**Ryan Littleton**

**DSC 501: Introduction to Data Science**

**Rainfall and Desertification in California: Expanding on a Public Concern**

**Abstract**

In assessing desertification worldwide, many approaches have made use of vegetation indexes, with the most successful framing vegetation in terms of change given precipitation to evaluate actual fertility of a region. This allows desertification to be defined in terms not just of where deserts are, but in how they change, and whether those changes are a substantial, long term and non-reversable change. This issue is of great public concern, as fertility in general and desertification in particular are connected both to food security and climate change. For this study here to cast light on the issue, the question was posed, “To what degree is desertification in California attributable to changes in precipitation.” To address this question, a linear regression was set up to test the significance of precipitation’s effect on vegetation in California and to model predictively vegetation given the other variables. A null hypothesis that precipitation does not influence vegetation was set up with an alternative hypothesis that it does influence precipitation. The significance chosen to reject the null hypothesis was p = 0.05. Precipitation failed to meet the significance threshold as a variable influencing vegetation, and instead the climate divisions of California and the Palmer Drought Index were found to be significant. Because the precipitation was not a significant variable in influencing vegetation, follow up research is indicated by the results of this study. An expansion of the data to include more areas, a longer timeframe, and more specificity can all be used to make follow up studies more robust and give a conclusive statement about what causes are influencing vegetation change in the region, whether they are human caused, related to climate change or are simply normal variability. Further studies may also draw comparisons to other geographic regions or focus more on semi-arid landscape to better capture specifically desertification rather than other vegetation variability.

**Introduction**

In the year 2006, the United Nations General Assembly observed the year of deserts and desertification. This highlighted to public attention the significance of the expansion of deserts and soil fertility loss on a worldwide scale, which was reinforced years later when the same General Assembly declared 2015 year of the soils. This is certainly an issue that is of importance to all people, as it quite directly effects both food security and climate change over time. As this topic is naturally very large, a small aspect of it that is more localized and specific was used for this study. To limit the scope, the timeframe and the geographic region to be considered were limited to the years from 2000 through the first month of 2010 and the state of California. California is a suitable subject for general studies of climate data due to its geographic diversity. It contains a variety of features including but not limited to coastal regions, river basins, mountains and deserts. This provides fertile ground for conclusions to be drawn about climate variations in such a region. In light of the high levels of human activity and climate change which are under public scrutiny whenever environmental issues are discussed, it is useful to investigate what other factors might influence landscape change so as to differentiate the levels to which different variables are responsible. Here, an aspect of this is the subject of investigation. Specifically, this study is to address to what degree desertification in California is due to changes in precipitation.

**Literature Review**

One feature seen throughout much of the existing literature regarding desertification is some disagreement and even uncertainty as to how desertification should be defined and evaluated. Even in climate news sources, this problem is acknowledged. The article “Defining Desertification”, published online by the Nasa Earth Observatory, outlines the problem and one approach taken to both defining and detecting desertification in the Sahel region of Africa. (Riebeek, 2007) The biogeographer Stephen Prince, whose work is described in the article acknowledges the main issue with the problem of a definition of desertification by explaining that definitions of desertification that use only landscape description fall short. As he states, “But if it starts to rain and vegetation returns, what do you call it?” From this, it is evident that desertification is only worth classifying if the vegetation loss is not reversable by a change in weather patterns. That is, it can be defined in a more useful way as fertility loss. In the study described in this article, Prince utilized rain levels and satellite vegetation indexes over time to show that the Sahel region was not on the whole losing vegetation but was rather fluctuating as drought increased and decreased. What his study did find was that certain specific areas in the region did exhibit a lack of regrowth as rainfall increased and he concluded that these areas were the areas to pinpoint for further investigation as potentially desertifying. As this was a feature article and not a scholarly piece, the details of this process were discussed but minimal. Even so, the concern for a definition and the outline of an approach to the problem were discussed.

In “Expansion and Contraction of the Sahara Desert from 1980 to 1990”, found in *Science,* the patterns of movement of the southern border of the Sahara Desert are again described, again using satellite gathered vegetation indexes, in this case the normalized difference vegetation index (NVDI). (Tucker et all, 1991) Much as in the study described in the previously mentioned article, this study identified that, on the whole, the desert boundaries increase and decrease in similar amounts, with variability in specific timeframe depending on the weather of the year in question. Areas not prone to great change in desert boundary in general were deemed useful for identifying specific exceptions, and furthermore the scope of a greater desert boundary change was estimated to be a case for a multiple decade study rather than merely observation of the boundary over the course of a short time. Unlike the other studies mentioned, this study did not go to great lengths to define the desert and desertification as anything other than the change of the areas deemed “arid”. It did, however, give further example of the utility of satellite vegetation imagery data as a metric and tool for evaluation.

The article “Assessing the effects of human-induced land degradation in the former homelands of northern South Africa with a 1 km AVHRR NDVI time-series” by K. Wessels, S. Prince, P. Frost, and D van Zyl from the publication Remote Sensing of Environment outlines a more detailed approach to this question. They begin by defining desertification as land degradation and fertility loss. (Wessels et all, 2004) This does away with the notion of defining desertification merely by the boundaries of desert landscape and introduces a more utilitarian definition from a conservation standpoint. To assess this state, multiple metrics were used to isolate a repeatable method of identifying land degradation for the purposes of their study. First, land was categorized into degraded and non-degraded segments based on the results of a survey done by professionals employed in South Africa’s survey of land degradation. Then precipitation data and the normalized difference vegetation index (NVDI) were used to determine the vegetation growth per unit of rainfall for the regions in question. Areas judged to be similar in land capability units, an agricultural measure of how fertile a region could be, considering its geography were then compared to determine the measurable differences between degraded and non-degraded areas. This study establishes a method, then, that makes use of satellite image derived vegetation growth data and precipitation indexes to identify land areas that could be deemed under threat of desertification. This offers a methodological precedent for the kind of inquiry this current study is aimed at.

Another study that addresses the question of how best to identify desertification using available climate data is “Assessing Desertification” by S.R.Veron, J.M. Paruelo and M. Oesterheld in the *Journal of Arid Environments.* (Veron et all, 2006) This publication is a historical meta-analysis of various definitions and methods used to determine the desertification status of an area, and specifically deals with how it is defined and detected. It broadly discusses methods used to assess the phenomenon of desertification, eventually drawing conclusions of what the problems with older methods are and settling on a method the author purports to be best, while acknowledging the problem is an ongoing one. The methods assessed are desert edge displacement, which looks at physical features associated with deserts and looks at where their boundaries change, field data indexes and cryptic matrices, which use a multivariate matrix of indexes to attempt to quantify the phenomenon and rain use efficiency, which looks at the effectiveness with which vegetation utilizes rainfall, compared to how much is lost to runoff, etc. The final key findings are that methods based off rain use efficiency are most effective and objective, especially combined with a separate index, precipitation marginal response, which adds the ability to measure the changes in plant type. The conclusions drawn here include, then, further evidence that vegetation indexes and precipitation indexes, when used in combination, offer an effective means of identifying fertility loss of land in a given region.

**Data**

For this study, two sources of data were used. First was a multivariable online output from the NOAA databases of climate data, constrained to the state of California and the time frame of 2000 to 2010, further divided on a month by month basis. (NOAA, 2018) This data was geographically divided into seven climate divisions, represented in the dataset by number codes 1 through 7. The corresponding names of those climate divisions are specific to California and are as follows: 1 is North Coast Drainage, 2 is Sacramento Drainage, 3 is Northeast Interior Basins, 4 is Central Coast Drainage, 5 is San Joaquin Drainage, 6 is South Coast Drainage and 7 is Southeast Desert Basin. For reference, a map of these climate divisions is included below.

**Figure 1: California Climate Centers Map (Climate Prediction Center Internet Team, 2006)**

Other variables included in this dataset are the precipitation (PCP), which measures rainfall in inches along with the standardized precipitation indexes for probabilities of precipitation in a given number of months (SPxx). The first is more generalized, and the second is divided into 1, 2, 3, 6, 9, 12 and 24-month ratings. A zero index represents the median, with -3 representing extreme dryness and +3 representing extreme wetness over the given period. Together, these variables capture the data relating to rainfall directly, but several variables are also included which capture ratings for periods of drought specifically. These variables include the Palmer Drought Severity Index (PDSI), the Palmer Hydrological Drought Index (PHDI) the Modified Palmer Drought Severity Index (PMDI) and the Palmer Z Index (ZNDX). The first three of these all fall on a scale of -7 to 7, with negative values indicating dry spells and positive values indicating wet spells. The range considered to be “typical” variation by the agency, however, is -6 to 6. Anything outside this range is extreme for this index. The ZNDX, on the other hand, is a measure of the z index of the other Palmer indexes, or the departure of them from the average for a given month.

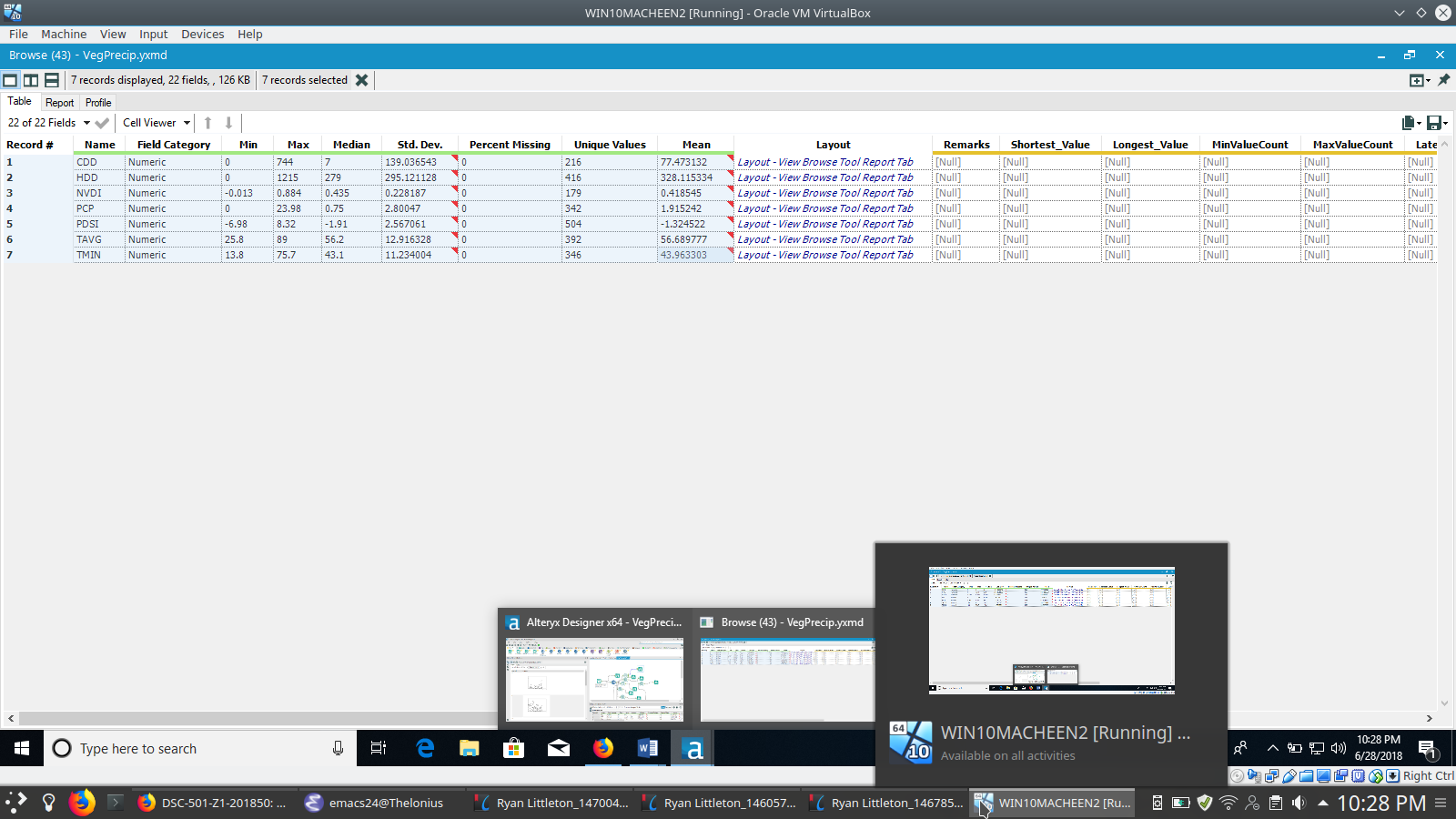
Other variables in this dataset relate to temperatures for a given month. These variables include the temperature average for the month (TAVG), the temperature maximum (TMAX), the temperature minimum (TMIN), cooling degree days (CDD), and heating degree days (HDD). The average, minimum and maximum are self-explanatory, while the heating degree days and cooling degree days represent the differences between the mean daily temperature for the period and a base value of 65 degrees Fahrenheit, with heating degree days representing a negative difference and cooling degree days representing a positive difference. So, a temperature of 63 degrees on one day would add 2 to that month’s HDD index and a temperature of 67 would add 2 to that month’s CDD.

The other data set used tables of normalized vegetation difference index (NVDI), gathered as a 1-month average of the index over the measurement period. (Alan & Stockli, n.d.) This index is a measure of vegetation levels as reported by light reflections gathered from satellite imagery. As discussed in the literature review, this variety of data has been used in other areas to assess changes in vegetation and give information regarding soil fertility differences associated with those changes. While this data exists in raw GIS format, for the purposes of this study, an output from NASA in csv format was used, designated by latitude/longitude pairs. A further narrowing was done in which the value at the location closest to the average latitude and longitude of each climate center per the NOAA metadata of the other dataset was chosen as the representative for vegetation data for that climate center. There are limitations to this approach, as it does not account for all locations in a geographic region and relies on the assumption that average location for a climate division correlates to the rest of the climate division, but given the purpose of climate divisions as groupings of areas with similar climate features, it serves to allow trends to be discovered within the computing power and time that was available for this study.

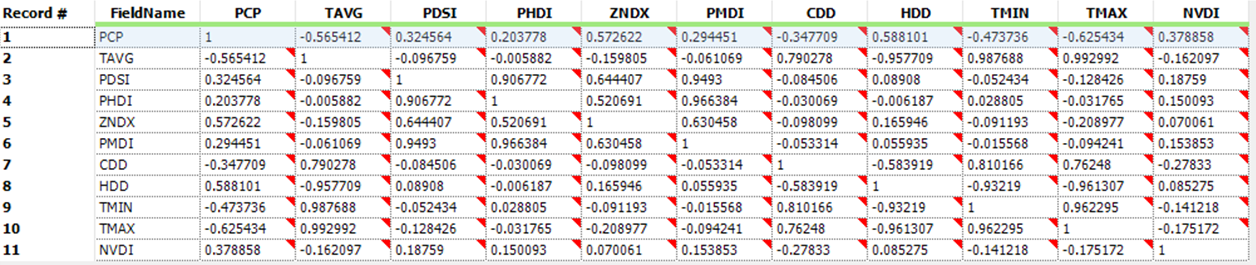
As both of these datasets are derived from government climatology institutions, there is a certain expectation of accuracy and lack of bias to them, however there are still several areas bias might be introduced. First, the NOAA climate summary datasets rely on information gathered from climate centers that is then aggregated per climate division. Spatial autocorrelation allows the assumption to be made that these can be used to make assumptions about these climate phenomena uniformly across these regions, but that also necessitates good sampling as far as locations are concerned as well as accurate readings from these climate centers. This data has a high likelihood of being accurate, then, but there is always the possibility for the introduction of error. Second, specific indexes are used, and the assumption is made that these are the best ways to aggregate this kind of data. Any time an index is calculated, some specificity is lost for the sake of readability and the ability to interpret. This is very useful, but it does involve a reliance on the specific choices made when designing the indexes. The algorithms for these indexes are publicly available and therefore subject to scrutiny, but they still must be relied on and assumed accurate in order to make use of this data. Next, as relating to the NVDI data, as mentioned above there was first a level of sampling done to reduce the dataset from raw GIS data and then another level of sampling done as part of this study to give single values for each climate division, allowing a comparison to be made. This introduces the possibility for sampling error as well.

**Data Analysis**

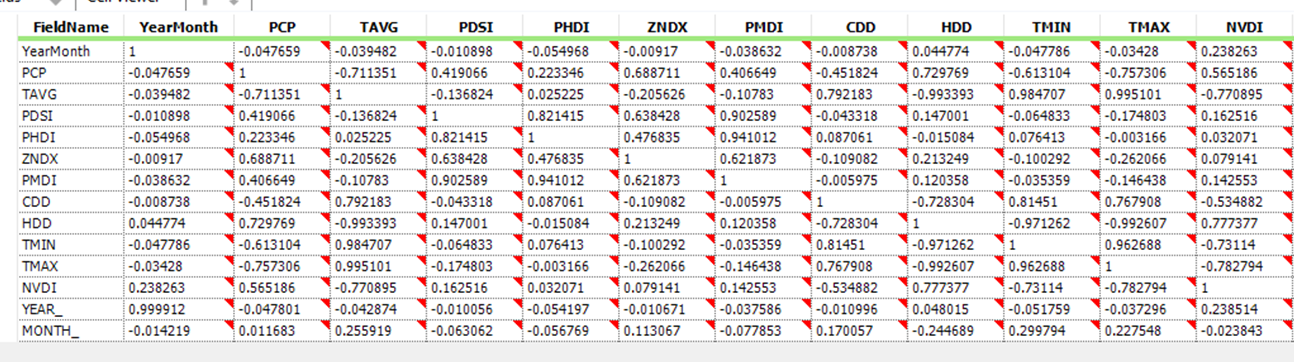
The exploratory data analysis and also the later statistical modeling was performed using Alteryx Designer version 2018.1.4.44311. At the beginning of this analysis, the decision was made to eliminate the SPxx date specific precipitation index variables from consideration, as they correlated highly with other precipitation and drought related variables and were largely describing the same phenomena over slightly different timeframes. To avoid multicollinearity, they were taken out at this stage. No variables contained missing values, but several did contain zero values. As can be seen from the minimums, these were the CDD, the HDD and PCP, all of which share the trait of having values for which zero is a likely long-term possibility. PCP could easily be zero during a period of sustained drought, and in fact, some might even define a drought by such a value. Likewise, the months where HDD or CDD were at zero values contained an abundance of the opposite, merely indicating a month where every day was either below or above 65 degrees Fahrenheit, respectively. As these are all explainable, measurable values rather than missing values, they do not require any special treatment. When the summary statistics are viewed, the CDD, HDD and PCP have medians which differ significantly from the means. This can be an indicator of outliers, which considering the nature of these variables, is not unexpected. The ability of these variables to have a real value of zero could drive down the median significantly while high values at different seasons could keep the mean at a different level. These same factors also account for the high level of variance among these variables. The NVDI has a range including a negative number, which is unexpected, as the index ranges from 0 to 1, however the negative number is extremely small, and unlikely to have significant influence on the analysis, considering the mean, median and maximum are all within expected values. Likewise, the temperature data all fall within ranges that are normal for temperatures and the Palmer drought index (PDSI) contains terms within the expected range, with minimum and maximum representing the extremes.

**Table 1: Summary Table Statistics**

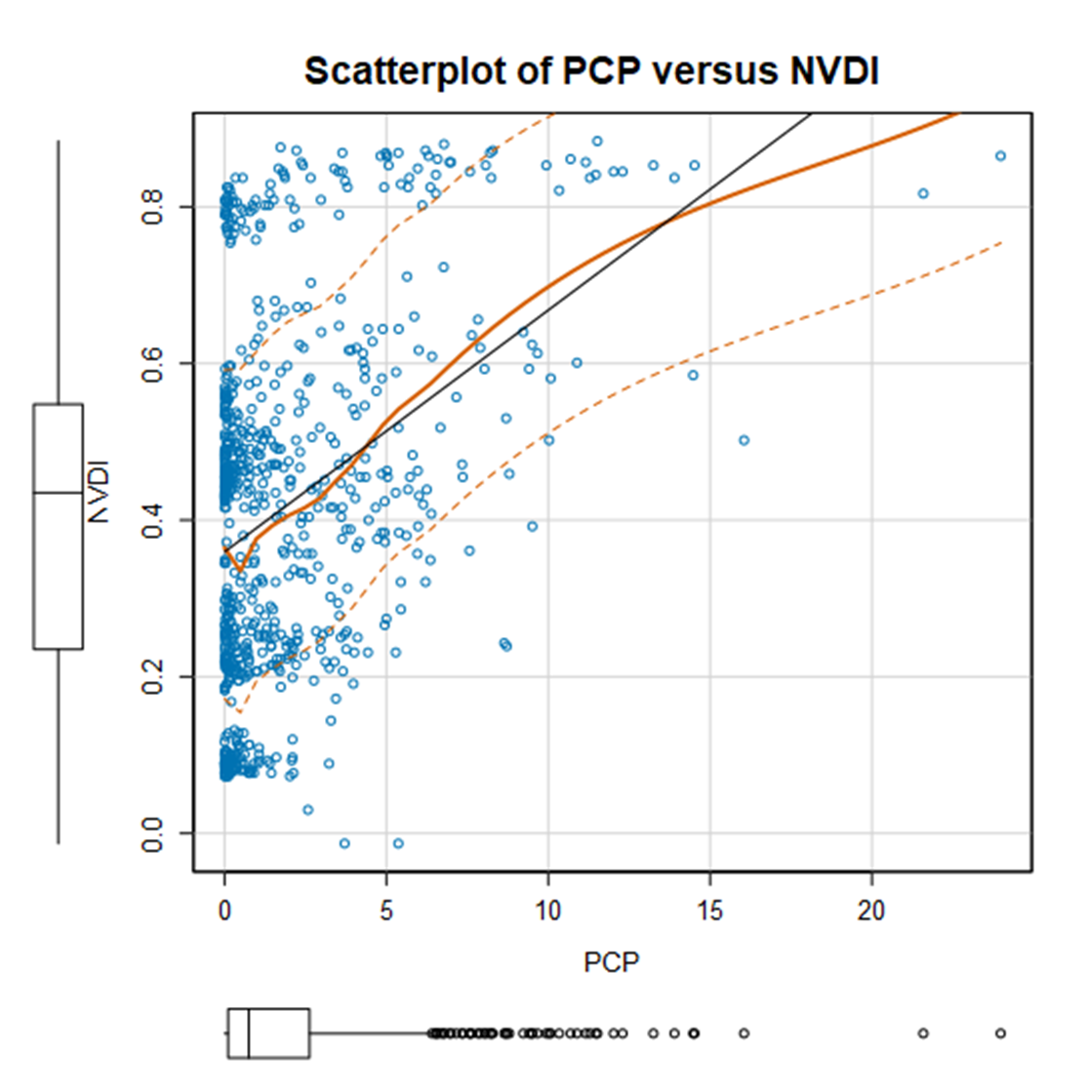
Observing the correlations among the variables, a high level of correlation is found between the drought indexes (PDSI, PHDI and PMDI), and so it is reasonable to reduce these as well to one, the PDSI to avoid the effects of multicolinarity. Similarly, the temperature maximum, minimum and average all correlate over .9, and in the final analysis only one will be needed. In this case TAVG was chosen as the variable under consideration. There are less strong but still significant levels of correlation between the PCP and the NVDI, which indicates an area worth investigating for this study, as the NVDI is the variable likely most useful for evaluation.

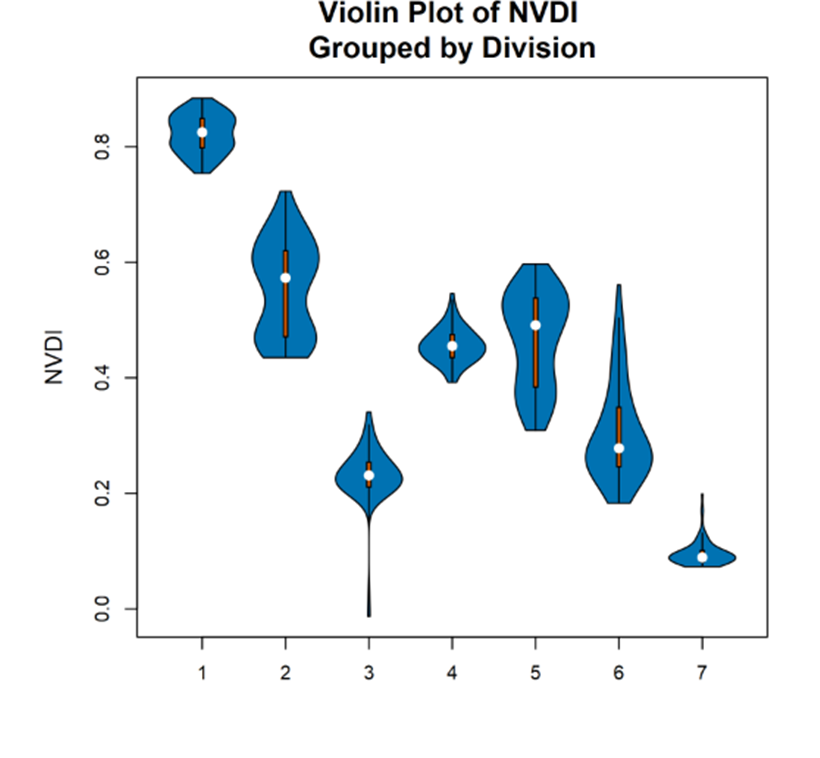
**Table 2: Correlation Matrix**

When the correlation tables were formed again on a division by division basis, it also was evident that some correlations were stronger in one climate division than another, or as compared to the average of all divisions. The exploratory analysis contained many such hints that there was a significant variation from one geographic location to another that merited investigation. One such example is the stronger correlation between the PCP and NVDI in climate division 3 (The Northeast Interior Basins).

**Table 3: Correlation Table, Division 3**

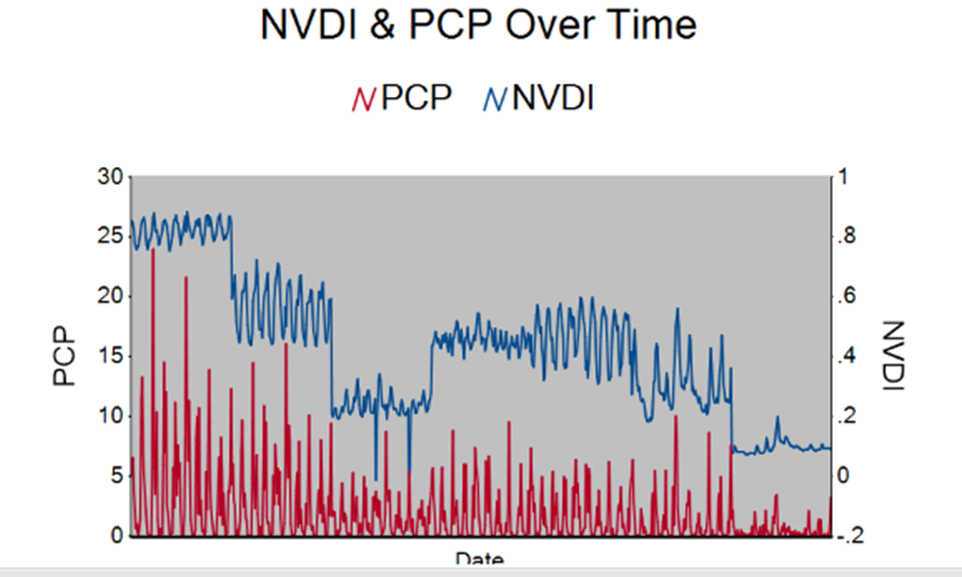
Following these investigations into correlations and the difference between the climate divisions, it bears investigating what sort of relationship the NVDI and the PCP have with one another, and how that connects to the differences between the climate divisions. As seen on the plot below, the NVDI and the PCP appear to have a linear relationship of some sort, but it is obscured by a clustering effect in certain groups of vegetation indexes. Further plotting of the means of the NVDI by the climate divisions show that these clusters are exactly the scale that some of the unique ranges of the climate divisions contain. It is also worth noting that some of these clusters of ranges show greater correlation than others, which emphasizes the earlier observed difference in correlation between NVDI and PCP among the climate divisions.

**Table 4: Plot of NVDI by PCP**

**Table 5: Violin Plot of NVDI Grouped by Division**

Another area that bears investigation is the change over time of these variables. Given the influence of long time frames on climate data, seeing any trends that might appear can give insight as to what findings are part of a normal variation as opposed to a real trend or change. When plotted across the entire timeframe of ten years, the NVDI shows periods where it increases and decreases around a particular center, followed by a shift. This results in a stairstep or plateau like appearance which is mirrored to a certain, though lesser extent in the overall trajectory of the precipitation. Both trend overall downward during the evaluation period. This underlies a potential bias in the data chosen for this study, as the timeframe is in terms of climate, relatively short. It is entirely possible that this period of measurement takes place over a naturally more arid or drought prone timeframe that has the potential to correct in the opposite direction over a longer period. It is also possible that there are other human factors causing some of the vegetation issues, like human influenced climate change or agricultural technique changes. Because the period is only 10 years, it is questionable whether this particular dataset is a representative climate sample.

**Table 6: NVDI and PCP Over Time**

****

**Methodology**

For the purposes of answering the question, “To what degree is desertification in California attributable to changes in precipitation?” several methodological steps were taken. Following the previously described data preparation and analysis in which several variables were removed for multicollinearity, the dependent variable of NVDI was selected to be evaluated based on the independent variables PCP, PDSI, Climate Division, HDD, CDD and TAVG. Of these, the climate division- being a categorical variable- was set up as a dummy variable with seven outputs, one for each climate division. A linear regression equation was set up to evaluate each of these variable’s effect on the dependent variable, with two goals. Prior to running the data through the regression equations, 20% was sampled and reserved for testing against the expected values generated by scoring of the model. First, to create a predictive model of the climate data to predict vegetation based on the other given variables, and second, to evaluate the null/alternative hypothesis pair in which the null hypothesis is that precipitation does not affect the NVDI and the alternative hypothesis is that precipitation affects the NVDI. This hypothesis testing was accomplished with significance testing with a threshold of p = 0.05.

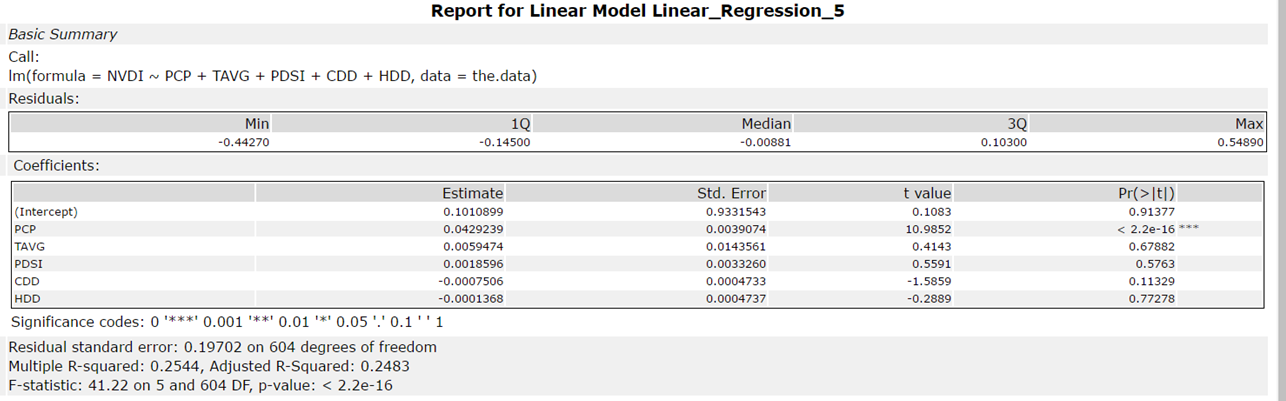
The advantage of the linear regression equation in this case is the ability to apply predictive power to a relationship that appears linear. As previously discussed, the relationship between the NVDI and PCP showed signs of linearity, making this an appropriate selection. One weakness is if the relationship is incorrectly identified as linear, the results would then be compromised.

**Findings**

The initial iteration of the linear regression was performed without including the climate division in the equation and found that the PCP passed the significance threshold- it was in fact the only variable in this model considered significant, with an equation also passing the threshold, initially appearing that the null hypothesis should be rejected. However, the R-squared was only 0.2483, implying a fairly low predictive power of the whole equation. This led to a search for a more optimal model. The regression equation for this initial model was:

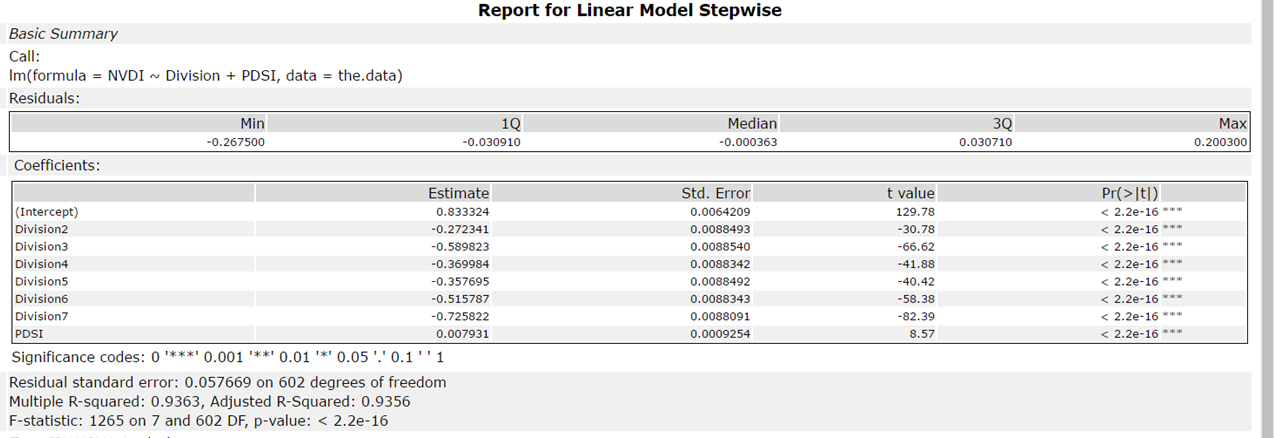
NVDI = PCP \* 10.9852 + TAVG \* 0.4143 + PDSI \* 0.5591 + CDD \* -1.5859 + HDD \* -0.2889

**Table 7: Linear Regression Output Without Climate Divisions**



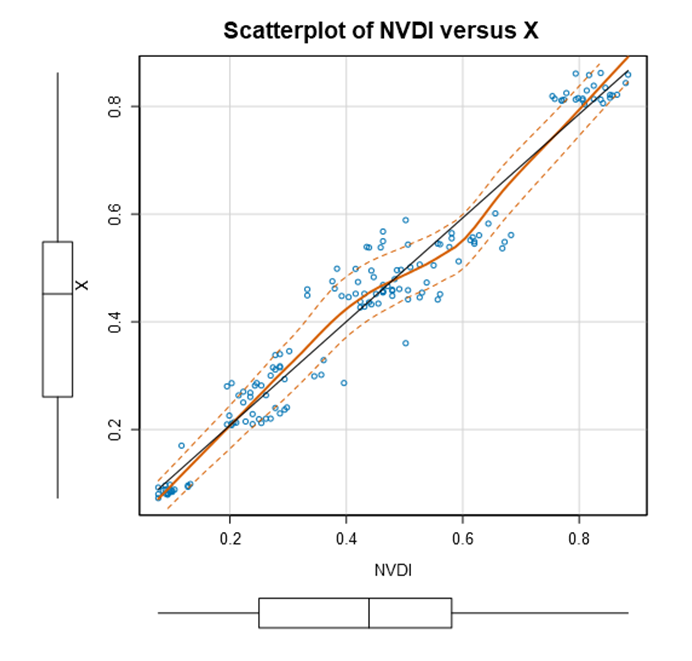
The next step in finding a more optimal model was including the climate divisions. This ended up making a much more effective model but changed the previously reported results from the first model to something quite different. When divisions were included, PCP was no longer a significant variable, but the coefficients for the divisions and the PDSI were. This suggests not rejecting the null hypothesis. Furthermore, this model had an R- squared value of 0.9356, indicating quite a lot more predictive power, and this equation also meets the significance threshold. Based on this, the null hypothesis should not be rejected. The regression equation for this model is:

NVDI = PDSI \* 8.57 – 30.78 (if Division 2) – 66.62 (if Division 3) – 41.88 (if Division 4) – 40.42 (If Division 5) – 58.38 (if Division 6) – 82.39 (If Division 7)

**Table 8: Linear Regression Output with Divisions**

As a follow-up to this output, two additional steps were taken. First, the dataset was divided by the climate divisions and separate regression equation was run with all variables included for each division separately. This was done to further validate that the results of this model, and that the model is in fact the best one for the data in these datasets. Second, as a safeguard against overfitting, a scoring tool was employed to give an output of the regression equation that was then tested against the 20% of the dataset that was earlier sampled and reserved. Then, the predicted values were compared with the actual values for accuracy. This also serves both to verify the predictive ability of the model. The when tested for correlation, the coefficient for the predicted and actual values was 0.97455, indicating that the model was in fact as predictive as expected. Below is a scatterplot of the predicted and actual values for a visual representation of this.

**Table 9: Scatterplot of Predicted (X) vs Actual (NVDI) Values**

**** These results were not without surprises, of course. The climate divisions, rather than any of the precipitation indices mentioned previously, dominate the model. This may be due to them being chosen based at least partially on vegetation type, among other geographic factors. It does stand to reason, for instance, that the Southeast Desert Basin (Climate Division 7) might have less overall vegetation than the other divisions. It is also rather unexpected that the PDSI had a positive rather than negative coefficient in the regression equation here generated. This seems to indicate that drought conditions rather than precipitation correlate positively with vegetation growth. This combined with precipitation levels themselves being found not significant gives a different picture than what was expected initially and during the exploratory phase of the research

It is worth noting that these findings are limited to California, which is a rather unique location. On the one hand, it has a wide variety of landscape features, which gives the appearance that it may be able to be generalized to many other areas, however it also has unique precipitation patterns, temperature variation and of course its own set of soil types. Also, unlike other studies cited in the literature review, this one did not focus only on areas neighboring deserts, so the conclusions have a different meaning regarding desertification. They might have implications regarding land fertility, drought and the variation of land types in California, but it would be erroneous to assume they can be applied in other geographic regions or that all variation in NVDI noted in this dataset specifically pertains to desertification.

**Summary**

In this study, to expand upon the question, “To what degree is desertification in California attributable to changes in precipitation,” climatological data including indexes of precipitation, drought and temperature was analyzed, divided by the climate divisions of the NOAA, and combined with the normalized difference vegetation index (NVDI) of the same areas to arrive at a greater understanding of the forces driving desertification in California. The way this study was approached was influenced by the previous work of Wessels, Prince, Frost and Zyl in northern South Africa, in which desertification was assessed as changes to a function of vegetation growth per unit of rainfall, (Wessels et all, 2004) and further by Veron, Paruelo and Oesterheld, whose meta-analysis identified several methods of identifying desertification and solidified it as a case of fertility loss rather than a function of the boundaries of arid regions alone. It is of great importance in any such study to establish this difference between a desert, which is a naturally occurring geographic area and desertification, which is a long-term change in fertility that leads to the expansion of arid vegetation types and densities. Based on these previous works, this study used significance testing and regression analysis to evaluate the hypothesis that precipitation in California significantly influences long term vegetation as opposed to the null hypothesis that it does not. The null hypothesis was not rejected, leading to the conclusion that precipitation is not the primary driver of vegetation differences in California. Furthermore, on building the linear model, it was found that the drought index and climate divisions themselves were of far greater significance.

The literature review and unexpected nature of these findings offers some insight as to where bias and some methodological weaknesses might have been introduced in this analysis. First, as desertification was what was being addressed, the assumption in this study that the vegetation indexes from all of the climate divisions in California were equally useful has the potential to be erroneous. It is possible that some climate divisions were not under threat of desertification at all, and only added noise to the signal of the study. Furthermore, the steps taken in data processing to facilitate the analysis limited the amount of vegetation data below what was available, and it is possible that the choices made in truncating the dataset were chosen in a fashion that was not representative of the regions as a whole. Another source of bias in the analysis is the view that California was under the threat of desertification. This presupposition influenced the whole study, and should that have been an inaccurate assumption, the study would be significantly less useful in addressing the question. Another source of bias lies in the selection of the specific dates involved. The assumption was made that these dates are more or less representative of climate in California, but that may well not be the case. For instance, the years immediately following the dates used in this study, 2011-2017 were identified by the Public Policy Institute of California as the driest period since records have been kept. (Hanack et all, 2016)

In order to responsibly use the conclusions of this study, it is imperative to understand that the results are first of all very much tied to the geographic region of California and second that they are not conclusive regarding the actual cause of vegetation change in the region, or even desertification in specific areas as discussed elsewhere. It is vital to acknowledge this, as issues relating to desertification and climate can influence the policy decisions that inform such activities as agriculture and water conservation. Acting inappropriately on such matters can cause distress to the livelihoods of people working in those sectors. On the other hand, failing to act should there be sufficient cause for action can result in long term consequences for food security and accelerate climate change. Given the gravity of these decisions, this study cannot ethically be taken on its own and must either indicate future research that is more detailed or be used alongside other studies that are effective in those ways. Furthermore, nothing here can ethically be taken outside the geographic area this study is in without significant further information tying these effects to those regions in question.

Many opportunities exist to expand upon the research done here. The time frame in question can be expanded, which is very much recommended for other uses of this data, especially considering the observation made by Tucker, Dregne and Newcomb that several decades might be necessary to make judgements regarding changing desert boundaries. (Tucker et all, 1991) In addition, the data can be made significantly more granular, as within the climate divisions, each climate center has its own output that is available to the public, and the NVDI data exists in raw GIS documents that are not limited by the scale of latitude/longitude coordinates to the extent that the CSV versions are. This could allow focusing on areas more likely to be under threat of desertification. An investigation into whether any areas have been classified as healthy versus degraded could be made, so that their behaviors can be compared, and similar data to that included here also exists for the rest of the United States as well as much of the rest of the world, so the geographic boundaries can also be expanded to make a study more generalizable.

**Works Cited**

Alan, J., & Stockli, R. (n.d.). Vegetation Index (1 month - Terra/MODIS). Retrieved June 03, 2018, from https://neo.sci.gsfc.nasa.gov/view.php?datasetId=MOD\_NDVI\_M&date=2001-04-01.

Climate Data Online Data Output. (2018, January). Retrieved May 26, 2018, from https://www1.ncdc.noaa.gov/pub/orders/CDODiv1044737646037.txt. doi:https://www7.ncdc.noaa.gov/CDO/cdodivisionalselect.cmd

Climate Prediction Center Internet Team. (2006, January 6). NOAA Climate Divisions with Counties- California [Digital image]. Retrieved June 28, 2018, from <http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/regional_monitoring/CLIM_DIVS/california.gif>

Hanack, E., Mount, J., & Chapelle, C. (2016.). California's Latest Drought. Retrieved June 29, 2018, from http://www.ppic.org/publication/californias-latest-drought/

Riebeek, H. (2007). Defining Desertification : Feature Articles. Retrieved from https://earthobservatory.nasa.gov/Features/Desertification/desertification2.php

Tucker, C. J., Dregne, H. E., & Newcomb, W. W. (1991). Expansion and Contraction of the Sahara Desert from 1980 to 1990. Science, 253(5017), 299-300. doi:10.1126/science.253.5017.299

Veron, S. R., Paruelo, J. M., & Oesterheld, M. (2006). Assessing Desertification. *Journal of Arid Environments,* *66*(4), 751-763. doi:https://doi.org/10.1016/j.jaridenv.2006.01.021

Wessels, K., Prince, S., Frost, P., & Zyl, D. V. (2004). Assessing the effects of human-induced land degradation in the former homelands of northern South Africa with a 1 km AVHRR NDVI time-series. Remote Sensing of Environment, 91(1), 47-67. doi:https://doi.org/10.1016/j.rse.2004.02.005