ABSTRACT

SODIYA OLAKUNLE EMMANUEL. Spatial Analysis of the Forest Products Industry in North Carolina. (Under the direction of Dr. Rajan Parajuli, and Dr. Bob Abt).

The forest products industry in North Carolina (NC) is the one of the leading manufacturing sectors, contributing about $34 billion to the state economy annually. The housing market crash in 2006 followed by the 2008 great financial crisis distorted the entire forest product industry in North Carolina, resulting in a decline in the numbers of mills and shifts in market demand toward new wood-based energy and biomaterials. This study explored the current spatial distribution of forest products manufacturers, and evaluated the factors influencing their choice of location in various parts of NC. Based on the Getis-Ord Gi* spatial statistical Optimized Hot Spot analysis and Inverse Distance Weighted (IDW) interpolation analysis, this study examined the spatial distribution of primary- and secondary forest products manufacturers, forest resources, and identified major hot spots. A count data model was employed to evaluate the factors influencing the location of forest products manufacturers in NC.

Results suggested that primary forest products manufacturers are clustered in proximity to the source of raw materials, particularly where there is the abundance of major species type used for their production. The primary hardwood manufacturers are found to cluster in western NC near the hardwood resources hot spot, while the primary softwood manufacturers are clustered in the eastern part of NC close to the softwood resources hot spot. However, most secondary forest products manufacturers were found not spatially tied to resource locations relative to primary forest products manufacturers, rather they clustered close to markets of their final goods, mainly around the piedmont region.

Regression results indicated that the presence of existing secondary manufacturers, cities with high population, availability of raw materials, and higher economic tiers positively impacted the location of primary forest products manufacturers in a county. Similarly, the presence of existing primary mills, cities with high population, high volume of sawlog harvests, availability of labor, and higher economic tiers positively influenced the location of secondary forest products manufacturers in a county. However, the presence of interstate highways, railroad, or port in a county did not significantly increase the likelihood of attracting either primary or secondary forest product manufacturers. Compared to secondary manufacturers, availability of labor did not significantly influence the location of primary forest products manufacturers in a certain county. The findings from this study serve as guidelines to policy and decision makers, and all stakeholders involved in managing forests and forest products manufacturing industries to foster cluster based economic development in NC.

Spatial Analysis of the Forest Products Industry in North Carolina

by
Olakunle Emmanuel Sodiya

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Forestry and Environmental Resources

Raleigh, North Carolina

2021

APPROVED BY:

Dr. Rajan Parajuli
Co-Chair of Advisory Committee

Dr. Bob Abt
Co-Chair of Advisory Committee

Dr. Josh Gray

DEDICATION

This master's thesis is dedicated to my dearest father and mother, Mr. & Mrs. Israel Oludotun and Motunrayo Janet Sodiya. Thank you for praying and supporting this dream into manifestation. Thank you for always putting your children's needs first. I hope I can live up to your expectations. My appreciation to you is far beyond words.

BIOGRAPHY

Olakunle Emmanuel Sodiya was born and raised in Abeokuta, the capital of Ogun State, in the southwestern region of Nigeria. He obtained his bachelor's degree in Forestry and Wildlife Management in 2014 from the Federal University of Agriculture, Abeokuta, Ogun, Nigeria. His desire to solve environmental related problems especially in the forest and forest product industry in Nigeria and across Africa, inspired him to garner more knowledge to function in this capacity. He proceeded to West Virginia University, Morgantown, WV in 2017 to study forestry at master's level but ended up studying Resource Economics and Management where he picked up much interest in economics and geographical information system (GIS) analysis. In 2019, Olakunle transferred to North Carolina State University to complete his master's degree where he got an opportunity to work on the project, "Wood Supply Assessment in North Carolina: Examining Market Dynamics, Resource Availability, and Sustainability" under the supervision of Drs. Rajan Parajuli and Robert Abt. Ultimately, he worked on the "Spatial Analysis of Forest Products Industry in North Carolina" for his master's thesis.

ACKNOWLEDGMENTS

To God be all the glory for the success of this research.

I would like to express my gratitude to my advisor, Dr. Rajan Parajuli for accepting me into his research lab and guiding me throughout this research. I immensely benefited from his wealth of knowledge and experience. Thank you for your unreserved support and encouragement at every step of the way. I also cherish the privilege to interact with Professor Emeritus, Bob Abt. Every moment spent with you was inspiring. Thank you for your deep and critical perspective about issues regarding this research. Your help with sourcing for data cannot be overemphasized, without which this research would not have been better, for this I am thankful. My encounter with Dr. Josh Gray was phenomenal. His willingness to serve on my committee and co-supervise was astounding. Thank you for your recommendations, ideas, and support toward the completion of the spatial analysis aspect of this research.

I would also like to thank the North Carolina Forest Service (NCFS) for providing my funding in the form of a research assistantship and for sponsoring this research project. I would like to specially thank Barry New and Robert Ross for providing an updated primary forest products manufacturers' data used in this research. Without these financial and data support, I would not have been able to complete my degree.

My acknowledgement also goes to the North Carolina State University data and visualization Librarian, Jeff Essic. Thank you for making time to meet via zoom to review the spatial analysis aspect of this research. Your contribution to the success of this research is invaluable.

To all the faculty and staff of the Department of Forestry and Environmental Resources that have made my stay so far smooth and worthwhile; Sarah Slover, Sam Cook, AnneMargaret Braham, Dr. Zakiya Leggett, Dr. Stacy Nelson, Dr. Myron Floyd, Dr. Erin Sills, to mention a few, I am very grateful. To Kenneth Dunn and other friends, whom I met in the course of my program who have made my stay at NC State memorable, I am thankful.

Finally, to the people I hold dearest to my heart, whose prayers and support keeps me going; Mr. & Mrs. Israel O. Sodiya, my siblings (Ayodeji, Mabayomije, Oluwapelumi, and Oluwakanyinsola), Mr & Mrs. Ikuomola, and members of the Hope Chapel Foursquare Church, I love you all.

TABLE OF CONTENTS

Chapter 4 Results

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

BACKGROUND

1.1 Introduction

North Carolina has about 18.14 million acres of timberland, primarily dominated by hardwoods in terms of number of species, acres of timberland, and live standing tree inventory with the largest area in the coastal plains followed by the piedmont and mountains (McConnell et al. 2016). Private individuals control 61% of the timberlands, private corporations or industry ownerships hold 24%, 14% is owned by the government while the remaining 1% is owned by Native Americans and non-governmental organizations. There are about 48.9 billion cubic feet of live standing timber inventory (35% softwood and 65% hardwood) of which 72% is sawtimber and 28% pulpwood (Brown and Vogt 2015). Majority of the softwood species are found in eastern NC while hardwood species predominate in western part. Forest management goals vary across the mountains, piedmont, and coastal plain regions of NC. According to Parajuli and Bardon (2020), forest management is mostly determined by landowner goals in these regions. In 2018, forestry and forest products activities employed about 150,000 people, paid $8.3 billion in labor income while providing additional 1.01 jobs in NC for every job created in the forest sector, and overall contributed $33.6 billion to NC’s economy (Parajuli and Bardon 2020).

A variety of wood products are produced by NC’s forest products manufacturers. Sawmills convert the delivered sawtimber-sized trees to lumber. High quality hardwood is often used as a primary production input in the secondary manufacturing industries such as the furniture industry while the low quality hardwood sawlog produces lumber for furniture frames, road boards, pallets, and mats (Parajuli and McConnell 2019). Also, pulp and paper manufacturing industries producee fluff pulp used as raw material for the production of absorbent and personal care items like diapers, box plants for paper production and other byproducts used in other industries. North Carolina's forest products are consumed by both domestic and foreign markets. In terms of port value, NC forest products exports were about 15.5% of the total exported forest products in the US south in 2013 (Figure 1).

Source: USDA Foreign Agricultural Service

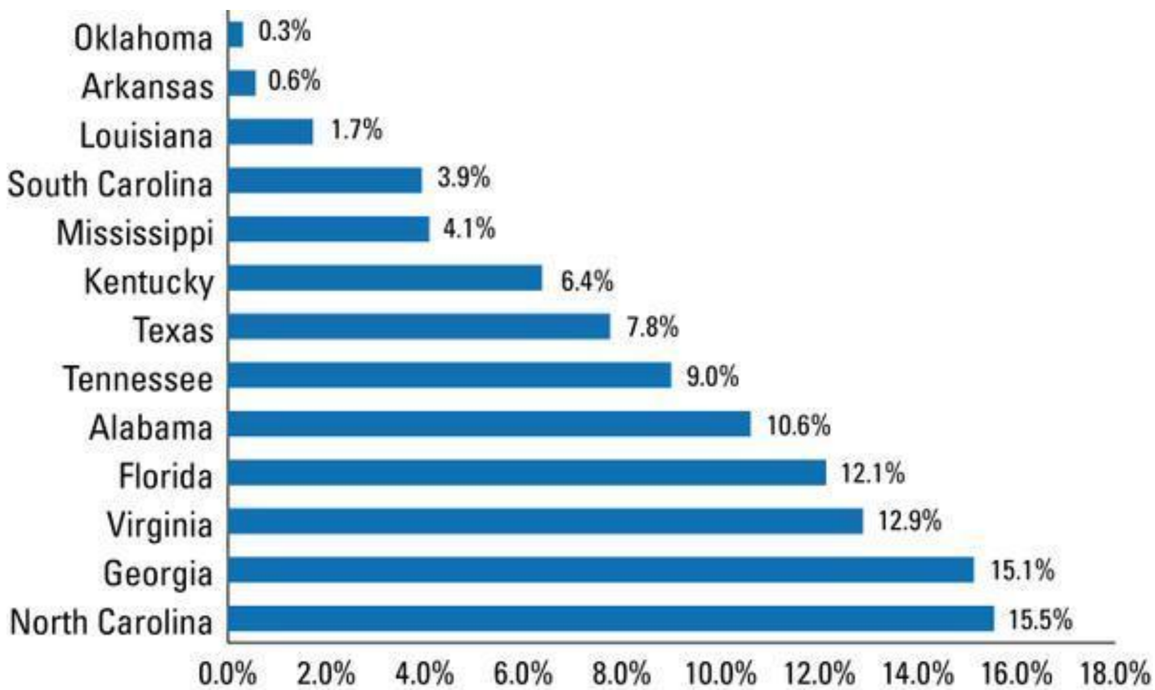

Figure 1: Foreign exports in the US South, 2013

The lumber production in NC constitutes about 5.2% of the total annual national lumber production, and about 2.8% of the total annual lumber consumption in the United States (Parajuli and McConnell 2019). Softwood species account for over two-thirds of the state's lumber production which comes from large southern pine sawmills located in the piedmont and coastal plain regions (east) of the state. A larger portion of the hardwood species of NC are found in the west and other hardwood species in the east. The last housing market crash in the US and global recession in 2008 largely impacted the nation's forest product industry. According to Parajuli and McConnell (2019), NC's annual lumber production averaged about 2.301 billion board feet (about 5.2%) of the total US annual domestic lumber production, while annual lumber consumption averaged approximately 1.575 billion board feet (about 2.8%) of the total annual lumber consumed in the US since 1998 (Parajuli and McConnell 2019). NC's lumber production and consumption suffered a decline of 1.048 BBF and 0.983 BBF respectively between 2005 and 2009. Although production and consumption trends have been rising after 2009, neither of the trends had returned to pre-recession levels as of 2016 (Parajuli and McConnell 2019). In addition, the decline in the state's wood furniture industry shifted the demand condition for hardwood lumber from domestic to foreign markets (e.g. to China). The US Department of Labor (2011) reported that the number of furniture and other related product manufacturing industries in NC declined from 1,323 in 2001 to 989 in 2010.

The observed shift in production and consumption trends, and changes in the number of mills have affected market demand for the forest product in NC. These changes affect wood prices, forest management goals, and the geographical distribution of mills. Several studies have investigated the spatial distribution of primary wood product manufacturers in the US south and factors driving these spatial patterns (Aguilar et al. 2009, Aguilar 2009, 2008). A number of previous studies have also examined the clusters within specific states in the US (Aguilar and Vlosky 2006, Hagadone and Grala 2012, Brandeis and Hodges 2018, Kaur et al. 2020, Richardson 2016, Michaud and Jolley 2019). Despite the recent increase in spatial analyses there are few studies that have specifically examined the forest products industry in NC. Goyette (1967) evaluated factors influencing the location of paper mills in the South using NC as a case study and Porter (1998) identified only one cluster of household-furniture cluster in NC. There have been no published analyses of spatial distribution of the forest products manufacturers in NC.

This study addresses this gap by investigating the current spatial distribution of the primary and secondary forest products manufacturers based on their incidence location and mill production capacity, and factors influencing the spatial location of mills in NC. Spatial distribution affects hauling distance, transaction cost, and can improve communication (Porter 1998) and efficiency in production and marketing (Braden et al. 1998).

The research questions addressed in the study are:

- Does proximity to sources of raw material impact the location of mills?
- Does the type of species utilized by primary forest products manufacturers influence their choice of location?
- Are mills sited based on accessibility to transportation networks?
- Do rural or urban areas affect where mills choose to locate?
- Does the presence of existing primary manufacturers influence the location of secondary manufacturers or vice versa?

Based on the literature reviewed, the following set of hypotheses were formulated and tested using different techniques and approaches presented in the subsequent chapters:

Hypothesis, H1: The spatial distribution of forest products manufacturers in NC is significantly influenced by the distribution of hardwood and softwood resources (raw materials).

Hypothesis, H2: The average annual production capacity of primary forest products manufacturers has a significant effect on their choice of location.

Hypothesis, H3: The presence of a final good market or high population influences the choice of location of forest products manufacturers.

Hypothesis, H4: The presence and access to interstate highways, railroad or ports have a significant effect on the location of forest products manufacturers.

Hypothesis, H5: The presence of primary forest products manufacturers in a county has a significant effect on attracting secondary forest products manufacturers.

Hypothesis, H6: Availability of labor has a significant influence on the choice of location of the new forest products manufacturers’ location.

Hypothesis, H7: The economic well-being and tier designation (rank) of a county has a significant effect on influencing the choice of the forest products manufacturers’ location.

The remainder of this study is structured into literature-review of previous studies, methods, results, discussion and conclusion sections.

CHAPTER 2

LITERATURE REVIEW

2.2 Background on Industrial Localization

The “cluster” concept is fundamental to this research. A review of literature shows that there has been a growing interest within economic geography, global economics, and business clusters, in the spatial distribution of industries. Towards the end of the nineteenth century, the seminal paper of Edward Ross (1896) on “The Location of Industries" is plausibly one of the first studies to highlight the causes and explain the reasons an industry is located at a certain place and not elsewhere. In an attempt to explain why, Ross focused on understanding the underlying fundamentals and patterns in industry location (Ross 1896).

According to Martin and Sunley (2003), in spite of the fact that there are copious literatures on industrial localization, the origin of the process of industrial localization, the reasons it started in certain places and not in others is still a question of concern. Marshall’s (1890) seminal work on industrial districts published 1920 was regarded as the departure point in literature on clusters (Rocha 2013). Marshall’s rationale for industrial districts – *a local industrial atmosphere characterized by a trio of external economies: availability of labor, growth of supporting ancillary trades, and development of local inter-firm division of labor in different stages and branches of production* - was associated with the role of the localization of industry in generating external economies of scale which are external to the firm but internal to the geographic area and improves the efficiency of individual firm (Marshall 1920). Much more, industrial districts legitimately emanated from the process of economic evolution; that is, the growth of an economy through further subdivision of units and intimate connections (Martin 2005).

However, the theory of competitiveness (Porter 1990) is considered to be the turning point of the recently renewed interest in clusters (Martin and Sunley 2003). Michael Porter defines cluster as "a geographically proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and complementarities" (Porter 1998: p. 199). The geographical concentrations of interconnected companies, specialized suppliers, service providers, firms in related industries, associated institutions (such as universities, agencies, and trade associations) in particular fields that compete but also cooperate are both vertical (buying and selling) and horizontal (technologies, similar input, labor). Porter argued that a firm's location is the main source of competitive advantage within a global economy (an economy characterized by rapid transportation, high-speed communication and accessible market). That is, productivity and regional or national competitiveness are highly influenced by the quality of the business environment (Porter 1998). In addition, the creation of the business environment is contingent on the interaction of four factors referred to as Porter's diamond: firm strategy, structure and rivalry; factor input (supply) conditions; demand conditions; and related and supporting industries (Porter 1990).

In Weber (1929) opinion, agglomeration occurs as a result of cost minimization decisions made by firms. If a location possesses certain economic advantages that help firms reduce cost and save more, such will attract firm agglomeration. Weber based the location of manufacturing production on three factors: cost, availability of labor, and agglomeration (Braunerhjelm and Carlsson 1999). Cost and labor are related to the primary requirements of manufacturers, that is, inputs (raw materials) and proximity to the market while agglomeration is defined by advantages derived at a certain location. These advantages can either be derived from (a) economies of scale in terms of production, or (b) proximity to suppliers, labor, or knowledge spillover, or (c) urbanization economies which result from concentration of firms in a location where they enjoy the low cost of energy and other infrastructures (Braunerhjelm and Carlsson 1999).

Industrialization in Europe, United States, China, Switzerland, Southern Germany, to mention a few, was one of the key characteristics of the nineteenth century (Asheim et al. 2006; Rocha 2013). However, many of the localized concentrations of specialized activities experienced a different geographically dispersed pattern of production which became the main basis of economic development due to a major shift in industrial structure, increasing mass production methods and the ascendancy of large, integrated firm; this event refers to the transition from Fordism to Post-Fordism era (Piore and Sabel, 1984). In spite of the fact that Gray (1998) and Reich (2001) reckoned a rapid transition towards a globalized and technology driven economy to have undermined the significance of location and spatial proximity for business operation and profitability, the apparent reality is that globalization and technological change has been advancing rather than reducing the importance of location in economic development (Ashiem et al. 2006).

Furthermore, cities and regions as centers of contemporary economic development and governance have emerged over the last four decades to be localized production systems of strategic industrial agglomeration (Sabel 1989; Krugman 1997; Porter 1998; Scott 1988; Ashiem et al. 2006, Rocha 2013). The emergence of localized production systems of specialized industrial agglomeration over four decades ago are proof of a part (or a more general) resurgence of regions and cities as centers of modern-day economic development and governance. Paul Krugman, the main author of "new economic geography" modeled agglomeration as a trade-off between agglomeration economies and transportation costs using multi-equilibrium models of imperfect competition (Asheim et al. 2006). The works of Fujita et al. (1999) and Fujita and Thisse (2002) indicated that the new economic geography has moved beyond industrial districts. The new economic geography is about "*a universal micro-economic theory of spatial agglomeration that covers everything from intra-urban patterns of economic specialization and spatial structure of cities right through industrial districts and clusters, to North-South dualism in the world economy*" - Asheim et al. (2006).

Krugman's perspective on industrial localization differs remarkably from that of forerunners like Porter. Krugman (1998) observed that the existing competitive advantages in a global economy are often heavily localized resulting from the concentration of highly specialized skills, knowledge, institutions, rivalry, related businesses and customers. The regional and local specializations arising from increasing global integration impacted transportation cost and trade barriers which enabled firms to agglomerate with other like-firms in order to benefit from local external economies of scale (Krugman 1991, Fujita et al. 1999) thereby aiding innovation and productivity growth (Martin and Sunley 1998). These activities and many others gained more attention and attracted further studies on "localization of the world economy" (Krugman 1997) and the rise of "global economies" (Scott 1998). Ashiem et al. (2006) anecdoted on localized concentrations of specialized activities has more detailed summary of numerous neologisms coined to describe various concentrations such as: 'industrial districts' in Italy - as one of the earliest types of industrial concentrations, the rise of 'new industrial spaces', 'local production systems', 'localized agglomeration' – local high-tech milieu, new economic geography, learning regions and the rise of industrial or business clusters.

2.2 Spatial Analysis of Forest Product Industry in US

Management of the forests and forest product industry in the US has been approached by utilizing spatial analysis especially in the last two decades to maximize efficiency and/ or minimize the cost of production. Anderson et al. (2011) utilized a geographic information system-based spatial analysis to study sawmill wood procurement and how procurement pressure may impact the future of the sawmill industry in northeastern United States and Canada. Using the mills characteristics and location information, it was observed that large softwood sawmills operating in Maine and across the border in Quebec and New Brunswick depend on industrial and investment-oriented forestlands to meet production requirements; which constitute a hot spot of procurement pressure relative to the lower procurement pressure in most part of the other study region (Anderson et al. 2011). Also, further sensitivity analyses suggested that procurement pressure in existing hot spots would escalate if the procurement range is limited by high costs of transportation. (Anderson et al. 2011).

Supply chain networks are a significant part of the forest product industry. Several studies have considered the relationship between the price of forest products (sawlogs, pulpwood, biomass, etc.), transportation cost and hauling distance, and how they impact forest management and the forest product industry respectively (Silva et al. 2019, Pokharel and Latta 2020). Han et al. (2018) and Pokharel et al. (2019a) estimated the transportation logistics for forest products utilization using different methods. These studies utilized mathematically optimal, fixed, and aerial distance for estimating cost effective means of transporting forest products, particularly forest biomass. Pokharel et al. (2019a) used a spatial logistic regression model to estimate the willingness of mills in the US south to use mill residues based on their location, existing transportation networks, other mills, forested areas, and city centers. Their findings suggested that mills were more likely to use logging residues if they are at a proximate location to a major road, sawmill, pulp, paper, and paperboard mill, and post mill. However, they found that mills close to a major river, other wood using mills (based on their classification), forest and city center were less willing to use woody biomass (log residues). By employing a mixed-integer programming model coupled with a network algorithm, Han et al. (2018) evaluated the production cost minimization for bioenergy and bioproducts; which optimally reduced the cost of logistics in the biomass supply chain up to 11%.

Also, the spatial allocation of forest products processing facilities, primary forest products, and processing infrastructure in the forestry supply chain has been studied by Pokharel and Latta (2020) in an attempt to identify forest merchantability limitations across the US. By analyzing the location and national road dataset of primary forest products manufacturers in the contiguous US at varying haul-time periods for two scenarios with respect to transportation costs, Pokharel and Latta (2020) identified the hot spots locations where an opportunity to merchandise forest products enables the flexibility in forest management, and cold spot where alternatives are limited. Major findings from their research subsumed an observation of severe merchantability issues in the public and western forests, substantial transportation costs significantly improved merchantability, and about 9% of forests have low merchantability for sawlogs, pulpwood, and biomass even at the highest paid price.

Several studies have assessed the emergence and development of wood products manufacturing clusters in the US. Most of them built on the works of Porter (1990, 1998, 2003), Piore and Sabel (1984), Krugman (1991, 1997), Marshall (1920), and Ross (1896). It has been shown that the emergence of clusters is not limited to urban areas and that manufacturing clusters in rural areas can grow and foster cluster based economic development. Braden et al. (1998) identified several clusters of wood products manufacturers in the Pacific Northwest to support the argument that clusters are opportunities to contribute to economic development of rural areas. Factors that contributed to the development of observed clusters were identified and how these factors impacted each cluster development was determined. Proximity and access to regional market, abundant supply of raw materials and potential customers, skilled labor, co-locating with related firms, low level of initial competition, innovation, small cluster size, government policies and other factors impacted the emergence and development of wood products manufacturing clusters in the Pacific Northwest (Braden et al. 1998).

In the US South, spatial analysis has been utilized to study the spatial distribution of the primary forest products industry. Aguilar et al (2009) found evidence of regional clustering in the primary wood products manufacturing industry in the US South by analyzing the incident locations of manufacturers using a bivariate kernel density function, F-hat and L-hat, and Chi-square statistics. In addition, location preference of primary mills and factors promoting clustering of these mills were identified and explored, respectively. Their results underpinned the findings of Braden et al. (1998) by suggesting that counties with availability of sufficient transportation infrastructure, and the presence of related industries were more likely to host new primary forest products manufacturers (Aguilar et al. 2009). Nevertheless, spikes in logs and energy prices can have a negative impact on spatial aggregation of wood-using mills in this region of the US.

In another study, Aguilar (2009) utilized an exploratory analysis to determine clustering, and a Bayesian spatial autoregressive approach to model the presence of lumber industry (sawmills) in the US South. Results showed that counties characterized by cost advantages, availability of labor, and transportation access were most likely to host a lumber industry. Among the counties projected to have a high probability of hosting lumber industry developments in the South, five were identified in NC (Aguilar 2009). Centrifugal (dispersive) and centripetal (clustering) forces were also identified in the softwood lumber industry in the US South (Aguilar 2008) by building on the concept of new economic geography (Krugman 1991, Fujita and Krugman 2004). Based on Aguilar's findings, clustering of softwood sawmills might be disadvantageous to the lumber industry because it engenders high costs of raw materials, congestion and undesirable competition. However, co-agglomeration between the primary and secondary forest product manufacturers could potentially aid a competitive industrial structure (Aguilar 2008).

Although a few principles seem to govern the development of clusters but the content and intensity of the forces at work tend to differ in space and time (Asheim et al. 2006, Braden et al 1998). Specific studies have also been conducted in some states in the US South to further understand the spatial distribution pattern of the forest products industry and factors that may have influenced the formation of the observed agglomeration pattern. In Louisiana, primary forest products manufacturers showed higher spatial dependency relative to secondary forest products manufacturers. Aguilar and Vlosky (2006) found that the primary manufacturers are located close to the source of raw materials while secondary manufacturers are found near major populated areas in Louisiana. Brandeis and Hodges (2018) found that the spatial distribution of sawmills in Tennessee changed between 1999 and 2011 from a random distribution to a clear localized pattern based on their wood consumption.

Furthermore, Hagadone and Grala (2012) identified spatial clustering in the primary and secondary forest products manufacturing industry in Mississippi. The amount of harvest sawlog significantly influenced the location of primary and secondary manufactures while the volume of pulpwood harvest had an effect on secondary mills. Other factors such as presence of railway, existing primary mills, and availability of labor in a county positively impacted the location of secondary mills while the presence of four-lane interstate highway decreased the likelihood of a secondary manufacturing locating in a county (Hagadone and Grala 2012). Using a spatial analysis, Kaur et al. (2020) identified two hot spot clusters that could sustain the redcedar bioproducts industry in Oklahoma. Their results indicated that the eastern redcedar industry manufacturing particleboard, mulch, and oil in Oklahoma could contribute $96 million per annum to the state's revenue and generate additional 319 employment opportunities (Kaur et al. 2020). Kies et al. (2009) identified several hot spots in German primary and secondary wood products industry while investigating regional employment patterns of the total forest sector using exploratory geostatistical methods for regionalized cluster mapping. The primary manufacturing industry hot spot locations showed a weak concentration pattern across Germany due to the small size of the industry. Conversely, the secondary manufacturing industry showed a stronger agglomeration pattern and had a higher impact on regional employment and demand for production inputs in the German forest sector.

CHAPTER 3

MATERIALS AND METHODS

3.1 Research Methods

The study investigated the spatial distribution of forest products manufacturers - primary and secondary forest products manufacturers - in North Carolina (NC). We first identified major hot spots (existing forest business clusters) and potential forest business clusters suitable for cluster-based economic development and examined the impact of factors influencing the location of forest products manufacturers in NC. We employed two approaches to study the geographic concentration of forest product manufacturers in NC. First, the optimized hot spot analysis, Getis-Ord Gi* statistics was employed to identify potential forest products manufacturers' clusters (hot spots) at county levels using incident point location and production capacity dataset. Secondly, a Poisson regression analysis was also conducted to determine the impacts of major factors related to resource availability and socio-economic condition on the location of forest products manufacturers in NC.

3.2 Z-Score

The z-score is a measure of statistical significance of every feature. it indicates whether the observed spatial clustering is more pronounced than one would expect in a random distribution of those same values. The z-score fields do not reflect any kind of False Discovery Rate (FDR) correction.

In spatial statistical analysis (pattern analysis), identifying the null hypothesis is always important, which is generally known to be Complete Spatial Randomness (CSR), either of the features or values associated with the features. The z-scores output from the pattern analysis tools in ArcGIS helps determine whether to reject the null hypothesis or not (Mitchell 2005). A higher z-score for a feature indicates a spatial clustering of high values. A lower negative z-score indicates a spatial clustering of low values. The higher (or lower) the z-score, the more intense the clustering. A z-score near zero indicates no apparent spatial clustering.

For instance, when the result of running a spatial pattern analysis shows that the features or values associated with the features being analyzed exhibit statistically significant clustering or dispersion, this indicates some underlying spatial processes at work, and we can therefore reject the null hypothesis of CSR and vice versa. In addition, z-scores are standard deviations and are associated with standard normal distribution (Mitchell 2005).

3.3 Spatial Distribution of Forest Products Manufacturers in North Carolina

Spatial analysis uses geographical and statistical theories to solve complex location-oriented problems and better understand what pattern is exhibited and where it is occurring in a given study area. It therefore goes beyond mere mapping to assess the characteristics of places and the relationships between them. Spatial analysis lends new perspectives to decision-making. It has been employed to determine the distribution pattern of the forest product industry statistically and geometrically in the southern US (Aguilar and Vlosky 2006; Aguilar 2007; Hagadone, and Grala 2012; Kaur et al. 2020).

Most spatial models are based on micro-economics and predict the spatial patterns occurring in a distribution such as the location pattern of forest products manufacturers given certain preconditions such as, profit minimization or maximization. In this study, the spatial distribution pattern of the forest products industry in North Carolina was examined to determine if the past geographical location pattern of forest product manufacturers exhibited either a random, clustered or dispersed distribution pattern. Industrial clustering was determined by exploring variations in the concentrations of these manufacturers and deviations from complete spatial randomness (spatial autocorrelation) in their geographic distribution (Mitchell 2005). Areas with high concentrations of manufacturers might constitute forest business clusters and could also be potentially suitable for cluster-based economic development (Frizado et al. 2009; Hagadone, and Grala 2012). The spatial analysis was conducted separately for the primary and secondary forest products manufacturers and altogether for the entire state of North Carolina to account for differing location patterns. All spatial data were projected and analyzed on the Lambert Conformal Conic, North American Datum (NAD) "1983_2011_StatePlane_North_Carolina_FIPS_3200_Ft_US" coordinate system.

3.4 Forest Products Manufacturers in NC

Forest products manufacturers can be categorized into primary and secondary manufacturers based on their primary output (Vlosky and Chance, 2001). Primary forest products manufacturers produce wood products directly from logs (small and large) or wood chips. For example, sawmills produce lumber from logs, and pulp mills produce pulp products from pulpwood. These categories of manufacturers correspond to codes 321 and 322 in the North American Industry Classification System (NAICS), respectively. In this study, the primary forest products manufacturers were further subdivided into “sawmills” and “other wood-using mills” (defined as biomass/energy plant, pulp and paper mill, log exporter, veneer and plywood mill, post mill, mulch, and pole mill). The clustering of related manufacturers captures the causal effect of similar production inputs based on the assumption that all firms in this subdivision have comparable production functions; an approach Aguilar et al. (2009) also employed in studying clustering in the primary wood products industry of the US South.

In contrast, secondary forest products manufacturers are defined as those manufacturers which depend on primary forest products manufacturers for their inputs in remanufacturing and value added. A furniture mill that uses lumber to produce furniture and a cardboard box plant that produces boxes from linerboard are both examples of secondary forest products manufacturers. This category of manufacturers corresponds to NAICS code 337.

3.5 Optimized Hot Spot Analysis

An optimized hot spot analysis was conducted to determine if local patterns differ from the global pattern using the Getis-Ord Gi* tool in ArcMap 10.7.1 (Getis and Ord 1992). Since we are interested in evaluating the spatial pattern (clustering) of the location of forest products manufacturers (incident points), the optimized hot spot analysis tool was employed to aggregate the incident locations of forest products manufacturers within counties in North Carolina in order to obtain incident counts as an analysis field; that is counties with their associated counts of manufacturers. The aggregation of incidents within a polygon is an appropriate strategy when points are associated with fixed political or administrative units such as states, counties, tracts or districts (Ord and Getis 1995; Esri 2017). Next, the optimized hot spot analysis tool employs the Incremental Spatial Autocorrelation tool to assess the intensity of clustering at increasing distances and select an appropriate scale of analysis at a peak distance where spatial processes promoting clustering are most pronounced (Mitchell 2005). An optimal fixed distance band was selected for the optimal scale of analysis.

Given a set of weighted point features (location of forest products manufactures), the optimized hot spot analysis tool calculates the Getis-Ord Gi* statistic for each forest products manufacturer in each county and identifies counties with statistically significant spatial clusters of high values (hot spots), and or low values (cold spots) based on their z-scores (Esri 2017). It is important to note that features with high values may not be statistically significant hot spots all the time. Therefore, in order to be a statistically significant hot spot, features must have high values and be surrounded by other features with high values (Getis and Ord 1992; Ord and Getis 1995; Mitchell 2005). Also, the hot spot analysis tool compares the local sum of each forest products manufacturers and their neighbors proportionally to the sum of all forest products manufacturers in computing the Getis-Ord Gi* statistics. The result of the calculation is automatically corrected for multiple testing and spatial dependence using the False Discovery Rate correction method (Caldas et al. 2006). Thus, the resulting Gi* statistic reports statistically significant z-score and p-value at 99 percent, 95 percent or 90 percent confidence level for each feature given that the null hypothesis for the spatial pattern analysis is Complete Spatial Randomness (CSR). Mathematically, the Getis-Ord Gi* local statistic is represented as (Esri 2017):

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \bar{X} \sum_{j=1}^{n} w_{i,j}}{S \sqrt{\frac{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j} \right)^2 \right]}{n-1}}} \qquad (1)$$

Where x_j is the attribute value (existing forest products manufacturer) for feature *j* (given county), $w_{i,j}$ stands for the spatial weight between feature *i* (NC counties) and j in order to account for their spatial interrelationship ($i \neq j$), *n* is the total number of features, that is, all counties in NC (*n* = 100), $\underline{X}$ and S are the mean of observed values and standard deviation, respectively, which are stated as:

$$\bar{X} = \frac{\sum_{j=1}^{n} xj}{n} \qquad (2)$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - (\bar{X})^2} \qquad (3)$$

As the positive z-score increases, the intensity of clustering of a high value (hot spot) also increases. Conversely, as statistically negative z-score decreases, the more intense the clustering of low values (cold spot). Hence, the clusters of forest products manufacturers were identified using the Getis-ord Gi* statistics and the z-score statistics at the 95 percent confidence level.

The optimized hot spot analysis was completed separately and jointly for both primary and secondary forest products manufacturers based on their individual incident point locations respectively. The hot spot analysis for the primary forest products manufacturers was also conducted by their individual production capacity values. We could not complete the optimized hot spot analysis for the secondary forest products manufacturers based on their production capacities as the production size of secondary mills were not available.

In order to gain a clearer understanding on how the spatial distribution of primary forest products manufacturers are tied to the forest resources distribution in NC, the incident data of primary manufacturers was further analyzed based on the types of wood species used for manufacturing purposes. Hence, the optimized hot spot analysis was completed for the net merchantable bole volume of growing-stock trees (at least 5 inches diameter at breast height, d.b.h.) of softwood and hardwood species to identify where the resources hot spot are located in NC and how they correspond with the location of forest product manufactures.

3.5.1 Optimized Hot Spot Analysis by the Production Capacity

Additional Optimized Hot Spot Analysis was conducted specifically for the primary forest products manufacturers using their production capacity. Since we were interested in analyzing data on production capacity values (used as field of analysis) associated with each primary forest products manufacturer at the county level, the resulting z-score values from the Optimized Hot Spot Analysis only helped in identifying where high and low production capacity values clustered with respect to each manufacturer's incident point location and did not aggregate hot spot into county locations. Due to this, an "Interpolation Analysis" was conducted using the Inverse Distance Weighted (IDW) technique in the Spatial Analyst toolbox in ArcGIS ArcMap 10.7.1 software under the Geospatial Analyst extension. The analysis helped examine where high and low primary forest products manufacturers' production capacity values were clustered by smoothing the z-score values from the optimized hot spot analysis of production capacity across all counties in NC.

3.5.2 Interpolation Analysis – Inverse Distance Weighted (IDW) Interpolation

Generally, interpolation analysis is utilized to predict values for cells in a raster from a limited number of sample data points; however, it is not restricted to raster data. It can be employed to predict unknown values for any geographic point data, such as noise level, elevation, disease concentration amongst other things (Watson and Philip 1985). The IDW interpolation is a type of raster interpolation that determines cell values using a linearly weighted combination of a set of sample points. The weight is a function of inverse distance generated from a surface (study area), that is, of a location-dependent variable. According to Watson and Philip (1985), IDW is recommended when sampling is sufficiently dense with respect to local variation as in the case of our production capacity data. The analysis specifically utilized the following parameters: z-score values from the Optimized Hot Spot analysis of the primary forest products manufacturers by production capacity as the numeric field (z-field) to be interpolated into a surface raster, an inverse distance raised to the exponential power of two in order to control the significance of surrounding points on the interpolated value, and a variable search radius of 12 in order to find a specified number of nearest input sample points to be used to perform interpolation (Watson and Philip 1985).

3.6 Data Sources

The data used in this research were obtained from various sources. Primary forest products manufactures' location data were obtained from the Timber Products Output (TPO) survey from the NC Forest Service (NCFS), TPO forest service data set, and NC forest service dataset. The secondary forest products manufacturers' location data were obtained from the Forest Products Locator's website developed by the Southern Group of State Foresters. The primary and secondary data were provided in Excel format with X and Y geographical coordinates. The coordinates were utilized to import the data before being converted into shapefiles for further spatial analysis. A total of 233 primary forest products manufacturers and 566 secondary forest products manufacturers located throughout NC were included in this study.

The interstate highway and railroad transportation data were obtained from North Carolina Department of Transportation, NC OneMap, while the port data were generated from Google Maps by searching every county within 100 miles from Virginia port, Wilmington port and Morehead City port. For the regression analysis the transportation data were coded as binary variables. Every county with the presence of an interstate highway was coded as 1 and 0 otherwise. Similarly, the presence of railroads in a county was coded as 1 and 0 otherwise. While every county within a 100 miles distance from any of the three ports locations were coded 1 and 0 otherwise.

Total merchantability bole volume of growing stock trees, sawlog harvest, and pulpwood harvest data were obtained from Forest Inventory and Analysis (FIA) as variables to account for the source of raw materials in our model. Unemployment data was obtained from the Bureau of Labor Statistics (U.S. Department of Labor, 2018) and population data was obtained from U.S. Census Bureau, 2019. The presence of a city in a county with a population greater than 50,000 people was coded as 1 and 0 otherwise, in order to assess the influence of the final market on the location of the forest products manufacturing industry. Economic tier ranking data was obtained from the NC Department of Commerce, 2020, to determine if the economic distress level of a county has an influence on the location of forest products manufacturers in NC. Table 1 shows the description of variables used in the regression analysis to determine their impact on the location of forest products manufacturers in NC.

Table 1. Description of variables used to model factors influencing the location of forest products manufacturers in North Carolina.

Explanatory variable	**Description**	**Source**
Primary forest products manufacturer	Number of primary forest products manufacturers in a county	Timber Products Output (TPO) survey from NC Forest Service (NCFS),
Secondary forest products manufacturer	Number of secondary forest products manufacturers in a county	Southern Group of State Foresters, https://secondary.forestproductslocator.org/
Road	Binary variable 1, represents interstate highway and national roads, 0 otherwise	North Carolina Department of Transportation, NC OneMap.
Rail	Binary variable equals 1 if the county is connected with railways, 0 otherwise	North Carolina Department of Transportation, NC OneMap.
Port	100 miles distance from port	Google Map
Unemployment	Total unemployed population in a county in 2018	Bureau of Labor Statistics

Table 1 (continued).

Total net merchantable volume	Total net merchantable bole volume of growing-stock trees (at least 5 inches d.b.h.), in cubic feet, on timberland	Forest Inventory and Analysis (FIA)
Annual Harvest Removals	Average annual harvest removals of merchantable bole volume of growing-stock trees (at least 5 inches d.b.h.), in cubic feet, on timberland	FIA
Sawlog harvest	County 2017 volume of harvested timber in thousand cubic feet	FIA
Pulpwood harvest	County 2017 volume of harvested pulpwood in thousand cubic feet	FIA
Population	Presence of a city in a county with a population greater than 50,000 people	U.S. Census Bureau, 2019
Economic Tier 2, 3	Binary variable equals, 1 if the county is in tier 2 or 3, 0 otherwise.	NC Department of Commerce, 2020

3.7 Factors Influencing the Location of Forest Product Manufacturers in NC

An econometric model was developed to quantify the effect of selected factors identified in literature on clustering within the forest products manufacturers across North Carolina. Using a count data model, this study explored the common attributes of sawmills and other wood-using mills that share similar production inputs (like biomass/energy plant, pulp and paper mill, log exporter etc.), and clustering factors within the forest products manufacturing industry. The empirical models used to determine the location like-likelihood of primary forest products manufacturers had the following specifications:

Primary mills = f (secondary mills, road, rail, port, population, total merchantable volume, unemployment, economic tier)

Secondary mills = f (primary mills, other primary, road, rail, port, population, sawlog harvest, pulpwood harvest, unemployment, economic tier)

Econometric Methods

The dataset has the count data dependent variable with a Poisson distribution of discrete events or occurrences over a specified interval or continuum (time, length, distance etc.). The Poisson distribution is represented as (Gujarati 2004, Silva and Tenreyro, 2010):

$$Pr(y_i = j|x_i) = \frac{\exp(-\mu)\mu^j}{j!}, \quad j = 0, 1, 2, \dots \tag{4}$$

and,

$$\mu = exp\ (\mathcal{X}_i{}'\beta) = exp\ (\beta_0 + \beta_1 \mathcal{X}_{1i} + \cdots)^2 \tag{5}$$

Where y represents the count of forest products manufacturers per county, x_i is a vector of factors influencing the number of forest products manufacturers per county (i), j stands for explanatory variables, and β is the vector of parameters of interest including an intercept. The Poisson distribution is determined entirely by the mean, which is equal to the variance, that is:

$$Var(y|x) = E(y|x) = \mu(x_i) \qquad (6)$$

Thus, given a random sample (*i = 1, 2,…, n*), the model (equation 10) was estimated by maximizing the log-likelihood function given by:

$$InL(\beta) = \sum_{i=1}^{n}[(y_i x_i{}'\beta) - exp(x_i'\beta) - In(y_i!)] \qquad (7)$$

Where we drop the term $[-In(y_i!)]$ because it does not depend on β. This model allows us to find the conditional probability for any of the selected factors (explanatory variables) affecting the location of forest product manufacturers. Since equation 10 is nonlinear in parameter, a nonlinear least square estimation method was used. However, nonlinear least squares do not account for heteroscedasticity exhibited by all standard count data distribution instead a quasi-maximum likelihood estimation method was utilized (Wooldridge 2010). The standard errors of the Poisson estimates were obtained after maximizing the log-likelihood function. Compared to the ordinary least squares (OLS) estimation, the count data model does not permit direct interpretation of the coefficients of the regressors, so we estimated marginal effects to interpret the magnitude of estimated coefficients (Greene 2002, Wooldridge 2010). The marginal effect of the x_j (explanatory variable) with respect to $E(y|x_j)$ is given by:

$$\frac{\partial E(y_i|x_i)}{\partial x_i} = exp\ (x_i'\beta)\beta_j \tag{8}$$

Where, $\bar{y}.\beta_j$ is the average marginal effect and $\bar{y}$ is the average number of manufacturers per county. For Poisson regression with an exponential mean function, the average of fitted values is the same as the average of the original outcomes on y_j. In other words, marginal effects account for the effects of changes (i.e., one-unit change) in the mean values of the regressors (Wooldridge 2010).

CHAPTER 4

RESULTS

4.1 Spatial Distribution Pattern of Forest Product Manufacturers in North Carolina

Based on the data obtained from NCFS, 233 primary forest products and 566 secondary forest products manufacturing facilities are located throughout the regions in NC. The primary forest products manufacturers are present in 79 counties in North Carolina. Similarly, the secondary forest product manufacturers are present in 86 counties in North Carolina. Based on the visual observation, the distributions of both primary and secondary forest products manufacturers seem to cluster in the northwest, southwest, and central regions in North Carolina (Figure 2).

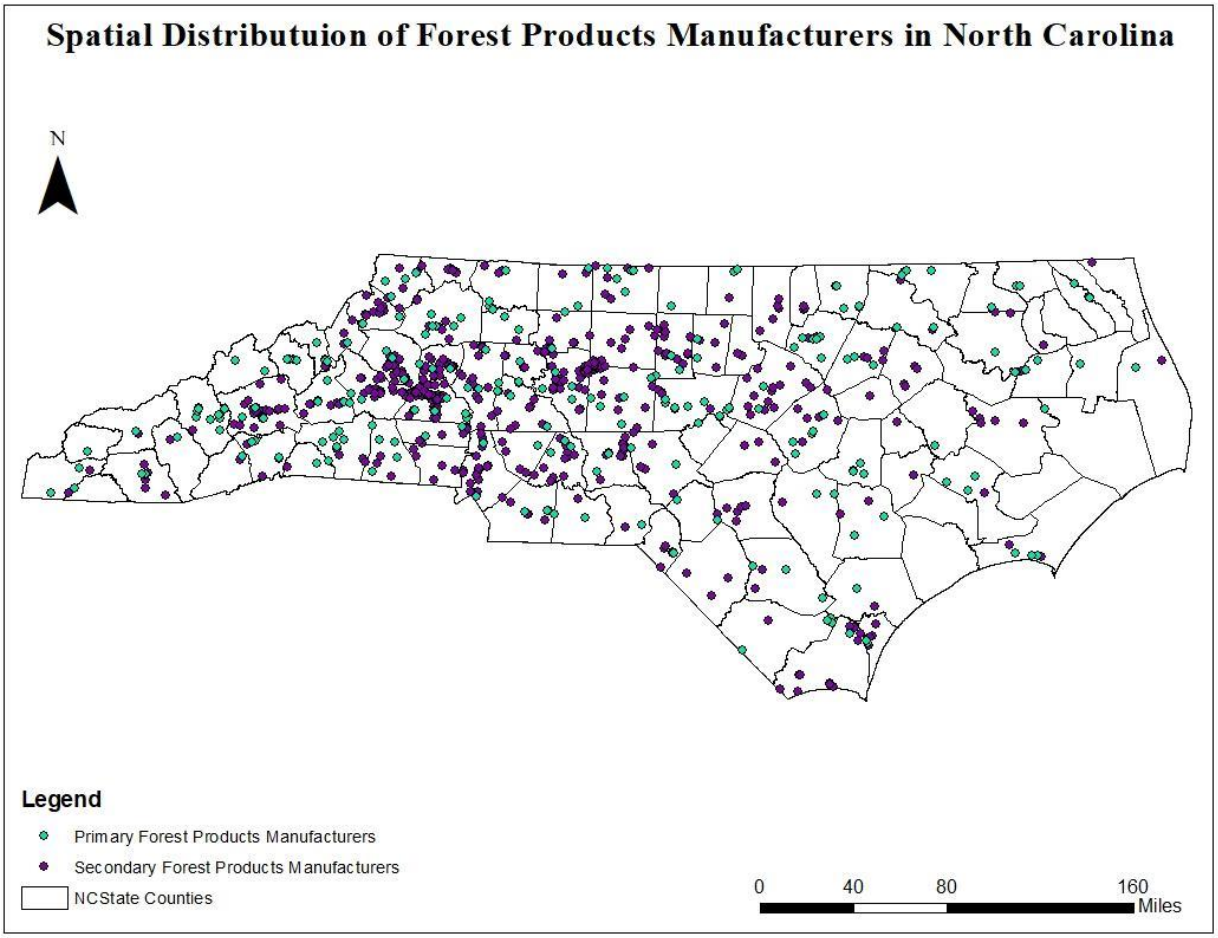

Figure 2. Spatial distribution of forest products manufacturers in NC.

4.2 Optimized Hot Spot Analysis

4.2.1 Optimized Hot Spot Analysis of Primary Forest Products Manufacturers by Location

Only 218 valid input features out of 233 primary forest products manufactures were considered in the optimized hot spot analysis of the primary forest products manufactures with no outliers in this study. This exclusion was due to missing data. Locational outliers are manufacturers that are much farther away from neighboring manufacturers than the majority of manufacturers in the dataset. Incident aggregation analysis performed on all polygons (100 counties) generated a minimum, maximum, and mean incident count values of 0, 9, and approximately 3 manufacturers in each county, respectively. The optimized hot spot analysis identified one major statistically significant hot spot of primary forest products manufacturers. This cluster of high concentration stretched across neighboring counties in the western and northwest regions in the state. Seven major (Gi* z-score > 1.96) and six minor (Gi* z-score, 1.65 > z < 1.96) statistically significant county hot spots were identified[1] at the 95 percent and 90 percent confidence levels, respectively. The major counties include Madison, Buncombe, Henderson, Rutherford, Burke, Caldwell, and Catawba. More so, the hot spot is dominated mainly by sawmills and some other wood-using primary mills (Figure 3A and 3B). Ultimately, there is evidence of spatial clustering amongst the primary forest products manufacturers across North Carolina's counties.

---

[1] Since our discussion and conclusions are based on hot spots identified at the 95 percent confidence level using the Getis-ord Gi* statistics and z-score statistics, only clusters identified within this hot spot range are explained in this study.

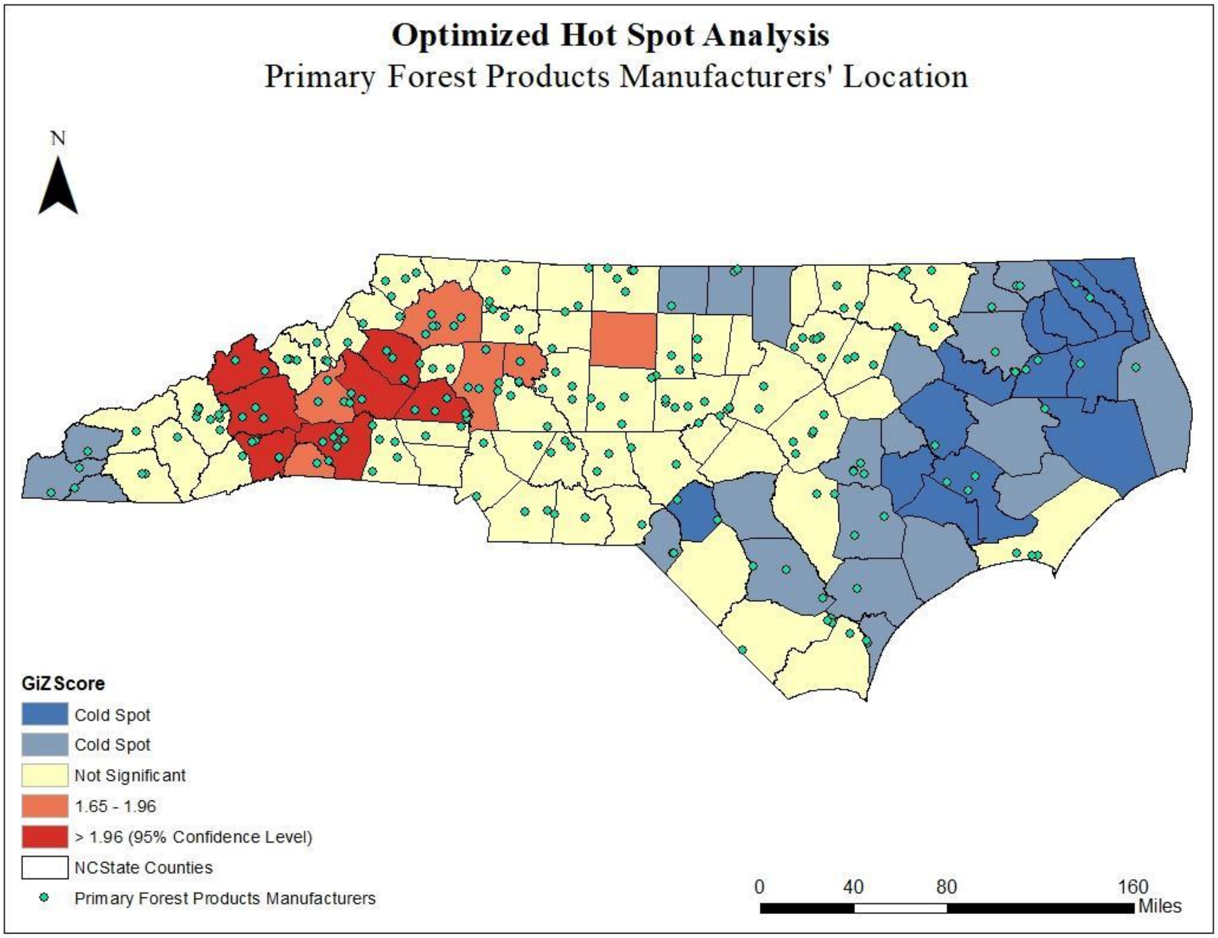

Figure 3A. Identified hot spot (red) and cold spot (blue) locations of primary forest products manufacturers in NC.

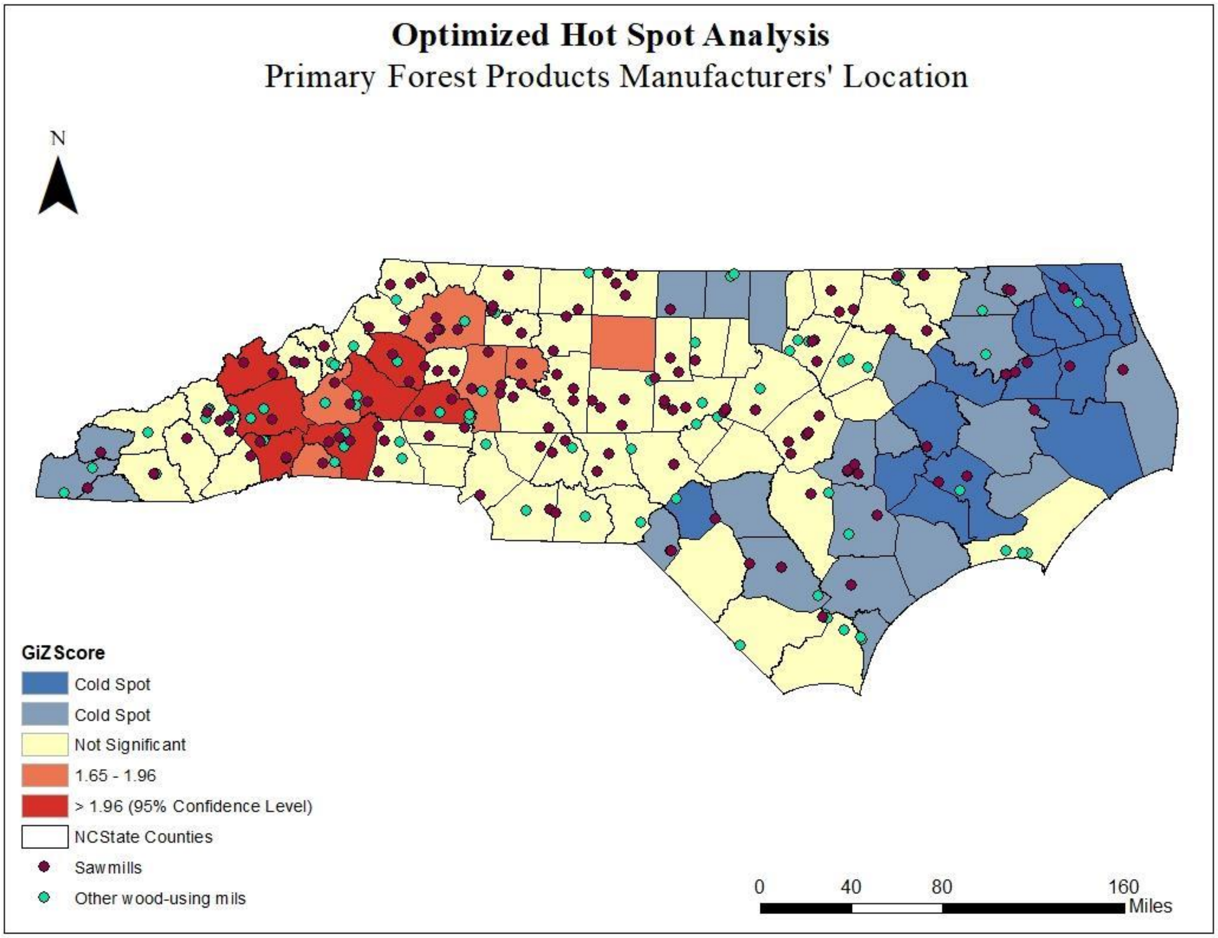

Figure 3B. Identified hot spot (red colored cluster) and cold spot (blue colored cluster) counties of primary forest products manufacturers (i.e., sawmills and other wood-using mills) in NC.

4.2.2 Optimized Hot Spot Analysis of Primary-Hardwood Forest Products Manufacturers by Location

The initial data assessment revealed that there were just 33 valid records of primary hardwood forest products manufacturers. Incident aggregation analysis performed on all polygons (100 counties) generated a minimum, maximum, and mean incident count values of 0, 4, and approximately 0 manufacturers in each county, respectively. The optimized hot spot analysis identified three major statistically significant hot spots (Gi* z-score > 1.96) of primary forest products manufacturers at a 95 percent confidence level. These hot spots consist of six counties (Davidson, Guilford, Alamance, Randolph, Nash and Warren) concentrated around the Piedmont, and Coastal Plain regions in NC. In addition, the hot spots also included three other statistically significant counties (Gi* z-score, 1.65 > z < 1.96) at the 90 percent confidence level. The identified hot spots counties are primarily dominated by sawmills (Figure 4). Hence, the distribution of primary-hardwood forest products manufacturers across NC exhibit deviation from complete spatial randomness.

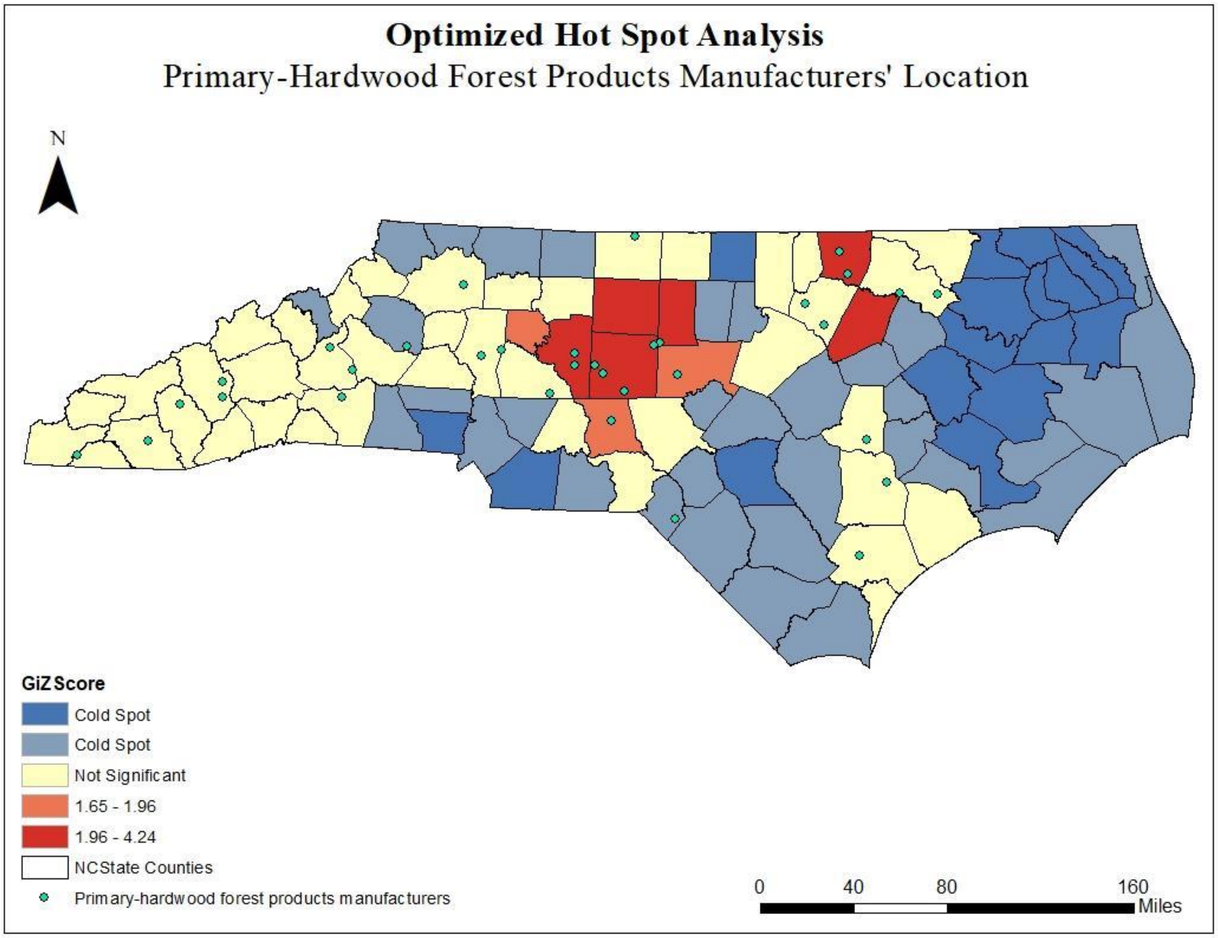

Figure 4. Identified hot spot (red colored clusters) and cold spot (blue colored clusters) locations of primary-hardwood forest products manufacturers in NC.

4.2.3 Optimized Hot Spot Analysis of Primary-Softwood Forest Products Manufacturers by Location

Only 32 valid primary-softwood forest products manufacturers were considered based on the initial data assessment test. Incident aggregation analysis performed on all polygons (100 counties) generated a minimum, maximum, and mean incident count values of 0, 3, and approximately 0 manufacturers per county, respectively. The result of the optimized hot spot analysis identified four major statistically significant hot spots (Gi* z-score > 1.96) of softwood forest products manufacturers at the 95 percent confidence level. These hot spots were identified in the Coastal Plain regions of the state. Also, nine other counties statistically significant (Gi* z-score, 1.65 > z < 1.96) at the 90 percent confidence level were identified around the same regions (Figure 5). Thus, the distribution of primary-softwood forest products manufacturers across North Carolina's counties followed a clustering spatial distribution pattern.

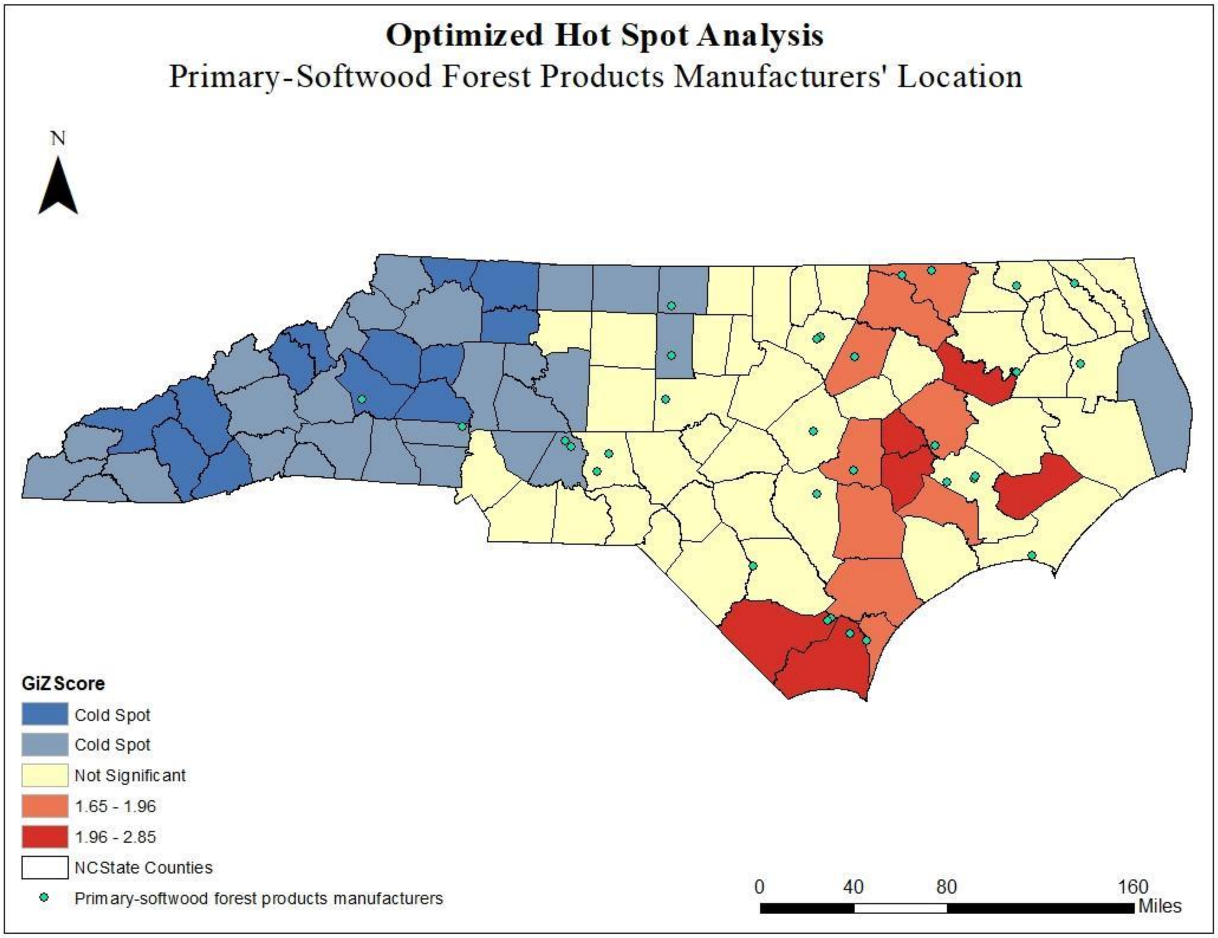

Figure 5. Identified hot spot and cold spot counties of primary-softwood forest products manufacturers in NC.

4.2.4 Optimized Hot Spot Analysis of Primary Forest Products Manufacturers Using Hardwood-Softwood Species by Location

Precisely 90 valid primary forest products manufacturers using both hardwood and softwood species were included in the optimized hot spot analysis. Incident aggregation analysis performed on all polygons (100 counties) generated a minimum, maximum, and mean incident count values of 0, 5, and approximately manufacturers per county, respectively. The result of the optimized hot spot analysis identified one large statistically significant (Gi* z-score > 1.96) hot spot of primary forest products manufacturers which uses both hardwood and softwood timber at the 95 percent confidence level. This hot spot consists of 19 counties concentrated around the western, northwest, and Piedmont regions of the state. Also, one county, statistically significant (Gi* z-score, 1.65 > z < 1.96) at the 90 percent confidence level was identified around the same region (Figure 6). Thus, the distribution of primary hardwood-softwood forest products manufacturers across North Carolina's counties showed deviation from complete spatial randomness; that is, followed a clustering spatial distribution pattern.

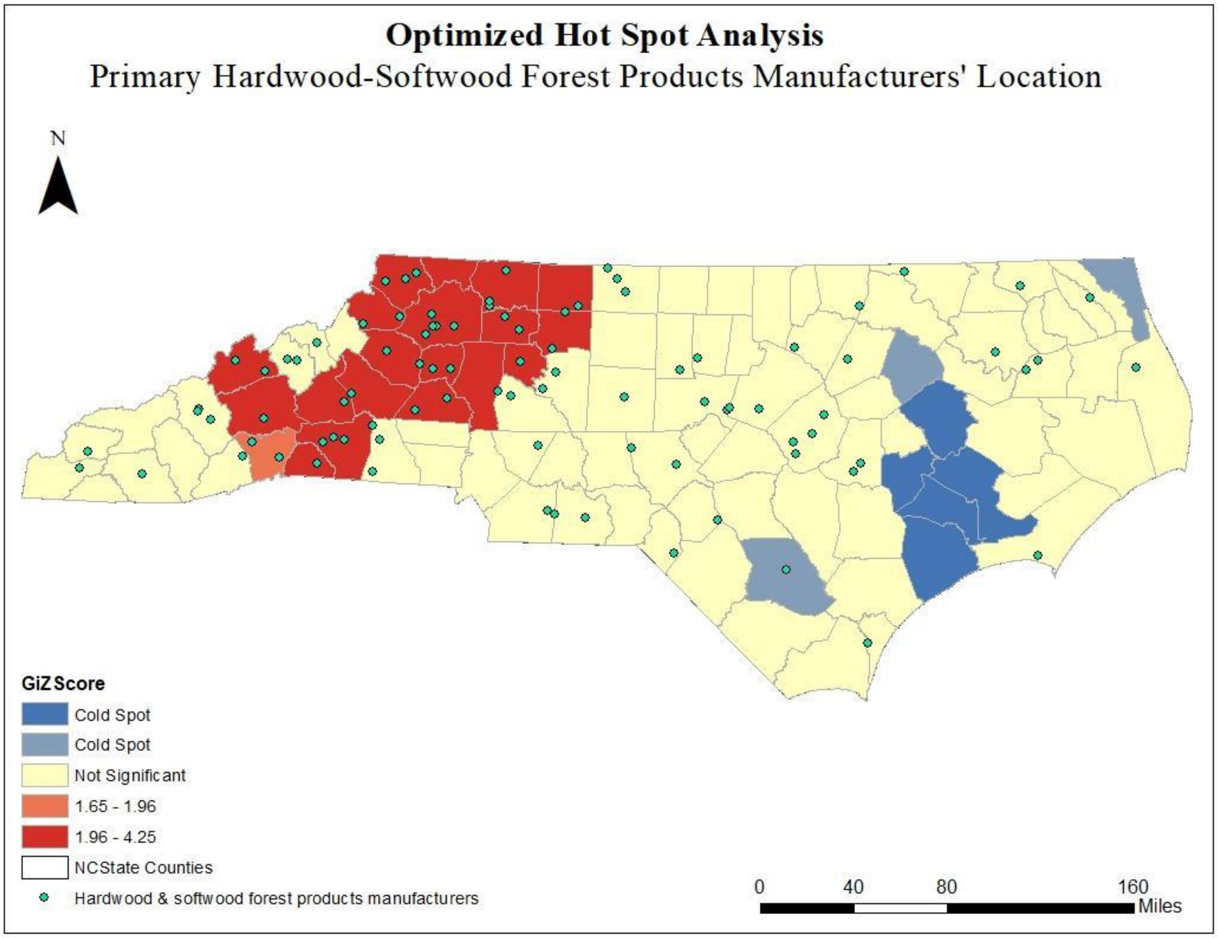

Figure 6. Identified hot spots and cold spots locations of primary hardwood-softwood forest products manufacturers in NC.

4.2.5 Optimized Hot Spot Analysis of Secondary Forest Products Manufacturers by Location

All 566 secondary forest products manufacturers' locations data passed the initial data assessment test and had no outliers. Outliers are the point locations that do not fall within the defined neighbors used in computing this analysis. The incident aggregation analysis performed on all polygons (100 counties) generated a minimum, maximum, and mean incident count values of 0, 59, and approximately 6 manufacturers per county, respectively. One statistically significant hot spot of secondary forest products manufacturers was identified around the northwest, southwest, and Piedmont region of the state. The hot spot comprises of 17 major (Gi* z-score > 1.96) and 5 minor (Gi* z-score, 1.65 > z < 1.96) statistically significant county clusters at the 95 percent and 90 percent confidence level, respectively (Figure 7). The major counties include Mitchell, Watauga, Wilkes, Caldwell, Burke, Cleveland, Gaston, Lincoln, Alexander, Catawba, Yadkin, Iredell, Mecklenburg, Davidson, Randolph, Stanly, and Rowan. Hence, the optimized hot spot analysis of secondary forest products manufacturers revealed a clustering spatial distribution pattern (deviation from complete spatial randomness). In Figure 7, the red colored features on the map represent the hot spot counties where high-incident counts are clustered while the blue colored features represent the cold spots counties where low-incident counts are clustered.

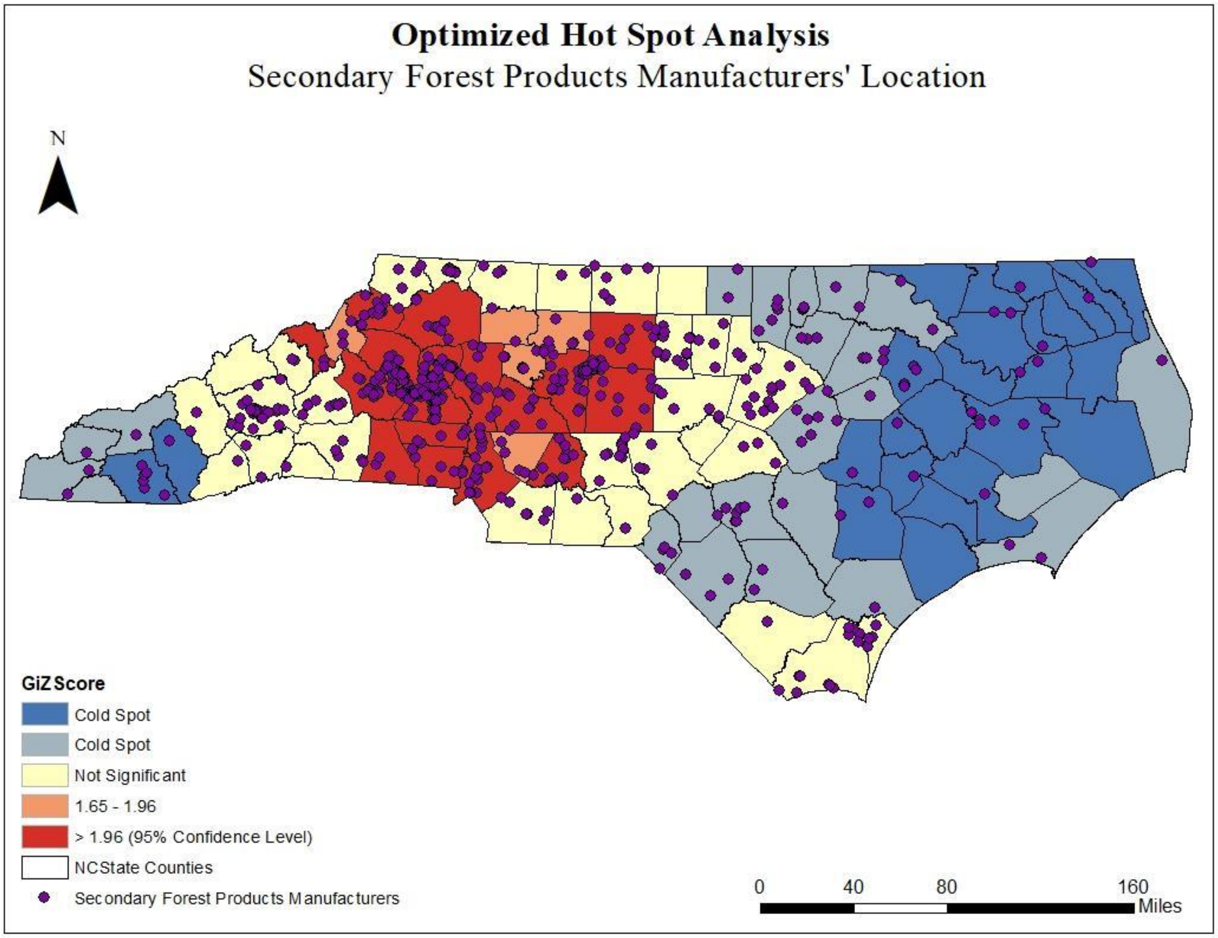

Figure 7. Identified hot spot and cold spot counties of secondary forest products manufacturers in NC.

4.2.6 Optimized Hot Spot Analysis of Primary and Secondary Forest Products Manufacturers by Location

A total of 784 primary and secondary forest products manufacturers' locations in 96 of 100 (96%) counties passed the initial data assessment test. The incident aggregation analysis performed on all polygons (100 counties) generated a minimum, maximum, and mean incident count values of 0, 65, and approximately 8 manufacturers per county, respectively. One large statistically significant hot spot of primary and secondary forest products manufacturers' clusters was identified around the west, northwest, southwest, and Piedmont region of the state. The hot spot consists of 42 counties with a Gi* z-score greater than 1.96 (statistically significant at the 95 percent confidence level) (Figure 8). Hence, the optimized hot spot analysis of the combined primary and secondary forest products manufacturers revealed that their distribution pattern is clustered, showing deviation from complete spatial randomness. In Figure 8, the red colored counties on the map represent the hot spots where high-incident counts of both manufacturers are clustered while the blue colored counties represent the cold spots where low-incident counts are clustered.

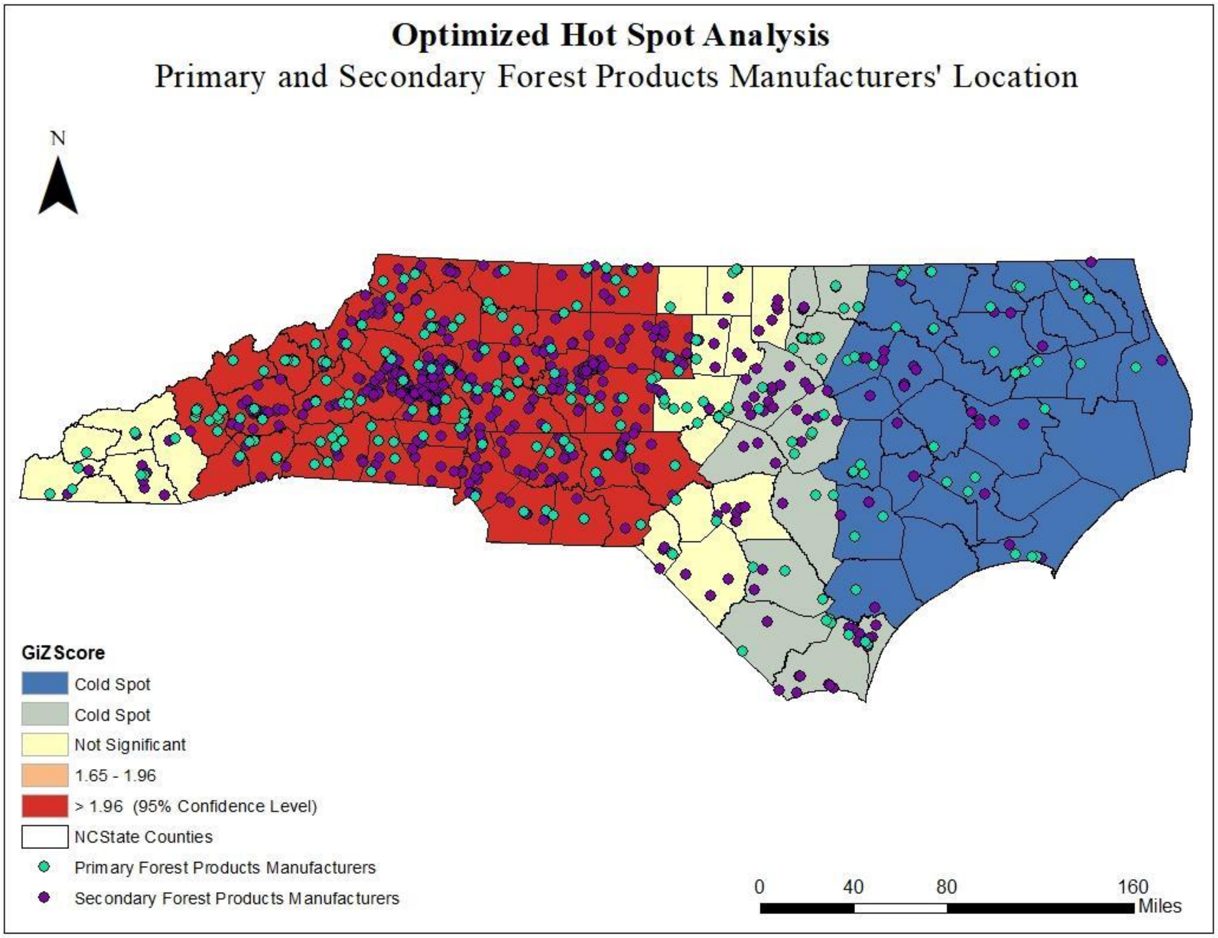

Figure 8. Identified hot spot and cold spot counties of primary and secondary forest products manufacturers in NC based on incident locations.

4.2.7 Optimized Hot Spot Analysis of Hardwood and Softwood Net Merchantable Bole Volume

The optimized hot spot analysis of the net merchantable bole volume of growing-stock trees on private timberland in North Carolina revealed a heterogeneous distribution pattern of hardwood and softwood species (Figure 9). On the map, the dark red colored hot spot counties around the northeast and southeast regions of the study area suggested spatial clustering of the net merchantable bole volume of softwood growing stock trees in NC. Similarly, the blue colored hot spot counties around the northwest and western regions of NC revealed spatial clustering of the net merchantable bole volume of hardwood growing stock trees (Figure 9). Only hot spots statistically significant at the 95 percent confidence level (Gi* z-score > 1.96) were considered in determining the counties where hardwood and softwood species are available and accessible, and for further analysis.

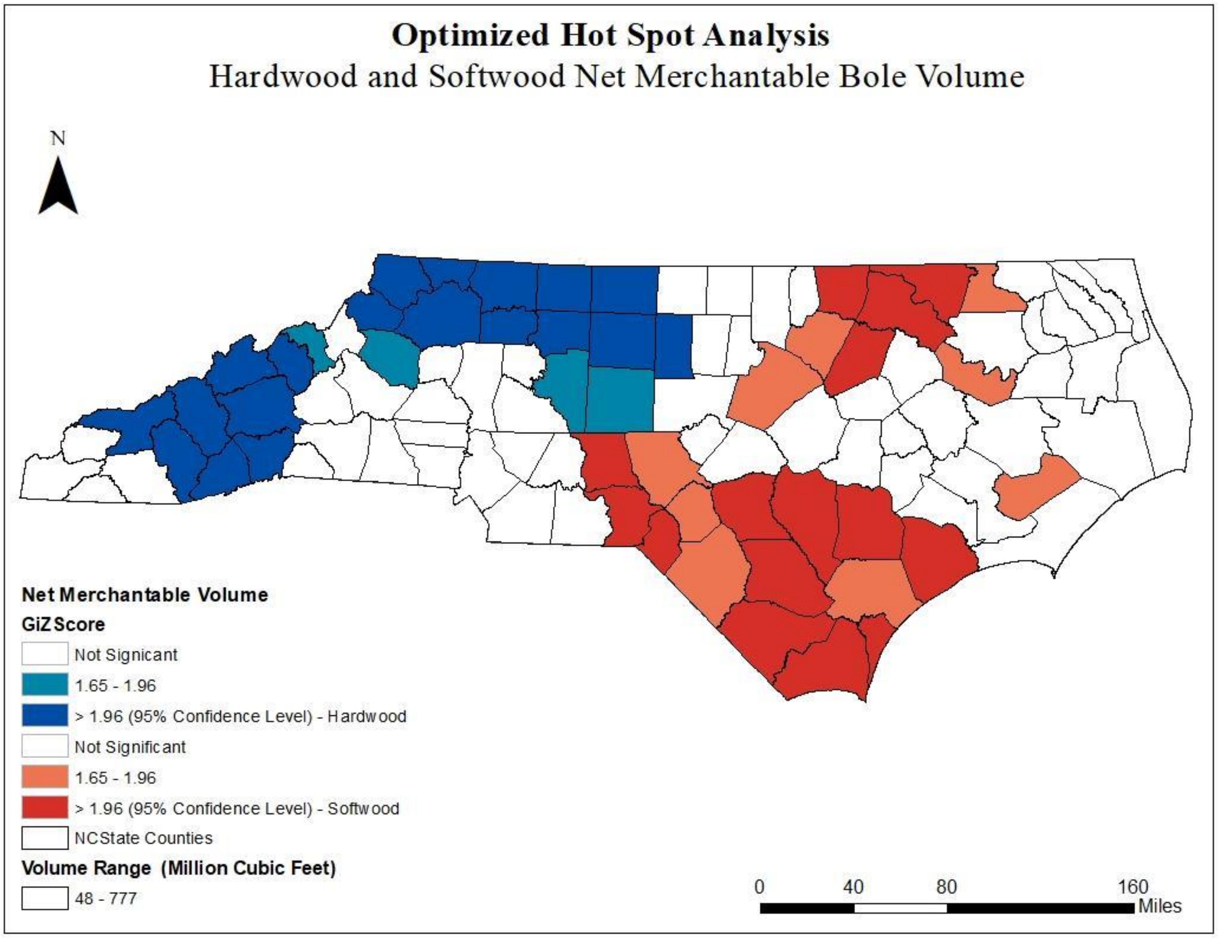

Figure 9. Identified hot spot counties of net merchantable bole volume of hardwood and softwood growing-stock trees in North Carolina.

4.2.8 Optimized Hot Spot Analysis - Hardwood and Softwood Average Annual Harvest Removals of Merchantable Bole Volume of Growing-Stock Trees

Only 90 counties had softwood harvest removals data while 95 counties had hardwood harvest removals data. The optimal fixed distance band of average annual harvest removals of softwood was calculated based on the peak clustering at 298,163 US feet. On the other hand, the optimal distance of average annual harvest removals of merchantable bole volume for hardwood was calculated based on the average distance to four nearest neighbors at 157,114 US feet.

Based on the optimized hot spot analysis of average annual harvest removals of merchantable bole volume of growing-stock trees on private timberlands in North Carolina, results showed that there is a homogeneous spatial distribution pattern of hardwood and softwood species. On the map, the dark red colored hot spot counties around northeast and southeast regions of NC indicate evidence of deviation from complete spatial randomness (clustering) of the average annual harvest removals of merchantable bole volume for softwood species (Figure 10). Similarly, the blue colored hot spot counties around the northeast and southeast regions of NC show evidence of deviation from complete spatial randomness (clustering) of the average annual harvest removals of merchantable bole volume for hardwood species (Figure 10). Martin county is a statistically significant hot spot of intersection of harvest removals merchantable bole volume of hardwood and softwood species in NC. Only hot spots statistically significant at the 95 percent confidence level (Gi* z-score > 1.96) were considered in determining the counties where harvest removals of merchantable bole volume of hardwoods and softwoods species are occurring, and for further analysis.

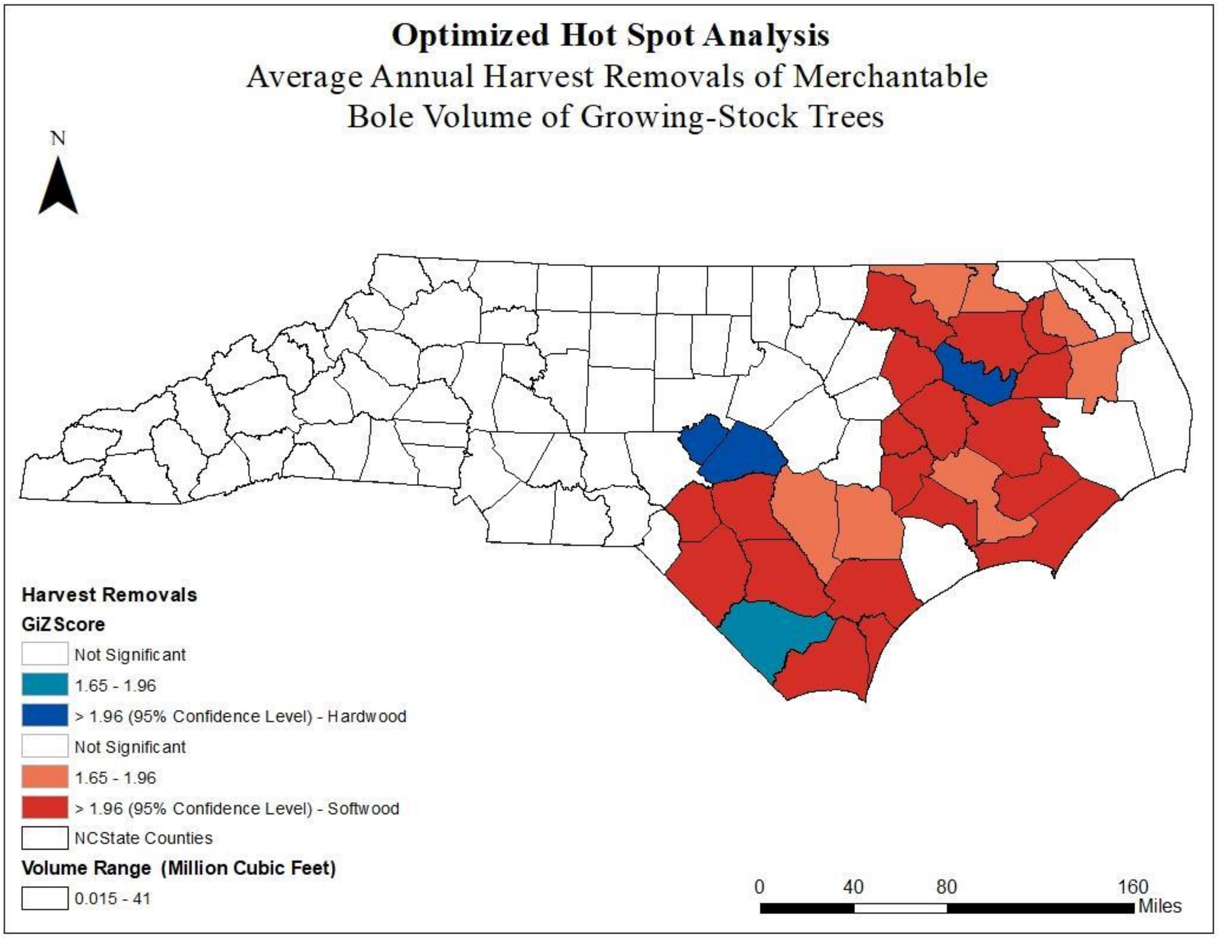

Figure 10. Identified hot spot counties where average annual harvest removals of merchantable bole volume of hardwood and softwood growing-stock trees in North Carolina.

4.2.9 Optimized Hot Spot Analysis of Sawmills and Other Wood-Using Mills (by Location and Their Production Capacity) and Interpolation Analysis Sawmills

4.2.9.1 Sawmills

Only 133 sawmills in our dataset had their annual production capacity data. Results of the optimized hot spot analysis of sawmill locations by their average annual production capacity values differed largely from the results of the optimized hot spot analysis of sawmills by only their incident locations. The hot spots areas (green colored) were identified in the Piedmont and Coastal plains (except Brunswick, and some parts of Columbus, Bladen, and Pender) regions of NC while the cold spots areas (blue colored) were identified in the Mountains. Hence, sawmills in NC tend to cluster around the Piedmont and Coastal Plains with respect to their production capacity values, relative to their clustering pattern by incident locations.

Since the average production capacity values of sawmills were analyzed, results of the analysis were particularly based on the production capacity values of each sawmill by location. Due to this, z-scores values generated from the optimized hot spot analysis were then interpolated across counties. Figure 11 shows the map result generated by overlaying the optimized hot spot analysis of sawmills by location on the interpolated optimized hot spot analysis of sawmills by their average production capacity values.

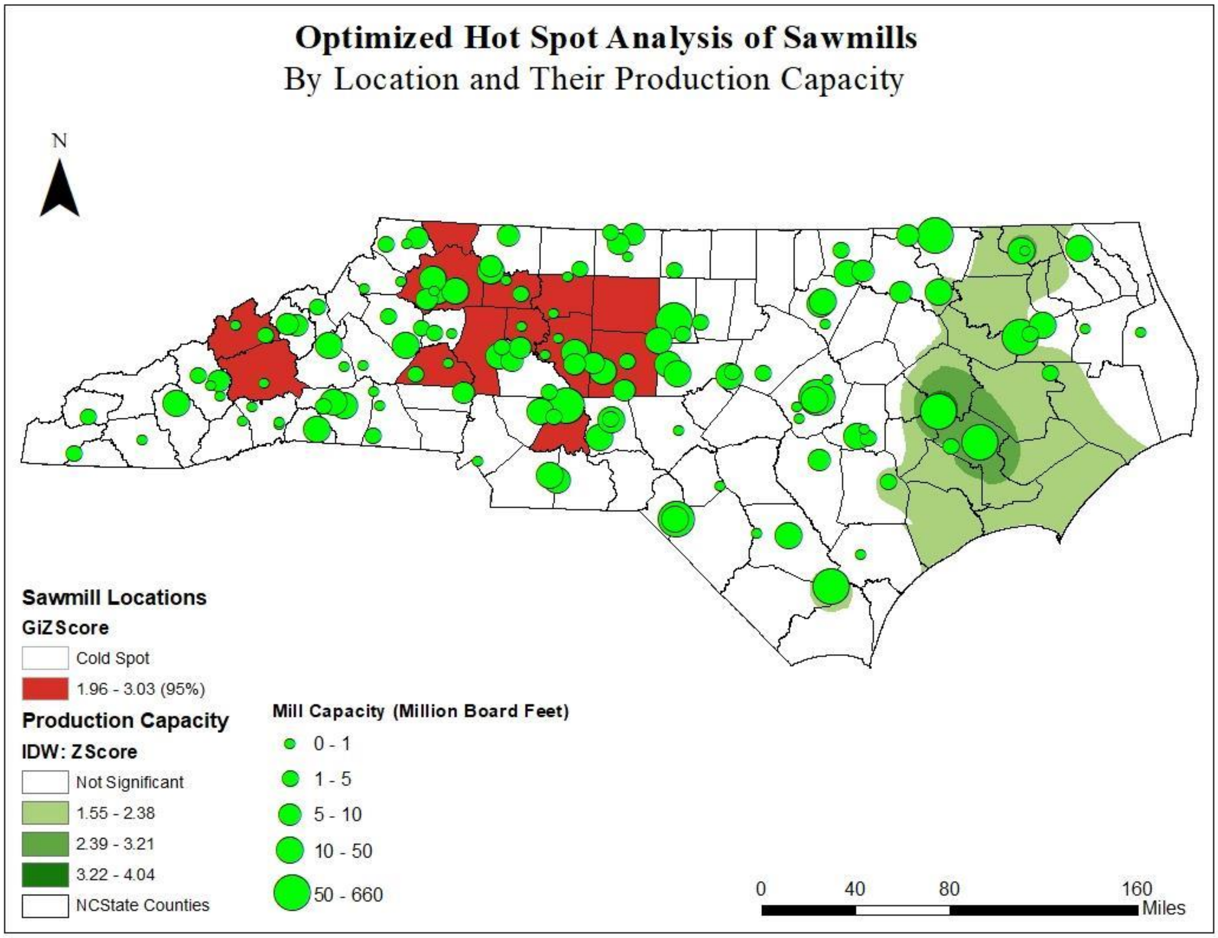

Figure 11. Identified hot spot of primary forest products manufacturers by location and production capacity.

4.2.9.2 Other primary wood-using mills

Other wood-using mills in this study are defined as all primary mills except sawmills including chip mills, log exporters, timber buyers, veneer mills, and woodyard. A total of 39 other wood-using mill locations' average production capacity data passed the initial data assessment test. The optimized hot spot analysis result of other wood-using mill locations based on their average production capacity values contrastingly differ from the results of the optimized hot spot analysis of other wood-using mills by their incident locations (Figure 12). The hot spots counties (green colored) are identified around the western, northwest and north central regions of NC and are statistically significant at the 90 percent confidence level. Based on our spatial pattern analysis, other wood-using mills in NC tend to cluster around the Mountain and northern parts of the Piedmont region with respect to their production capacity sizes, relative to their clustering pattern by incident locations.

Furthermore, since the average production capacity values of other wood-using mills were analyzed, results of these analyses were particularly based on the production capacity values of each other wood-using mills by location. Due to this, z-scores values generated from the optimized hot spot analysis were then interpolated across counties. Figure 12 shows the map result generated by overlaying the optimized hot spot analysis of other wood-using mills by location on the interpolated optimized hot spot analysis of other wood-using mills by their average production capacity values and their production capacity sizes.

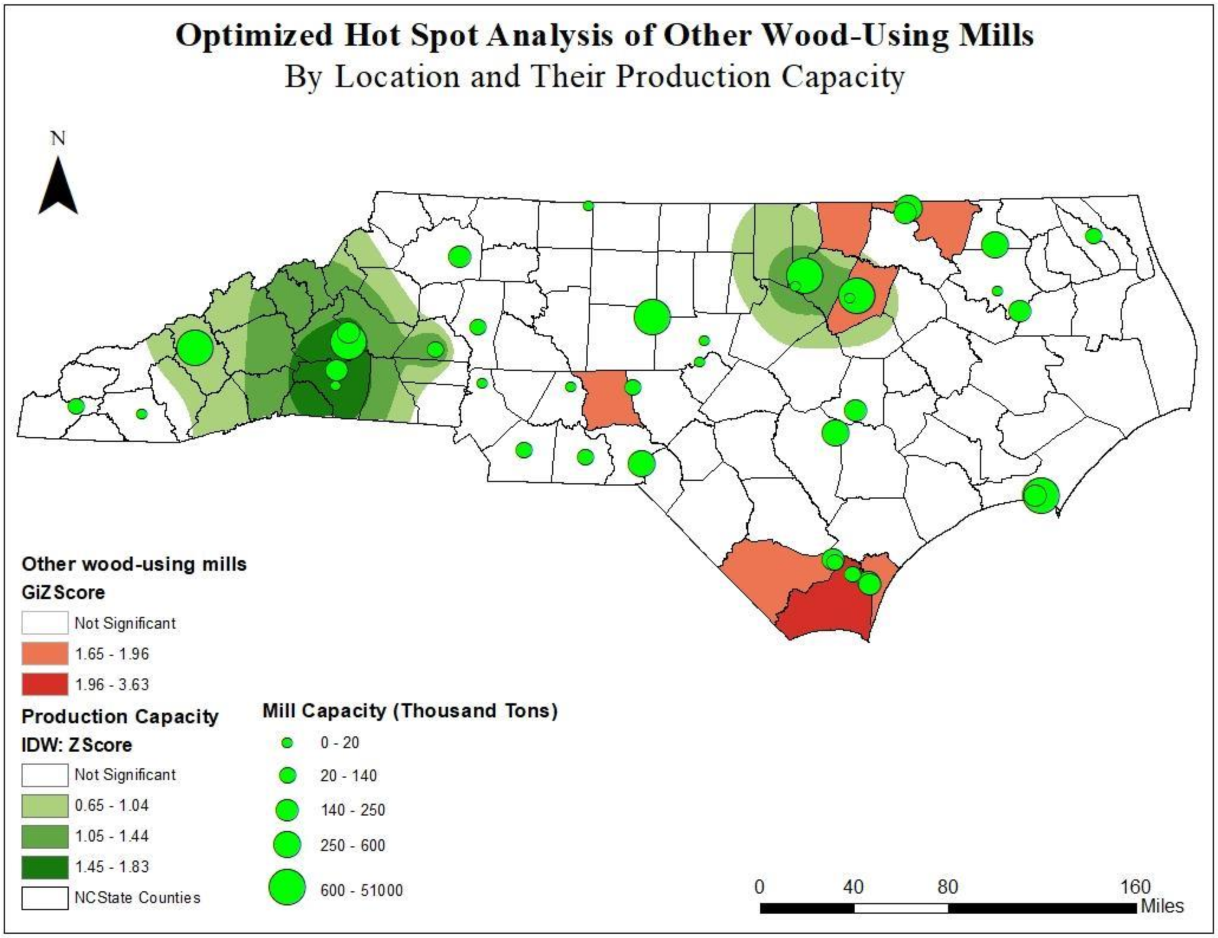

Figure 12. Identified hot spot of primary forest products manufacturers by location and production capacity.

4.3 Factors Influencing the Location of Forest Product Manufacturers in NC

Factors influencing the location of forest products manufacturers in NC were assessed using the Poisson regression model. Table 2 shows the descriptive statistical summary of variables used in the model. The number of primary forest products manufacturers in a county varied from 0 to 9 with an average value of 1.94. In addition, 79 of 100 counties had at least one primary forest products manufacturer within the county. Similarly, the number of secondary forest products manufacturers in a county varied from 0 to 64 with an average count of 5.67. Also, 86 of 100 counties hosted at least one secondary forest products manufacturer. Figure 13 and 14 show the frequency distribution of primary and secondary forest products manufacturers' data used in modeling the factors influencing the location of forest products manufacturers in NC.

4.3.1 Results of the Poisson Regression models examining the Factors Influencing the Location of Primary Forest Products Manufacturers in North Carolina

Results of the likelihood ratio test of alpha (over-dispersion parameter) provide evidence of a better fit of Poisson regression over negative binomial regression with the coefficient of alpha not significantly different than zero (Table 3). The Poisson maximum likelihood estimation results for primary forest products manufacturers are reported in Table 4. A maximum likelihood ratio of -149.73 shows a strong statistical significance of the explanatory variables ($\chi^2 < 0.0000$). The average predicted number of primary forest products manufacturers per county in NC was 1.94 (yhat). All factors were statistically significant ($p \leq 0.01$, 0.05 and 0.10) except *road* and *economic tiers* in influencing the location of primary forest products manufacturers in a county. The positive and negative effect of each factor on the location of primary forest products manufacturers in a county is given by the coefficient signs in Table 3.

Table 2. Descriptive statistics summary of variables used in analyzing factors influencing the location of forest products manufacturers in NC (n = 100).

Variable	Mean	Std. Dev.	Min	Max
Primary mills	1.94	2.03	0	9
Other primary mills	0.32	0.47	0	1
Secondary mills	5.67	9.04	0	64
Road	0.45	0.50	0	1
Rail	0.85	0.36	0	1
Port	0.11	0.31	0	1
Population greater than 50k	0.56	0.50	0	1
	390825.2	185133.2		
Total mech. volume	0	0	56712	832206
Unemployment	1984.23	3318.25	93	23252
Economic tiers	1.79	0.76	1	3
Sawlog harvest	3373.91	2870.11	0	15015
Pulpwood harvest	3794.77	3742.39	0	16948

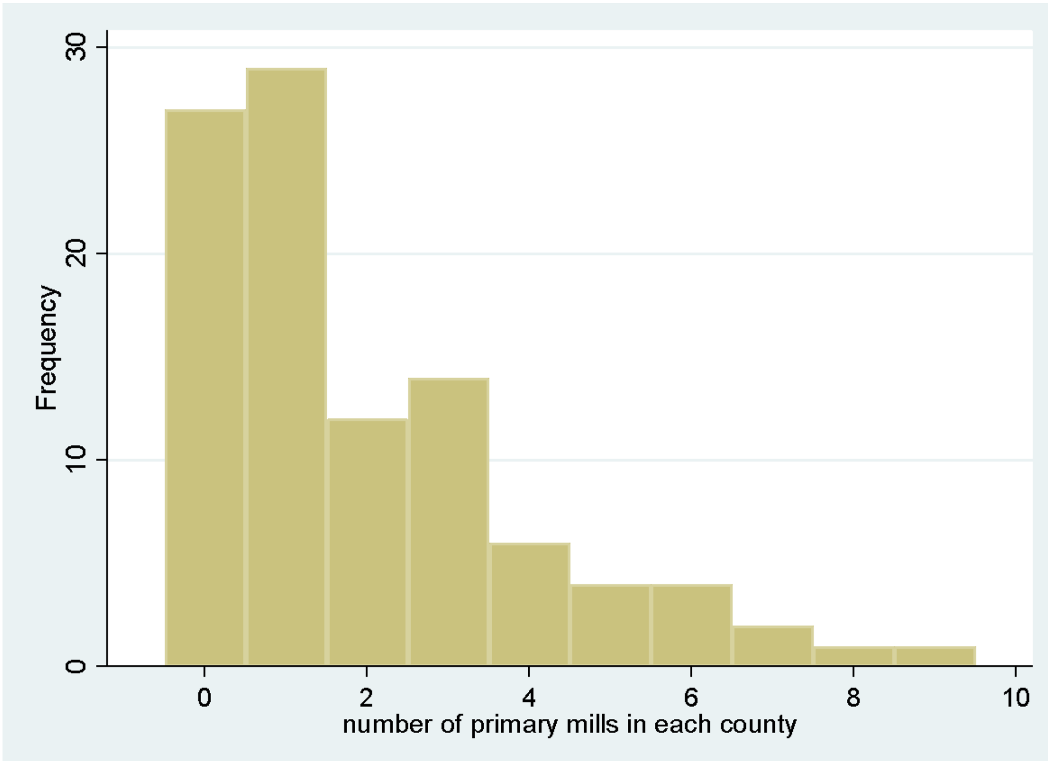

Figure 13. Distribution of primary forest products manufacturers in the dataset used for the Poisson regression.

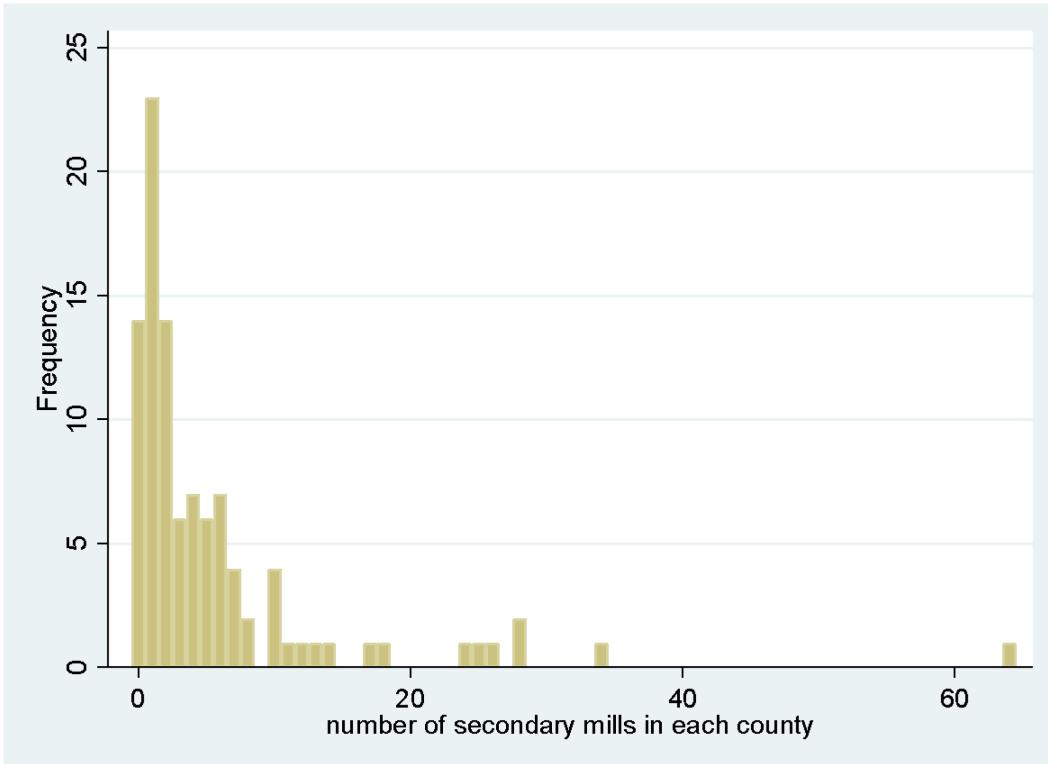

Figure 14. Distribution of secondary forest products manufacturers in the data set used for the Poisson regression.

Table 3. Negative binomial regression used in testing model fitness.

Factor	Coefficient	Std. Err.	z	P>z
Secondary mills	0.6372	0.2978	2.14	0.032
Road	0.1555	0.3946	0.39	0.694
Rail	0.1550	0.2671	0.58	0.562
Port	-0.4613	0.2766	-1.67	0.095
Population greater than 50k	0.3810	0.1904	2.00	0.045
Unemployment	0.0000	0.0000	-1.18	0.238
Total merch. volume	0.0013	0.0004	3.02	0.002
Economic tiers				
2	0.2712	0.1734	1.56	0.118
3	0.4178	0.2086	2.00	0.045
Constant	-0.8895	0.5160	-1.72	0.085
/lnalpha	-2.6284	1.0863		
alpha	0.0722	0.0784		

Log-likelihood Ratio test of alpha = 0: χ^2 (01) = 1.11 Prob >= χ^2 = 0.146

Table 4. Poisson regression results examining the factors influencing the location of primary forest products manufacturers in NC.

Factor	Coefficient	Std. Err.	z	P>z	IRR
Secondary mills	0.62	0.28	2.19	0.028**	1.8684
Road	0.15	0.37	0.40	0.691	1.1585
Rail	0.16	0.25	0.64	0.523	1.1718
Port	-0.48	0.26	-1.87	0.062*	0.6162
Population greater than 50k	0.39	0.18	2.24	0.025**	1.4809
Unemployment	0.00	0.00	-1.34	0.179	0.9999
Total merch. volume	0.00	0.00	3.30	0.001***	1.0013
Economic tier					
2	0.27	0.16	1.70	0.089*	1.3108
3	0.44	0.19	2.33	0.020**	1.5487
Constant	-0.878	0.49	-1.80	0.07	0.4150

* P < 0.1, ** P < 0.05, *** P < 0.01. n (counties) = 100. Poisson model: χ^2 (df 9) log - likelihood ratio test = -180.05, probability > χ^2 < 0.0000. Average predicted count of primary forest products manufacturers per county = 2.32

Furthermore, the results of the Poisson regression analysis of location of primary forest products manufacturers in North Carolina indicated that the presence of secondary mills, cities with population greater than 50,000, total merchantable volume of timber in a county, and a county with a distress ranking of 2 or 3 were the most important factors influencing the location of primary forest product manufacturers in NC (Table 4). The existence of one or more secondary forest products manufacturers in a county increased the expected probability of hosting primary forest products manufacturers by 87%. Similarly, the presence of a city with a population of at least 50,000 people in a county increased the expected number of primary forest product manufacturers by 15.85%. A thousand cubic feet increase in the average merchantable bole volume of timber in a county increased the predicted number of primary manufacturers by a small effect of 0.13%. Also, any county with a distress ranking of 2 or 3 is expected to experience an increase in the number of primary forest products manufacturers by 31.08% and 54.87%, respectively. However, the presence of ports within a 100 miles drive distance from a county actually decreased the expected number of primary forest products manufacturers by 61.62%. Our results indicated that the presence of interstate highway (road), railways, and unemployment had no statistically significant effect on the location of primary manufacturers in a particular county in NC.

4.3.2 Results of Poisson Regression Estimation of Factors Influencing the Location of Secondary Forest Products Manufacturers in North Carolina

The Poisson maximum likelihood estimation results for secondary forest products manufacturers are reported in Table 5. A maximum likelihood ratio of -335.31 shows a strong statistical significance of the explanatory variables ($\chi^2 < 0.0000$). The average predicted number of secondary forest products manufacturers per county in NC was 5.67 (yhat). All selected factors but railways were statistically significant ($p \leq 0.01$, 0.05 and 0.10) in influencing the location of secondary forest products manufacturers in a county. The positive and negative effects associated with factors related to the location of secondary forest products manufacturers in a county are given by their coefficient signs (Table 5). Holding all other factors constant, the presence of existing primary forest products manufacturers in a county increased the expected number of secondary forest product manufacturers in a county by 11.01%. Every city in a county with a population of at least 50,000 people is predicted to increase the presence of a secondary manufacturer by 2.84 times (or 168%). Similarly, a high volume of sawlog harvest, and higher unemployment increased the predicted number of secondary manufacturers in a county by 8.80% and 0.01% respectively. In addition, compared to the distress ranking of 1, any county with a ranking of distress ranking rate of 2 and 3 increased the predicted number of secondary manufacturers in a county by 2.36 times (or 72%) and 67%, respectively. In contrast, the presence of roads as captured by interstate highways, ports, and pulpwood harvest decreased the predicted number of secondary manufacturers in a county by 38%, 44.67%, and 91.65%, respectively.

Table 5. Negative binomial regression used to model factors influencing the location of secondary forest products manufacturers in NC.

Factor	Coefficient	Std. Err.	z	P>z	IRR
Primary mills	0.1044	0.0575	1.82	0.069	1.1101
Road	-0.9677	0.4182	-2.31	0.021	0.3800
Rail	0.3581	0.3246	1.10	0.270	1.4306
Port	-0.8058	0.3922	-2.05	0.040	0.4467
Population greater than	1.0438	0.2607	4.00	0.000	2.8401
Sawlog harvest	0.0843	0.0440	1.92	0.060	1.0880
Pulpwood harvest	-0.0871	0.0366	-2.38	0.017	0.9165
Unemployment	0.0001	0.0000	1.75	0.080	1.0001
Economic tier					
2	0.8604	0.2456	3.50	0.000	2.3642
3	0.5147	0.3104	1.66	0.097	1.6731
Constant	0.6539	0.4710	1.39	0.165	1.9230
/lnalpha	-0.5296	0.1981			-0.5296
alpha	0.5889	0.1167			0.5889

Log-likelihood Ratio test of alpha = 0: χ^2 (01) = 235.68. Prob >= χ^2 = 0.000 * P < 0.1, ** P < 0.05, *** P < 0.01. n (counties) = 100. Poisson model: χ^2 (df 11) log likelihood ratio test = -335.31, probability > χ^2 < 0.0000. Average predicted count of secondary forest products manufacturers per county = 5.67

CHAPTER 5

DISCUSSION AND CONCLUSION

5.1 Discussion

Identifying clusters aids to enhance a clear understanding of contemporary patterns, industrial transformation processes, competition, and regional economic development (Hallencreutz and Lundequist 2003). Understanding the spatial distribution pattern of the forest products industry is crucial to the efficient and sustainable management of forest resources, and their movement through the supply chain network. Several past studies examined the spatial clustering of industries (Marshall 1920, Weber 1929, Porter 1998, Hallencreutz and Lundequist 2003, Fujita and Krugman 2004), and specifically the forest products industry (Porter 2003, Aguilar and Vlosky 2006, Aguilar 2009, Aguilar et al. 2009, Kies et al. 2009, Hagadone and Grala 2012, Brandeis and Hodges 2018, Kaur et al. 2020). These factors influencing industrial clusters include access to- and cost of transportation, proximity to consumers' location or market, nearness to sources of raw materials, availability of labor, demand conditions, presence of related and supporting industries, technological developments, and raw materials consumption capacity (Aguilar et al. 2009, Hagadone and Grala 2012, Aguilar 2008).

The knowledge of factors impacting forest products manufacturers' location, concentration and agglomeration can help inform policy and decision makers, and guide stakeholders on what, where and how to plan towards siting and operating the forest products industry sustainably. Ross (1896) explained that the power of a locality to sustain an industry far exceeds the original power of attraction. It is therefore central that a locality must excel enough not only in attracting an industry but also in sustaining the continuous operation of an industry. This study examined the spatial distribution of forest products manufacturers, major business hot spots, potential business clusters, and factors influencing the location of primary and secondary forest products manufacturers' in NC.

5.2 Primary and Secondary Forest Products Manufacturers

Co-agglomeration of primary and secondary forest products industry in NC showed high concentration of interrelated companies in the northwest, southwest and Piedmont region of NC (Figure 7). However, there are fewer dispersed primary and secondary manufacturers in the eastern part of the state. This industry includes sawmills, timber buyers, veneer mills, log exporters, plywood mills, and dominated largely by furniture stores and manufacturers.

Spatial co-agglomeration pattern has the ability to support regional supply chains as a potential source of competitive advantage (Kies et al. 2009). Geographical concentration of primary and secondary manufacturers that encourages collaboration can foster competitive industrial structure (Aguilar 2008). The identified hot spot of primary and secondary manufacturers consists of 42 counties that are connected at close proximity to major interstate highways (I-40, I-77, I-73, I-85, I-285), and railroad lines. Availability and access to transportation, regional market, labor, and knowledge exchange are therefore most likely to have an effect on the observed spatial arrangement as a source of competitive advantage in the hot spot region (Figure 7). According to Murray (1995), spatial co-agglomeration of primary and secondary manufacturers can enhance gains in efficiencies and cost reductions. For instance, sawmills can benefit from economies of scale.

Identification of primary and secondary forest products manufacturers' hot spots by incident location analysis validates the locational hot spots of forest resources (Figure 8 and 9). The primary mills are clustered close to the western region where raw materials are abundant, and the secondary mills are concentrated majorly around the highly populated counties near their consumers. This finding provides strong evidence of deviation from complete spatial randomness in the primary and secondary forest products industry. However, since our scale of analysis was set at county level, further analysis is required at the census tract level to identify factors peculiar and most favorable to each county within the identified hot spots on the maps. Also, our results reinforce some of the features that foster cluster development as explained by Braden et al. (1998). Thus, this information can serve as a guide to the forest products industry in making decisions that could potentially lead to successful forest industry economic development in North Carolina.

Also, the regression analysis of the primary forest products manufacturers suggested that the presence of secondary forest products manufacturers in a county affected the location of primary forest products manufacturers. Similarly, counties with an active presence of primary manufacturers were more likely to host secondary forest products manufacturers. This finding reinforced the co-agglomeration of primary and secondary manufacturers as seen in Figure 7. Although a raw material dependent industry like the lumber industry would prefer to locate near raw materials instead of the near final market (Weber 1929), Porter (1998) noted that the creation of a business environment is contingent on the demand condition amongst other factors. The interaction between the primary and secondary forest products manufactures' is therefore important for the sustainable flow of supply in the forest product industry. Several studies did not account for how the presence of secondary forest products manufacturers in a region can influence the location of primary forest products manufacturers in their models (Aguilar and Vlosky 2006, Aguilar 2009, Aguilar et al. 2009, Hagadone and Grala 2012). Although Aguilar and Vlosky (2006) found evidence of a geographical coincidence between primary wood products industry and final market based on population density, their study was limited in determining the connection between primary and secondary manufacturers.

5.3 Spatial Distribution of Primary Forest Products Manufacturers in NC

Based on the locational pattern analysis, results suggested that the primary forest products manufacturers' location exhibited spatial clustering in the western and northwest regions and are largely dispersed in other parts of the state of NC. The identified hot spots in Figure 2A and 2B (consist of 7 counties: Madison, Buncombe, Henderson, Rutherford, Burke, Caldwell, and Catawba) are dominated by sawmills and a few other primary wood-using mills that mainly utilize hardwood species for their manufacturing operations. Mapped hardwood merchantable volume hot spots corroborated the locational analysis findings (Figure 8). These revealed that primary manufacturers clustered lightly towards western NC where there is an abundance of hardwood tree species which is near the identified hot spots of hardwood merchantable volume. Also, it indicates that primary forest products manufacturers tend to locate close to the source of raw materials which aids in reducing the cost of transportation (Porter 2003).

Our results align with the findings of Aguilar et al. (2009) which identified spatial clustering in the primary wood products industry in the US South and highlighted their dependence on proximity to forest resources. The observed spatial pattern is also consistent with the clustering pattern identified in the primary forest products manufacturing industry in southwest Mississippi (Hagadone and Grala 2012). Hagadone and Grala (2012) argued that the southwest counties in Mississippi offered favorable conditions that might have resulted from abundance of input resources, positive business environment, and suitable infrastructure or historical presence of the industry. Similarly, Aguilar and Vlosky (2006) also found the evidence of spatial dependency of primary forest products manufacturers in Louisiana.

The results of primary-hardwood, primary-softwood, and both hardwood and softwood using primary forest products manufacturers' locational analyses are discussed in the following subheadings with respect to net merchantable bole volume, and average annual harvest removals of merchantable bole volume hot spots respectively.

5.3.1 Primary-Hardwood Forest Products Manufacturers

None of the three identified hot spots in Figure 3 intersected with the hot spots in Figure 2A and 2B. This suggests that primary-hardwood forest products manufacturers are clustered more around counties (Davidson, Guilford, Alamance, Randolph, Nash and Warren) in the Piedmont than in the Mountains region of NC. Two of the six counties within the identified hot spots have no primary-hardwood manufacturers within the boundary of those counties. This implies that those two counties might most likely have potential for the development of primary forest products industry.

Harvesting is quite challenging in the western NC. Logging, particularly in the undulating terrain of the mountain region, is less accessible and riskier for loggers to log and haul the harvested products (Powell 2006). In addition, the Mountain region in the western NC has more public and reserved forestland, where logging operations are less frequent. Emerging environmental policies, high level of recreational use of public forestlands retained within urban landscape, restrictive water quality standards, concerns for aesthetics, and smoke restrictions are generally identified as factors limiting access to timber harvests in NC (Bardon et al. 2002).

Although on a per acre basis, most standing inventory is in the Mountains, however, the Piedmont has the largest area of timberland next to the Coastal Plains (McConnell et al. 2016). About 74 percent of the timberland in the Piedmont is hardwood forest (Powell 2006), including oak-hickory, oak-pine, and oak-gum-cypress combinations which are among the main hardwood species used by the primary hardwood manufacturers. The map of the annual harvest removals of merchantable bole volume (Figure 9) is consistent with findings in Figure 3; one hot spot (Lee County) at the 95 percent confidence level, and two hot spots (Warren and Chatham county - not mapped) at the 90 percent confidence level were identified in the Piedmont region. Other factors such as transportation cost, limited availability of accessible raw materials, and legacy effects may also have an impact on the spatial dependent pattern observed in primary forest products manufacturing industry in NC (Goyette 1967).

Higher growth rate and better market access in the Piedmont than in the Mountains may also have an effect on these spatial distribution patterns of hardwood manufacturers (McConnell et al. 2016).

When the interstate highway and railroad shapefile(s) were overlaid over Figure 3, the identified hot spots intersected with several interstate highways (I-285, I-74, I-40, I-85, I-95, I-73, and I-785) and railroad networks. In addition, the identified clusters are in the highly populated counties in NC; Davidson, Randolph, and Alamance have population size of over 140,000 people with the highest population in Guilford with 537,174 residents (U.S. Census Bureau 2019).

However, the regression analysis showed that the presence of interstate highway, railroad, and unemployment have positive coefficients but did not significantly have an effect on the location of primary forest products manufacturers in a county. Hagadone and Grala (2012) noticed that the presence of four-lane interstate highways did not significantly impact the location of primary forest products manufacturers. However, Aguilar et al. (2009) found that the presence of highway interstate had a significant and direct relationship with the presence of sawmills in counties across the US South lumber industry. Possible reasons for mixed results might be due to the fact that the interstate highway variable was coded as binary and did not account for closest interstate highway to a primary manufacturer. In other words, there is also a possibility that some counties which do not have an interstate and were coded as zero might be within a close distance of interstate highways in neighboring towns or county, or even have other roads facilitating transportation of logs not accounted for by our data.

Also, railroad might be insignificant due to the fact that, in the west, the majority of the primary manufacturers are small in size and the cost of moving logs by railroad may not be economical for their cost of production. About 50 to 70 percent of wood arrived at mills by rail in the post-world war era in NC for several reasons (Goyette 1967). Goyette (1967) explained that this was due to the large production capacity size of the mills and variation in the amount of wood required, increasing procurement distance, mountainous and challenging topography in the West, and the need for consistent volume flow of wood to the mills. Thus, the size of the mills and amount of raw material required by a mill may have an effect on the influence of railroad on the dependent variable.

The presence of ports within 100 miles drive distance from a county decreased the likelihood of hosting primary manufacturers in such a county. Since the majority of primary forest products manufacturers are located in the Mountains and Piedmont regions away from the coast, the presence of ports might not have had much effect on the location of these manufacturers. The spatial analysis findings indicated that the hardwood mills are clustered near the hardwood resources which happens to be away from the ports. Also, most of the mills were not established for exports; this might also be a possible explanation for the significant negative effect of the port variable. Austin and Darr (1975) reported that the share of logs transported to the forest industry by water has declined steadily. Aguilar (2008) considered proximity to port as a less important variable affecting the location of the softwood lumber industry in the US.

5.3.2 Primary-Softwood Forest Products Manufacturers

The primary-softwood forest products manufacturers exhibited a spatial distribution pattern contrasting to the pattern observed when all primary forest products manufacturers were analyzed jointly. All four hot spots were identified in the northern and southern Coastal Plains consisting of six counties in NC. Since the Coastal Plains have the largest area of timberland dominated by softwood species in NC (McConnell et al. 2016), this suggests that primary-softwood manufacturers tend to locate at close proximity to the sources of raw material**.** All identified primary-softwood manufacturers hot spots intersected with counties where the average annual harvest removals of merchantable bole volume are significant at the 95 percent confidence level (Figure 9). In other words, the primary-softwood forest products manufacturers are located in counties where the average annual harvest removal of softwood timber is high relative to other counties. Similarly, the intersection of the identified primary-softwood hot spots with softwood net merchantable bole volume (Figure 8) confirms that the observed spatial pattern exists in proximity to a promising source of raw material in the Coastal Plains.

Our spatial analysis also intimated that the location of high raw materials dependent primary-softwood industry is most likely influenced by the location of raw materials than final consumers' location (Figure 4 and 10). The spatial pattern exhibited by the primary-softwood forest products manufacturers supports the suggestion of Weber (1929) in that the manufactures operating within the identified hot spots have high production capacity size even in less populated counties. For instance, the primary-softwood manufacturers in Columbus and Brunswick counties have an average production capacity size of 190 and 26.67 million board feet, respectively, while those in Martin County have an average production capacity size of 209 million board feet. In addition, all six counties have a population size less than 56,000 people except Brunswick with a population size of 142,820 people (U.S. Census Bureau 2019). This position is also supported by Aguilar and Vlosky (2006) which found no relationship between the location of primary forest products manufacturers and market location. Our result is also consistent with the findings of Brandeis and

Hodges (2018) which identified a hot spot of sawmill industry where the location exhibiting increased procurement volume of softwoods matched with the areas of mill concentration in West Central Tennessee. It is also important to note that the eastern region is not surrounded by many secondary manufacturers as in the west despite the competitive advantage of easy access to the export market.

All three of six counties within the identified hot spots currently have no presence of any type of primary forest products manufacturers. This might be an indication based on the availability of forest resources that these counties can host a primary-softwood forest products industry sustainably. However, further research is needed to evaluate the implication of the introduction of primary-softwood industry for forest management since new spatial distribution patterns could impact market availability to landowners. Also, none of these counties tally with any of the five counties identified by Aguilar (2009) which is most likely because their study did not distinguish between softwood and or hardwood using mills. Other factors such as procurement distance and cost of transportation may also be a reason for the spatial pattern seen in figure 4.

5.3.3 Primary Hardwood-Softwood Forest Products Manufacturers

We also examined the locational distribution of primary forest products manufacturers which utilize both hardwoods and softwoods species for their manufacturing operations. The identified hot spot, consisting of 19 counties (Figure 5), coincides with the 6 counties identified in the joint locational analysis of all primary manufacturers in Figure 2A and 2B. In spite of the challenging geography in the western NC, the location of primary hardwood-softwood forest products manufacturers revealed deviation from complete spatial randomness (clustering). This pattern is attributed largely to the types of species used by these manufacturers. Some of the common timber species used by the primary hardwood-softwood manufacturers such as red oak, white oak, hickory, yellow poplar, sycamore, white pine, sweetgum, soft maple, to mention a few, cover significant acreage in the Mountain region in NC. According to Powell (2006), although 90 percent of the timberland in the Mountains is covered by hardwood forest, white pine forests dominate in the areas covered by softwoods. Also, the primary hardwood-softwood forest products manufacturers' hot spot falls in the same region as the hot spot of hardwood net merchantable bole volume of growing stock trees since the majority of the mills within the hot spot use more hardwood than softwood species. Two of the identified counties in the hot spot (Figure 5) corresponds with 2 of 5 counties identified by Aguilar (2009) where new developments in the softwood lumber industry could likely occur in NC. It is therefore most likely that the observed pattern is caused by proximity to the source of the dominant species (raw materials) used by these manufacturers.

The significance of the total merchantable volume of growing stock trees in our regression analysis (Table 4) corroborated our spatial analysis findings as expected. An explanation for the small size of the coefficient might be because our regression accounted for the availability of resources in the entire state and not by geographical region. Nonetheless, this provides supporting evidence suggesting that the primary forest product manufacturers' location is raw material dependent and would likely locate close to the source of raw materials in a county (Aguilar and Vlosky 2006). Primary forest products manufacturers enjoy reduced cost of transportation when located close to a consistent log supply area. Sufficient supply of logs and proximity to the log supply area were reported by Aguilar (2008) as one of the most important factors influencing the location of mills in the US Softwood lumber industry. Hagadone and Grala (2012) found that primary manufacturers have the tendency of been attracted to counties abundant in sawlogs.

Availability of labor may have played an important role in the formation of the hardwood-softwood forest products manufacturers' cluster in the Mountains. Considering the population size of the 19 counties within the hot spot in Figure 5, four counties have a population size greater than 100,000 people, nine have a population size less than 50,000 people while the other six counties' population size falls in between the range of 50,000 and 100,000 people (U.S. Census Bureau 2019). Thus, the availability of labor in terms of the population size of each county seems like a strong factor that may have influenced the observed spatial distribution pattern in western NC. This is consistent with the study by Aguilar et al. (2009) who identified access to labor as one of the factors that significantly affect the clustering of primary wood products industry in the US South. Availability and access to labor was explicated by Braden et al. (1998) to be critical to the formation and sustainability of manufacturing industrial clusters.

Contrary to the expectation, the unemployment variable (proxy for availability of labor) in the regression analysis was insignificant with a coefficient of approximately zero in our model. This might be because average unemployment rate is one of the four development factors used in calculating county economic tiers in NC[2]. Since our economic tier variables are both significant,

[2] Labor & Economic Analysis Division Memo – Changes to County Development Tiers for 2019, https://files.nc.gov/nccommerce/documents/files/2020-Tiers-memo_asPublished_120219.pdf

they can be interpreted to account for availability of labor as well. According to our results (Table 4), counties in tier 2 and 3 distress rankings have a higher likelihood of hosting primary forest products manufacturers. That is, every county with an average unemployment rate, median household income, percentage growth in population, and adjusted property tax base per capita, estimated and designated as less and least distress (NC Department of Commerce 2020) counties respectively, have a higher tendency of attracting primary forest products manufacturers. This tier system is incorporated into various state programs to encourage economic activity in the less prosperous areas of the state.

Furthermore, the spatial concentration of primary hardwood-softwood manufacturers in the western NC might be due to competitive advantage derived from firm interactions. The identified hot spot is characterized by clusters of some small sized sawmills, other wood-using primary manufacturers, and a few, albeit larger other sawmills. Their agglomeration may have impacted the observed concentration of mills in Figure 5. Brandeis and Hodges (2018) also noticed the clustering of small and large mills in their analysis of the 2011 mill incident data in the sawmill industry in Tennessee. According to the study done by Duranton and Overman (2005), small and large firms have the capacity to influence mill localization in an industry. In our study, the presence of smaller mills in the west suggests that smaller mills are more likely the stronger driver of mill localization in western, NC.

5.4 Secondary Forest Products Manufacturers

Results of the locational analysis of the secondary forest products manufacturers in NC showed that they exhibit deviation from complete spatial randomness (evidence of spatial clustering) at county level, consistent with Hagadone and Grala (2012) in Mississippi's forest products industry. The identified hot spot consists of 22 counties with high concentration of secondary forest products manufacturers. Only 17 of them are statistically significant at the 95 percent confidence level; 5 in the Mountains, and 12 in the Piedmont regions (Figure 6). This spatial distribution shows that a lot of the secondary forest product manufacturers are disconnected from resource locations, compared to primary forest products manufacturers. This is because although secondary forest products manufacturers depend on primary manufacturers for their production input however, they are not spatially tied to the source of raw materials and can either source their raw materials from other regions (within or outside NC) or by importation (Aguilar and Vlosky 2006). In contrast to logs, it is easier to transport processed wood such as lumber for a long distance. Figure 8 and 9 also substantiate that the hot spot of secondary forest products manufacturers does not match with the hot spots of net merchantable bole volume, and average annual harvest removals of merchantable bole volume of growing-stock trees, respectively.

In addition, the high density of secondary forest products manufacturers follows that the observed spatial pattern may be due to better market access in the Piedmont than in the Mountains (McConnell et al. 2016). Of all 17 counties identified in this region, eight have a population size greater than 100,000 people, seven have a population size greater than 50,000 but less than 100,000 people, and the last two have a population size less than 50,000 people (U.S. Census Bureau 2019). The top three (3) most populated counties in the cluster are Mecklenburg (1,110,356), Guilford

(537,174), and Gaston (224,529) in a descending order. This helps to explain the fact that secondary forest products manufacturers are located near their final consumers in well populated areas of the state. In fact, the popular High Point Market, Furnitureland South, and some outlets of major manufacturers like Ashley Furniture are located within the identified hot spot region (Google Map[3]).

Similarly, the regression analysis of the secondary forest products manufacturers provides supporting evidence for our spatial analysis findings in that the proxy variable (population) for final market was highly significant ($p = 0.01$, and $p \leq 0.05$) compared to sawlog harvest which is not significant at any of these p-values, $p = 0.10$ (Table 5). This further stressed the fact that although the availability of resources in a county might likely attract secondary forest products manufacturers, proximity to the ready market would likely attract more secondary manufacturers. Thus, the need for secondary manufacturers to be closer to their customers far exceeds the need to be close to resource locations (Hagadone and Grala 2012, Aguilar and Vlosky 2006).

The advantages derived from such localization pattern include reduced cost of transportation, proximity to where their products are needed, shared knowledge, access to labor, and other benefits derived from economies of scale. Aguilar and Vlosky (2006) expounded that a secondary forest product industry with the above characteristics could foster cluster based economic development. Braden et al. (1998) observed that clusters were characterized by proximity and access to regional markets. In Oregon, the mouldings and millwork industry started by locating along railroad lines,

[3] https://www.google.com/maps/search/

the Bitterroot Valley in Montana is located along interstate highways and near a market for log homes (Braden et al. 1998).

Other factors such as availability of labor and county economic status (economic tier) may have influenced the development of the observed secondary manufacturers' cluster except access to railroad transportation. Unemployment, economic tier- 2, and 3 were significant in our secondary manufacturers' model. However, access to interstate highways and railroad did not significantly influence the likelihood of hosting secondary manufacturers in a county. The same reason given for a similar situation in our primary model applies here. That is, the inability to account for variation in road distance might be responsible for the negative signs of the road variable. Also, majority of the secondary manufacturers in the identified hot spots are not within 100 miles distance from port locations.

5.5 Spatial Distribution of Sawmills and Other Wood-Using Mills by Production Capacity

Figure 10 shows the contrast between the spatial distribution patterns of sawmills by incident point location, and by their average annual production capacity which helps understand variations in their spatial distribution. While sawmills are clustered in the western part based on their incidence numbers, they tend to cluster in the Coastal Plains based on their production capacity. These patterns may be due to the fact that most of the hardwood mills are smaller mills and counting each mill as an observation can be skewed towards sources of hardwood resources (Figure 10). However, most of the bigger mills are softwood mills and are located near the regions where softwood harvests are prevalent in the eastern NC. Sawmills in the Coastal Plains, even if fewer and more dispersed, are larger in terms of their production capacity. So, mills are concentrated around a lot of locally accessible softwood timber resources.

In addition, there is a higher probability of siting a big mill around the identified hot spot and expect to keep it working sustainably because of availability of locally accessible softwood. However, it is important to note that placing too many sawmills around the capacity hot spot can be detrimental and must therefore be controlled to avoid unwanted competition. Aguilar (2008) found that clustering of softwood sawmills resulted in higher log prices and labor cost which could potentially lead to road congestion and undesirable competition.

In the eastern part of NC, the presence of port for easy access to the export market was expected to have an influence on the spatial pattern of sawmills by incident point location, however, no hot spot was identified in the eastern region. Based on the date of establishment information in our data (especially the sawmills), we suspect that a few of the mills might have been established in the past without any plans for the end uses such as export markets. Goyette, (1967) noted that after world war II ended, shift in the locational pattern of new mills reflected the greater availability of wood and limited availability of coastal sites due to increased wood consumption in the U.S. South by 43 percent. This suggested that legacy-effect may have also influenced the pattern of mill distribution in North Carolina.

Other wood-using mills hot spots were identified in the Mountains and northern part of the Piedmont (Figure 11). Counties within the hot spot include Rutherford, McDowell, Franklin, Wake, Nash and Granville. These counties are dominated by bigger other wood-using mills with high production capacity (greater than 250 thousand tons). Hardwoods (Hemlock, Oak, Hickory, Sweetgum, Yellow poplar etc.) which are the most common species used by the other wood-using mills may have an effect on the observed spatial pattern. Proximity to source of raw materials seems to be an important factor driving the location of other wood-using mills in NC.

5.6 Conclusion

The spatial distribution of primary and secondary forest products manufacturers in NC follows a clustering distribution pattern. Based on the incidence point location, different clustering patterns were identified within the primary forest products industry with respect to resource location hot spots. The primary- hardwood, and hardwood-softwood mills' hot spots were identified in the Mountains and Piedmont region near hardwood resources hot spots while the primary softwood mills' hot spot was identified in the Coastal Plains around softwood resources. In terms of production capacity, a large number of small sized sawmills were identified to cluster in the west, and a few albeit large in capacity sized sawmills clustered in the east. Similarly, other wood-using primary mills by their production capacity were identified to cluster mainly around the west and lightly around the northeast part of the Coastal Plains. What was common to all primary mills analyzed by production capacity was that they showed higher spatial dependency with respect to resource location hot spots. However, our analysis suggested that secondary forest products manufacturers gravitate more towards where their products are needed than forest resources. The optimized hot spot analysis revealed evidence of spatial clustering in the secondary forest products industry and also in the joint concentration of primary and secondary forest products industry. Overall, the spatial distribution of forest products manufacturers in NC is significantly influenced by the distribution of hardwood and softwood resources, and the average annual production capacity size of primary forest products manufacturers had a significant impact on their choice of location.

In addition, this research identified other factors influencing the location of forest product manufacturers in NC. The presence of a ready market or high population influences the choice of location of forest products manufacturers. But the location of high resource dependent primary softwood manufacturers does not seem to be highly impacted by proximity to the final market based on our spatial analysis.

Although our spatial analysis suggested correlation between primary and secondary forest product manufacturers' location, and availability of transportation systems in a county, our count data model did not find statistically supporting evidence for these findings. This may be due to the inability of our data to account for distance or other roads facilitating log movement to mill locations. Therefore, further analysis would consider accounting for the road distance and other roads that facilitate transportation of forest products and resources.

Evidence from our spatial analysis suggested that the presence of primary forest products manufacturers favors the establishment of secondary manufacturers in a county. The econometric analysis of the secondary forest products industry further confirms this finding as important to sustainability of the supply chain network in the forest product industry. The coefficient of the primary forest products manufacturers had the expected sign.

Also, the unemployment variable was used to assess the effect of labor on the choice of location of forest products manufacturers. The regression analysis showed that although the unemployment variable does not significantly influence the location of primary manufacturers compared to its effect on the secondary manufacturers, the economic tier variables were able to capture the effect of average unemployment rate as one of the factors used in determining county tier rankings in NC. The economic tier variables were both significant in the primary and secondary models. It is also important to note that the economic tiers also accounted for the impact of adjusted property tax base per capita, amidst other factors, on the location of forest products industries in NC.

The spatial analysis of the forest products industry in NC helped to understand the spatial arrangement of the existing manufacturers in the state. The identified hot spots, especially in counties that do not currently host a forest products industry, are potential locations for new upcoming mills. Further analysis is, however, required to specifically assess how the introduction of a new mill will affect the spatial distribution of forest resources and forest management in these counties. This is important in order to prevent a high market demand condition that could lead to undesirable pressure on the forest resources and increased cost of logs. More investment in the transportation system particularly in the Mountains and Coastal Plains is recommended as the infrastructure could substantially boost the development of the forest product industry and the economy of these counties. Counties with good access to transportation networks would have an edge in attracting new forest product manufacturing firms.

REFERENCES

Aguilar, F. X. (2008). Effect of centrifugal forces on cluster patterns in the softwood lumber industry of the United States. *Forest Science*, *54*(2), 242-249.

Aguilar, F.X. (2007). Factors Influencing the Spatial Distribution of Natural Resource Based Industries: The Case of the Softwood Lumber Industry in the United States South. Doctoral Dissertation, Louisiana State University.

Aguilar, F.X., Grala, R.K. and Bratkovich, S.M., (2009). Use of georeferenced data to study clustering in the primary wood products industry of the US South. *Canadian journal of forest research*, *39*(12), pp.2494-2504.

Aguilar, F.X., and Vlosky, R.P., (2006). Spatial analysis of forest product manufacturer clusters in Louisiana. Wood and Fiber Science 38 (1), 121–131.

Anderson, N. M., Germain, R. H., and Bevilacqua, E. (2011). Geographic information system-based spatial analysis of sawmill wood procurement. *Journal of Forestry*, *109*(1), 34-42.

Asheim, B., Cooke, P., and Martin, R. (2006). The rise of the cluster concept in regional analysis and policy. *Clusters and regional development: critical reflections and explorations*, 1-29.

Brandeis, C., and Hodges, D. G. (2018). Sawmill Industry in Tennessee: Assessing Location Pattern Changes and Their Effects on Sawlog Procurement Distribution. *Forest Science*, *64*(3), 280-289.

Braden, R., Fossum, H., Eastin, I., Dirks, J., and Lowell, E. (1998). The role of manufacturing clusters in the Pacific Northwest forest products industry. *Center for International Trade in Forest Products. Seattle, WA. Working paper*, *66*, 43.

Duranton, G., and Overman, H. G. (2005). Testing for localization using micro-geographic data. *The Review of Economic Studies*, *72*(4), 1077-1106.

Braunerhjelm, P., and Carlsson, B. (1999). Industry Clusters in Ohio and Sweden, 1975--1995. *Small Business Economics*, *12*(4), 279-293.

Caldas de C., Marcia, and Burton H. S., (2006). "Controlling the False Discovery Rate: A New Application to Account for Multiple and Dependent Test in Local Statistics of Spatial Association." *Geographical Analysis* **38**, pp 180-208.

Ebdon, D. (1985). *Statistics in Geography: A practical approach-revised with 17 programs*. Wiley-Blackwell, Oxford, UK.

Esri, 2017. How hot spot analysis (Getis-Ord gi*) works. http://pro.arcgis.com/en/proapp/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatialstati.htm. (Accessed June 20, 2020).

Frizado, J., and Smith, B. W., Carroll, M. C., & Reid, N. (2009). Impact of polygon geometry on the identification of economic clusters. *Letters in Spatial and Resource Sciences*, *2*(1), 31-44.

Fujita, M., Krugman, P. and Venables, A. (1999). The Spatial Economy: Cities, Regions and International Trade, *Cambridge, Mass., MIT Press.*

Fujita, M., and Krugman. P. (2004). The new economic geography: Past, present and the future. Papers Reg. Sci. 83(1):139 –164.

Fujita, M. and Thisse, J.F. (2002). Economics of Agglomeration: Cities, Industrial Location and Regional Growth. *Cambridge, Cambridge University Press*

Gray, J. (1998). *False Dawn: The Delusions of Global Capitalism*, London, Granta Books.

Getis, A. and J.K. Ord. (1992). "The Analysis of Spatial Association by Use of Distance Statistics" in *Geographical Analysis* 24(3).

Goyette, A. V. (1967). *Factors influencing the location of paper mills in the South: the example of North Carolina mills* (Doctoral dissertation, University of North Carolina at Chapel Hill).

Gujarati, D.N., 2004. Basic Econometrics. McGraw-Hill, New York.

Hagadone, T. A., and Grala, R. K. (2012). Business clusters in Mississippi's forest products industry. *Forest Policy and Economics*, *20*, 16-24.

Hallencreutz, D., and Lundequist, P. E. R. (2003). Spatial clustering and the potential for policy practice: Experiences from cluster-building processes in Sweden. *European Planning Studies*, *11*(5), 533-547.

Han, H., Chung, W., Wells, L., and Anderson, N. (2018). Optimizing biomass feedstock logistics for forest residue processing and transportation on a tree-shaped road network. *Forests*, *9*(3), 121.

Kaur, R., Joshi, O., Will, R. E., and Murray, B. D. (2020). Sustainable management of unused eastern redcedar: An integrated spatial and economic analysis approach. *Resources, Conservation and Recycling*, *158*, 104806.

Kies, U., Mrosek, T., & Schulte, A. (2009). Spatial analysis of regional industrial clusters in the German forest sector. *International Forestry Review*, *11*(1), 38-51.

Krugman, P. (1991). Increasing returns and economic geography. *J. Polit. Econ. 99*(3)*:*483– 499.

Krugman, P. (1997). *Pop Internationalism*, Cambridge, Mass., MIT Press.

Marshall, A. (1890). Principles of economics Macmillan. *London (8th ed. Published in 1920)*.

Marshall. A. (1920). Principles of Economics. Macmillan, London. 820 p.

Martin, R. L. and Sunley, P. (1998). Slow convergence? The new endogenous growth theory and regional development, *Economic Geography*, 74, 201–27.

Martin, R. L. and Sunley, P. (2003) Deconstructing clusters: chaotic concept or policy panacea? *Journal of Economic Geography*, 3, 5–35

McConnell, T.E., Jeuck, J., Bardon, R., Hazel, D., Altizer, C., and New, B. (2016). North Carolina's Forest and Forest Products Industry by the Numbers. https://content.ces.ncsu.edu/north-carolinas-forest-and-forest-products-industry-by-the-numbers.

Mitchell, A. (2005). *The ESRI Guide to GIS Analysis, Volume 2.* ESRI Press, Redlands, California.

Ord, J.K. and Getis, A. (1995). "Local Spatial Autocorrelation Statistics: Distributional Issues and an Application" in *Geographical Analysis* 27(4).

Murray, B. 1995. Measuring oligopsony power with shadow prices: U.S. markets for pulpwood and sawlogs. *Rev. Econ. Statist.* 77:486–499.

Piore, M. and Sabel, C. (1984). *The Second Industrial Divide*, New York, Basic Books.

Pokharel, R., Grala, R. K., Grebner, D. L., and Cooke, W. H. (2019). Mill willingness to use logging residues to produce electricity: A spatial logistic regression approach. *Forest Science*, *65*(3), 277-288.

Pokharel, R., and Latta, G. S. (2020). A network analysis to identify forest merchantability limitations across the United States. *Forest Policy and Economics*, *116*, 102181.

Porter, M. E. (1990). *The Competitive Advantage of Nations,* New York, The Free Press.

Porter, M. E. (1998). *On Competition*, Boston, Mass., Harvard Business School Press.

Porter, M. E. (2003). The economic performance of regions. *Regional studies*, *37*(6-7), 549-578.

Powell, W. S. (2006). *Encyclopedia of North Carolina*. University of North Carolina Press.

Reich, R. B. (2001). *The Future of Success: Work and Life in the New Economy*, London: Random House.

Rocha, H. O. (2013). *Entrepreneurship and regional development: The role of clusters*. Springer.

Ross, E. A. (1896). The location of industries. *The Quarterly Journal of Economics*, *10*(3), 247 268.

Sabel, C. (1989). Flexible specialisation and the re-emergence of regional economies, in Hirst, P. and Zeitlin, J. (eds) *Reversing Industrial Decline: Industrial Structure and Policies in Britain and Her Competititors*, Oxford, Berg, 17–70.

Scott, A. (1988) *New Industrial Spaces: Flexible Production Organization and Regional Development in North America and Western Europe* (London: Pion Press).

Silva, J. S., and Tenreyro, S. (2010). On the existence of the maximum likelihood estimates in Poisson regression. *Economics Letters, 107*(2), 310-312.

Silva, B. K., Cubbage, F. W., Gonzalez, R., and Abt, R. C. (2019). Assessing market power in the US pulp and paper industry. *Forest Policy and Economics*, *102*, 138-150.

Tobler W. (1970) "A computer movie simulating urban growth in the Detroit region". *Economic Geography*, 46(Supplement): 234–240.

U.S. Census Bureau. 2019. County Population Totals: 2010-2019 [accessed May 29 2019].

Vlosky, R. P., and Chance, N. P. (2001). Employment structure and training needs in the Louisiana value-added wood products industry. *Forest Prod. J.* 51(3):34–41.

Watson, D. F., and Philip, G. M. (1985). A refinement of inverse distance weighted interpolation. *Geoprocessing*, *2*(4), 315-327.

Wong, D. W. S., and Lee, J. (2005). *Statistical analysis of geographic information with ArcView GIS and ArcGIS* (No. G 70.212. L43 2005). Hoboken, NJ: John Wiley & Sons.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.