Extracting Grass Contribution of NDVI

For Dan Rubenstein et al.

Max Gotts[1,2,3] ✉

[1] Princeton University, Princeton, NJ 08544, USA

[2] Upwell Turtles, Monterey, California 93940, USA

[3] Lost Years Institute, Mill Valley, California 94941, USA

✉ mgotts@princeton.edu

We intend to extract grass-contribution to NDVI from LANDSAT data since the level of greenery of an area's grasses is a key aspect of the life-dinner model for habitat use.

Random forest models were used on raw LANDSAT band data to classify the North Mpala Research Area into three distinct categories, each characterized by % tree cover: open bush (OB, $\sim 10\%$), light bush (LB, $\sim 50\%$), and medium bush (MB, $\sim 90\%$). We also have NDVI from the same LANDSAT data, so we have enough information to estimate the grass contribution for each pixel of NDVI.

Consider that the NDVI of a pixel can be given by the expression:

$$\text{NDVI}(\text{pixel}) = \sum_{z \in \text{pixel}} \frac{\text{ndvi}(z)}{\text{area}(\text{pixel})}$$

Where the function $\text{NDVI}(\cdot)$ is the average NDVI of a pixel, whereas the functiopn $\text{ndvi}(\cdot)$ has virtually infinite resolution. Let's say that the resolution of this function is a "plot," like a pixel, but as small as we want. We will use this hypothetical function as a guide so we can semi-formally construct a model of NDVI-contributions. Our first assumption is that tree NDVI is more constant than grass NDVI. In fact, we assume tree NDVI will be constant through an area during a certain time period (i.e. during the 16-day period in which the NDVI pictures in question were taken).

Then let's say that $\text{gndvi}(\cdot)$ is the NDVI of grass, equivalent in all ways to the function $\text{ndvi}(\cdot)$, but we want to separate it out. This is the function we are solving for.

$$\text{ndvi}(z) = \begin{cases} T \times \text{area}(\text{plot}) & \text{if } z \text{ contains only trees} \\ \text{gndvi}(z) & \text{if } z \text{ contains only grass} \end{cases}$$

Now notice that

$$\text{area}(\text{pixel}) = \Big(\#\text{ of plots}\Big)\Big(\text{area}(\text{plot})\Big)$$

We now plug all of this into the original formula:

$$\begin{aligned} \text{NDVI}(\text{pixel}) &= \frac{(\#\text{ trees plots})T\text{area}(\text{plot}) + \sum_{z \in \text{pixel}, z \text{ contains only grass}} \text{gndvi}(z)}{\Big(\#\text{ of plots}\Big)\Big(\text{area}(\text{plot})\Big)} \\ &= (\%\text{ tree cover}(\text{pixel}))T + \text{mean grass NDVI contribution}(\text{pixel}) \end{aligned}$$

Which means we can say,

$$\text{mean grass NDVI} = \frac{\text{NDVI}(\text{pixel}) - (\%\text{ tree cover})T}{1 - \%\text{ tree cover}}$$

By simply scaling up our previous value to imagine what would happen if grass covered the entire pixel.

To do this practically, we need to know the value of T and know the percent tree cover. We will perform a regression to obtain T, and we will use our habitat data for percent tree cover, making general estimates for each value.

We know that T will approximate

$$T \approx 1/90\% \times (\text{mean MB NDVI})$$

We get $T \approx 7.238281 \times 10^{-5}$.