EXPLAINING PATTERNS OF BIODIVERSITY ACROSS NEON SITES USING LANDSAT-BASED DISTURBANCE METRICS.

Jasper Van doninck[1], Annie C. Smith[2], Jon Knott[1], Sydne Record[3], Phoebe L. Zarnetske[1]

[1] Michigan State University, Department of Integrative Biology, Spatial & Community Ecology Lab;
[2] Washington State Department of Natural Resources; [3] Bryn Mawr College, Department of Biology

ABSTRACT

Disturbance regimes can strongly influence geographic patterns of biodiversity. The types of disturbances and their frequencies can have varying impacts on different dimensions of biodiversity and taxonomic groups, and their influence can also vary with spatial scale. Yet disturbance layers are lacking at sufficiently high spatial resolution and extent to uncover these relationships with biodiversity. We detected disturbances for the conterminous United States from Landsat time series using the established LandTrendr temporal segmentation with a novel secondary classification that incorporates spatial context. We then included these disturbance layers, aggregated to metrics at different temporal and spatial scales, into model of species richness at National Ecological Observatory Network sites.

***Index Terms*—** Landsat, NEON, biodiversity, time series analysis, disturbance detection

1. INTRODUCTION

Quantifying the influences of biodiversity drivers across multiple contexts and scales is a primary goal of ecology. Leveraging the rich data on biota and environmental conditions from the spatially-nested sampling scheme of the National Ecological Observatory Network (NEON) offers a unique opportunity to address the spatial scale-dependence of biodiversity drivers. Across multiple sampling locations within sites, NEON collects hundreds of thousands of individual-level trait measurements on organisms spanning taxonomic groups.

Although NEON measures many abiotic drivers of biodiversity (e.g., climate, soil attributes), it lacks spatially explicit data on past land use, as well as anthropogenic and natural disturbances. Without these key data, it is challenging to interpret biodiversity patterns and capture important cross-scale interactions resulting from anthropogenic processes. For example, the disturbance history of a location has profound influences on biodiversity [1], ecosystem processes and structure [2,3], species distributions and abundances [4], and biotic interactions [5].

In this study, we derive disturbance layers across NEON sites for the last decades using the Landsat archive. We then use this disturbance dataset to model one of the dimensions of biodiversity, taxonomic species richness, at the NEON site level.

2. STUDY SITE, MATERIALS AND METHODS

4.1. Terrestrial NEON sites and species richness data

NEON is a continental-scale network of 47 terrestrial and 34 aquatic sites strategically located across the U.S. within 20 eco-climatic domains. At each site, sentinel taxa and a diverse suite of environmental variables are sampled using standardized methods designed to enable scaling from the plot, reach, or airshed scale to the site, domain, and, ultimately, continental scale. NEON's publicly available datasets for these taxa include sites that have been sampled consistently since 2013.

We calculated site-level species richness for six taxonomic groups (trees, herbaceous plants, birds, beetles, small rodents and mosquitoes) for 45 core and relocatable terrestrial NEON sites within the conterminous United States (CONUS). The number of years over which the NEON sites have been sampled varies over sites and taxa (Fig. 1). In order to minimize the effects of this unequal sampling, we calculated cumulative species richness per site over all available sampling years. Preliminary analysis showed that species accumulation curves flatten off quickly, allowing comparison between sites with different sampling effort.

4.2. Environmental variables at NEON sites

We derived site-representative values of four environmental variables for all NEON sites: mean annual temperature, mean annual precipitation, terrain ruggedness index (TRI) [6] and elevation. These variables were chosen because of their known relationships with dimensions of biodiversity [7]. All four variables were derived for the individual plots within each NEON site, and then averaged to obtain a site-representative value.

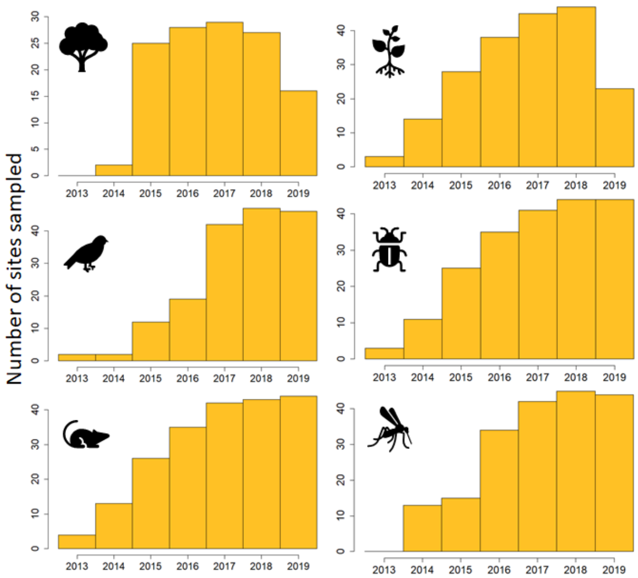

Fig. 1: Number of sites sampled for the different taxa and sample years. Sampled taxa (clockwise from top left: trees, herbaceous plants, beetles, mosquitoes, small mammals and birds).

4.3. Landsat disturbance metrics

We calculated disturbance metrics using the LandTrendr algorithm for Landsat time series segmentation implemented in Google Earth Engine [8] in combination with a secondary classification incorporating spatial structure [9]. Provided an annual Landsat pixel time series for a spectral band or index, LandTrendr returns a limited number of linear segments. We calculated a set of three metrics for each of these segments: segment start value, segment magnitude, and segment duration. These segment metrics were then assigned to each year of the respective segment. When done for each 30-m pixel in the CONUS, this analysis provides for each year in the Landsat time series (1984-2020), wall-to-wall layers of the three temporal parameters. We then derived a set of three spatial parameters for each of these temporal parameters, based on a circular neighborhood around each pixel: median value, standard deviation and ratio of focal value and neighborhood median.

In the secondary classification, the wall-to-wall layers of temporal and spatial parameters were combined with a reference dataset in a Random Forests classifier. As reference data we used the LCMAP (Land Change Monitoring Assessment and Projection) Reference Data Product [10]. This dataset contains annual land cover and change labels for the period 1984-2018 at 25,000 randomly selected 30 m by 30 m plots across the CONUS, assigned using the TimeSync software tool [11]. We extracted the temporal and spatial variables at the 25,000 reference plots from the wall-to-wall data layers, resulting in a reference dataset of time series metrics and corresponding reference disturbance pairs.

We calculated the temporal and spatial LandTrendr parameters for four Landsat bands/indices: Normalized Burn Ratio (NBR), Tasselled Cap Brightness (TCB), Tasselled Cap Greenness (TCG) and Landsat TM band 5–equivalent shortwave infrared 1 (SWIR1). We then calibrated and validated Random Forests models for these four bands/indices individually and combined, and using temporal parameters alone and temporal and spatial parameters combined. The accuracy of these ten models ((4 bands/indices + 1 ensemble) times 2 (spatial or spatial+temporal parameters)) was validated using 10-fold cross validation with random folds with 10% of the reference data used as validation data. This 10-fold cross-validation was repeated 20 times. Applying the calibrated Random forest classifier based on all reference data on the entire image dataset provides, for each year in the time series, a binary classification whether or not a pixel underwent a disturbance in that year.

4.3. Species richness modeling

We build a baseline Generalized Linear Model (GLM) for each of the six terrestrial taxa with NEON site-level species richness as response variables and the four environmental variables (mean annual temperature, mean annual precipitation, TRI, elevation). Additionally, we build GLMs for each taxon using the baseline predictive variables plus one disturbance metric. The disturbance metric was defined as the fraction of 30-meter by 30-meter pixels that were identified as having undergone at least one disturbance during a certain time period over a certain spatial scale. We used three timeframes as temporal scales: the 5 years, 10 years and 20 years leading up to the year 2017, which is the year in which data collection had begun for almost all of the sites for the six taxa (Fig. 1). We considered four spatial scales to calculate the fraction of disturbed pixels during each of these temporal scales: the site boundaries as defined by NEON, and circular neighborhoods with radius of 5 km, 10 km and 25 km around the NEON site centroid.

The performances of the different models (one baseline model and twelve using a combination of spatial and temporal scales) for each taxon were compared using Akaike's information criterion with a correction for small sample sizes (AICc). We identified the two models with the largest AICc for each taxon. This analysis was performed in R software using the *stats* and *MuMin* packages.

3. PRELIMINARY RESULTS AND DISCUSSION

3.1. Landsat disturbance detection

Figure 2 summarized the Kappa coefficients of agreement derived from the 10-fold cross validation of disturbance detection. These showed that adding spatial information increases disturbance detection accuracy compared to when temporal variables obtained from LandTrendr time series segmentation are used alone, and this for all spectral bands/indices except SWIR1. This confirms previous results comparing temporal and spatial parameters for disturbance detection, where the increased accuracies have been attributed to a better detection of low-severity disturbances [9].

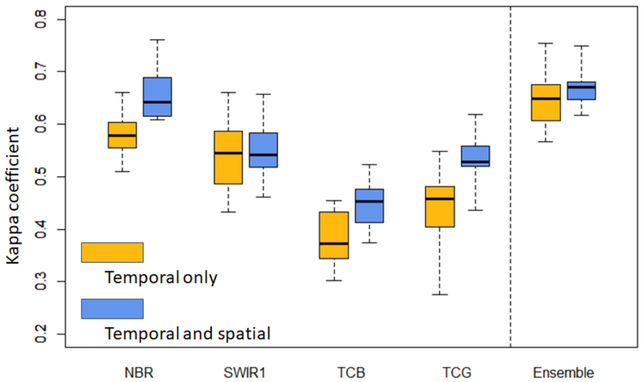

Fig. 2: Kappa coefficients for 20 runs of the 10-fold cross-validation for the different spectral bands/indices and the ensemble method, and using temporal variables alone and temporal and spatial variables combined.

Validation of the disturbance detection (Fig. 2) also showed that the use of an ensemble model that incorporates multiple spectral bands/indices performs better than the models based on a single band/index. This behavior has been observed in other studies for forest disturbance detection [12], indicating that different bands and indices carry complementary information, and ensemble methods can exploit this information in spectral space.

3.2. Species richness modeling

Figure 3 provides an overview of the two most informative models for the six taxa. The baseline model, which uses no disturbance metrics as predictive variables, was never identified among the best two models. Models including metrics of recent (5 years) disturbances best explained species richness for almost all taxa. Furthermore, Fig. 3 shows that these recent disturbances had a negative effect on species richness, i.e., increased disturbances were associated with a lower cumulative number of observed species. An exception to this were two models for birds and herbaceous plants, where cumulative disturbance over a longer scale (20 years) had a positive effect on species richness. This could imply that vegetation disturbances over longer scales allow the creation of more diverse niches, while recent disturbances cause removal of species without allowing the settlement of new successional species.

Trees are the only taxon for which disturbances over longer time scales result in a decrease of species richness at NEON sites. This could be explained as a sampling artifact, because only trees with a diameter at breast height (DBH) larger than 10 cm are sampled. Consequently, it will take several years for new trees to reach this size after disturbances.

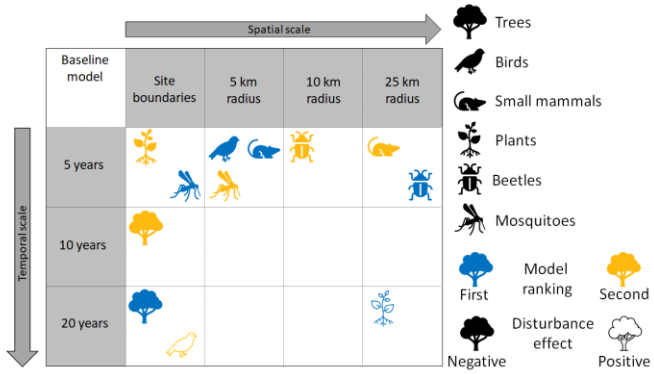

Fig. 3: Identification of the two models with highest AICc among thirteen models (one baseline model and twelve models including the baseline variable and one disturbance metric) for the six terrestrial sampled at terrestrial NEON sites.

In general, disturbance metrics calculated within the NEON site boundaries or in a circular neighborhood with radius of 5 km around the site's centroid provided the most informative models. This could be expected for less mobile taxa such as trees and herbaceous plants. However, this was also the most informative scale for some of the more mobile taxa (e.g., birds, mosquitoes), where we had expected a more important contribution of disturbances over a larger spatial scale.

4. CONCLUSIONS

In this research, we detected disturbance events across the CONUS using long (1984-2020) Landsat time series, based on the LandTrendr temporal segmentation with a secondary classification that includes spatial context. Using an existing reference dataset of 25,000 visually interpreted pixel time series, we found that disturbance detection accuracy was improved by incorporating spatial context and by combining data from several bands and indices into an ensemble. Further research will focus on attribution of change agent to detected disturbances.

Generalized linear models were more capable of modeling species richness at the NEON site level when the included a disturbance metric as predictive variable in addition to temperature, precipitation, elevation and terrain ruggedness variables. For most taxa, including disturbance metrics calculated for the lower end of the temporal and spatial scales resulted in the most informative models, and disturbances had a negative effect on species richness. Further research will include more advanced disturbance metrics, and will investigate the influence of disturbance agent on species richness and other biodiversity metrics.

5. REFERENCES

[1] J. S. Harding, E. F. Benfield, P. V. Bolstad, G. S. Helfman, and E. B. D Jones, "Stream biodiversity: The ghost of land use past," PNAS, vol. 95, no. 25, pp.14843–7, Dec 1998.

[2] J. L. Dupouey, E. Dambrine, J. D. Laffite, and C. Moares, "Irreversible Impact of Past Land Use on Forest Soils and Biodiversity," Ecology; vol. 83, no. 11, pp. 2978–84, 2002 .
[3] R. Seidl, W. Rammer, and T. A. Spies, "Disturbance legacies increase the resilience of forest ecosystem structure, composition, and functioning," Ecological Applications, vol. 24, no. 8, pp. 2063–77, Dec 2014.
[4] S. J. Goring, and J. W. Williams, "Effect of historical land-use and climate change on tree-climate relationships in the upper Midwestern United States," Ecology Letters, vol. 20, no. 4, pp. 461–70, 2017.
[5] P. G. Hahn, and J. L Orrock, "Land-use legacies and present fire regimes interact to mediate herbivory by altering the neighboring plant community," Oikos, vol. 124, no. 4, pp. 497–506, 2015.
[6] S. J: Riley, S. D. DeGloria, and R. Eliot, "A Terrain Ruggedness Index that quantifies topographic heterogeneity," Intermountain Journal of Sciences, vol. 5, no. 1-4, pp. 23-27, 1999.
[7] Q. D. Read et al., "Beyond counts and averages: Relating geodiversity to dimensions of biodiversity," Global Ecology and Biogeography, vol. 29, no. 4, pp. 696–710, 2020.
[8] R. E. Kennedy et al., "Implementation of the LandTrendr Algorithm on Google Earth Engine," Remote Sensing, vol. 10, no. 5, Art. no. 5, 2018.
[9] A. C. Smith, Z. Liu, P. L. Zarnetske, and A. P. Ballantyne, "A new approach to ecosystem disturbance detection and attribution (disturbr) using temporal and spatial information". In review.
[10] B. W. Pengra, et al., LCMAP Reference Data Product 1984-2018 land cover, land use and change process attributes: U.S. Geological Survey data release, doi.org/10.5066/P9ZWOXJ7, 2020.
[11] W. B. Cohen, Z. Yang, and R. Kennedy, "Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — Tools for calibration and validation," Remote Sensing of Environment, vol. 114, no. 12, pp. 2911–2924, Dec. 2010.
[12] W. B. Cohen, Z. Yang, S. P. Healey, R. E. Kennedy, and N. Gorelick, "A LandTrendr multispectral ensemble for forest disturbance detection," Remote Sensing of Environment, vol. 205, pp. 131–140, Feb. 2018.