbance history (Table 1). It should be noted that plots were revisited several years

¹ Research Physical Scientist (GCL), Research Forester (CWW) and Forester (BFW), U.S. Forest Service, Northern Research Station, 1992 Folwell Ave, St. Paul, MN 55108; Biological Scientist (SAG), U.S. Forest Service, Rocky Mountain Research Station. GCL is corresponding author: to contact, call 651-649-5192 or email at gliknes@fs.fed.us.

Table 1.—Summary of analysis methods, response variables, climate predictor variables, and forest stand metrics used to assess the relationships between climate means, climate variability, and tree mortality and regeneration

	Source	Variable	Analysis Method	
			Random Forests	Correlation
Response variables	FIADB	volume mortality	x	x
	FIADB	seedling abundance change	x	x
Climate predictor variables	NARR	mean annual temperature	x	x
		interannual temperature standard deviation	x	x
		mean annual precipitation	x	x
		interannual precipitation standard deviation	x	x
	PRISM	mean annual temperature	x	x
		interannual temperature standard deviation	x	x
		mean annual precipitation	x	x
		interannual precipitation standard deviation	x	x
Stand predictor variables	FIADB	stand age	x	
		relative density	x	
		live biomass	x	
		dead biomass	x	
		Lorey's height	x	
		disturbance code 1	x	
		disturbance code 2	x	
		disturbance code 3	x	

apart (mode = 5 years, range = 1 to 10 years), and it is therefore not possible to know when mortality occurred (akin to censoring in survival analysis) or how changes to seedlings abundance were distributed across the time interval.

Monthly climate data were used from the PRISM dataset (Daly 2002), including minimum and maximum daily temperature (Tmin and Tmax, respectively) and total precipitation. Tmin and Tmax were averaged to obtain a monthly average temperature. In addition, temperature and precipitation data from the North American Regional Reanalysis Project (NARR) (Mesinger et al. 2006) were assembled in order to create a second, comparable monthly climate dataset.

To investigate the relative explanatory power of mean climate and climate variability, mean and interannual standard deviation were calculated from the monthly data for both total precipitation and

average temperature. As previously stated, the exact timing of tree mortality and changes in seedling density is unknown relative to plot observations, and furthermore, forest responses are known to lag climatic conditions, sometimes by a number of years. As a result, six different lag periods were investigated, ranging from 1 to 6 years preceding the initial plot observation (Fig. 1). The endpoint of each climate averaging interval was the midpoint of the FIA observation interval. Given the unknown timing of mortality or seedling change, this endpoint is a compromise and can result in a climate time interval that is either offset (precedes) or overlaps with the forest response. Climate means and standard deviation were calculated for each of the six averaging periods for both temperature and precipitation, resulting in 24 climate variables from both the PRISM and NARR datasets. In turn, correlation coefficients were calculated for each climate variable in relation to volume mortality and seedling abundance change (for 96 variable pairs).

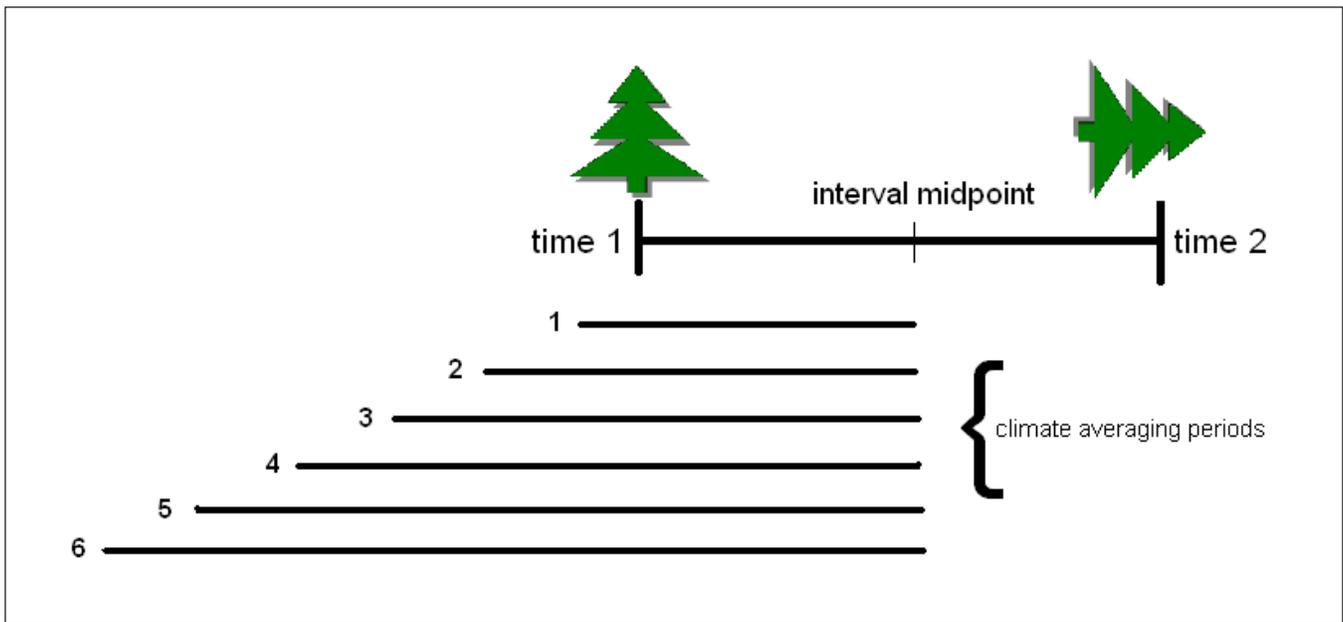


Figure 1.—Depiction of time periods used to average climate variables relative to repeated forest inventory observations. Six different lag periods ranging from 1 to 6 years preceding the first inventory observation were used to calculate annual means and interannual standard deviation.

All variables, except seedling abundance change, were found to be non-normal and were transformed using the Box-Cox method (R v. 2.13, “car” package, bcPower function; Weisberg 2005). Pearson product correlations were calculated for the transformed data. Plots with non-zero mortality were included in the mortality correlation analysis (59 percent of all available plots). Plots with less than 10 cm of floodwater or snow cover were included in the seedling analysis in order to avoid observation bias; outliers and plots with zero change were removed, and 61 percent of all available plots remained for the seedling analysis.

While considerable debate exists between proponents of classical statistical methods and newer algorithmic methods (e.g., Breiman 2001b), we considered such methods to be a useful exploratory tool to investigate the relative importance of climate predictor variables. The Random Forests algorithm has appealing features, such as automatic consideration of interactions between variables, accommodation of nonlinear relationships, and built-in analysis of variable importance. We built predictive models of mortality

and seedling density change using Random Forests (R v. 2.13, “randomForest” package) (Liaw and Wiener 2002). A total of eight models were constructed (all possible combinations of mortality vs. seedling change, NARR vs. PRISM, and temperature vs. precipitation, Table 1).

RESULTS

Correlation analysis resulted in very few statistically significant coefficients relating climate variables to tree mortality, with interannual standard deviation of precipitation the only exception (PRISM dataset; 3, 4, 5, and 6-year lag periods; coefficients range from 0.017 to 0.018). While nearly all seedling/climate correlation coefficients were significant (exceptions were interannual standard deviation of temperature for both PRISM and NARR; 1, 2, 3, and 4-year lag periods), none exceeded 0.11. A few other patterns are apparent, such as a general insensitivity of correlation coefficients to the lag period, a higher correlation of seedling change to precipitation than to temperature, and inconsistency in the relative magnitude of

correlations for mean and variability climate metrics. It should be noted, the climate mean variables followed a bi-modal distribution and a more robust transformation may be required for an adequate comparison.

Importance measures from the Random Forests analysis indicate standard deviation of both temperature and precipitation are frequently better predictors than the corresponding mean (Table 2). Additionally, the top climate predictors often provide better or at least comparable information to the three highest ranked FIA stand metrics. The most important stand metrics for the four mortality models and the two temperature/seedling models were consistently Lorey’s height, stand age, and live biomass (in various orders), while the most important stand metrics for the seedling change models involving precipitation included disturbance history rather than stand age. The difference in importance between mean and standard deviation of climate variables derived from PRISM was consistently larger than the difference derived using NARR data.

DISCUSSION

Our intent was to investigate whether climate variability data provides more information about tree mortality and regeneration in forested ecosystems when compared climate mean data. We present initial results toward that aim, using more than 23,000 remeasured inventory plots spanning a variety of climatic regimes in the eastern United States. While correlation analysis did not indicate that variability provides more information than mean values, the Random Forests algorithm does indicate interannual standard deviation of both temperature and precipitation may be better predictors of mortality and seedling regeneration than annual mean values.

Future work will include a classic regression approach to modeling tree mortality and seedling regeneration in an attempt to reconcile the seemingly contradicting results from the correlation and Random Forests analyses. In addition, more robust time series methods will be applied in order to quantify variability at numerous time scales. For example, interannual

Table 2.—Variable importance measures for eight Random Forests models of tree mortality and seedling abundance change using climate mean, climate standard deviation (SD), and forest inventory data (stand metrics) as predictors. Temperature (Temp) and total precipitation (PPT) data were derived from two different datasets (PRISM and NARR). Importance is measured by the increase in mean square error when a variable is removed from the model. Values are based on the climate variables with the highest importance score and the three most important stand metrics.

% Increase in MSE						
Tree Mortality						
Climate Variable	Mean	SD	Lorey’s Height	Stand Age	Live Biomass	
PRISM Temp	15.17	24.58	24.35	19.03	18.66	
NARR Temp	16.16	17.25	26.77	18.39	18.66	
PRISM PPT	17.21	22.10	20.59	24.74	16.05	
NARR PPT	18.91	22.18	20.85	26.87	17.04	
Seedling Change						
Climate Variable	Mean	SD	Lorey’s Height	Stand Age	Live Biomass	Disturbance Code 1
PRISM Temp	7.88	17.03	14.24	10.64	11.19	----
NARR Temp	12.37	13.74	22.86	12.75	10.78	----
PRISM PPT	9.26	23.60	9.88	----	9.48	10.59
NARR PPT	18.83	16.34	10.09	----	9.27	7.95

seasonal variability could be an informative predictor, as well as decadal or even intra-annual variability. A nationwide, consistent protocol for gathering forest inventory information across a large climate gradient in the United States presents great opportunity for unlocking the complexity of climate/forest interactions. The FIA program is beginning to assemble remeasurement data on a regular return interval (e.g., 5-10 years). This work represents an early attempt to connect temporally rich climate data with observations of change in the Nation's forested ecosystems.

ACKNOWLEDGMENTS

The authors thank Oregon State University's PRISM Climate Group and the National Climatic Data Center of the National Oceanic and Atmospheric Administration for making PRISM and NARR data available.

LITERATURE CITED

- Breiman, L. 2001a. **Random Forests**. Machine Learning. 45(1): 5-32.
- Breiman, L. 2001b. **Statistical modeling: the two cultures**. Statistical Science. 16(3): 199-231.
- Daly, C.; Gibson, W.P.; Taylor, G.H.; Johnson, G.L.; Pasteris, P. 2002. **A knowledge-based approach to the statistical mapping of climate**. Climate Research. 22: 99-113.
- Liaw, A.; Wiener, M. 2002. **Classification and regression by randomForest**. R News. 2(3): 18-22.
- Mesinger, F.; DiMego, G.; Kalnay, E.; Mitchell, K.; Shafran, P.R.; et al. 2006. **North American regional reanalysis**. Bulletin of the American Meteorological Society. 87: 343-360.
- Weisberg, S. 2005. **Transformations. Chapter 7**. In: Applied linear regression. 3rd ed. Hoboken, NJ: John Wiley and Sons: 147-166.
- Woudenberg, S.W.; Conkling, B.L.; O'Connell, B.M. LaPoint, E.B.; Turner, J.A.; Waddell, K.L. 2010. **The Forest Inventory and Analysis database: database description and users' manual version 4.0 for Phase 2**. Gen. Tech. Rep. RMRS-GTR-245. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 336 p.

The content of this paper reflects the views of the author(s), who are responsible for the facts and accuracy of the information presented herein.