# Analysis of Combining Multiple Road Centerline Datasets in Order to Improve Geocoding Spatial Accuracy and Match Rates of Valid Addresses

Ross H. Kleiner[1, 2]
[1] *Department of Resource Analysis, Saint Mary's University of Minnesota, Winona, MN 55987;* [2]*Applied Data Consultants Inc., Eau Claire, WI 54703*

*Keywords:* GIS, Geocode, Geocoding issues, Addresses, Street Data, Reference Datasets, Source Datasets, Road Centerlines

## Abstract

The validity of all geographic analysis is directly related to the accuracy of all geographic datasets representing real-world phenomena involved within a given study. Spatial data of customers or test subjects are often obtained through address geocoding. GIS users run the risk of producing unsatisfactory geocoding match rates due to discrepancies in the source data, reference data, or both datasets. Research conducted in this project examined the improvements in geocoding match rates when combining two updated multiple road centerline datasets. A Geographic Information System (GIS) geocoded records from source address datasets representing multiple test areas with two competing spatial reference datasets of road centerlines. Comparative statistics between the two reference datasets were created for analysis. Investigations of the geocoded output datasets revealed a projected improvement in match rates when combining the two road centerline datasets into a hybrid reference dataset.

## Introduction

As the population continues to grow and technology rapidly advances, the precise knowledge of the spatial location and distribution of people and cultural phenomena becomes increasingly relevant. GIS has become the most widely used tool to collect, analyze, and store such data.

*Address geocoding* can be defined as a function for which a GIS is used, adds a point to a virtual map at a calculated "X,Y" location which represents the coordinates of an address that is listed in an events table (Mills, 2002). It is an important feature of a GIS, utilized by research organizations, private businesses, and government offices in order to transfer important records from a tabular format to a spatial format. The quality of a GIS user's completed spatial product is directly related to geocode match rates and spatial accuracy. Geocoding, however, is rarely a perfect application. *Match rate* represents the amount of records from the source data (address dataset in tabular form) that have successfully geocoded to the reference data (road-centerlines in spatial format with attributes). Match rates are often

Kleiner, Ross H. 2008. Analysis of Combining Multiple Road Centerline Datasets in Order to Improve Geocoding Spatial Accuracy and Match Rates of Valid Addresses. Volume 10, Papers in Resource Analysis. 9 pp. Saint Mary's University of Minnesota Central Services Press. Winona, MN. Retrieved (date) from http://www.gis.smumn.edu

displayed as a percentage (Hassan et al., 2004). Low percentages can result from numerous discrepancies within the chosen reference dataset, which can cause complex spatial and analytical inaccuracies in the final output of geocoded points.

Improvements to the source data and adjustments to the geocode interpolation technique are often made in order to improve match rates. The research conducted in this study investigated the differences in geocode match rates between road-centerline datasets from two of the nation's top competing GIS base map data vendors, as well as the improvements in geocode match rates when the two updated reference datasets are combined to create a hybrid reference dataset. Address records from several regions in the United States were delineated by ZIP code, classified by population density, and used as a source dataset. Updated road-centerline datasets from data vendors TeleAtlas and NAVTEQ were used as a combined reference dataset. Statistical analysis of the match rates in the output spatial datasets were conducted after geocoding each address dataset normally to the competing streets.

## Methods

The three components that represent the input data for this study were the updated source address data in tabular form, the TeleAtlas Dynamap 2000 road-centerline data, and the NAVTEQ road-centerline data.

### *Source Data*

The source address data used in this study was updated annually by one of the nation's top telephone companies for the purpose of distributing telephone books to residences and businesses. The addresses have been declared "postally valid" for the year 2007. The only fields available for this study was; a unique ID field, address prefix, address number, address suffix, address second number, street name, alternate street name, ZIP code, ZIP +4, city, and state. The address records were grouped by postal ZIP codes which were selected at random. The ZIP Codes containing address data for this study were classified regionally and placed into four categories: East, Midwest, South and West.

In the East region, the address records for this study were from ZIP codes located in Massachusetts, New York, Maine, Rhode Island, and West Virginia. In the Midwest region, the address records have been extracted from ZIP codes located in Minnesota, Wisconsin, Missouri, Illinois, and Indiana. The address records from the South region have been extracted from the ZIP codes located in Texas, Arkansas, Oklahoma, and Georgia. Finally, in the West region, address records have been extracted from ZIP codes located in California, Nevada, and Arizona.

The addresses from each regional category were then placed into sub-categories classified by ZIP Code population density displayed in Table 1.

Table 1. A table displaying the number of address records used in this study delineated by region and ZIP code population density.

| Region | East Records | East ZIPs | Midwest Records | Midwest ZIPs |
|---|---|---|---|---|
| Rural | 8450 | 10 | 6217 | 8 |
| Fringe | 46920 | 10 | 60211 | 10 |
| Urban | 106538 | 10 | 104501 | 10 |
| **TOTAL** | **161908** | **30** | **170929** | **28** |
| **Region** | **South Records** | **South ZIPs** | **West Records** | **West ZIPs** |
| Rural | 191631 | 9 | 64379 | 11 |
| Fringe | 124795 | 8 | 119236 | 11 |
| Urban | 87410 | 10 | 157572 | 13 |
| **TOTAL** | **231836** | **27** | **341203** | **35** |

The first classification ranged from a ZIP Code population density of zero to 79 residents per square mile. This sub-classification represents rural addresses in each region. The second classification represents the suburban or fringe addresses in each region and was classified by a population density of 80 to 899 residents per square mile. The third, representing the urban addresses in each region, was classified by 900 to 117,474 residents per square mile. All data extraction, manipulation, and database merging were completed using database tools from ArcToolbox in ArcGIS 9.2 (Figure 1).

Figure 2 displays how these ZIP codes are often distributed throughout the U.S. Urban classified ZIP codes often appear to take the place of a bull's-eye in a target pattern, surrounded completely by a boundary of fringe ZIP codes. Finally, the fringe ZIP codes boundary runs adjacent to rural ZIP codes which expand outward until it meets a similar pattern.

***Reference Data***

The TeleAtlas Dynamap 2000 GIS base map data is a comprehensive dataset consisting of addressed streets, unaddressed streets, census, and postal data of US-based geographic applications in spatial and attribute form. Updates to the dataset consist of field collection via GPS and backing up information from over 50,000 "global resources". Updates to Dynamap 2000 are made available quarterly (TeleAtlas, 2007). The Dataset used for the research in this study was updated and released in October of 2007. The NAVTEQ street dataset is based on comprehensive streets with and without addresses along with census and postal data. NAVTEQ will drive streets with a GIS and gather information from other resources (like TeleAtlas); however, NAVTEQ also collects reports submitted online from paying and non-paying clients who discover additions or discrepancies while in the field (NAVTEQ, 2007).

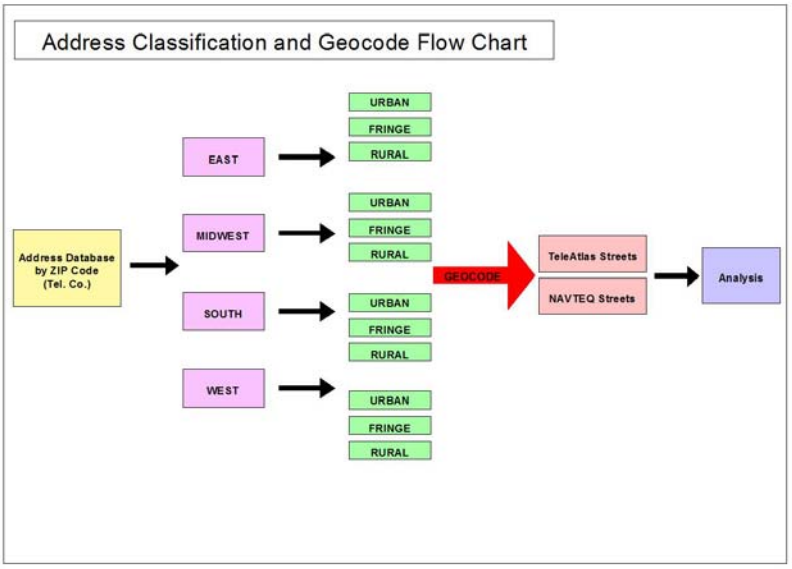


Figure 1. A Flow Chart displaying how the addresses were extracted from the telephone company's database. Regionally at first, then broken down into classes by population density prior to geocoding and analysis.

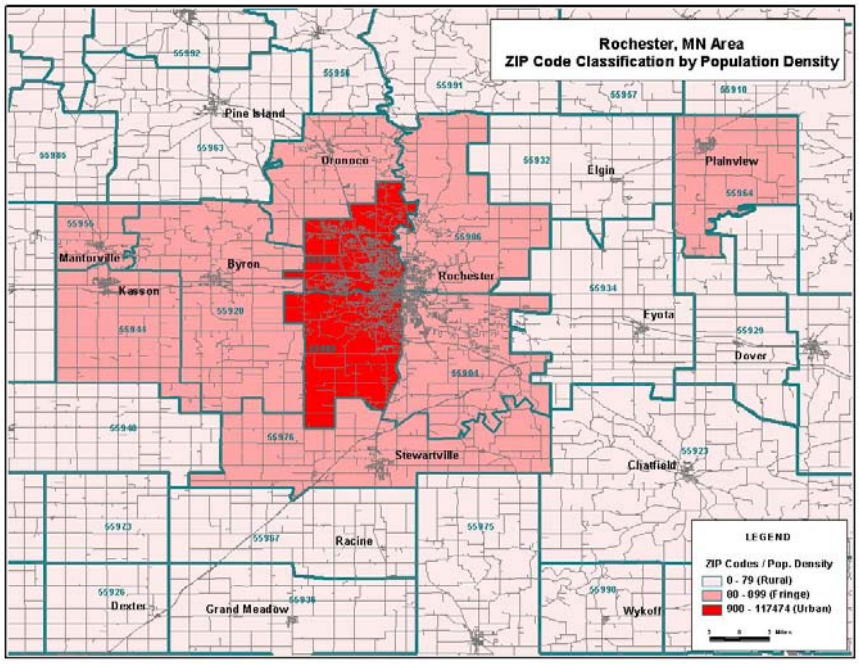


Figure 2. A thematic map of the Rochester, MN Area displaying the population density by ZIP code.

This is a significant difference considering NAVTEQ currently provides their data to millions of customers for GPS features on cell phones and routing data for on-board vehicle navigation devices (for example, the Garmin and Tom-Tom GPS Navigators). The NAVTEQ dataset used for this research project was released in November of 2007.

### *Interpolation Technique*

The interpolation technique and algorithms used in this study included the geocoding feature in ArcView 3.3 by Environmental Systems Research Institute (ESRI) combined with additional AVENUE-based programming from Matt Rantala, a Senior GIS Programmer and Consultant at Applied Data Consultants Inc. in Eau Claire, WI. The geocoding algorithms developed by ESRI remain, while Mr. Rantala's code provides further matching of the source data to the reference data. Furthermore, Mr. Rantala's additions created attribute fields within the geocoded dataset indicating the reference dataset to which the address data has successfully geocoded, and the amount of spatial accuracy. An address can geocode perfectly to a given road segment matching the corresponding address range precisely. Also, an address could possibly geocode to a street without a matching range to use as reference. A spatial representation of the address could be created, however it would not be as spatially accurate as a perfectly geocoded point.

### *Research Analysis Methods*

The research topics that this project addressed included:

- Analysis of geocode match rates of postally valid addresses to reference data provided by TeleAtlas Dynamap 2000 to urban, suburban, and rural areas in the U.S.

- Analysis of geocode match rates of postally valid addresses to reference data provided by NAVTEQ to urban , suburban, and rural areas

- Statistical analysis of the match rates between the two reference datasets and postally valid addresses

- Statistical analysis of the improvements in match rates of address points to a combined reference data set consisting of the most updated street data from TeleAtlas and NAVTEQ

Each address dataset classified by region and population density was created in duplicate. The first datasets were geocoded directly to the TeleAtlas Dynamap 2000 Streets using the AVENUE-based geocoding program in ArcView GIS 3.3 by ESRI and enhanced by Matt Rantala. The second batch of identical address datasets was then geocoded to the NAVTEQ Streets using the same geocoding program in ArcView 3.3. Comparative descriptive analysis of geocoding match rates was then conducted. Furthermore, addresses that were successfully geocoded to the NAVTEQ streets but not the TeleAtlas streets were identified along with their matched roads. Analysis comparing the improvements made to match rates when using multiple reference datasets was conducted. Percentage of increase by ZIP Code and classification was observed. Thus, conclusions were made as to how beneficial this method is, and what applications may benefit from using multiple datasets.

**Results**

Geocoding results in tabular form are displayed by region in Tables 2 - 5. The first regional table displays the statistics from geocoding with the TeleAtlas Dynamap 2000 streets. The second table displays the statistics from geocoding with the NAVTEQ streets. The third table by region displays the statistics resulting from geocoding the unmatched records extracted from the TeleAtlas dataset and re-geocoding them to the NAVTEQ streets. Thus, resulting in statistics correlated to a hybrid reference dataset where all addresses are either geocoded to the TeleAtlas or NAVTEQ streets. The fields below titled “Perfect TELE” or “Perfect NAV” display the records that geocoded precisely to the corresponding reference dataset. The fields titled “Geocode TELE” or “Geocode NAV” represent the records that geocoded to the arcs on the reference data where small spelling discrepancies resulted in a match that was not precise. Thus, the spatial accuracy cannot be considered perfect. An example of this could be where a source data record for “112 Peachtree Rd” did not find a match in the reference data but was placed as a point on the arc containing “112 Peach Tree Rd.”

***East Region***

Research in this study yielded the following results: In the East region, displayed in Table 2, the match rates using the TeleAtlas streets resulted in

higher match rates only in the rural addresses by 4.24%. The match rates for the fringe addresses were higher with the NAVTEQ streets by 0.55%, as well as the urban addresses by 0.93%. Geocoding using both reference datasets resulted in a 2.91% increase from the TeleAtlas streets for the rural addresses, a 1.6% increase from the NAVTEQ streets for the fringe addresses, and only a 0.23% increase from the NAVTEQ streets for the urban addresses. After totaling up all three classifications in the East region, there was a 1.00% increase when combining both reference datasets.

Table 2. Match Rate Statistics for Addresses in the East Region.

| **East TeleAtlas** | | | |
|---|---|---|---|
| | Total Records | Geocode TELE | Perfect TELE |
| Rural | 8450 | 103 | 7666 |
| Fringe | 46920 | 95 | 37587 |
| Urban | 106538 | 366 | 103044 |
| **Total** | **161908** | **564** | **148297** |
| | Total Geocode | Match Rate | Ungeocoded |
| Rural | 7769 | 91.94% | 681 |
| Fringe | 37682 | 80.31% | 9238 |
| Urban | 103410 | 97.06% | 3128 |
| **Total** | **148861** | **91.94%** | **13047** |
| **East NAVTEQ** | | | |
| | Total Records | Geocode NAV | Geocode NAV |
| Rural | 8450 | 82 | 7329 |
| Fringe | 46920 | 369 | 37571 |
| Urban | 106538 | 2192 | 102213 |
| **Total** | **161908** | **2643** | **147113** |
| | Total Geocode | Match Rate | Ungeocoded |
| Rural | 7411 | 87.70% | 1039 |
| Fringe | 37940 | 80.86% | 8980 |
| Urban | 104405 | 98.00% | 2133 |
| **Total** | **149756** | **92.49%** | **12152** |
| **East Combined** | | | |
| | Total Records | Geocode TELE | Perfect TELE |
| Rural | 8450 | 103 | 7666 |
| Fringe | 46920 | 95 | 37587 |
| Urban | 106538 | 366 | 103044 |
| **Total** | **161908** | **564** | **148297** |
| | Geocode NAV | Perfect NAV | Total Geocode |
| Rural | 0 | 246 | 8015 |
| Fringe | 29 | 1008 | 38719 |
| Urban | 177 | 1060 | 104647 |
| **Total** | **206** | **2314** | **151381** |
| | Match Rate | Ungeocoded | |
| Rural | 94.85% | 435 | |
| Fringe | 82.52% | 8201 | |
| Urban | 98.23% | 1891 | |
| **Total** | **93.50%** | **10527** | |

### *Midwest Region*

In the Midwest region, displayed in Table 3, the match rates with the NAVTEQ streets resulted in a 3.43% higher match rate than the TeleAtlas data for the rural addresses. The differences in the fringe and urban source data were not as large, resulting in the TeleAtlas data yielding higher match rates in both by 0.12% and 0.51%. Geocoding using both reference datasets resulted in a 3.43% increase in match rate for the rural addresses, a 2.21% increase for the fringe addresses, and only a 0.64% increase for the urban addresses. Overall the hybrid reference dataset did yield a 1.41% increase in match rates for all three classifications in the Midwest.

Table 3. Match Rate Statistics for Addresses in the Midwest Region.

| **Midwest TeleAtlas** | | | |
|---|---|---|---|
| | Total Records | Geocode TELE | Perfect TELE |
| Rural | 6217 | 50 | 4446 |
| Fringe | 60211 | 361 | 56683 |
| Urban | 104501 | 220 | 102680 |
| **Total** | **170929** | **631** | **163809** |
| | Total Geocode | Match Rate | Ungeocoded |
| Rural | 4496 | 72.32% | 1721 |
| Fringe | 57044 | 94.74% | 3167 |
| Urban | 102900 | 98.47% | 1601 |
| **Total** | **164440** | **96.20%** | **6489** |
| **Midwest NAVTEQ** | | | |
| | Total Records | Geocode NAV | Geocode NAV |
| Rural | 6217 | 355 | 4330 |
| Fringe | 60211 | 1653 | 55318 |
| Urban | 104501 | 1718 | 100651 |
| **Total** | **170929** | **3726** | **160299** |
| | Total Geocode | Match Rate | Ungeocoded |
| Rural | 4685 | 75.36% | 1532 |
| Fringe | 56971 | 94.62% | 3240 |
| Urban | 102369 | 97.96% | 2132 |
| **Total** | **164025** | **95.96%** | **6904** |
| **Midwest Combined** | | | |
| | Total Records | Geocode TELE | Perfect TELE |
| Rural | 6217 | 50 | 4446 |
| Fringe | 60211 | 361 | 56683 |
| Urban | 104501 | 220 | 102680 |
| **Total** | **170929** | **631** | **163809** |
| | Geocode NAV | Perfect NAV | Total Geocode |
| Rural | 76 | 326 | 4898 |
| Fringe | 105 | 1228 | 58377 |
| Urban | 88 | 585 | 103573 |
| **Total** | **269** | **2139** | **166848** |
| | Match Rate | Ungeocoded | |
| Rural | 78.78% | 1319 | |
| Fringe | 96.95% | 1834 | |
| Urban | 99.11% | 928 | |
| **Total** | **97.61%** | **4081** | |

### *South Region*

In the South region, displayed in Table 4, the statistics show a dramatic difference in a match rate of 17.50% for the rural addresses in the NAVTEQ streets favor. NAVTEQ also yielded higher geocode rates for the fringe addresses matching 3.62% better than the TeleAtlas streets. Again, the match rates for the urban addresses in the South region had only a small difference of 0.75% in the TeleAtlas streets favor. When combining the street datasets for

the rural addresses, the match rate improved by 2.49% from the NAVTEQ geocode. Match rates with the combined reference datasets increased by 1.11% for the fringe addresses and only 0.50% for the urban addresses in the South region. Overall, the increase for all three classifications with the combined reference dataset was 1.28%.

Table 4. Match Rate Statistics for Addresses in the South Region.

| **South TeleAtlas** | | | |
|---|---|---|---|
| | Total Records | Geocode TELE | Perfect TELE |
| Rural | 19631 | 117 | 14438 |
| Fringe | 124795 | 432 | 109898 |
| Urban | 87410 | 63 | 85597 |
| **Total** | **231836** | **612** | **209933** |
| | Total Geocode | Match Rate | Ungeocoded |
| Rural | 14555 | 74.14% | 5076 |
| Fringe | 110330 | 88.41% | 14465 |
| Urban | 85660 | 98.00% | 1750 |
| **Total** | **210545** | **90.82%** | **21291** |
| **South NAVTEQ** | | | |
| | Total Records | Geocode NAV | Geocode NAV |
| Rural | 19631 | 523 | 17468 |
| Fringe | 124795 | 3201 | 111652 |
| Urban | 87410 | 3209 | 81795 |
| **Total** | **231836** | **6933** | **210915** |
| | Total Geocode | Match Rate | Ungeocoded |
| Rural | 17991 | 91.65% | 1640 |
| Fringe | 114853 | 92.03% | 9942 |
| Urban | 85004 | 97.25% | 2406 |
| **Total** | **217848** | **93.97%** | **13988** |
| **South Combined** | | | |
| | Total Records | Geocode TELE | Perfect TELE |
| Rural | 19631 | 117 | 14438 |
| Fringe | 124795 | 432 | 109898 |
| Urban | 87410 | 63 | 85597 |
| **Total** | **231836** | **612** | **209933** |
| | Geocode NAV | Perfect NAV | Total Geocode |
| Rural | 121 | 3803 | 18479 |
| Fringe | 178 | 5735 | 116243 |
| Urban | 30 | 410 | 86100 |
| **Total** | **329** | **9948** | **220822** |
| | Match Rate | Ungeocoded | |
| Rural | 94.13% | 1152 | |
| Fringe | 93.15% | 8552 | |
| Urban | 98.50% | 1310 | |
| **Total** | **95.25%** | **11014** | |

### *West Region*

In the West region, displayed in Table 5, match rates for both TeleAtlas and NAVTEQ were at their best, however it was the NAVTEQ streets that resulted in higher match rates in all three classifications. Geocoding with strictly NAVTEQ yielded a difference of 2.17% for the rural addresses, 2.88% for the fringe addresses, and only 0.35% for the urban addresses. Geocoding using both reference datasets resulted in 1.78% increase from the NAVTEQ streets for the rural addresses, a 0.80% increase for the fringe addresses, and only a 0.44% increase for the urban addresses. The overall increase for all three classifications when combining the reference datasets was 0.82%.

Table 5. Match Rate Statistics for Addresses in the West Region.

| **West TeleAtlas** | | | |
|---|---|---|---|
| | Total Records | Geocode TELE | Perfect TELE |
| Rural | 64395 | 396 | 57360 |
| Fringe | 119236 | 887 | 110370 |
| Urban | 157572 | 392 | 154822 |
| **Total** | **341203** | **1675** | **322552** |
| | Total Geocode | Match Rate | Ungeocoded |
| Rural | 57756 | 89.69% | 6639 |
| Fringe | 111257 | 93.31% | 7979 |
| Urban | 155214 | 98.50% | 2358 |
| **Total** | **324227** | **95.02%** | **16976** |
| **West NAVTEQ** | | | |
| | Total Records | Geocode NAV | Geocode NAV |
| Rural | 64395 | 1025 | 58127 |
| Fringe | 119236 | 4987 | 109701 |
| Urban | 157572 | 2565 | 153196 |
| **Total** | **341203** | **8577** | **321024** |
| | Total Geocode | Match Rate | Ungeocoded |
| Rural | 59152 | 91.86% | 5243 |
| Fringe | 114688 | 96.19% | 4548 |
| Urban | 155761 | 98.85% | 1811 |
| **Total** | **329601** | **96.60%** | **11602** |
| **West Combined** | | | |
| | Total Records | Geocode TELE | Perfect TELE |
| Rural | 64395 | 392 | 57360 |
| Fringe | 119236 | 887 | 110370 |
| Urban | 157572 | 392 | 154822 |
| **Total** | **341203** | **1671** | **322552** |
| | Geocode NAV | Perfect NAV | Total Geocode |
| Rural | 166 | 2379 | 60297 |
| Fringe | 205 | 4181 | 115643 |
| Urban | 26 | 1207 | 156447 |
| **Total** | **397** | **7767** | **332387** |
| | Match Rate | Ungeocoded | |
| Rural | 93.64% | 4098 | |
| Fringe | 96.99% | 3593 | |
| Urban | 99.29% | 1125 | |
| **Total** | **97.42%** | **8816** | |

## Discussion

With the exception of the East region, match rates were greater for the fringe addresses than the rural addresses. Furthermore, match rates were greater for the urban addresses than the fringe addresses. This should be the trend in almost every situation simply because street data becomes more accurate and up to date when representing streets in higher populated places. More accurate street data in higher populated areas are due to the presence of older, more established streets along with the location of businesses and consumers that are considered valuable by the data vendors.

The NAVTEQ data yielded higher match rates than the TeleAtlas streets in most cases during this study. The only case where the match rates using the TeleAtlas streets were considerably higher was the rural addresses in the East region. NAVTEQ has made reporting discrepancies in their data quite easy for any consumer with access to the internet. This factor may be one of the main reasons why more addresses geocoded to NAVTEQ streets than TeleAtlas.

Geocoding using multiple data sets yielded higher match rates in all 12 of the test geocode routines conducted in this study. Improvements were definitely greater in the rural areas where percentages were smaller when using only one dataset. Improvements in the fringe addresses exceeded 2% only once and never breached 1% with the combined streets. Thus, the question of whether or not combining reference datasets is necessary for geocoding addresses in densely populated areas comes into consideration. Furthermore, the methodology of creating the hybrid street dataset used in this study could be possibly improved upon for yielding even higher match rates. In this case, the TeleAtlas streets were used for the initial geocode prior to matching them with the NAVTEQ streets. Since NAVTEQ yielded higher percentages than TeleAtlas on its own, it is possible that NAVTEQ should be used as the reference dataset for the initial geocode with TeleAtlas as the secondary set.

### *Project Parameters*

Limitations in the research of this project that might have hindered the results to the point where conclusions may have been inaccurate include any unforeseen errors or inaccuracies in the tabular address datasets, or any errors with the geocoding algorithms embedded within the interpolation technique. Furthermore, the fact that the address records for this study were only from 16 states could mean that the results did not reflect the geocoding trends that would occur in all 50 states. Ultimately, access to valid addresses throughout the entire U.S. would have resulted in the most accurate conclusions.

## Conclusions

This study was conducted in order to identify which of the two most popular base map data vendors result in higher address geocode match rates for selected regions of the United States. Furthermore, the results of this research was to be used to analyze what the expected improvements in match rates could be when using a hybrid reference dataset comprised of both TeleAtlas and NAVTEQ street data. The results of this study were successful in finding the above conclusions.

In most cases, the street data from NAVTEQ will give a GIS user higher match rates when address geocoding. The NAVTEQ datasets appear to be more complete and updated than the TeleAtlas Dynamap 2000 street datasets in most rural and fringe areas in the United States. The differences in match rates between the two datasets were extremely similar when geocoding urban addresses. It was less than 1% in every case. If a GIS user was required to geocode addresses that are only located in heavily populated areas, then either reference dataset would yield exceptional match rates. However, the NAVTEQ streets would be a much

better product when geocoding rural addresses.

Combining the TeleAtlas and NAVTEQ streets in order to successfully geocode more points in a source address dataset does result in higher match rates. However, the question of which street dataset is used first or for what types of addresses becomes a major factor. It can be concluded that NAVTEQ, the more complete and updated road-centerline dataset should be used for the initial geocode, while the TeleAtlas road-centerlines would be used for the secondary match routine. Again, this tool would be successful when geocoding large numbers of rural addresses where the validity of accurate street data becomes less of a factor to consumers and businesses. Any research, analysis, or map production of addresses in rural areas would become more accurate if a GIS user was to combine these two reference datasets into his or her geocode routine. Match rates increased by a very small margin for addresses in urban areas. Therefore, the use of two reference datasets may not be necessary.

**Acknowledgements**

I would like to thank Mr. John Ebert and Mr. Patrick Thorsell, professors from the Department of Resource Analysis at Saint Mary's University of Minnesota for providing guidance and assistance when it was so greatly needed throughout the course of this research project. I would also like to thank Mr. Matt Rantala, Senior GIS Consultant at Applied Data Consultants, Inc. for providing his support on this project by troubleshooting any issues with the creation of the source datasets and reference datasets in ArcGIS 9.2 Also, Matt assisted in running the modified geocoding program in ArcView 3.3.

**References**

Hassan K. A., Ducik M., Rasdorf W. 2004. Evaluation of Uncertainties Associated with Geocoding Techniques. *University of Pittsburgh, Department of Information Sciences and Telecommunications.* Retrieved February 2007, from http://www.blackwell-synergy.com/links/doi/10.1111%2Fj.1467-8667.2004.00346.x

Mills, J. 2002. Tain't Necessarily So: Address Geocoding in the Real World. *University of Texas at Tyler, Office of Research Services/Geographics Information Systems Lab*. Retrieved February 2007, http://gis2.esri.com/library/userconf/proc99/proceed/papers/pap623/p623.htm

*NAVTEQ Data.* 2007 NAVTEQ. Retrieved May 7, 2007 from http://www.navteq.com/about/data.htm

*TeleAtlas Dynamap*. 2007 TeleAtlas. Retrieved June 17, 2007 from http://www.teleatlas.com/OurProducts/MapData/Dynamap/index.htm