2006

Correlation between Pollen Dispersion and Forest Spatial Distribution Patterns in the Southeastern United States

I-Kuai Hung
Arthur Temple College of Forestry and Agriculture, Stephen F. Austin State University, hungi@sfasu.edu

CORRELATION BETWEEN POLLEN DISPERSION AND FOREST SPATIAL DISTRIBUTION PATTERNS IN THE SOUTHEASTERN UNITED STATES

Peter P. Siska
School of Agriculture and Geosciences, Austin Peay State University, TN 37044

I-Kuai Hung
College of Forestry, Stephen F. Austin University, TX 75962

Vaughn M. Bryant Jr
Department of Anthropology, Texas A&M University, TX 77843

ABSTRACT

The pollen that falls to the surface at any given point is called the pollen rain. For most regions of the world the pollen rain provides a fairly reliable record of the plants that produce and disperse airborne pollen within a radius of about 30 km from the sampled location. To some extent the local pollen rain can also reflect limited information about the insect-pollinated plants living in a region. For some regions of North America, existing studies of the pollen rain and the regional vegetation associated with those data demonstrate a reliable relationship between these two vegetation aspects. For other regions of North America pollen rain studies exist but they have not been linked or correlated with the regional vegetation. In many others areas of North America there are no existing pollen rain studies.

One objective of this project is to develop a method using geographic information systems to correlate existing pollen rain data with remote sensing based on classified vegetation patterns, especially in the forested biomes of North America. In addition, spatial interpolation methods will be used in GIS to predict the pollen rain in other regions where remote sensing data is available but no pollen rain data currently exist. Once completed, these correlations can be used to produce actual and projected pollen rain distributions for many regions of North America. Understanding the relationships between pollen rain data and the vegetation biomes they represent will then enable researchers and practitioners to use existing fossil pollen records to map past environmental changes in forested regions of North America and to predict future global changes of the biosphere. A secondary benefit of this research is that it will provide actual and projected pollen rain maps for North America. Those maps will permit law enforcement agencies to use pollen as a geographical marker and powerful forensic tool in their effort to solve crimes and catch potential terrorists before they can commit violent acts of destruction.

In Prisley, S., P. Bettinger, I-K. Hung, and J. Kushla, eds. 2006. Proceedings of the 5th Southern Forestry and Natural Resources GIS Conference, June 12-14, 2006, Asheville, NC. Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA.

KEYWORDS. Pollen rain, biomes, spatial analysis, splines, zonal statistics, oak and pine pollen, correlation.

INTRODUCTION

Geographic Information Science (GIS) is one of the most rapidly developing fields in the natural sciences with an extremely broad spectrum of applications in related fields such as the social sciences, and engineering. During the last decade GIS applications have begun to be applied to the study of palynology. Plant pollen plays a critical role in fertilization and genetic transfer; therefore adequate pollen distribution plays an important role in maintaining healthy, natural ecosystems. Pollen and the nectar in flowers also play important roles because they become a source of food for many insects and some bird and mammal species. The spatial distribution of pollen can be captured plotted using GIS and a significant body of surface pollen distributional data has recently been compiled and is now available (Whitmore et al., 2005). These surface pollen distributional data (also known as the *pollen rain*) provide valuable clues about the local and regional climate and vegetation. In regions where favorable soil and microbial environments exist, pollen from gymnosperms and angiosperms, combined with dispersed spores from local cryptogrammic plants can be preserved for thousands or millions of years and becomes one of our most important tools for reconstructing paleoenvironments.

Pollen grains and spores, are spatial variables and their geographic distribution has be the focus of study by palynologists (pollen analysts), plant geographers, botanists, geologists, ecologists, and entomologists since the emergence of the field of pollen analysis in 1916. Beginning in the 1960s a number of palynologists began using surface and later subsurface fossil pollen data to produce regional isopol maps showing the known and projected spatial variations of pollen distribution. As more pollen data became available, new isopol maps plotting the distribution of different pollen taxa began being published for different regions, nevertheless because most available pollen data was available for areas of Eastern North America, that region received the greatest detail of study and emphasis (McAndrews and Webb III 1976). One of the first books on this topic was the one by Delcourt et al., (1984) that produced a large number of maps using summaries of pollen data. A number of research papers began focusing on various types of computer simulations and programs for modeling pollen dispersal and pollen accumulation rates over the landscape and their relationship to changes in vegetation and climate. A few examples of these include the efforts by Broström et al. (1999) and Jackson and Lyford (1999). One of the first efforts to apply GIS techniques to pollen distribution patterns was the attempt by Siska et al. (2001). In that effort, the geostatistical modules in GIS were used to generate patterns of predicted pollen deposition over a large region of the Big Bend National Park in west Texas. The effort to construct a pollen distribution atlas using GIS techniques for *Ficus* and a few related species was recently published by Boyd and Jago (2003).

Traditionally one of the most important uses and applications of pollen analysis is in the reconstructing past vegetation patterns with the intent of using those data to infer past climatic changes. This has been most frequently applied to the time period of the late Quaternary, especially the time around the end of the last major glacial period and the following Holocene period (Bartlein at al., 2004). Because vegetation composition thrives best in the ecological zones most favorable to their needs, and because we can assume that evolutional changes in plant taxa are relative slow, we are able to use fossil pollen assemblages as evidence of past vegetational dynamics and by inference climate.

Geographic Information Systems' capabilities support the input and analysis of multilevel spatial information and therefore are useful for analyzing the special relationships between pollen distribution and vegetational composition. For example, Fyfe (2005) used coarse resolution of digital elevation models to determine the influence of topographic relief on pollen distribution and deposition using landscape simulations. The site analysis of this example, based on relief properties, appeared to play a significant role in the reconstruction of landscapes and for testing it against archeological and paleoecological theories.

Data and Objectives

The purpose of this project is to study the spatial relationship between two important pollen taxa (*Pinus* and *Quercus*) and recent forest patterns in the southeastern United States. During the last several centuries the forest ecosystem has been severely affected by anthropologic activities that resulted in the significant fragmentation of the forest and consequent changes in gene flow patterns. This has also been documented in other parts of the United States (Apsit at al., 2002). The trees that are wind-pollinated are an important component of the natural ecosystem in this region and maintaining the genetic variability of individual plants is important for the survival of the entire biomes. Destruction of continuous vegetation cover and fragmentation of forests contribute to areas of isolation and genetic restrictions, which are caused by inbreeding and can lead to genetic drift and restricted gene flow (White at al., 1999, Boshier at al., 2002). Therefore, a major purpose of this research is to evaluate the efficiency of forest pollen distribution by looking at the spatial relationships between forest vegetation fragmentation and forest pollen distribution.

The collection of these data from the forests in the southeastern part of United States can then be expressed on maps indicating the maximum and minimum amount of pollen distribution and accumulation. Figure 1 shows the study area as the shaded region. The dots represent existing pollen distributional data for North America, including the distribution of airborne pollen types in the study region. Another aspect of our research will be to develop a method, based on GIS that would evaluate the co-regionalization of pollen distributional patterns with forest vegetation patterns. In this respect, our planned work will be similar to the research of Jackson and Williams (2003) in their analysis and comparison of modern vegetation gradients (derived from remote sensing data) with pollen distributional patterns in North America.

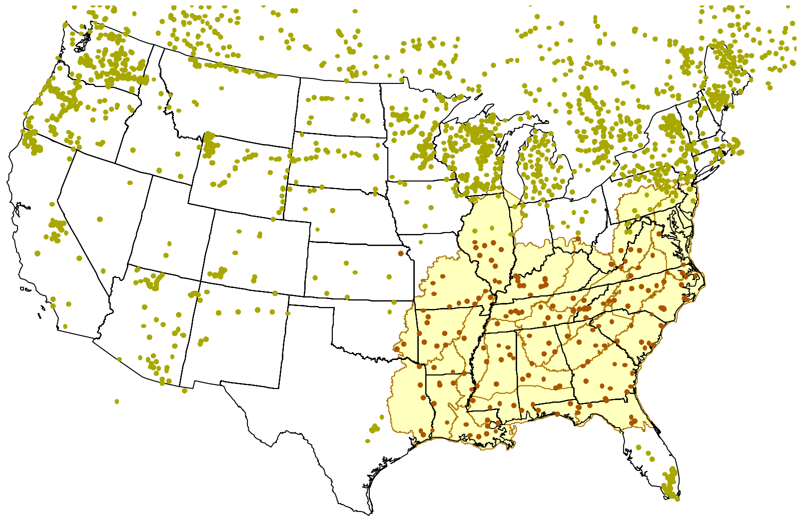

Figure 1. Study area and distribution of pollen samples.

The data used by Jackson and Williams (2003) were acquired from the North American and Greenland database compiled and developed by Whitmore et al., (2005). This database contains 4,634 pollen distributional records, derived from the analysis of surface samples for 134 plant taxa found in the flora of North America and the ice-free regions of Greenland. The total number of pollen grains examined in the 4,634 samples was 2,231,415. Besides recording pollen grain counts, this database also contains many other attributes of spatial relevance including bioclimatic data, broadleaf and needle leaf forest percentage cover and more.

For this study we imported the data into ArcGIS 9.3 and further analyzed it. Data on the forest cover and forest type were retrieved from the National Atlas website (www.nationalatlas.gov). These data are available in geo-referenced raster format with a spatial resolution of one kilometer. Figure 2 shows the distribution of major classes of general forest cover as well as water and non forest lands in the southern United States.

Figure 2 was produced using Advanced Very High Resolution Radiometer (AVHRR) composite images and Landsat Thematic Mapper (TM) data from the 1991 growing season (Zhu and Evans, 1994). For the purpose of this project the AVHRR data were reclassified for the southeastern part of the United States into six categories including pine, oak, elm-ash-cottonwood, spruce-fir, lakes and non-forested area. The mesophytic broadleaf deciduous forest is located mostly in the northern part of this region followed by a zone of oak-pine found mostly in piedmont region, areas of northern Alabama, Georgia and the Carolinas. The southern evergreen forest, which is dominated by *Pinus palustris, P. taeda, P. echinata* and *P. elliottii* and is accompanied by oak (*Quercus*) hickory (*Carya*), magnolia (*Magnolia*), sweetgum (*Liquidambar*), ash (*Fraxinus*) and maple (*Acer*) occupies the Deep South all the way to the Gulf of Mexico.

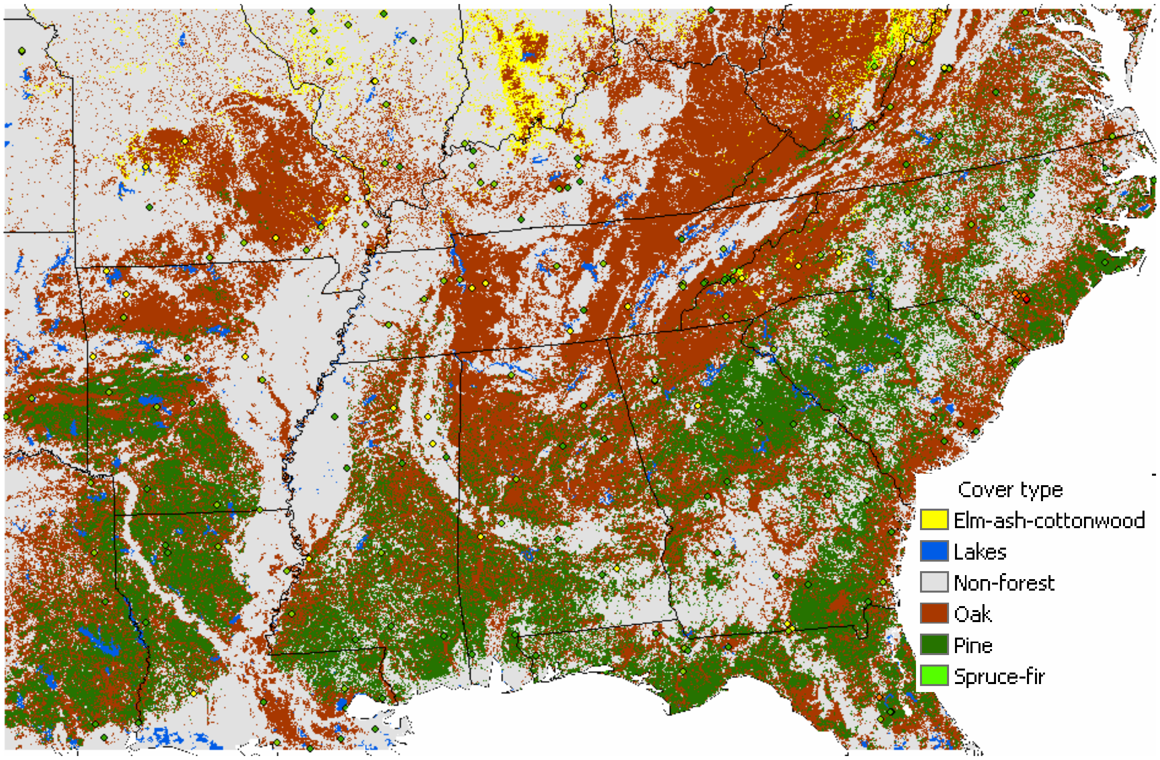

Figure 2. Distributions of forest cover in the southeastern US.

METHODS AND RESULTS

From the data set of 4,634 pollen samples collected for the North American and Greenland project, a subset the 204 samples were selected for our use in this project. All 204 samples are located within the study area of the southeastern region of the United States, which we plan to examine. The pollen data reveal that the distributions of pine and oak pollen frequencies are the most abundant in the southeastern region (a total of 31,792 pollen grains from pines versus 22,078 pollen grains for oak). It is important to note that for the purpose of this study the pine pollen was combined and represents the composite of diploxylon, haploxylon and other pine pollen grains that could not be assigned to either of the two major morphological groups. However, in terms of the entire North American data set, 23.14 % of all the oak pollen occurs in the southeastern US region yet only 6.7% of all the North American data set pine pollen accumulated in the southeastern US study area.

There are other major airborne pollen taxa fir (*Abies*), birch (*Betula*), beech (*Fagus*), spruce (*Picea*), (*Quercus*), hemlock (*Tsuga*), and grasses (Poaceae) that have overall lower pollen distribution percentages in the southeastern region than the pine pollen (8.76%). *Picea* is a prominent pollen type in some regions of the North American yet in the study area *Picea* pollen is insignificant and is reflected by only approximately 0.03 % of its total in the data base. Another example of low representation in the southeastern region is *Fagus*, which is represented by less than 2% of its total. In terms of covariance values (i.e., variation of pollen amounts of different pollen taxa across the North American continent and the study area) Pearson correlation coefficients determined that the correlation between the different pollen taxa in the southeastern part of United States maintain the same values, with minor differences, as the one in the entire North American continent. The strongest correlation in the southeastern region, as determined by Pearson correlation coefficients is between *Betula* and *Tsuga* pollen r = 0.72 (i.e. an increase or decrease of *Betula* pollen is accompanied also by an increase or decrease of *Tsuga* pollen in the southeast region). The relationship between pine and oak pollen is overall very weak (r = 0.04) in the southeast. On the other hand, oak and beech pollen correlation is significantly stronger (r = 0.36) (i.e., the amount of beech pollen is related to the oak pollen in the study area). Based on environmental factors the Pearson values indicate that both hemlock and birch are strongly influenced by elevation. The correlation value of 0.8 for birch indicates that in the southeastern region elevations are very important factor in explaining pollen distribution and by inference vegetation. As the relief becomes more diverse and overall elevation increases the birch pollen becomes more abundant. This applies also for hemlock where its correlation with elevation reaches 0.67.

The point-sampled data represent an isolated form of information that is difficult to interpret spatially. This includes pollen data that are usually collected from point locations and then analyzed in a laboratory. To evaluate spatial distributions of natural phenomena, and its relationship to other environmental factors, discrete forms of information must be converted into a continuous surface by using sophisticated interpolation algorithms such as the inverse distance method, and the family of best linear unbiased estimators such as kriging or splines. For this project completely regularized spline interpolator was selected to model the distribution of pollen taxa (pine and oak) in the studied area. The use of splines has proven effective for the study of large areas (Burrough and McDonnell 1998). The regularized splines are similar to thin plate spline but the basis function has a different form:

$$\frac{1}{2\pi}\left\{\frac{d^2}{4}\left[\ln\left(\frac{d}{\tau}\right)+c-1\right]+\tau^2\left[K_0\left(\frac{d}{\tau}\right)+c+\ln\left(\frac{d}{2\pi}\right)\right]\right\}$$

where τ is the weight, d is the distance between the sample point and the point being interpolated, c = 0.577215, and $K_0(d/\tau)$ is the Bessel function that can be approximated by a polynomial. Due to the piece-wise nature of spline functions the points in the local neighborhood are used for interpolation and therefore the spatial locations between sample points are close to the sample values provided so that measurement errors are small. The “smoothness” characteristics of splines guarantee that mathematical derivatives can be easy computed for the direct analysis of surface geometry and topology (Kang-tsung 2004).

The continuous surfaces of the two major pollen taxa (pine and oak) are represented in Figures 3 and 4. The results indicate that oak and pine pollen have distinctly different distribution patterns. The pine pollen tends to accumulate near the coastal areas. There are two major concentrations in the southeastern region. The first concentration is in eastern and southeastern areas of North Carolina and the northeastern part on South Carolina approximately between the cities of Raleigh and Columbia moving towards the Atlantic coast. As the prediction map indicates, the total pine pollen present in the counts

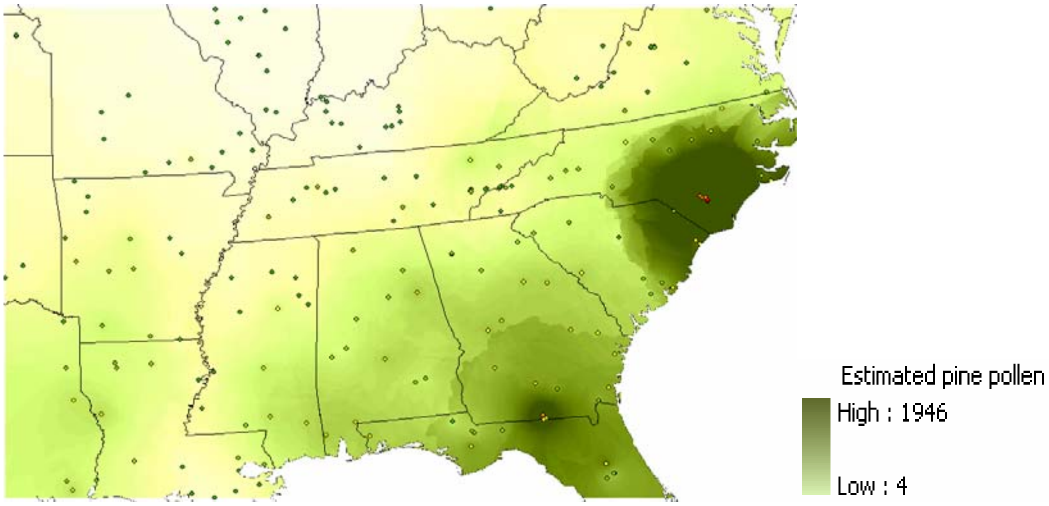

Figure 3. Spatial distribution of pine pollen in the southeastern US.

from the combined surface samples reach over 2000. The second important patch of pine pollen density is located on the boundary between Georgia and Florida; however, the extend here is not as pronounced as the previous one. The density of pine pollen gradually decreases towards Tennessee and to the north and west directions, except for the southern and westernmost regions of the Southeastern US in Louisiana and east Texas. In both of those areas the amounts of pine pollen slightly rises.

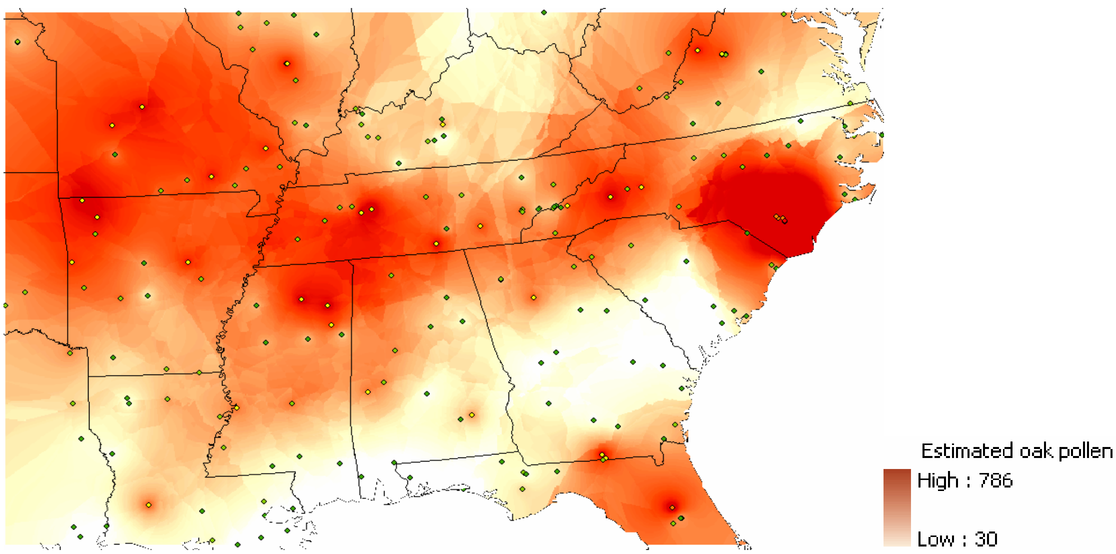

Figure 4. Spatial distribution of oak pollen in the southeastern US.

The spatial pattern of oak pollen in the coastal area follows a similar pattern. Similar to the distribution of pine pollen, oak pollen frequencies have two distinct high concentration values in North and South Carolina and a weaker area between southern Georgia and northern Florida. These two extremely large concentrations (up to a total of 800 oak pollen grains from the combined counts) are separated by low concentrations where the pollen counts of oak sink below 20 grains. Contrary to previous pine pollen distributions, oak pollen counts significantly increase towards the west and north portion of the region in areas of middle and western Tennessee, the northwest areas of Alabama, the northeast region of Mississippi, northern Arkansas, and the southern and Middle Missouri regions.

SYNTHESIS

Next, we want to compare pollen distribution patterns with the spatial distribution of oak and pine forests (Figure 2). The map in Figure 2 was constructed from the National Land-Cover Database (Homer at al., 2004) and was reclassified for the purpose of this project. Later this map was overlaid with spline-interpolated pollen distribution patterns to produce composite synthesis of vegetation and pollen distributions (Figures 5 and 6). As one can see from these maps, the distribution of major pollen taxa and the corresponding mother-forests covariate only to a limited degree. For example, the highest concentration of pine pollen described previously does not correspond well with known pine forests in this region, except for the southern coastal area of North Carolina. On the other hand, the continuous pine forests in central and western South Carolina and Georgia do coincide with higher pine surface pollen concentrations. A similar correlation is evident for the Gulf Coast region in the south. We were surprised to see that the thick, continuous pine forests west of the Mississippi River in Louisiana and in the areas of east Texas are represented by very low occurrences of pine pollen counts (Figure 5).

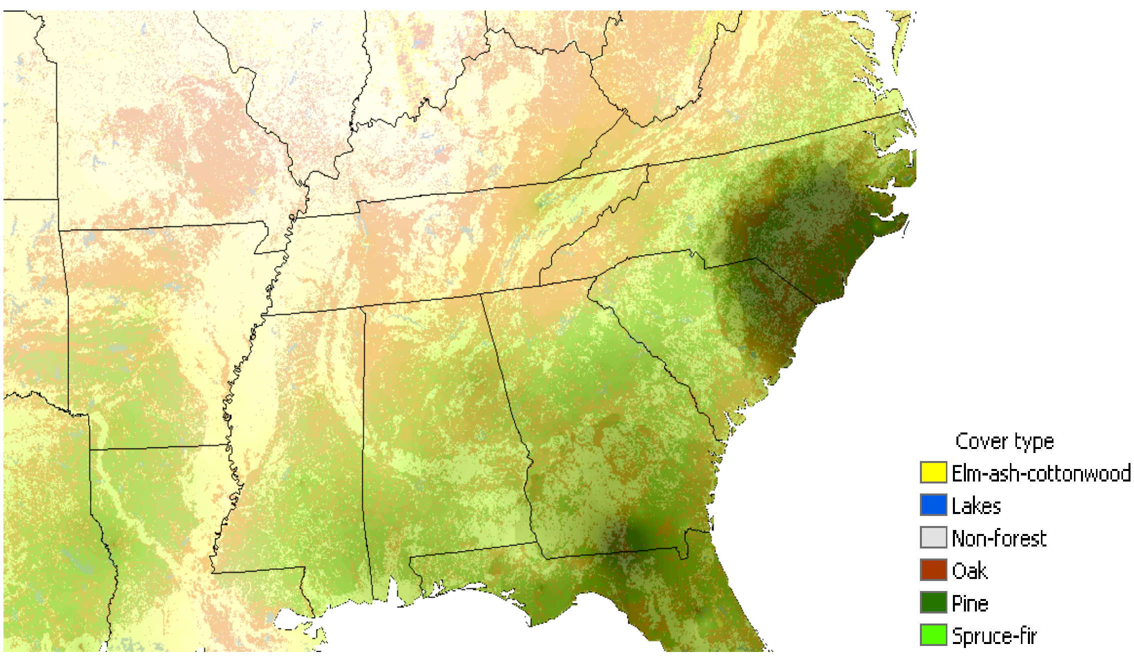

Figure 5. Spatial synthesis of pine pollen and forest cover.

Unlike pine pollen distributions, the higher frequencies of oak pollen in North Carolina do not coincide with the distribution of oak forests in that same region. This also applies to the region of northern Florida. On the other hand, the medium levels of oak pollen concentrations in the Appalachian Mountains correspond well with the major southwest to northeast orientation and clearly follows the Appalachian structure. The oak forest towards the west from the southern part of Appalachian Mountains becomes more and more fragmented, especially in the Mississippi valley region (i.e., west Tennessee and the eastern part of Arkansas). The oak pollen concentrations, however, continues high in the east to west direction with high amounts of surface pollen (over 700 grains in the composite samples).

The southeast area of the United States, especially areas in the east and central south portions of South Carolina, Georgia, Alabama, Mississippi and Louisiana reflect extremely low amount of oak pollen, which coincides with the area of pine dominated forests and non-forested areas. There are, however, small patches of oak forests scattered in this area that would potentially benefit from higher oak pollen flow.

<u>Zonal Statistics</u>
In an effort to determine the spatial relationship between the frequencies of the two pollen taxa and the spatial co-distribution of forest vegetation, we selected the zonal statistics function in Spatial Analyst ArcGIS module to capture the quantities of pine and oak pollen taxa in each forest types. In addition, non-forested areas and water (lakes) were also included in this analysis. As the results indicate, the majority of pine pollen is located within the areas of pine forests. The second largest forested area in the southeastern United States receiving some pine pollen are the oak forests. The small water bodies, mostly lakes, are the third largest reservoir of pine pollen with an estimated total of over 100 pollen grains in the combined counts. A lot of the pine pollen is deposited in non-forested areas, which are mostly urbanized regions in the southeastern area. Finally, the smallest amount of pine pollen is deposited in the areas dominated by spruce-fir forest and the elm-ash-cottonwood forest biome. The composite map of pine pollen and vegetation shows the overlapping areas between these two components, which were described previously. In addition, the areas of strongest forest destruction and fragmentation can be compared and evaluated with respect to amount of pollen that each vegetation type receives during the spring pollination period. This is one of the useful features of our analysis because it can suggest gaps in the pollination efficiency of forest genera, which could affect the successful transport of genetic material between isolated forest patches.

The zonal statistics of oak pollen vs. oak land cover indicates a different pattern of spatial match-mismatch (Figure 7). Based on the pollen data and spatial analysis of the oak pollen distribution over the area of forests and other land cover categories seems to be more uniform than some other tree taxa pollen types. Oak pollen distribution is fairly high in the areas assigned to spruce-fir forests, which is mostly a factor of long distance transport of the oak pollen from nearby forested areas. The areas in the Southeast that are mostly pine forests receives the lowest amounts of oak pollen, which is to be expected because pine trees produce and disperse far greater amounts of pollen than do oak trees. In addition, the total amount of pine pollen produced in normal pine forested regions will total dominate any other pollen type dispersed either in the pine forest or from nearby vegetation outside the pine forested areas.

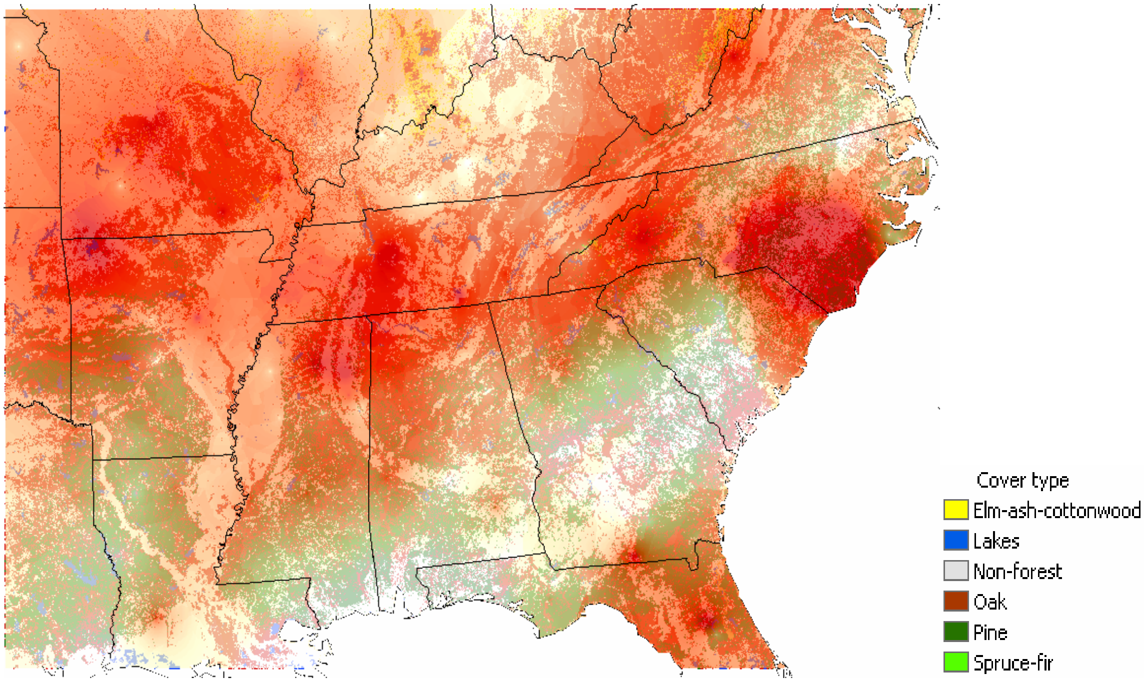

Figure 6. Spatial synthesis of oak pollen and forest cover.

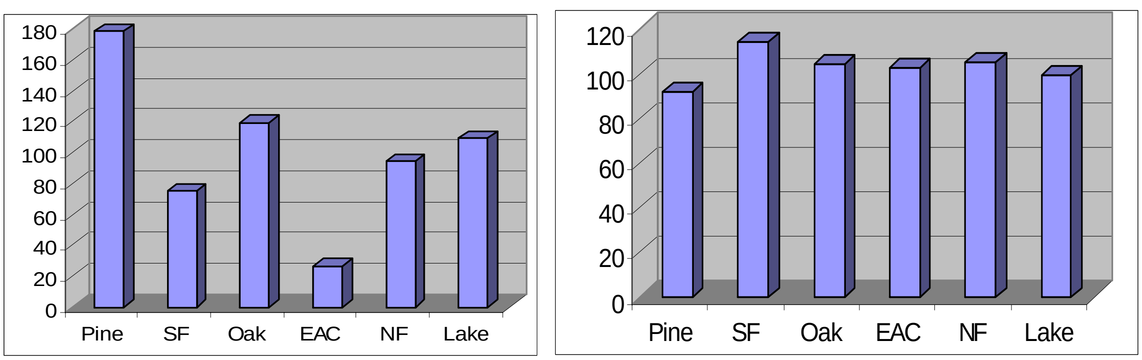

Figure 7. Comparisons of pine (left) and oak (right) zonal statistics. Vertical axis determines the mean estimated pine pollen in corresponding land cover categories (SF: spruce-fir, EAC: elm-ash-cottonwood and NF: is non-forested).

<u>New Method for Forest Cover and Pollen Dispersion Interpretation</u>

The previous analysis provides a general evaluation of the spatial co-behavior of pollen taxa with the plant sources of dispersion and the various categories of land cover. In an effort to refine the spatial covariation between these two components of an ecosystem, we are beginning a project to address this issue with the help of GIS. The new method is based on geometric spatial units (Dirichlet-Voronoi polygons) generated around each pollen sample value and the interpolated estimated values of pollen that would fill each of the polygons. The biometric values, such as the percent of cover and imperviousness will be extracted from a NLCD database. Linear regression will then be used to develop a linear regression model for pollen values as a function of vegetation parameters.

CONCLUSIONS

Pollen dispersion and vegetation are both important components of a natural ecosystem. Healthy forested areas in the southeastern US are composed of many wind-pollinating species that produce large amounts of pollen grains. These high production amounts combined with normal wind dispersion methods guarantee effective pollination and the production of seeds with high genetic diversity. During the last two centuries human activities have severely disrupted the natural ecosystem in many regions of North America, including the southeastern region. The results of these modifications have been the fragmentation of continuous forests and the formation of forest patches in many areas. The reduction of vegetation cover also increases the amount of carbon dioxide that is released into the atmosphere, which contributes to global warming (Murty at al., 2002, Lal, 2001).

There is still a large amount of forested areas in the southeastern part of the United States, as determined from satellite images. Those images reveal over 550,000 square kilometers of oak forest (an area equivalent to the country of France) and almost 300,000 square kilometers of pine forest still remaining in southeastern part of United States. Those forested regions, however, are rapidly becoming more fragmented. More disturbing is that the current image analyses indicate that the non-forested areas in the southeastern U.S. now cover the largest portion in the region (678,000 sq. km). When examined together, these two aspects (i.e., remaining forest cover vs. expanding areas of non-forested areas) make the study of pollen dispersal and distribution, the determination of genetic plasticity within forest patches, and the criteria that will keep the forested ecosystems healthy in the southeastern U.S. become studies of critical importance. One of the first essential tasks to address this important issue is to develop effective methods of mapping the variables of pollen distribution vs. forest cover. The production of reliable maps of this type will help those in the lumber business, those wanting to protect areas of old growth forests, and researchers in forestry plan effectively for the future.

REFERENCES

Apsit, V.J., F.W. Davis, R.J. Dyer, M. Fernandez, B. Kuhn, P.E. Smouse, and V.L. Sork. 2002. Pollen movement in declining populations of California Valley oak, Quercus lobata: Where have all the fathers gone? *Molecular Ecology*. 11: 1657.

Bartlein, P.J., P.L.Leduc, B.N. Shuman, T. Webb III, and J.W. Williams. 2004. Late-quaternary vegetation dynamics in North America: Scaling from taxa to biomes. *Ecological Monographs*. 74: 309-334.

Boshier, D.H., W. Powell, and G.M. White. 2002. Increased pollent flow counteracts fragmentation in a tropical dry forest: An example from Swietenia humilis Zuccarini. *PNAS*. 99: 2038-2042.

Boyd, W.E. and L.C. Jago. 2003. A GIS atlas of the fossil pollen and modern records of ficus and related species for island south-east Asia, Australasia and the Western Pacific. *Australian Geographical Studies*. 41: 58-72.

Broström, M. Gaillard, and S. Sugita. 1999. Landscape openness and pollen records: A simulation approach. *The Holocene*. 9: 409-421.

Burrough, P.A. and R.A. McDonnell. 1998. *Principles of Geographical Information Systems*. New York: Oxford University Press.

Homer, C., C. Huang, L.B. Yang, B. Wylie.and. M. Coan. 2004. Development of a 2001 national land-cover database for the United States. *Photogrammetric Engineering & Remote Sensing*. 70: 829-840.

Delcourt, H.R., P.A. Delcourt, and T. Webb III, 1984. Atlas of mapped distributions of dominance and modern pollen percentages for important tree taxa of eastern North America. *The American Association of Stratigraphic Palynologists*. 14: 131.

Fyfe, R. 2005. GIS and the application of a model pollen deposition and dispersal: A new approach to testing landscape hypotheses using the POLLANDCAL models. *Journal of Archaeological Science*. 20: 1-11.

Jackson, S.T. and M.E. Lyford. 1999. Pollen dispersal models in Quaternary plant ecology: Assumptions, parameters, and prescriptions. *Botanical Review*. 65: 39-75.

Jackson, S.T. and J.W. Williams. 2003. Palynological and AVHRR observations of modern vegetational gradients in eastern North America. *The Holocene*. 13: 485-497.

Kang-tsung, C. 2004. *Introduction to Geographic Information Systems*. 2nd ed. New York: McGraw-Hill.

Lal, R. 2002. Soil carbon dynamics in cropland and rangeland. *Environmental Pollution.* 116: 353-362.

McAndrews, J.H. and T. Webb III. 1976. Corresponding patterns of contemporary pollen and vegetation in central North America. *Geological Society of America Memoir*. 145: 267-299.

Murty, D., M. F. Kirschbaum, R.E. McMurtrie, and H. McGilvray. 2002. Does conversion of forest to agricultural land change soil carbon and nitrogen? a review of the literature. *Global Change Biology.* 8: 105-123.

Siska, P.P., V.M. Bryant, Jr., and J.G. Jones. 2001. The spatial analysis of modern pollen rain in Big Bend National Park. *Palynology.* 25: 199-216.

Zhu, Z. and D.L. Evans., 1994. U.S. Forest types and predicted percent forest cover from AVHRR Data. *Photogrammetric Engineering and Remote Sensing.* 60(5): 525-531.

White G. M., D.H. Boshier, and W. Powell. 1999. Genetic variation within a fragmented population of Swietenia humilis Zucc. *Molecular Ecology.* Vo. 8: 1899 -1905.

Whitmore, J., K. Gajewski, M. Sawada, J.W. Williams, B. Shuman, P.J. Bartlein, T. Minckley, A.E. Viau, T. Webb III, S. Shafer, P. Anderson, L. Brubaker. 2005. Modern Pollen Data from North American and Greenland for Multi-scale Paleoenvironmental Applications. *Quaternary Science Reviews* 24: 1828-1848.