The Geospatial Distribution of Employment: Examples from the Bureau of Labor Statistics Quarterly Census of Employment and Wages Program

By
Sheryl Konigsberg, David Talan, Richard Clayton
Division of Administrative Statistics and Labor Turnover
Office of Employment and Unemployment Statistics
Bureau of Labor Statistics

Introduction

The advent of powerful computing capabilities and mapping software now allows more sophisticated analysis of new and existing problems through the visual display of information. The centerpoint of these new features is the ability to provide pinpoint locations for geographic features; defined by precise latitude and longitude coordinates, called geocodes. In any geocoding system involving businesses, the key is to have accurate physical location addresses.

This article discusses the background of the Quarterly Census of Employment and Wages (QCEW) program, the definition of geocoding and its current and potential uses, examples of existing applications using labor market information, new ways of presenting these data, and early efforts, during a pilot project, to obtain and use geocodes for the Bureau of Labor Statistics (BLS) business establishment list. Finally, confidentiality, future uses, and how BLS plans to complete the geocoding project are profiled.

Quarterly Census of Employment and Wages

The Quarterly Census of Employment and Wages (QCEW) program, commonly known as the ES-202 program, is a product of the Unemployment Insurance (UI) system and is managed in a Federal/State cooperative environment. This program releases comprehensive tabulations of employment and wage information for workers covered by State UI laws and Federal workers covered by the Unemployment Compensation for Federal Employees (UCFE) program. BLS provides policies, standards and funding while states collect, edit, tabulate and publish the data.

The QCEW program serves as a non-farm business census and constitutes the most complete set of monthly employment and quarterly wage information by six-digit North American Industry Classification System (NAICS) at the national, State, Consolidated Metropolitan Statistical Area (CMSA), Metropolitan Statistical Area (MSA), and county levels.

Under the laws of each state, businesses are required to report the number of employees for all three months, total wages, taxable wages, contributions, and other related data on a quarterly basis. After these UI reports are collected and entered by the State, they are passed to the state QCEW unit for review, editing, and publication. These data are also used as the business register.

In addition to the UI reports, BLS funds two other collections to support the needs of its users. The first is the Annual Refilling Survey (ARS) that, over a three-year period, contacts all businesses to update or complete industry information (NAICS codes), and addresses. This is the primary method for updating physical location addresses within the business register. The second is the quarterly Multiple Worksite Report (MWR) that collects data for each individual establishment of a multi-unit business. The combination of information from these three sources comprises the resulting QCEW Covered Employment and Wages program.[2] Its comprehensiveness results in accurate data with substantial industry and geographic detail.

Geocoding

Geocoding is the process of adding geographic information such as longitude, latitude, county, and census tract to a file or database for use in a Geographic Information System (GIS). A GIS is an organized collection of computer hardware, software, geographic data, and personnel designed to capture, store, update, manipulate, analyze, and display spatial information.

Geocoding may use either a point or polygon approach. In a point-based approach, record information is linked to longitude and latitude coordinates. This information allows for locations to be plotted on a map. In a polygon-based approach, record information is linked to the center of a polygon that represents a reference layer such as census block group, census tract, or county. With the information, a user can associate all types of data that may be collected. The QCEW microdata file contains a rich set of geographic information that can be geocoded and applied to answer questions about the labor market.

There are two methods for deriving a geocode. The first and most accurate method is to use the physical location address. Addresses are geocoded by BLS using commercial software that accesses U.S postal data files. This software estimates the location of each address record from an input file and standardizes it. These standardized addresses are then matched against a Geographic Base File (GBF), which contains directories of street segment records. The second method is to approximate the location using zip codes. If the geocoding software is unable to make a match against an address, it will attempt to geocode at the zip centroid. A centroid is the calculated center of an area. The zip centroid assigns all locations to the geographic center of the zip code. These matching processes assign geographic codes to address records, establishing their spatial location.

Current and Potential Range of Uses

Currently, the QCEW program provides economic data at the national, State, and county level. Data from the QCEW serve as an important input to many BLS programs as well as other Federal and State programs. These data are used as a benchmark for the Current Employment Statistics and Occupational Employment Statistics. The QCEW also is used by the Bureau of Economic Analysis for GDP and personal income estimates.

Geocoded data are used extensively in government, business, and research for a wide range of applications including environmental resource analysis, land use planning, locational analysis, tax appraisal, utility and infrastructure planning, real estate analysis, marketing and demographic analysis, and habitat studies.

At the most detailed levels or, geocoded business addresses are valuable to transportation planning where approximate locations or higher level county aggregations are inadequate. For this purpose, the side of the street, the location along the block, and the exact corner of an intersection are critical to optimal planning of bus lines and other public transportation.

Geocoding data presents labor market information in a new dimension. Demands for more local data are providing an incentive to provide data for cities, towns, and even smaller areas. With the availability of geocoded data, BLS can potentially develop lower levels of aggregations, including cities, postal zip codes, census tract, census block, and natural boundaries such as floodplains.

Data Presentation

The conventional way of presenting economic data is two-dimensional, through tables and graphs. If tabular data is geocoded, it can be used to create a three-dimensional drawing. With the rise of the Internet and growing technology, GIS has made it possible to plot economic data to create abstractions and publish in the form of maps. This can be done by using geographic information, computers and geographic software to read the information and create spatial data visually.

As an example, the Quarterly Census of Employment and Wages (QCEW) program produces an annual bulletin with tabular data aggregated by State and Industry. Table 1 is a section from the 2002 QCEW publication, Private industry by State and 6-digit NAICS industry: Establishments, employment, and wages, 2002 annual averages, information sector.

The data in Table 1 is a standard way of presenting labor market information that has been in practice for many years. When looking at this table, it can be challenging to interpret what is being conveyed.

Table 2 is another example of tabular data and demonstrates how geocoded data can be displayed at the sub-county level. In this figure, employment and wage data by industry are shown for the city of Cleveland, Ohio. With the traditional way of displaying data, the lowest level of aggregation by boundaries is by county.

Without the latitude and longitude information, these data could not have been aggregated at this fine detail. While useful, these data can be used in even richer applications, which are illustrated in further examples of this article.

In September 2003, Hurricane Isabel was a major hurricane in the Atlantic Ocean, which made landfall on the east coast of North Carolina. Table 3 displays industries in the floodplain areas of Brunswick and New Hanover counties. With the use of geocoding and QCEW data, the average number of employees and employers that potentially were affected by this recent phenomenon are shown.

These examples illustrate the traditional method of displaying data in a tabular format. The following examples demonstrate the power of a GIS and how it conveys information quickly. A GIS also provides a strong visual that was previously unavailable.

In Map 1, the positions of the locations presented convey an immediate visual impression that is not seen in table presentations. Map 1 displays information about employment in the floodplain areas in North Carolina. Just by looking at the numbers in Table 2, it is difficult to determine the impact the hurricane had on business establishments and employment. It is very apparent in the map below; a picture gives a whole new dimension to the data.

Map 2 shows the 2004 hurricanes that hit Florida- Hurricanes, Charley, Frances, Ivan, and Jeanne. Florida tracks the path of each hurricane with a 20 mile buffer to show the impact on employment within the affected areas.

Map 3 shows the impact of the recent fires of October 2003 in San Diego, California. The State of California was able to use the QCEW data from second quarter 2003 to visually display employment and wages within the fire affected areas of Southern California and their relationship with the labor market.

In Map 4, Minnesota is able to display employment around major highways by using a thermal density map. With this type of map, data can be displayed by showing areas with a high concentration of employment without displaying confidential information.

It is apparent that map displays show how "a picture paints a thousand words." Information is processed more quickly by the reader, therefore, allowing for more timely decisions and conclusions to be made about a particular set of information.

Geocoding Pilot Project

In March 2003, the QCEW program completed a geocoding pilot project with the following 15 States: California, Connecticut, the District of Columbia, Florida, Hawaii, Maine, Maryland, Minnesota, Missouri, North Carolina, Ohio, Oregon, South Carolina, Texas, and West Virginia. These States published data based on the geocodes derived from the QCEW data. This study was used to help refine programs and plans for implementing geocoding in all States.

A great deal of time was spent working on address refinement for this project. The States' primary resource for locating addresses was the Internet. They also used other sources such as telephone books and phone calls to employers to obtain addresses. These last two sources proved to be less reliable and more time consuming for most States.

Obtaining government physical location addresses was a major obstacle for all States that participated in the pilot project. Governments tend to provide county-wide reports and finding a geocodeable address can be difficult.

Lastly, non-disclosure is an issue. Many States were unsure if they could publish sub-county data and to what extent. Some questions that arose during the project were:

- Does a dot on a map constitute confidential data based on address?
- Does a dot on a map constitute confidential data based on employment? and
- Does a dot on a map constitute confidential data based on industry?

Some States came to the conclusion that they could publish this type of information while others could not since non-disclosure laws vary from State to State.

Confidentiality

The issue of confidentiality is limiting the full potential of the geocoded data, just as it does for the basic economic data. Under BLS policies, the name and address of a business, as well as the industry code, and other data are considered confidential. The confidentiality of this information is essential for the full faith in the Bureau's data collection efforts. Geocoding and the use of GIS allow for the manipulation and display of spatial data. The QCEW data set contains individual establishment and employment listings. Uses of these data can be very important for research and planning. For that reason, it is very important to be able to display data while protecting confidentiality. In order to maintain confidential information, State agencies have been asking their legal departments for answers. How much information can be displayed? How far can you go? These answers are not always clear and are different in each State. One example of mapping data without breaching confidentiality previously outlined in this paper is use of density thermal mapping, where no individual establishment or employment can be determined, but high or low levels of concentration are visible.

Future Uses

The work that has been done with geocoding for the past two years has provided insight on the techniques for improving the accuracy of QCEW physical location addresses. These techniques have involved extensive work researching and updating the Bureau's existing business establishment data set. BLS will now be able to make data available for future research.

By using geocoded data from the QCEW, new economic information can be created such as sub-county estimates. There also is the potential to standardize addresses and reduce mailing costs for sample users.

Another use of geocoded QCEW data includes improving the Business Employment Dynamics (BED). BED is a set of statistics generated from the QCEW. These quarterly data series consist of gross job gains and gross job losses statistics from 1992 and subsequent years. These data help to provide a picture of the dynamic state of the labor market. Most data in these series are linked across time using a process that matches establishments by a unique number-the State Employment Security Agency identification numbers. Records that are not linked by this process are linked by various other means, one of which is a weighted match. The weighted match involves creating blocks to match data by mutiple criteria. With geocoded data, longitude and latitude information can be used in these blocks to create more accurate matches, thus allowing for better gross job gains and job losses data.

Lastly, GIS technology and spatial data play an important role in emergency response and preparedness. Large scale emergencies that impact humans and land are unpredictable and hard to envision. Two types of hazards are natural disasters and human-induced disasters. Natural disasters include events such as, earthquakes, volcanoes, landslides, wildfires, floods, and hurricanes. Human-induced disasters include events such as man-made fires, toxic spills, war, and bioterrorism. Using GIS, a great deal of time in decision making and evaluating the impact of a disaster can be saved before and after it occurs.

Getting the QCEW Fully Geocoded

The QCEW database contains approximately 8.5 million establishments with about 130 million employment. For the first quarter of 2005, the QCEW data file had 82 percent of its records and 93 percent of employment geocoded. The remainder of the units are mostly new firms, small firms, or government units that do not provide QCEW data by worksite. Over 1 million new businesses are started annually, placing a huge burden on the UI and QCEW to obtain accurate industry, geographic, and address information. Thus, every new business will always be a challenge. A small number of large reporters also do not provide QCEW data by worksite. The largest group of these is Professional Employer Organizations (PEO's). PEO's are very large employers that take over the actual workers and responsibility of handling the administrative side of employment such as hiring, firing, establishing and maintaining employee files, processing payroll, filing payroll taxes, and other human resource support so a firm can focus on the day-to-day tasks involved in maintaining and growing their buisiness. While many PEO's provide detailed reports covering each of their clients, many do not. These cause signifigant gaps in the ability to correctly allocate employment to the correct county and industry and to geocode the client locations. BLS continues to work with these firms to obtain the detail needed to support accurate data by county and industry. When obtained, these reports will also allow BLS to geocode the location and industry.

Conclusion

In summary, geographic data are an asset in data analysis, especially with the QCEW. Since BLS has implemented this aspect to the QCEW, the original two-dimensional data can be combined or used to create a three-dimensional way of looking at data. Using a geographic information system such as geocoding and mapping software, many datasets can be combined into one picture, thus saving time in reviewing data results and providing new insights which previously were unobserved. This article has provided a few examples of how the QCEW can benefit from the use of geocoded data.

Appendix: Tables and Maps

Table 1

Table 10. Private industry by State and 6-digit NAICS industry: Establishments, employment, and wages, 2002 annual averages — Continued

State	Average establishments	Annual average employment	Total annual wages (in thousands)	Annual wages per employee	Average weekly wage	Average establishments	Annual average employment	Total annual wages (in thousands)	Annual wages per employee	Average weekly wage	Average establishments	Annual average employment	Total annual wages (in thousands)	Annual wages per employee	Average weekly wage
	22132 Sewage treatment facilities					22133 Steam and air-conditioning supply					Information				
Total U.S.	751	8,898	$360,921	$40,707	$783	143	2,109	$120,086	$56,932	$1,095	150,107	3,364,485	$188,758,526	$56,103	$1,079
Alabama	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	1,782	34,206	1,483,340	43,365	834
Alaska	3	27	1,247	45,752	880	(2)	(2)	(2)	(2)	(2)	365	7,076	317,971	44,937	864
Arizona	16	99	3,684	37,334	718	–	–	–	–	–	2,369	51,875	2,218,526	42,766	822
Arkansas	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	1,202	20,367	723,446	35,521	683
California	27	653	27,799	42,580	818	9	131	10,418	79,479	1,528	22,265	499,881	35,051,307	70,147	1,349
Colorado	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	3,677	93,397	5,900,532	63,177	1,215
Connecticut	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	1,671	41,145	2,310,682	56,160	1,080
Delaware	(2)	(2)	(2)	(2)	(2)	–	–	–	–	–	334	7,745	393,936	50,864	978
District of Columbia	(2)	(2)	(2)	(2)	(2)	–	–	–	–	–	1,129	25,446	1,934,773	76,029	1,462
Florida	74	1,163	43,146	37,087	713	31	136	4,545	33,396	642	8,751	177,973	8,212,392	46,144	887
Georgia	4	85	3,986	46,980	903	3	43	1,693	39,747	764	4,492	132,432	7,563,572	57,113	1,098
Hawaii	19	80	3,367	42,039	808	(2)	(2)	(2)	(2)	(2)	691	11,701	505,167	43,173	830
Idaho	(2)	(2)	(2)	(2)	(2)	–	–	–	–	–	713	9,162	305,019	33,291	640
Illinois	27	364	13,612	37,413	719	12	61	3,586	58,853	1,126	6,454	145,409	7,667,873	52,733	1,014
Indiana	27	576	22,117	38,381	738	–	–	–	–	–	2,178	42,528	1,657,356	38,971	749
Iowa	6	13	281	22,335	430	–	–	–	–	–	1,743	35,193	1,225,782	34,830	670
Kansas	6	43	1,166	27,172	523	–	–	–	–	–	1,485	50,745	2,377,331	46,849	901
Kentucky	12	144	4,188	29,015	558	–	–	–	–	–	1,767	31,745	1,120,354	35,293	679
Louisiana	16	86	2,294	26,799	515	5	55	1,972	36,192	696	1,659	29,018	1,098,531	37,857	728
Maine	6	18	484	26,523	510	–	–	–	–	–	736	11,546	429,314	37,186	715
Maryland	6	40	1,463	36,268	697	6	123	8,206	66,628	1,281	2,914	53,449	3,010,295	56,321	1,083
Massachusetts	31	397	19,515	49,186	946	(2)	(2)	(2)	(2)	(2)	4,521	99,989	6,645,535	66,463	1,278
Michigan	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	3,977	73,480	3,467,610	47,191	908
Minnesota	(2)	(2)	(2)	(2)	(2)	9	217	13,654	63,102	1,215	3,036	67,161	3,199,455	47,639	916
Mississippi	13	409	10,783	26,342	507	–	–	–	–	–	1,047	16,070	569,159	35,418	681
Missouri	27	106	3,802	35,955	691	4	89	4,646	52,029	1,001	3,174	70,699	3,177,280	44,814	862
Montana	(2)	(2)	(2)	(2)	(2)	–	–	–	–	–	830	7,780	255,185	32,798	631
Nebraska	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	1,013	24,690	1,053,470	42,667	821
Nevada	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	921	16,967	766,774	45,193	869
New Hampshire	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	917	12,821	701,327	54,703	1,052
New Jersey	24	224	10,714	47,833	920	–	–	–	–	–	4,058	112,163	7,602,398	67,780	1,303
New Mexico	(2)	(2)	(2)	(2)	(2)	–	–	–	–	–	971	16,864	537,844	31,894	613
New York	11	101	4,600	45,729	879	4	26	1,465	56,226	1,081	11,713	295,415	19,665,362	66,569	1,280

Table 2

City of Cleveland Geocoded data on Establishments, Average Employment and Total Wages Paid by Industrial Sector as Covered under the Ohio and Federal Unemployment Compensation Laws First Quarter 2002 (a)			
Industrial Sector	**Number of Establishments**	**Average Monthly Employment**	**Total Wages (in Thousands of Dollars)**
Total covered under Ohio UC Law (b)	**9,365**	**279,396**	**$2,958,645**
Private Sector	**9,273**	**230,658**	**2,441,135**
Agriculture, forestry, fishing and hunting	4	21	94
Mining	7	231	3,295
Utilities	14	1,042	15,761
Construction	511	6,198	73,192
Manufacturing	1,138	31,964	333,537
Wholesale trade	678	12,229	159,430
Retail trade	1,247	13,458	73,763
Transportation and warehousing	243	4,116	34,241
Information	163	7,151	99,972
Finance and insurance	500	23,046	420,040
Real estate and rental and leasing	311	2,696	20,157
Professional and technical services	1,162	21,367	312,719
Management of companies and enterprises	53	6,418	93,517
Administrative and waste services	446	15,624	99,983
Educational services	73	9,960	92,572
Health care and social assistance	768	46,598	445,038
Arts, entertainment, and recreation	108	5,150	56,995
Accommodation and food services	898	14,112	47,337
Other services, except public administration	949	9,276	59,492
State & Local Government	**92**	**48,738**	**517,510**
State Government	21	3,940	46,608
Local Government	71	44,798	470,902
Federal Government (c)	18	8,213	103,766

Table 3

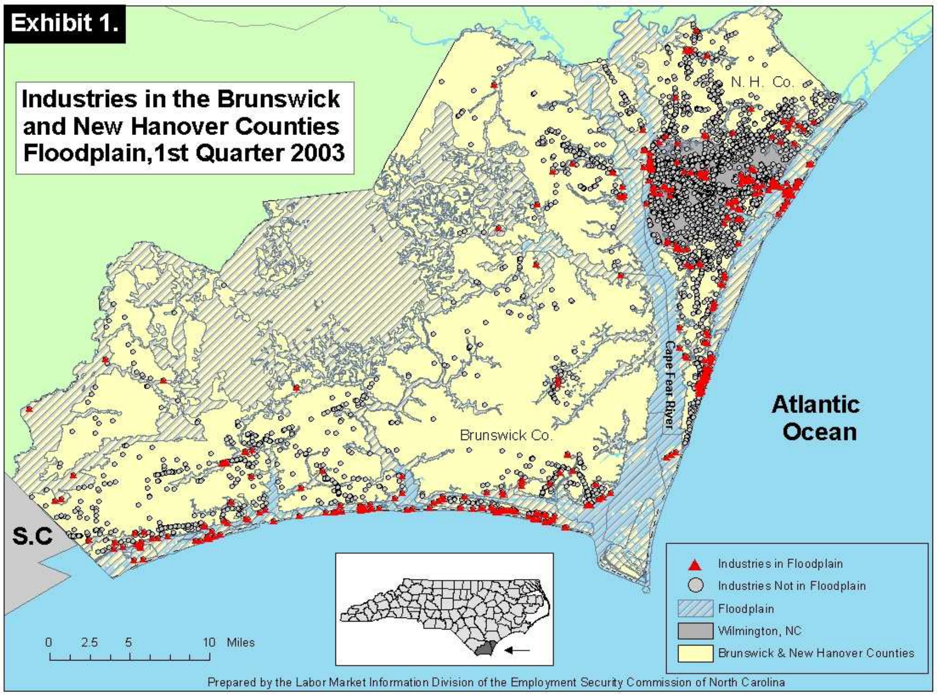
1st Quarter

2003

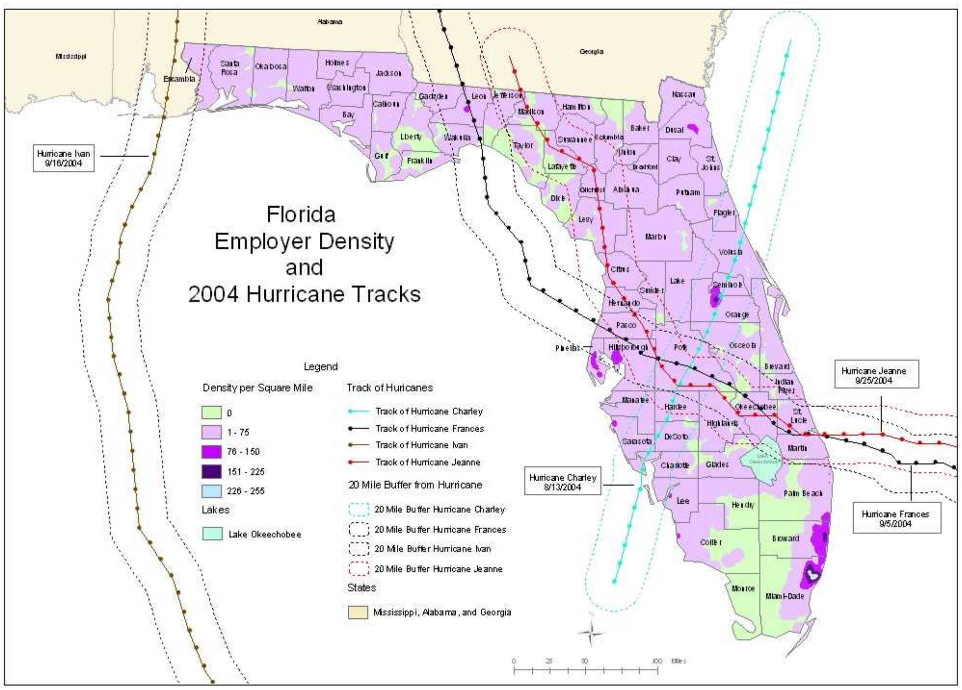
Brunswick and New Hanover Counties

Sector	**QCEW Units**	**QCEW Avg Emp**	**Units GeoCoded**	**% Emp GeoCoded**	**% Units In Floodplain**	**% Emp In Floodplain**
11	36	222	77.78%	94.14%	10.71%	2.39%
21	Suppressed	*	*	*	*	*
22	Suppressed	*	*	*	*	*
23	1,287	8,680	78.09%	91.28%	11.84%	10.67%
31-33	288	7,564	83.33%	98.19%	4.58%	11.88%
42	420	3,006	80.71%	92.18%	8.26%	6.64%
44-45	1,313	15,871	86.75%	97.92%	9.22%	6.15%
48-49	217	3,474	77.42%	89.87%	11.31%	3.68%
51	121	1,718	87.60%	94.82%	9.43%	1.60%
52	406	2,907	82.51%	85.55%	5.37%	5.31%
53	374	2,244	73.53%	79.55%	17.82%	29.19%
54	841	5,639	79.07%	81.72%	11.13%	11.85%
55	31	1,365	90.32%	96.56%	10.71%	0.68%
56	545	4,888	75.23%	88.09%	8.78%	4.39%
61	124	9,446	88.71%	99.84%	10.00%	17.17%
62	655	15,537	85.34%	92.19%	5.01%	1.79%
71	166	2,528	77.71%	89.60%	17.83%	9.58%
72	698	11,766	83.67%	99.46%	20.21%	18.16%
81	692	3,276	76.73%	87.97%	8.10%	7.11%
92	116	5,655	26.72%	54.94%	29.03%	7.02%
99	148	157	0.00%	0.00%	0.00%	0.00%
Total	**8,478**	**107,273**	**78.88%**	**91.56%**	**10.64%**	**10.36%**

Map 1

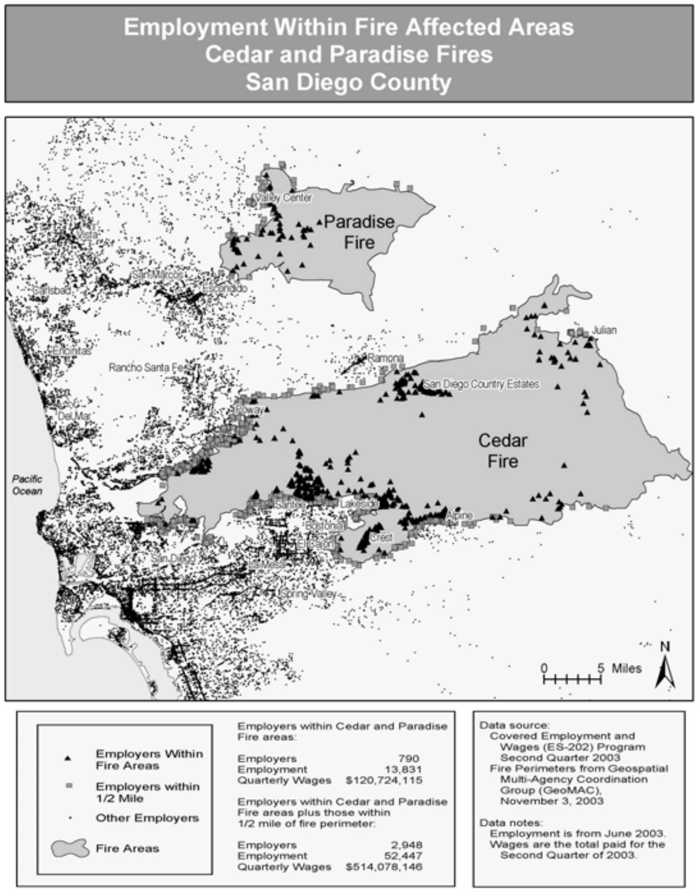


Map 2



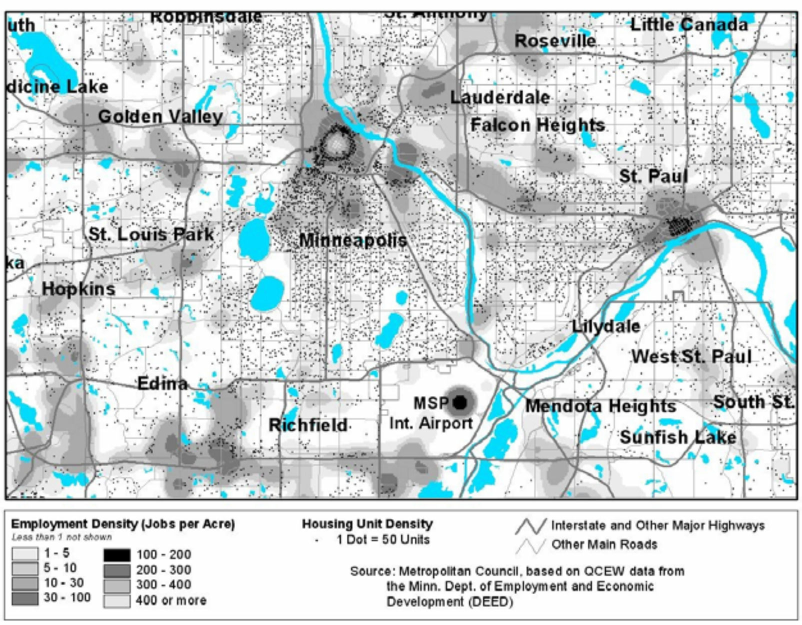
Note: Hurricane Charley made landfall August 13, 2004. Hurricane Frances made landfall September 5, 2004. Hurricane Ivan made landfall September 16, 2004. Hurricane Jeanne made landfall September 25, 2004.
Source: Agency for Workforce Innovation, Labor Market Statistics; Quarterly Census of Employment and Wages; and Florida Department of Environmental Protection.
Prepared by: Agency for Workforce Innovation, Labor Market Statistics, November 2004.

Map 3



Map 4

Employment and Housing Density, Twin Cities Metropolitan Area, 2000



Notes

Clayton, Richard, “Geocoding the Business Register at the Bureau of Labor Statistics”, *15th International Rountable on Business Survey Frames*, Washington, DC, October 22-25, 2001.

Establishments, employment, and wages, 2002 annual averages, Information“, *Employment and Wages Annual Averages, 2002*, U.S. Department of Labor; Bureau of Labor Statistics, November 2003, 3-7.

“Table 10. Private industry by State and 6-digit NAICS industry: Establishments, employment, and wages, 2002 annual averages, Information“, *Employment and Wages Annual Averages*, *2002*, U.S. Department of Labor; Bureau of Labor Statistics, November 2003, 496.

“Challenges for GIS in Emergency Preparedness and Response”, *An Esri White Paper*, *Enivironmental System Research Institute*, 2000.

“GIS Aids Emergency Response”, *Arcuser*, July-September 2001.

Page, Dan, “GIS Played a Critical Role in the Aftermath of September’s Terroist Attacks”, *Government Technology*, Jaunary 2003.

“Understanding GIS: The ArcInfo Method”, *Redlands*, *CA: Enivironmental System Research Institute*, 1990, 1-2.