Modelling spatial patterns of vegetation activity and climatological parameters in the U.S. Great Plains: a satellite bioclimatology case study

S. Tan [a,*]

[a] Department of Geography, University of Cambridge, Downing Place, Cambridge CB2 3EN, U.K. – syt23@cam.ac.uk

Abstract – This study attempts to model vegetation dynamics through a spatio-temporal analysis of NDVI and climate data in the semi-arid region of the U.S. Great Plains. NDVI, monthly temperature, and monthly precipitation records from 305 stations were collected from 1982 to 2000. Correlation analysis, principal components analysis, and cluster analysis captured the seasonal response of NDVI to climate variability, supporting temperature as the dominant climate regime. Cluster analysis was used to develop a climate regionalization scheme based primarily on temperature, and NDVI characteristics were compared. The study resulted in a detailed understanding of climate and vegetation interactions in the Great Plains.

Keywords: Satellite bioclimatology; NDVI; Climate change; Land cover; Great Plains

1. INTRODUCTION

The spatial and temporal dynamics of vegetation are generally dependent upon environmental factors including various climatic conditions. Temperature and precipitation are two of the most important climatic factors that control differences in the Earth's vegetation cover, affecting growth rate, plant reproduction, and frost damage (Wang et al., 2001). However, the relationships between ecoclimatic conditions and vegetation growth are often complex and indirect. Consequently, a better understanding of vegetation-ecoclimatological interactions and their dynamics in space and time is required for modelling regional and global atmosphere-biosphere processes (Martin, 1993).

Previous research has demonstrated that satellite images, such as those obtained by the Advanced Very High Resolution Radiometer (AVHRR), provide unique opportunities for mapping, monitoring, and evaluating vegetation (Goward, 1989). The Normalized Difference Vegetation Index (NDVI) is the most widely used vegetation index, computed from the first two channels of the AVHRR. The NDVI is broadly indicative of plant photosynthetic activity and aboveground primary production, and has been widely used for assessing vegetation phenology and estimating landscape patterns of primary productivity (Sellers, 1985; Tucker and Sellers, 1986). However, as an indicator of vegetation activity, spatial patterns of NDVI are responsive to climatological variables, such as precipitation and temperature.

In this study, the influences of climate variables, such as temperature and precipitation, on spatial patterns of vegetation biomass were examined by integrating satellite (NDVI) and meteorological datasets. The U.S. Great Plains was an appropriate study area, since it represents a semi-arid region highly responsive to interannual climate fluctuations and long-term climate trends (Reiners, 1995). Being highly sensitive to climate variability and seasonal extremes, this area provided a template for studying critical interactions between climate and vegetation productivity.

2. MATERIALS AND METHOD

2.1 Study Area

The majority of the Great Plains region is classified as semi-arid. This region is characterized by a temperate continental climate with a strong east-west precipitation gradient and a north-south temperature gradient. Mean annual precipitation ranges from less than 450 mm toward the west to more than 1,200 mm toward the south-east; whereas mean annual temperature varies from less than 11°C toward the north-west to more than 15°C toward the south-east. Accordingly, vegetation ranges from tall-grass prairie and gallery forest in the east to short-grass prairie in the west (Reiners, 1995).

2.2 Data Sources

The data used in this study consisted of: (i) monthly AVHRR maximum value composite (MVC) 8-km resolution NDVI images covering the Great Plains, from 1982 to 2000, and (ii) precipitation and temperature data collected from surface meteorological measurements at various weather stations located throughout the study area over the same time period. For the purposes of this investigation, the study area was restricted to the U.S. portion of the North American Great Plains, extending from 26.38°N to 48.97°N latitude and 114.15°W to 90.27°W longitude.

Climate data were available from the United States Historical Climatology Network (USHCN) serial temperature and precipitation data sets. Climate data for 305 stations corresponding to the Great Plains study site were extracted and individual data sets for total monthly precipitation (hundredths of inches) and monthly mean temperature (hundredths of degrees fahrenheit) were used for analysis. Digital maps of greenness (NDVI) were derived from meteorological satellite data collected from the NOAA/NASA Pathfinder AVHRR Land (PAL) program, extracted from the PAL Global 8 km product with temporal coverage for the 1982 to 2000 time period. Since three 10-day composites were available for each month, it was necessary to compute the maximum value composite (MVC) value for each of the 228 months available. This was achieved by extracting the maximum NDVI value for each pixel of each of the images scenes of a given month.

* Corresponding author. Tel.: +44 1223 333399; fax.: +44 1223 333392.

2.3 Statistical Analyses

The purpose of this bioclimatological analysis was to quantify the spatial distribution and dynamics in vegetation activity produced by climate variability on a sub-regional scale. Statistical analyses employing correlation analysis, principal components analysis (PCA), and cluster analysis captured the seasonal response of NDVI to climate variability.

Using the available NDVI and climate data, the analyses focused on monthly periods over the 1982 to 2000 time period. Since a significant amount of data was missing for the 1994 and 2000 data sets, these years were excluded in several of the statistical tests that were not able to account for missing values (i.e. PCA). It was also necessary to remove the seasonality in the data, otherwise much of the patterns detected would have been attributed to the well-known seasonal variations that exist between vegetation response and annual climatic phenomena. The transformation used in this study was the monthly Z-score, which involved taking the set of values for a given month, subtracting the mean, and then dividing by the standard deviation (Yarnal, 1993).

3. RESULTS

3.1 Correlation Analysis

Pearson correlation coefficients between NDVI and monthly temperature and between NDVI and monthly precipitation were calculated for each of the 305 weather stations. Since a time lag exists between climatic events and the response of vegetation to such events, the time interval was taken into account by computing the NDVI/temperature (C_{NT}) and NDVI/precipitation (C_{NP}) correlation coefficients using time lags of 0, 1, 2, and 3 months. The maximum values occurring at each station were then selected as the C_{NT} and C_{NP} for each corresponding station. An overwhelming majority (52%) of the stations resulted with maximum C_{NT} at lag 0, whereas maximum C_{NP} was observed at a variety of lag lengths ranging from 0 to 3. This finding suggests that the NDVI response to changes in precipitation may be more delayed than responses to changes in temperature in the Great Plains. This may be due to more time required for water uptake by plants or percolation of water through soil to reach plant roots.

The average C_{NT} for the 305 stations was 0.73, with a maximum value of 0.95, which was statistically significant for the majority of weather stations in the Great Plains (290 of the 305 stations at 1% level of significance). Strong correlation was anticipated, since temperature serves as an indirect measure of the heat and energy available for plant development. As shown in Fig. 1, a distinct spatial distribution of C_{NT} resulted with low values observed in the south and maximum values occurring in the northern portion of the study area, suggesting that vegetation was more temperature limited in the north. Seasonal variations tend to be more pronounced in the north with shorter growing seasons, imposing limitations on plant growth. In contrast, temperatures in the south tend to be more consistent and milder throughout the year, with longer growing seasons.

The correlation plots essentially identified two separate zones of high and low vegetation-temperature dependence around the 37ºN mark. In other words, changes in vegetation, as reflected in NDVI, at northern latitudes were more highly correlated with deviations in temperature over the 19-year period. The finding also suggests that future changes in mean temperature or length of growing season may result in greater impacts on vegetation cover in northern latitudes of the Great Plains than in southern areas. For example, higher atmospheric temperatures due to global warming trends may induce stronger effects in the northern Great Plains than in the south, where the NDVI-temperature association was determined to be the weakest.

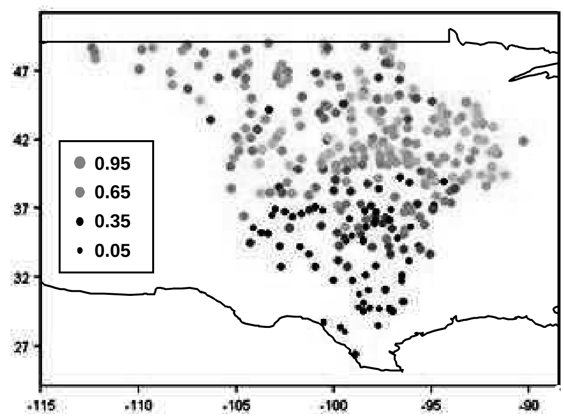

Figure 1. Map of NDVI/temperature correlation coefficients at 305 stations in the U.S. Great Plains. Axes denote latitude and longitude geographic coordinates.

In comparison, the average C_{NP} correlation was low at 0.38, although significant at the 1% level for most stations. The maximum correlation observed was 0.62, which was significantly lower than that of C_{NT}. Unlike the spatial distribution previously observed for C_{NT}, the north-south transition zone between low to high C_{NP} values was gradual and a clear separation zone could not be identified decisively (figure not shown here). Therefore, this finding suggests that temperature may exert a more direct impact on vegetation growth in comparison to the precipitation effect.

3.2 Principal Components Analysis

In order to further test the hypothesis of temperature as the dominant climate regime in the Great Plains, principal components analysis (PCA) was performed as a variable-reduction technique for exploring the structure of spatial variation in the NDVI data. Since the climate and NDVI datasets involved a large number of observed variables, it was useful to simplify the analysis by considering a smaller number of linear combinations of the original variables.

The spatial principal components for NDVI were computed from weather station data from 204 months, with years 1994 and 2000 excluded due to a significant amount of missing data. The optimal number of principal components (PC) to retain for accounting most of the variation in the dataset was determined by a scree test, which indicated that the first and second PC's were the optimal eigenvectors to retain. The first principal component (PC1) accounted for 43.1% of the variance in the NDVI dataset, while the second principal component (PC2) accounted for 30.2% of the total variance. Fig. 2 shows an interpolated surface plot of the principal component scores for PC1. A north-south trend in the map pattern of the spatial PC1 was strikingly similar to the spatial variation of mean temperature in the Great Plains. The pattern corresponded with the mean temperature north-south gradient, as well as exhibiting a steep transition zone at 37ºN latitude. This suggests that PC1 may be representative of temperature effects in the study area, accounting for the majority of the variation (i.e. 43.1%) in NDVI.

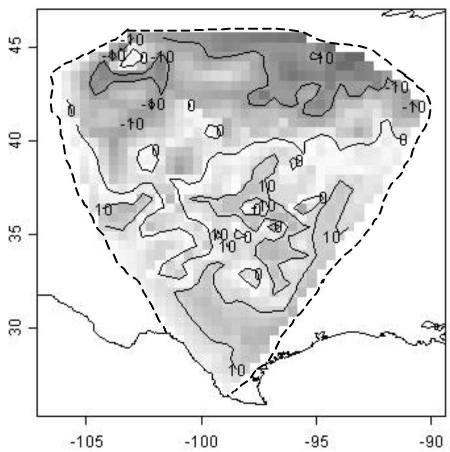

Figure 2. Interpolated surface plot of the first spatial principal component of NDVI.

In comparison, the surface plot of PC2 exhibited an east-west trend (figure not shown here), suggesting a similarity with the spatial distribution of the Great Plains precipitation regime. Since PC2 accounted for 30.2% of the overall variance, this suggests that precipitation exerts a smaller effect on vegetation than temperature. This result was consistent with the findings from the previous correlation analysis, further suggesting that temperature plays a more dominant role in the vegetation-climate relationship in the Great Plains area.

3.3 Climate Regionalization Scheme based on Seasonal Climate Cycles

Since the previous principal components analysis indicated that temperature was the dominant climate regime controlling the variation in vegetation cover in the Great Plains, the purpose of this section was to develop a climate regionalization scheme based primarily on temperature data. This analysis was focused on clustering temperature data for 1982 to 2000 and examining the NDVI-ecoclimatic relationships within each classified sub-region, in order to assess the adequacy of the resulting regionalization scheme.

In general, the rationale for cluster analysis is to find subgroups of data within a larger dataset, while maximizing the similarity of "within-cluster" observations and differences between clusters (Yarnal, 1993). The method involves dividing datapoints into groups of points that are considered "close" to each other. The clustering function continues to aggregate groups together until one large group remains. A necessary measure is to select the final number of clusters for the analysis. For this study, a simple method for deciding the optimal number of clusters to retain was used. The iterative process involved testing cluster sizes from 2 to 25 based on the temperature seasonal cycle, applying each cluster number to the temperature data, performing an analysis of variance (ANOVA), and retaining the residual error from each ANOVA test. By plotting the reduced residuals with rising cluster number, the optimal clusters were identified at the inflection point. In this analysis, the optimal cluster number determined for temperature seasonal data was four clusters.

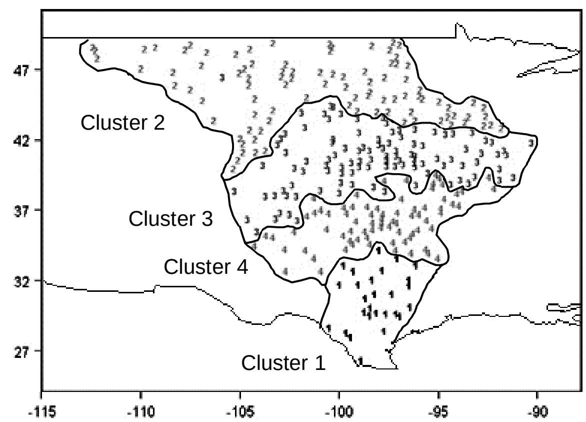

Figure 3. Geographical distribution of clusters classified by the temperature seasonal cycle. The four clusters are labeled and solid lines indicate boundaries of homogeneous clusters.

The temperature seasonal cycle was defined as the variation of monthly temperature over an eight-month period from March to October. Fig. 3 shows the result of clustering the temperature data only. The climate regionalization scheme with an optimal number of four clusters to retain resulted in distinct and coherent clusters with an anticipated north-south distribution. The distribution of clusters was broadly latitudinal with divisions occurring at 35ºN, 40ºN, and 45ºN.

Fig. 4 shows a monthly mean NDVI plot for all four clusters derived from the monthly temperature seasonal cycle. An inversion of the geographical distribution of mean NDVI was observed during the months from June to September, when lowest values were observed in southern clusters (Cluster 1) and highest values were observed in northern clusters (Clusters 2 and 3). The trend of decreasing NDVI during summer months for southern clusters may be indicative of drought conditions or temperatures exceeding limits of vegetation, which damage plant growth. However, the NDVI temporal trends for northern clusters were unaffected, exhibiting a characteristic symmetric bell-shaped distribution with a distinct green-up phase, reaching a maximum value in summer, and followed by a slope down due to senescence.

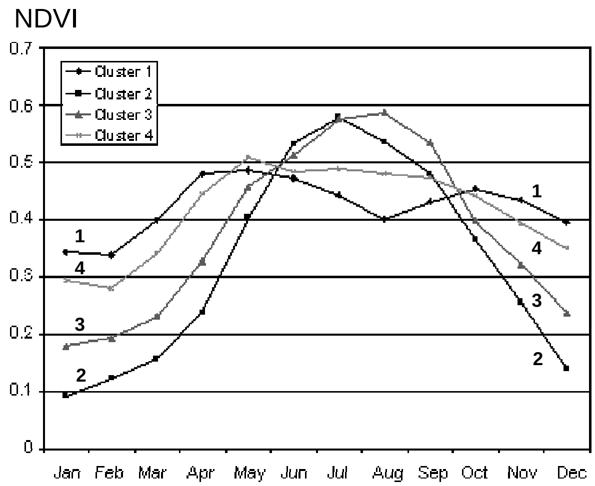

Figure 4. Monthly mean NDVI plot for the four clusters derived from temperature seasonal cycle data.

These findings suggested that the stations in clusters above 40ºN (i.e. Clusters 2 and 3) were marked with the presence of a stronger NDVI seasonal cycle. The NDVI response from stations occurring below this latitude tended to be more uniform throughout the year or actually showed a decreasing NDVI during summer months. This demonstrated the effectiveness of the regionalization scheme based on temperature alone in dividing the study area into regions with distinct NDVI characteristics.

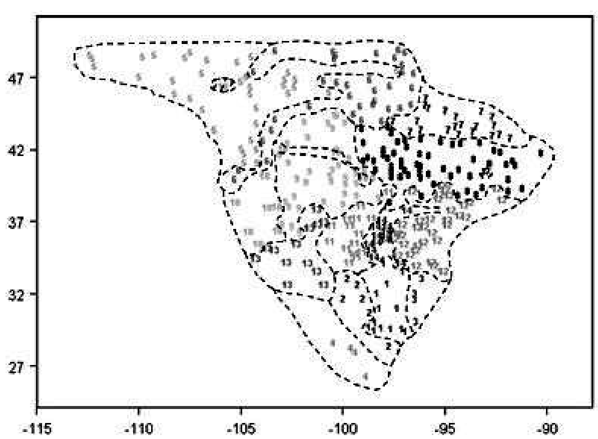

Figure 5. Geographical distribution of clusters with primary classification by temperature seasonal cycle and secondary classification based on precipitation seasonal cycle.

The climate regionalization scheme was further developed by secondary clustering using the precipitation data within each primary cluster formed from temperature data. Both stages of clustering were identical in terms of clustering technique and process of identifying optimal number of clusters to retain. Clusters derived from temperature data were subsequently clustered using a corresponding precipitation dataset with the results shown in Fig. 5. Resulting clusters were relatively compact and variable in size, although northern secondary clusters appeared to be more dispersed and mixed than southern clusters below 37ºN. The clustering technique employed was effective in producing a climate regionalization scheme based primarily on temperature as the dominant climatic factor affecting vegetation cover in the Great Plains.

4. CONCLUSIONS

This study focused on the critical interactions between climate and vegetation productivity in the highly dynamic system of the Great Plains. A variety of data sources were integrated, such as remotely sensed data and weather station records, covering a relatively long time period, particularly with respect to NDVI time series.

When considering monthly temperature and precipitation, the results of the analyses implemented in this study showed that interactions between climate variability and vegetation productivity in the Great Plains were significant and dynamic over time. Results from the statistical analyses employed supported temperature as the dominant climate regime in the Great Plains, in comparison to precipitation. The stratified climate regionalization scheme developed in this study based on primary clustering of temperature and secondary clustering of precipitation data facilitated the comparison of corresponding NDVI values.

However, it is recognised that considerable potential exists in incorporating other ecoclimatological factors, which could further account for variability in vegetation activity in this region. Additional research is also needed to model such relationships in different geographic locations to test consistency of results. Extending conclusions drawn from local or regional studies to other regions at a variety of spatial scales could lead to developing a more sensitive measure of global climatic change impacts and predicting the future distribution of vegetation in response to such impacts.

REFERENCES

Goward, S.N., “Satellite bioclimatology,” J Climate, vol 2, p.p. 710-720, 1989.

Martin, P., “Vegetation responses and feedbacks to climate: a review of models and processes,” Clim Dynam, vol 8, p.p. 201-210, 1993.

Reiners, W.A., Great Plains Climates and Biota: Past, Present, Future, Kluwer Publishers, Boston, USA, p.p. 11-32, 1995.

Sellers, P.J., “Canopy reflectance, photosynthesis, and transpiration,” Int J Remote Sens, vol 6, p.p. 1335-1371, 1985.

Tucker, C.J., and P.J. Sellers, “Satellite remote sensing of primary production,” Int J Remote Sens, vol 22, p.p. 3827-3844, 1986.

Wang, J., K.P. Price, and P.M. Rich, “Spatial patterns of NDVI in response to precipitation and temperature in the central Great Plains,” Int J Remote Sens, vol 22, p.p. 3827-3844, 2001.

Yarnal, B., Synoptic Climatology in Environmental Analysis, Bellhaven Press, London, UK, p. 195, 1993.