Detecting Latitudinal Shifts in Plant Functional Types Distribution within the North American Great Plains

Paul Lin

Thesis Advisors: Brent Helliker · Jane Dmochowski

University of Pennsylvania

Abstract

Grass-based biomes provide numerous crucial functions and services, yet remain as one of the most vulnerable ecosystems today. In particular, the distribution of Plant Functional Types (PFT) between C_3 grasses and C_4 grasses are expected to shift dramatically due to grass' sensitivity to environmental factors. This study analyzed the potential shifts in the spatiotemporal distribution of grass PFTs and its correlations with environmental factors within the North American Great Plains. A set of twelve annual phenological metrics were extracted from the seasonal Normalized Difference Vegetation Index (NDVI) profiles and are used to predict site PFT class along five training clusters and three testing transects between the years of 2010 and 2020. Amongst the three latitudinal transects analyzed, only the East Transect demonstrated a statistically significant northward migration in PFT distribution. However, this distribution shift appeared to be independent of environmental factors. These preliminary findings demonstrated that the degree of northward migration of grass PFTs along latitudinal transects varied across geographical regions.

Background

Plant Functional Types (PFT): C_3 vs C_4

The photosynthetic process is performed as $6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2$. Due to the structural similarities between O_2 and CO_2, the carboxylating enzyme RuBisCO frequently mistakenly binds O_2 instead of CO_2 under a process called *photorespiration.* Photorespiration decreases a plant's internal CO_2 concentrations and reduces its photosynthetic potential. C_4 species have evolved physiological adaptations to reduce photorespiration levels by maintaining high CO_2 concentrations, at the cost of extra energy expenditures.

Global PFT Distribution

C_4: high temperatures, dry environment, low CO_2 levels

C_3: low temperatures, moist environment, high CO_2 levels

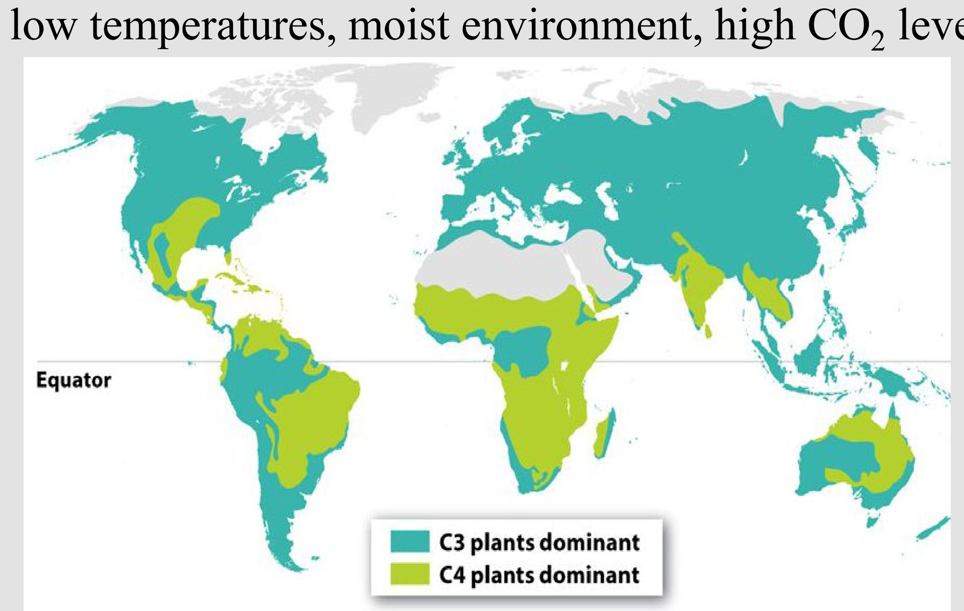

Figure 1. Global PFT Distribution[1]

Methods

Study Area: North American Great Plains

Figure 2. Training Clusters

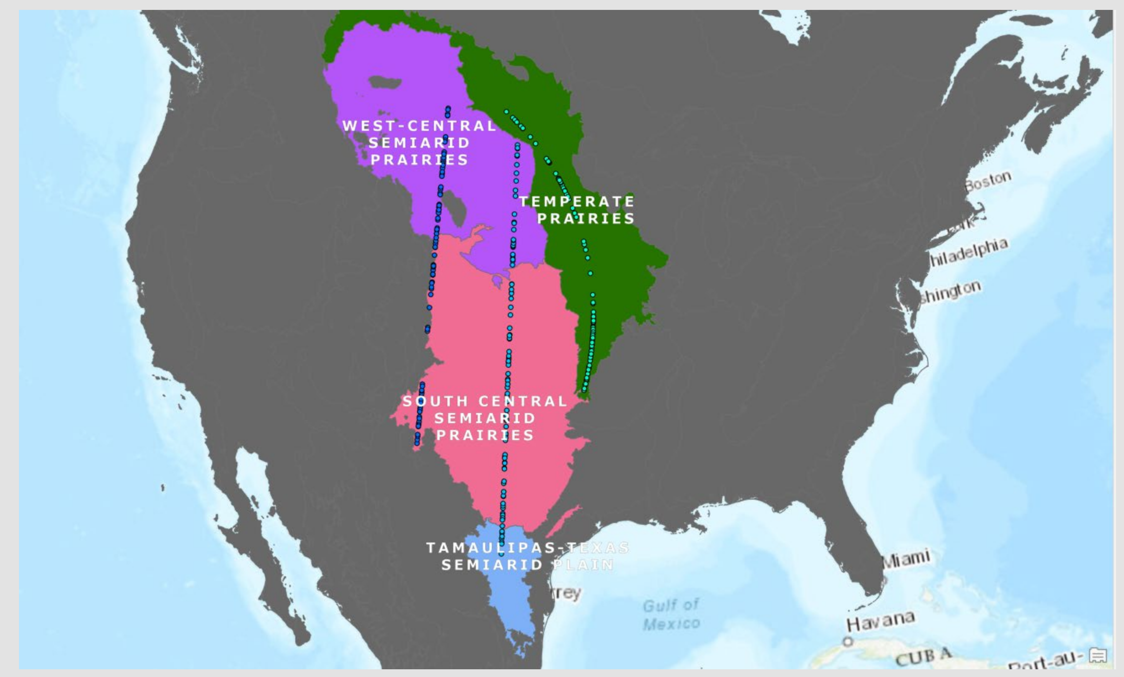

Figure 3. Testing Latitudinal Transects

Method Phases & Steps

1. **Site Generation**
   - Generate herbaceous land cover mask
   - Training Data: filter non-herbaceous training plots; identify training plot labels as C_3-majority or C_4-majority
   - Testing: select 100 herbaceous points along each transect
2. **Phenological Metric Feature Extraction**
   - Generate NDVI time-series at each training and testing site
   - Apply smoothing and interpolation to NDVI time-series
   - Extract phenological features (**Figure 4**)
3. **Building Model with Training Data**
   - Split training data into training and validation data
   - Build 6 models (Logistic Regression: LR, Support Vector Machine: SVM, k-Nearest Neighbors: k-NN, Random Forest: RF, Adaboost: Ada, Gradient Boost: GBM) with features from training sites
   - Evaluate model accuracy with cross validation
4. **Input Model with Testing Data**
   - Input testing site feature into learned model
   - Obtain class label for testing sites (**Figure 5**)
5. **Correlation Analysis**
   - Pearson's Correlation: PFT distribution vs year, PFT distribution vs temperature, PFT distribution vs precipitation

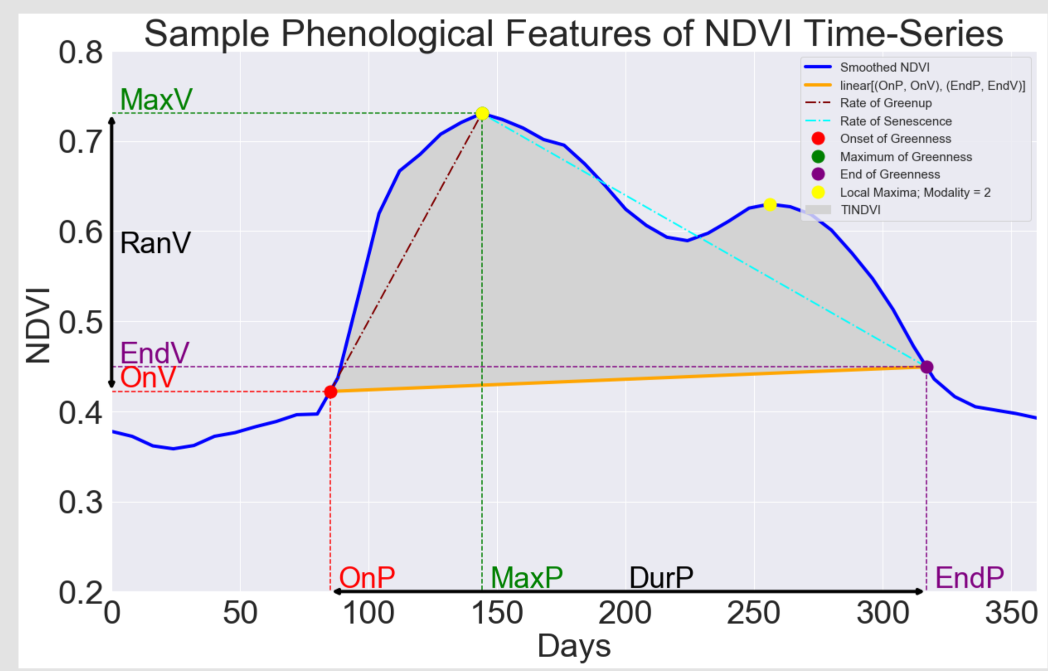

Figure 4. Extracted of phenological metrics from NDVI profile

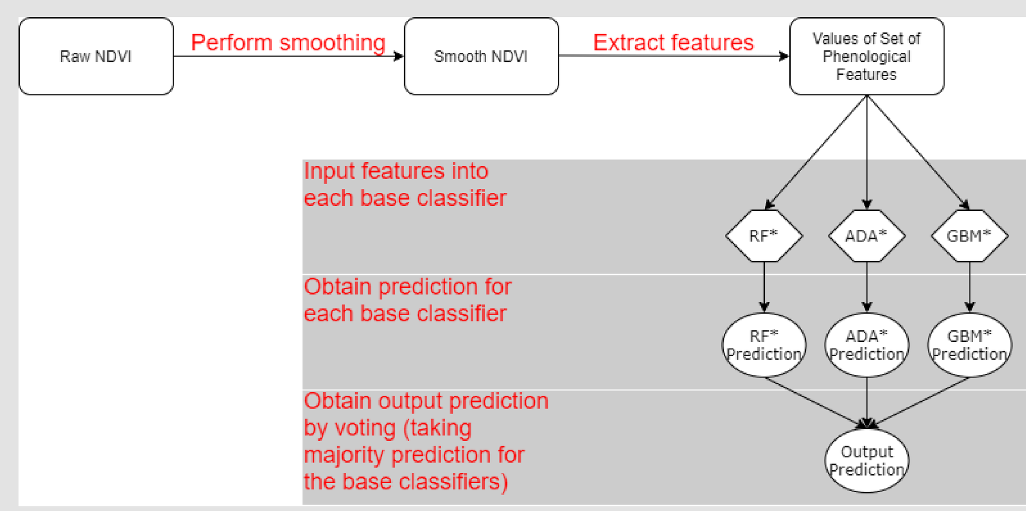

Figure 5. Pipeline transforming NDVI time-series profile into PFT prediction

Results

Figure 6. Confusion Matrices base classifiers. Grid numbers within the grid represent the number of samples with a known actual class (y-axis) and the model's predicted class (x-axis).

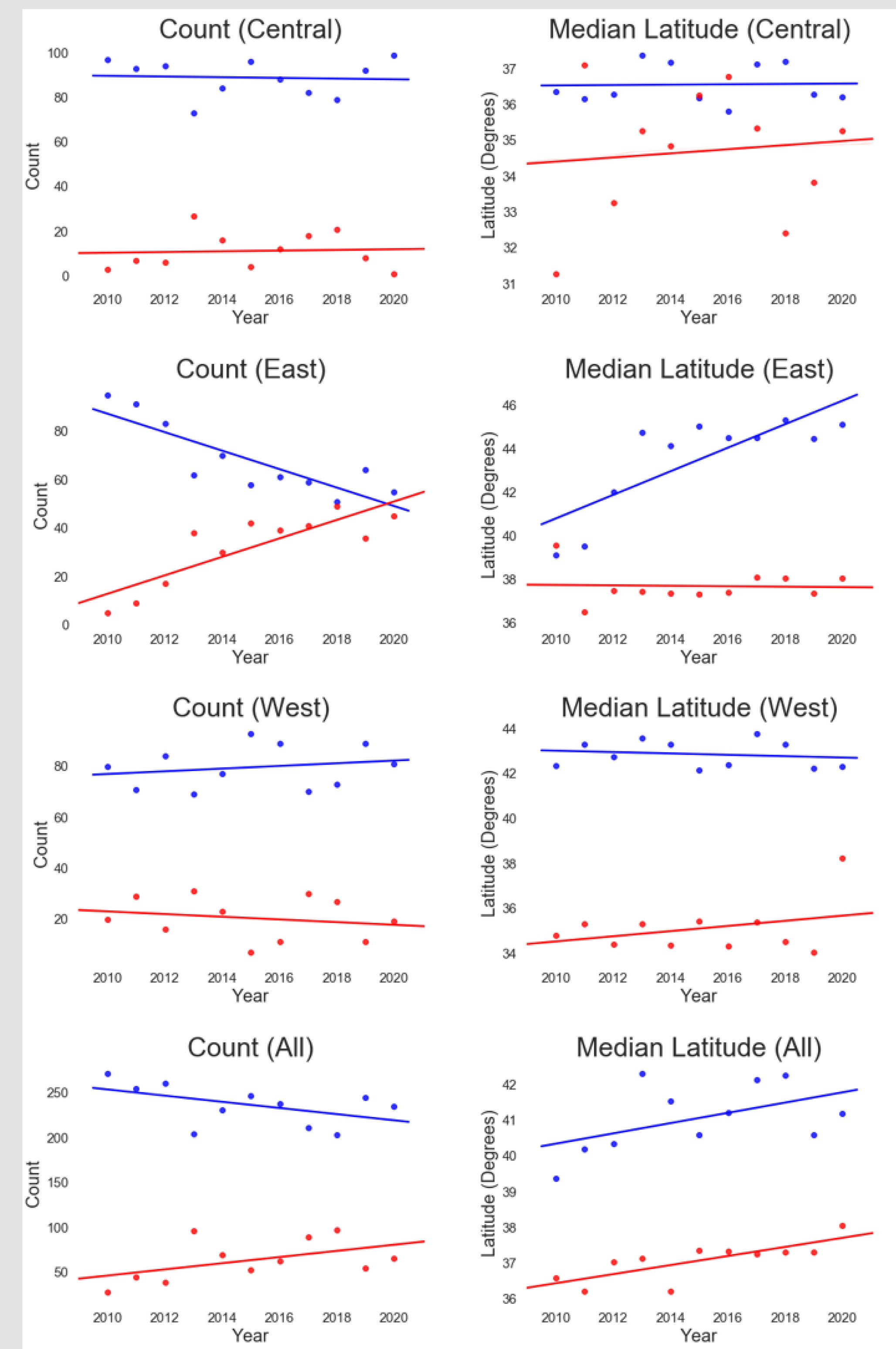

Figure 7. PFT Distribution Vs Year. C_3 in blue and C_4 in red.

Conclusion

Correlation Analysis

- No correlations between PFT distribution and temperature
- No correlations between PFT distribution and precipitation
- Significant correlation between PFT distribution and year at East Transect and aggregate of all transects
- C_3 centroid shifting north at 0.54 degrees / year at East
- C_4 centroid shifting north at 0.12 degrees / year at Aggregate

Further Work

- Increase training site diversity (5 clusters → 10 clusters)
- Increase temporal range of study (10 years → 20 years)
- Analyze regression with other environment variables (temperature, precipitation → soil, CO_2 levels, fire, etc.)
- Transform model type (Binary Classification: C_3/C_4 → Regression: *X%* C_3, *Y%* C_4)

References

1 Phelan, J. (2015). *What Is life?: A guide to biology with physiology.*