A new index for quantification of the spatial orderedness of natural and cultivated vegetation

**Introduction**

Natural vegetation is fundamental to the health of any watershed; it filters sediment and nutrients from runoff, stabilizes stream banks, controls sediment erosion, shades and cools water, and provides habitat and food for a diverse array of terrestrial organisms. On the other hand, cultivated vegetation is a poor soil stabilizer, provides limited habitat and implies the application of potentially detrimental pesticides, herbicides and fertilizers. Because of this, quantifying both the extent and character of vegetation in a watershed is of great interest to hydrologists, watershed planners, agronomists and other stakeholders.

The advent of aerial and satellite imagery has been crucial in allowing rapid characterization of landcover and vegetation on huge scales, but the techniques for analyzing this data have been primarily non-spatial: pixels are evaluated using the spectra collected at that particular point, but data from nearby pixels are ignored. Thus, these techniques have limited ability to distinguish natural and cultivated vegetation which may have similar spectral signals despite having obviously different spatial arrangements. Some more comprehensive techniques use convolution to extract “textural” (locally aggregated) statistics or spatial Markov chains to try to differentiate natural and cultivated vegetation with some success.

I propose a new, vector-based method to distinguish natural and cultivated vegetation. By evaluating the linearity of clusters of vegetation, an “orderedness index” can be calculated and therefore used to classify whether vegetation is natural. The purpose of this project is to explore the optimal implementation of this technique, its effectiveness compared to other techniques, and its general limitations. While natural vegetation is highly irregular, cultivated vegetation is expected to show a high degree of linearity due to the regular planting patterns preferred by agriculturists.

**Background**

**As appropriate for the discipline, either provide a brief review of the work that has already been done in the project area (together with complete references in the appropriate professional style) or provide a description of the context within which the project fits. This section should also include any personal information about the student, which would indicate to the reviewers the student's qualifications to successfully complete the project. A resume or curriculum vitae (as appropriate for the discipline) may be attached to the proposal to supplement this section.**

*Spectral and temporal methods*

*Textural methods and Markov Chains*

*Entropy, lacunarity, computer vision*

*Uniqueness of this approach*

*My background in geospatial modeling*

**Purpose**

The goal of this project is to develop a set of instructions (an algorithm) that can be used to distinguish natural vegetation from cultivated vegetation purely on the basis of its spatial relationship to other vegetation. Thus, the project will entail testing different aspects of the proposed methods, quantifying its predictive power, and comparing the results to spectral-based techniques.

**Methods**

Identifying vegetation using aerial imagery or LiDAR-derived digital height models is well-studied, and not the focus of this project. Rather, this project is focused on using the relative spatial position of vegetation to classify it as natural or cultivated. To explore this, a study area in middle Tennessee has been selected. The study area encompasses an apple orchard, a natural forest, and an area of cultivated coniferous trees. By using a digital height model, individual tree canopies can be identified.

Once tree locations are identified, the proposed method to identify a given tree as natural or cultivated is as follows: find all other trees within a certain radius. Then, use a clustering algorithm to identify natural groups of trees. Finally, find the line that best fits each cluster, and use each tree’s deviation from its cluster’s line to determine if the cluster is irregular or linear. Linear clusters imply the tree is cultivated (rows of trees in an order), while irregular clusters imply that the tree is natural.

Techniques such as hierarchical clustering and stochastic optimization may be used within the model and to adjust model parameters.

The proposed method is interesting because it takes into explicit consideration the spatial arrangement of vegetation, in contrast to the non-spatial spectral methods that currently dominate landcover classification schemes. If this method proves successful in the proposed study area, then it might also be applied to herbaceous vegetation, which shows a similar spatial dichotomy between row crops and natural herbaceous vegetation. Other methods such as spatial entropy and lacunarity-based analyses may also be explored as they are currently understudied for the purpose of vegetation classification.

**Timeline**

8/26: Download study area data, generate canopy model, vectorize tree locations and type

Xx: Build prototype model, explore effectiveness

Xx: Explore alternate methods (spatial entropy, lacunarity)

Xx: Compare models against each other, including traditional machine-learning spectral models

Xx: Write report

**Mentor Collaboration**

Dr. Henrique Momm will be my mentor for this project. We are already collaborating together on a project that is focused on the identification of riparian vegetation using LiDAR-derived data. While working on this project, we realized that our model was unable to differentiate natural and cultivated vegetation which limited the model’s effectiveness. The proposed project, though separate from our ongoing work, would be able to address this limitation. Dr. Momm and I already have ongoing weekly meetings where we discuss the status of our work. Because Dr. Momm had a critical role in developing the texture-base methods discussed elsewhere in this proposal, he is well-suited to advise on the proposed project.