JOURNAL OF
EMERGING INVESTIGATORS

Article

Presence of vegetation in relation to slope in Yosemite Valley, California

Colin Saltzgaber[1], **Julia Kelson**[2]
[1]The Nueva School, San Mateo, California
[2]Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, Michigan

SUMMARY

Vegetation is an essential part of our world and we rely on the presence of trees, plants, etc. in our current society, however vegetation itself also relies on other aspects of the Earth to thrive, for example, its growth conditions and habitat. Here we examined the relationship between the slope of a terrain and vegetation, measured by the normalized difference vegetation index (NDVI). We used satellite data from Yosemite Valley to test for statistically significant differences in NDVI values amongst terrains of different slope angles. We hypothesized that lower slope ranges would be more supportive of vegetation growth than higher slope ranges. We used confidence intervals and hypothesis tests to estimate the population mean NDVI for terrains with different slope angles. We report a statistical significance between the 30–45° slope range and the 75–90° slope range, which supports our hypothesis that lower sloped terrain is more supportive of vegetation greenness. Additionally, our analysis showed that no slope (even as extreme as 85–90°) prohibits the growth of vegetation completely; even the steepest slopes examined contain plant life. With our planet constantly evolving and growing, we are always looking for ways to improve efficiency. Knowing that steep slopes can still support plant life, we can begin to explore and start planting additional crops and plants at these extreme slopes.

INTRODUCTION

In a world that benefits immensely from the use of plants and trees, for example, through farming, we wanted to explore the presence of plant life at various slopes to determine the growth and thrive potential of green life at these various degree levels. Examining this relationship could provide new insight with regards to farming and improving efficiency with growing and cultivating crops. Here we explored the relationship between the slope of terrain and the presence of vegetation in Yosemite National Park to understand the role that slope plays in supporting vegetative communities. Vegetation presence is controlled by several factors including elevation, parent material, slope, and aspect (the compass direction that a slope or terrain faces). Here, we investigated the importance of slope. High degree slopes are commonly thought to hinder or prohibit vegetation growth; however, even cliffs can be supportive of vegetation. Previous work returned surprising and unexpected information about the presence of vegetation among cliffs, revealing that many cliffs were vegetated and supported woody taxa (1). Similarly, another research team found vegetation at high slopes; however, they also discovered that the steep terrain that the plant life was present on disturbed the growth process, resulting in "bent trees and widespread pit-and-mound microtopography" (2).

We hypothesized that cliffs would be less supportive of vegetation compared to lower sloped areas. We investigated the relationship between the slope of the terrain in Yosemite Valley and the vegetation that is present. We quantified this relationship using the normalized difference vegetation index (NDVI). NDVI represents the density of green within a certain pixel of a map, using two specific spectral bands, 842nm and 665nm (B8 and B4, near infrared and visible red wavelengths, respectively) measured remotely from a satellite (3). We analyzed data from Yosemite National Park because the location contains a large variety of slopes within a relatively small area. The slopes range from sheer granite cliffs (such as the famous Half Dome cliff or El Capitan), to flat valleys and meadows (such as Yosemite Valley) (4). Additionally, because the bedrock of Yosemite is primarily granitic, we avoided introducing variation due to rock type. Lastly, since we collected the data in a relatively small region in California, we avoided major differences in climate. After conducting our analysis of this relationship, we found statistical significance between all observed transects, which suggests a difference in vegetation presence in higher sloped areas compared to lower sloped areas. However, none of our observed transects and ranges for high sloped areas failed to show presence of vegetation, meaning that no degree of slope from 0 to 90 prohibited plant life in Yosemite Valley. This demonstrates that even at the steepest slopes on cliffs and mountains, plant life is still possible, which is an encouraging finding to our world which is always looking to become more efficient with growing crops and planting wildlife.

RESULTS

To analyze the relationship between slope and vegetation density, we measured the slope angle and greenness in the four transects in Yosemite, then tested for statistical

Table 1. Means and confidence intervals of each transect and the three ranges measured within.

Transect + Range	Mean NDVI	Lower Bound	Upper Bound
#1 – 75–90°	0.219	0.181	0.257
#1 – 30–45°	0.489	0.464	0.515
#1 – 0–15°	0.402	0.387	0.421
#2 – 75–90°	0.045	0.042	0.048
#2 – 30–45°	0.344	0.313	0.376
#2 – 0–15°	0.461	0.439	0.484
#3 – 75–90°	0.167	0.089	0.245
#3 – 30–45°	0.405	0.368	0.441
#3 – 0–15°	0.545	0.531	0.560
#4 – 75–90°	0.220	0.162	0.278
#4 – 30–45°	0.438	0.415	0.461
#4 – 0–15°	0.487	0.471	0.502

significance between the two variables. We drew four transects that encompassed a variety of slopes (both high and low degree terrain), allowing us to compare the NDVI values of respected ranges. The hypotheses that we used for our statistical analysis and tests are as follows: H_O (null hypothesis): the mean value of NDVI for the 75–90° interval is equal to the mean value of NDVI for the 30–45° interval; H_A (alternative hypothesis): the mean value of NDVI for the 75–90° interval is less than the mean value of NDVI for the 30–45° interval, at the 95% confidence level ($p < 0.05$).

We examined three slope intervals: 0–15°, 30–45°, and 75–90°. For these three slope intervals we calculate mean NDVI and the confidence intervals (**Table 1**). Confidence intervals of the means were calculated using a t-score (because the population standard deviation was unknown). For the hypothesis tests, we chose the interval of 75–90° degrees as the higher sloped data range and the interval of 30–45° as the lower sloped data range. Although we have NDVI values for the range of 0–15°, after examining the satellite imagery we found that terrains of 0–15° contained many roads, gravel patches, buildings, and dry meadows, all of which would ultimately affect the estimation of natural vegetation in those specific areas (e.g., if the grass is dead) (**Figure 1**). Thus, we chose to use the 30–45° range as our lower sloped data to compare to the higher sloped data in the hypotheses tests; the *p*-values are presented in **Table 2**. The hypothesis test for the difference between the means used a t-score with a corresponding *p*-value obtained from a one-tailed test.

Both the comparison of confidence intervals and the results of the hypothesis tests supported our hypothesis that slope is an influencing factor on NDVI values, more specifically that higher slopes were less supportive of vegetation than lower slopes (**Tables 1** and **2**). We tested the hypothesis using two methods to determine if there was statistical significance between the ranges of slope. We began with a statistical test to compare the ranges of the pairs of confidence intervals (one side of the pair being the NDVI values from high slopes (75–90°) and the other side being the NDVI values from lower slopes (30–45°)). If two confidence intervals fail to overlap (with their respected ranges), there is no data shared between the two, indicating statistical significance. All four pairs of confidence intervals failed to overlap one another, which means that no data was shared between the respective pairs of confidence intervals, meaning that those results were statistically significant (**Table 1**, **Figure 1**). The mean NDVI value for high sloped areas (75–90°) was consistently lower than those of the lower sloped areas (30–45°), which showed there was consistently less vegetation at higher sloped areas compared to lower sloped areas, supporting our original hypothesis.

Table 2. Hypothesis test results comparing the means of the two intervals measured for NDVI.

Transect	Degrees of Freedom	t-score	*p*-value
#1, 30–45° vs. 75–90°	140	10.006	0.00001
#2, 30–45° vs. 75–90°	107	15.717	0.00001
#3, 30–45° vs. 75–90°	105	5.0396	0.00001
#4, 30–45° vs. 75–90°	166	6.1596	0.00001
All/Combined Data	521	12.903	0.00001

We performed a second test for statistical significance which was a standard hypothesis test for the difference between means, which were represented through the t-scores for each transect as well as the *p*-values obtained (**Table 2**). All the *p*-values were lower than the alpha value (0.05), indicating that the samples showed statistically significant differences between the two ranges of slope. This means that the mean of NDVI for 30–45° was not equal to the mean of NDVI for 75–90°. This demonstrated the significant differences in the NDVI values of these two observed ranges of slopes, and that there was more greenness and dense vegetation present in lower sloped areas than higher sloped areas, given that our hypothesis test was a one-tailed test. In addition, we combined the data from each of the 75–

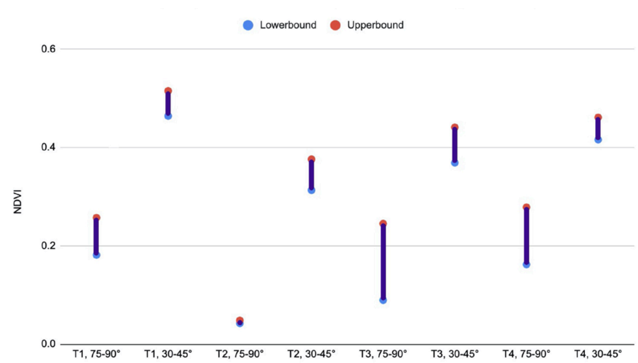

Figure 1. NDVI mean and confidence intervals for each transect and slope category from Yosemite. The mean NDVI, and the resulting confidence intervals, were calculated for each transect (T1, T2, T3, and T4) and slope (75-90° and 30-45° range) category from Yosemite. Eight confidence intervals are shown, two for each of the four transects that were analyzed (upper bound: red, lower bound: blue).

90° intervals as well as the data from each of the 30–45° intervals and ran a difference between means t-test. This test resulted in a *p*-value of 0.0001, giving us the same result as the previous test; the mean of NDVI for 30–45° was not equal to the mean of NDVI for 75–90°. This suggests that our hypothesis (alternative hypothesis) holds true not only for selected transects (which were drawn at random, but only offered a limited amount of data points), but also for larger areas (i.e., the majority of Yosemite National Park).

Interestingly, the distributions of NDVI in the 30–45° interval to the 75–90° interval were different from each other (**Figure 2**). The distribution of NDVI in the lower sloped terrain was left skewed, meaning that for this range of slope, it was more common to observe dense green patches of vegetation than seeing dull/colorless patches of vegetation; the distribution of NDVI in the higher sloped terrain was right skewed, meaning that there were more patches of colorless vegetation than dense green patches of vegetation. These results showed that for the 30–45° range, NDVI values lower than the mean were less common, whereas for the 75–90° range, NDVI values higher than the mean were less common. Pearson's coefficient of skewness for the combined values of all data within the 30–45° range and 75–90° range was -0.295 and 0.997, respectively. The two coefficients were different, in that one was positive and the other was negative, demonstrating the skewness direction: positive implies a right tailed skewness data distribution while a negative value implies a left tailed distribution. The value of 0.997 for the right skewed distribution for the 75–90° range means that the distribution for that range was more skewed than the distribution of the 30–45 range. This result indicates that in higher sloped areas, there were fewer data points with NDVI values higher than the mean, which is why it was right skewed and centered around a smaller value of NDVI. This supported our conclusion that higher sloped areas are less supportive of vegetation. In the higher sloped range, there were less observed patches/data containing high vegetation, implying that slope is an influencing factor for vegetation growth.

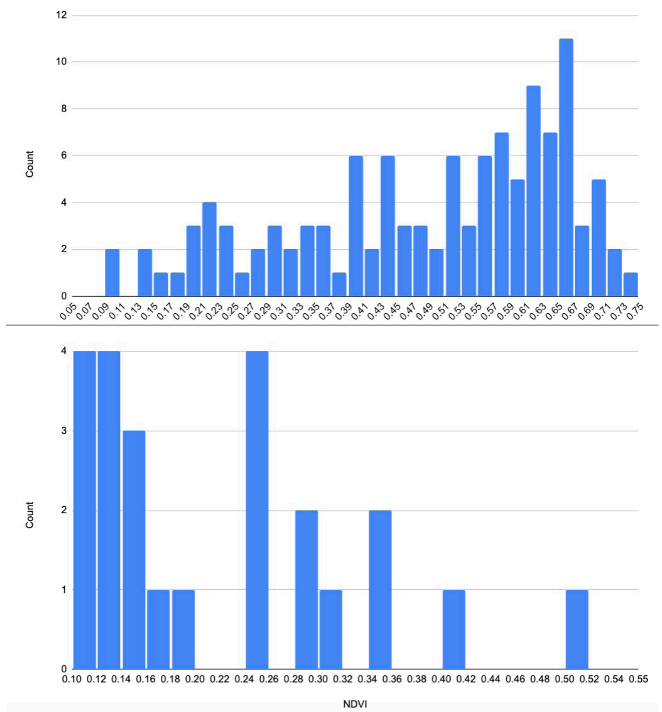

Figure 2. Histogram of NDVI from Transect 1. The upper panel shows the distribution of NDVI for the slopes in the 30-45° range. The lower panel shows the distribution of NDVI for the slopes in the 75-90° range.

DISCUSSION

Although our analysis suggested that lower sloped terrain is more supportive of vegetation growth compared to higher sloped terrain, we found that plant life exists even in those steeper areas. For all data points in the 75–90° range, all NDVI values were still greater than 0, which reveals that cliffs still support vegetation and that there is no point observed where slope hinders or prohibits plant life completely. This supports the findings of Larson et al., whose research revealed that even the steepest slopes could still support vegetation and plant life (1). It is also worth noting that because the satellite data was captured at a 10m resolution, the NDVI value reflects an average over 10m, and does not capture individual trees (< 10m).

However, because these tests used data from only four transects, there are potential sources of error. The first is variability in the data or the possibility that our four transects were not the best representatives of the population of Yosemite Valley and all its terrain. We acknowledge that there are some differences in climate parameters, like rainfall and temperature, that are related to changes in elevation. This may have affected the accuracy of our results to a certain extent; however, we believe that this was not a major problem because all the transects and samples indicated that there was statistical significance between the two slope ranges examined.

For potential further research, we believe that a two-way regression, incorporating another influencing factor for NDVI values (soil moisture, aspect, temperature, etc.) would further illuminate the environmental factors that control vegetation growth. Although we did our best to control many variables by limiting our study to a small area, there is a possibility that other factors may have influenced vegetation levels. This two-way regression would incorporate another factor that could address that slight variance in our original project (i.e., mean values differ slightly amongst transects). This would create a more thorough and complete way of modeling which variables impact NDVI. In this work, we treated the slope angle as the independent variable, but it is worth considering that vegetation can influence slope stability (and perhaps steepness, by extension); such feedback between slope steepness and vegetation have been long observed by geomorphologists (5). In the future, it would be worth exploring how vegetation can impact the slope of a terrain, perhaps through root structure and the mass of plants. Lastly, this procedure could be applied to evaluate the relationship between these two variables to other landscapes (both with similar and different characteristics such as climates). Our research and findings also have implications and can be applied towards other places around the world that have geologic features similar to Yosemite. These include other glaciated landscapes in the Rockies and in the Sierras, and globally. Investigating this relationship outside of Yosemite Valley would be important to explore in future studies.

METHODS

Datasets

To analyze the relationship between the slope of a terrain and its corresponding NDVI values, we examined two datasets. The slope dataset that we used for Yosemite Valley is LiDAR (light detection and ranging) acquired data, which was processed and completed by the National Center for Airborne Laser Mapping (NCALM) (6). This specific dataset was downloaded through *Open Topography*, which contains an elevation model at a 1m resolution that we used to calculate the degree of the slope. The dataset that we used to extract NDVI (an approximation of vegetation, as it measures the density of green on a patch of land) of Yosemite Valley was downloaded through *Sentinel Hub* (Measuring Vegetation 2000) (7). This dataset is based on satellite data captured at a 10m resolution, containing two specific spectral bands, B8 and B4 (near infrared and visible red wavelengths, respectively), which we used to calculate NDVI values using the equation: NDVI = index (B8, B4) = (B8-B4)/(B8+B4) (**Figure 3**) (5). This satellite image was recorded on July 11, 2017.

Dataset Processing and Transects

We overlaid both datasets, projected the two layers to the same resolution (10m) and to the Universal Transverse Mercator (UTM) system, ensuring that for every slope data point, there was a corresponding NDVI data point. We examined a portion of the valley which incorporated a large range of slopes, allowing us to analyze this relationship (**Figure 4**). Yosemite Valley is an east-west striking valley surrounded by granite mountains and cliffs on both sides. We created four total transects that spanned and encompassed these two different geographic features that are present (3). We extracted the corresponding data for each of the four specific lines which returned two columns of data: one for slope and the other for NDVI. The coordinates of the four transects that we extracted data from are as follows, with the format (UTM X, UTM Y): Transect 1, start: (270,404.51, 4,176,919.91), end: (271,549.30, 4,181,989.68); Transect 2, start: (277,978.33, 4,179,687.85), end: (278,747.21, 4,184,198.57); Transect 3, start: (273,728.40, 4,177,506.20), end: (274,797.07, 4,183,205.78); Transect 4, start: (271,049.85, 4,176,165.11), end: (272,387.42, 4,182,351.84).

Statistical Tests

Finally, to analyze and determine whether slope influenced NDVI, we used confidence intervals and hypothesis tests to test for statistical significance between two ranges of slope degrees. The specific hypothesis test was a standard hypothesis test for the difference between means using a t-score. Our α value was 0.05 which correlated with our threshold for statistical significance ($p < 0.05$). We divided the slopes into five categories of 15° intervals (e.g., 0–15°, 15–30°, etc.) up to 90°. We chose to examine the data in the 30–45° range vs. the data in the 75–90° range. Choosing these two ranges allowed us to compare the distributions of NDVI as a function of slope. We used Google Sheets to analyze and calculate the confidence intervals as well as our *p*-values. We also used Pearson's coefficient of skewness to measure the strength of the relationship and distribution of our data. In addition, our figures were made from both Google Sheets and QGIS.

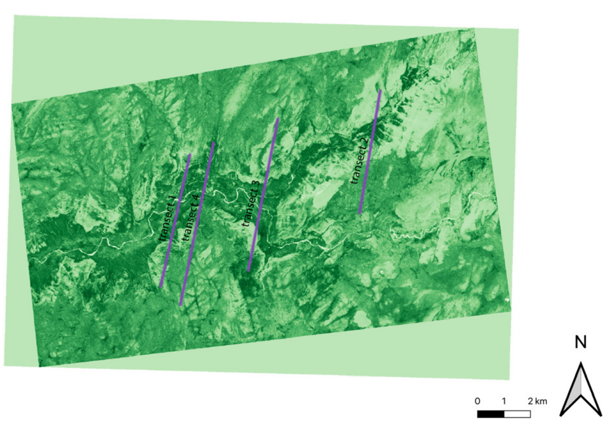

Figure 3. NDVI of Yosemite Valley and surrounding area. In the valley, darker green shades can be observed (high NDVI values). Low NDVI values are light green. In addition, the transects (T1, T2, T3, and T4) we examined are highlighted in purple.

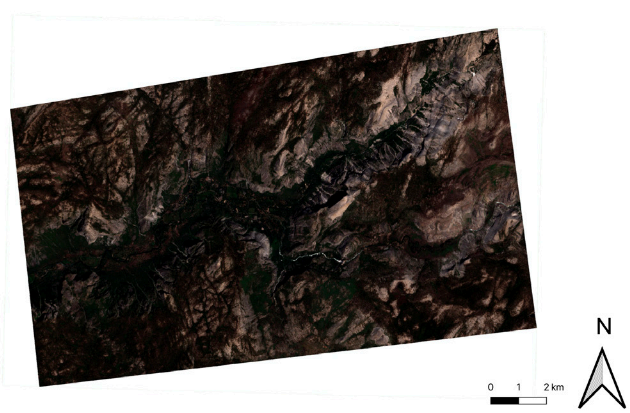

Figure 4. Satellite image of Yosemite Valley and surrounding area. An aerial image of Yosemite Valley captured from satellite on July 11, 2017. We drew the transects perpendicular to the general direction of the valley (E-W) which captured many angles of slope.

ACKNOWLEDGEMENTS

I would like to thank Spencer Eusden as well as Headwaters Science Institute for giving me the opportunity to conduct scientific research and pursue this passion of mine over the summer, as well as my parents for supporting me throughout this journey.

Received: October 10, 2020
Accepted: May 12, 2021
Published: September 11, 2021

REFERENCES

1. Larson, D. W., *et al*. "Evidence for the widespread occurrence of ancient forests on cliffs." *Journal of Biogeography.,* vol 27, no. 2, 2000, pp. 319-331.
2. Pawlik, Łukasz, *et al*. "Surface processes and interactions with forest vegetation on a steep mudstone slope, Stołowe Mountains, SW Poland." *Catena*, vol. 09, 2013, pp. 203-216.
3. Fragoso-Campón, Laura, *et al*. "Overstory-understory land cover mapping at the watershed scale: accuracy enhancement by multitemporal remote sensing analysis and LiDAR." *Environmental Science and Pollution Research*., vol. 27, no. 1, 2020, pp. 75-88.
4. Huber, N. King, *et al*. "Geologic Map of Yosemite National Park and Vicinity, California. U.S." *Geological Survey*, 1989. Map.
5. Gilbert, Carl Grove, and Clarence Edward Dutton. "Report on the Geology of the Henry Mountains." *US Government Printing Office*, 1880.
6. Stock, Greg (2012). Yosemite National Park, California 2007 Airborne Lidar Mapping (NCALM), distributed by OpenTopography, https://doi.org/10.5069/G9W66HPH
7. Sentinel Hub, "Measuring Vegetation 2000", Sinergise Ltd. https://apps.sentinel-hub.com/eo-browser/?zoom=15&lat=39.11714&lng=-120.23824 (NDVI imagery for world).