Centrus GeoCoder for ArcGIS User Guide

Software Release 1.6

April 2007

For Windows

GROUP 1 SOFTWARE, INC
DOCUMENTATION DEPARTMENT
4200 PARLIAMENT PL STE 600
LANHAM MD 20706-1844

USPS Notice

Group 1 Software, Inc. holds a nonexclusive license to publish and sell ZIP+4 databases on optical and magnetic media. The price of the Group 1 product is neither established, controlled, nor approved by the U.S. Postal Service.

A Reflection of Our Commitment to You... The Group 1 Software Mission Statement

"The first and foremost objective of Group 1 Software is the total, unreserved satisfaction of each Group 1 Client — satisfaction with Group 1 products and product support, with Group 1 people and with all other facets of the Client relationship. Every other Corporate objective is subordinate to and is addressed by this one."

DKCE16APUG

Contents

Preface

This preface provides an introduction to Centrus GeoCoder for ArcGIS and the *Centrus GeoCoder for ArcGIS User Guide*. It contains the following information:

- Who Should Read This Manual
- About This Guide
- Contacting Technical Support

Who Should Read This Manual

This manual is intended for users of ESRI's ArcGIS products (including ArcMap and ArcCatalog) who wish to cleanse and geocode addresses to a high degree of accuracy.

To use the Centrus GeoCoder for ArcGIS, you should be familiar with the ArcGIS product suite, and the data you wish to geocode.

About This Guide

This section provides the following information to help you use this guide:

- Typographic Conventions
- Chapter Descriptions

Centrus GeoCoder for ArcGIS for ArcGIS uses several data files, which Group 1 Software provides on the *Data Products Suite CDs*. This user guide does not provide the locations of the data files on the *Data Products Suite CDs*. For descriptions and locations of these files, refer to the *Release Notes for the Centrus Data Products*.

Typographic Conventions

To distinguish text that appears in Centrus GeoCoder for ArcGIS's user interface, this guide uses the following typographic conventions:

- Menu items and buttons are shown in **bold**, as are functions and dialog box options.
- Sections of dialog boxes are shown in *italics*.

The notation, **Geocoding > Geocode Addresses,** indicates that you should click the **Geocoding** menu and choose **Goecode Addresses** from the resulting drop-down menu.

Chapter Descriptions

This guide contains the following chapters:

- Chapter 1, "Introduction to Address Standardization and Geocoding," provides a quick overview of address standardization and geocoding.
- Chapter 2, "Installation," describes system requirements and how to install Centrus GeoCoder for ArcGIS.
- Chapter 3, "Accessing GeoCoder for ArcGIS," describes the process of accessing Centrus GeoCoder for ArcGIS through ArcMap and ArcCatalog.
- Chapter 4, "Using GeoCoder for ArcGIS," describes how to use Centrus GeoCoder for ArcGIS, and Centrus GeoCoder for ArcGIS's tools for finding records and manipulating processed data.
- Appendix A, "System Messages and Codes," describes how to interpret the codes Centrus GeoCoder for ArcGIS generates to provide detailed information about the matches retrieved, such as whether or not a match was found, information about the type of match found, information about why no match was found, and information about the geocode assigned.
- Appendix A, "Reference," provides reference information about the Query dialog box.

This guide also contains a Glossary and an Index.

Contacting Technical Support

If you have any technical questions regarding the use of Centrus GeoCoder for ArcGIS, please contact Centrus Technical Support:

Group 1 Software, Inc.
4200 PARLIAMENT PL STE 600
LANHAM, MD 20706-1844

Phone: 1-800-367-6950

Web: *www.g1.com/support*

Visit *www.g1.com/support* for patches, FAQs, product manuals, and a comprehensive industry glossary.

Downloads and technical documentation are password protected on the *www.g1.com/support*. See the `readme.txt` file for user name and password information. The username and password are case-sensitive and change bimonthly.

When contacting Centrus Technical Support, please provide the following information so that Group 1 Software can be more effective in helping you:

- Product name
- Company name under which the product is registered
- Company name with which you are affiliated
- Telephone number at which you may be contacted
- Development environment (if applicable)
- Thorough description of the problem

Centrus products make every effort to return accurate matches whenever possible, while providing the fewest false positive matches. No product can match with 100% accuracy or handle every data error. Some inaccuracies are inherent in the software, as trade-offs are made in the match logic to balance high match rates with low false positive matches.

The manuals available on www.g1.com/support supersede the printed documents and should be considered the current version. If there are any questions regarding current information, refer to the manuals available online.

Chapter 1

Introduction to Address Standardization and Geocoding

This chapter gives a basic overview of how address standardization and geocoding functionality works.

Topics include:

- About Address Standardization and Geocoding
- How Address Standardization and Geocoding Works

About Address Standardization and Geocoding

The Centrus GeoCoder for ArcGIS provides address standardization and geocoding capabilities that you can use with any data set containing address information. The Centrus GeoCoder for ArcGIS performs the following tasks:

- Identifies and corrects data entry errors in addresses.
- Standardizes addresses to United States Postal Service (USPS) standards.
- Appends new data about your address information, including ZIP+4 and Delivery Point barcodes.
- Identifies poorly formed or undeliverable addresses.
- Reduces delivery costs and increases delivery speed by supplying correct USPS information.
- Using address geocoding, assigns Latitude, Longitude, and Census ID information to the Block level.

- Using best-centroid geocoding, assigns a geocode when address geocoding is unavailable. ZIP+4 centroids return Census ID information to the Block Group level.

The Centrus GeoCoder for ArcGIS cleanses addresses to the highest level, based upon USPS CASS standards. Over 20 million addresses change every year. Regular use of the Centrus GeoCoder ensures the integrity and effectiveness of your address data.

Best results are achieved when addresses are formatted according to the USPS guidelines outlined in "Publication 28, Postal Addressing Standards." This document is available free of charge from the USPS. For more information, contact the USPS National Customer Support Center in Memphis, TN at 1-800-238-3150.

Address cleansing is an important early step in data transformation. Addresses must be tagged with geocodes, such as latitude and longitude coordinates or a Census ID, before they can be matched with household and business demographic data, spatially mapped, or pegged as in or out of a defined area.

How Address Standardization and Geocoding Works

The Centrus GeoCoder compares the street addresses you provide to the records in the USPS ZIP+4 Directory and the enhanced street network files. If the address is located in the USPS files, the address is standardized and a ZIP+4 (and all other USPS information) can be returned. If the address is also located within the street network files, the Centrus GeoCoder can return a very accurate latitude and longitude. If the address was not found in the enhanced street network files, location and census information is then taken from the ZIP+4, ZIP+2, or ZIP Code centroid files.

The overall standardization and geocoding success rates are determined by the quality of the addresses provided to Centrus GeoCoder. The Centrus GeoCoder can correct most minor misspellings as well as missing or incorrect directionals, street types, and ZIP Codes. However, if an address has an incorrect street number, or if the address contains a number of errors throughout, the Centrus GeoCoder might not be able to make a successful match.

Chapter 2

Installation

This chapter describes system requirements and how to install Centrus GeoCoder for ArcGIS on the Windows platform. It includes the following information:

- Installation Requirements
- Licensing
- Installation Procedure

Installation Requirements

Centrus GeoCoder for ArcGIS has the following minimum system requirements:

- A computer running Windows 2000/NT/XP
- ArcGIS 8.3 or 9.0 must be installed prior to installing the Centrus GeoCoder for ArcGIS plug-in.
- The plug-in has not been tested with any other versions of ArcGIS.
- Pentium processor or higher
- 16 megabytes of memory
- A hard disk with a minimum of 10 megabytes of free disk space (additional disk space required if copying database files from the CD to your hard drive)
- 3 1/2″ floppy disk drive

- CD-ROM drive
- Mouse or other pointing device

Note: Hard disk space requirements vary with the Centrus GeoCoder for ArcGIS and data files licensed. A fast disk drive, fast processor, and more memory improves performance.

Licensing

Centrus GeoCoder for ArcGIS uses several data files, which Group 1 Software provides on the *Data Products Suite* CDs. The data files you are licensed to use are installed during the data installation process. Refer to the *Release Notes for the Centrus Data Products* for data installation instructions.

Installation Procedure

The installation procedure depends on whether you are installing Centrus GeoCoder for ArcGIS for the first time, or you have a previous version already installed on your machine. If you have a previous version of Centrus GeoCoder for ArcGIS, the installation program detects the previously installed version.

The procedure for installing Centrus GeoCoder for ArcGIS includes:

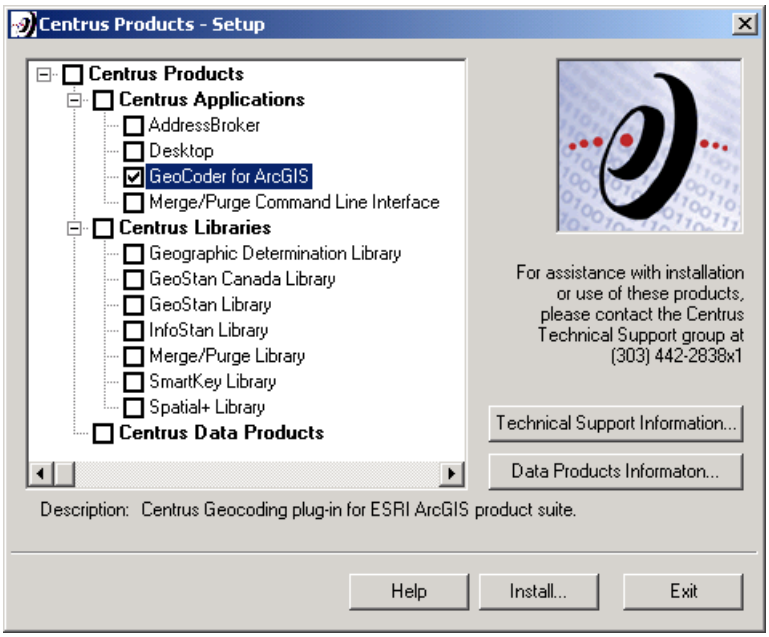
- Step 1– Install the Product
- Step 2– Install the Data Files

Step 1– Install the Product

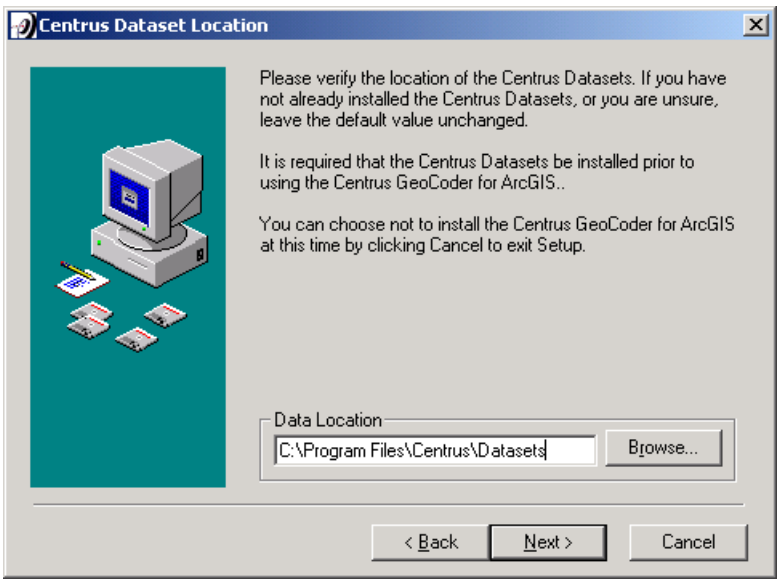
1. Close all Windows programs.
2. Place the *Software Installation* CD in the CD-ROM reader.

   The installation program starts automatically if you have AutoPlay enabled. If the installation program does not start automatically, double-click QMSstart.exe in the root directory of the CD-ROM to start the main installation application.

3. Select **GeoCoder for ArcGIS** in the **Centrus Products - Setup** dialog box.



4. Press the **Install** button to begin the installation process.
5. Click **Next** in the **Welcome** dialog box if all Windows programs are closed. Click **Cancel** if you need to quit Setup and close programs.
6. Read the license agreement and click **I Agree** if you accept the terms of the license agreement.
7. Read the Readme file and click **Next**.
8. Click **Next** to advance to the *Centrus Dataset Location* page.



9. If you have already installed the Centrus Data Products, specify the location to which the data was installed. Otherwise, you may specify the location to which you intend to install the Centrus Data Products, or leave the default value unchanged.
10. Click **Next** to advance to *License Information* page. Select the **Update License File** check box if this is the first time you have installed GeoCoder for ArcGIS.

    Note: For first time installations, this check box may not be unchecked.
11. Click **Next** to advance to the *Start Installation* page to begin the license installation process.
12. If this is a first time install or if you request to update your license file, you are prompted to specify the location of the new license file.
13. Once the Centrus GeoCoder for ArcGIS installation is complete, be sure to install the Centrus Data Products if you have not already done so.

Step 2– Install the Data Files

Note: If you plan to use WinSplit, you do not need to complete the steps provided here. The *Centrus Utilities Manual* provides all the steps necessary to split and install the data files.

The data files required for processing must reside on your machine, or the CD-ROMs must be directly accessible to your machine. For descriptions of the data files you may need to complete the desired processing, click **Data Products Information** on the **Centrus Products - Setup** screen. The *Release Notes for the Centrus Data Products* included in the installation package also describe the data files.

To copy data to your local drive:

1. Click **Centrus Data Products** on the **Centrus Products - Setup** dialog box and click **Install**.
2. Click **Next** on the **Welcome** dialog box.
3. Read the license agreement and click **I Agree**.
4. Read the data products Readme file and click **Next**.
5. Browse for the directory you want to copy the data to and click **Next**.
6. Select one or more data sets to install.
7. Click **HTML Docs** or **Data Info** to review data set information and data processing requirements, and then click **Next**.
8. Click **Next** to start installing the data.

   Insert the appropriate data CDs when prompted.
9. Add the data files path to Centrus GeoCoder for ArcGIS's Search Path.

Chapter 3

Accessing GeoCoder for ArcGIS

GeoCoder for ArcGIS may be accessed by the following:

- ArcMap
- ArcCatalog

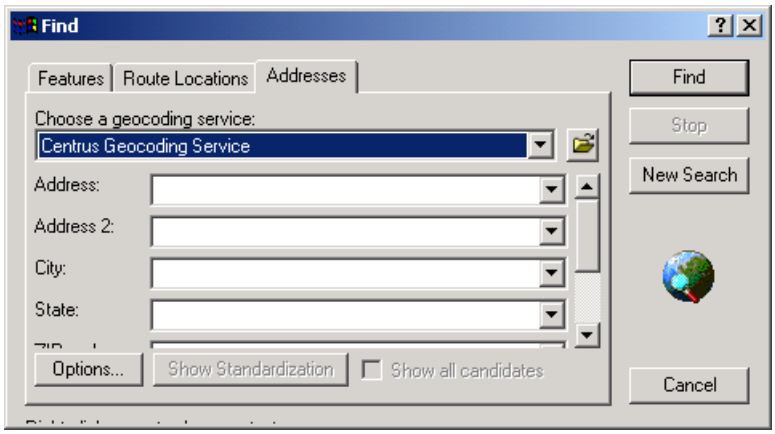
ArcMap

With ArcMap, you can geocode a single address or a table of addresses.

Geocoding a Single Address

1. Select **Edit > Find…** from the ArcMap window.

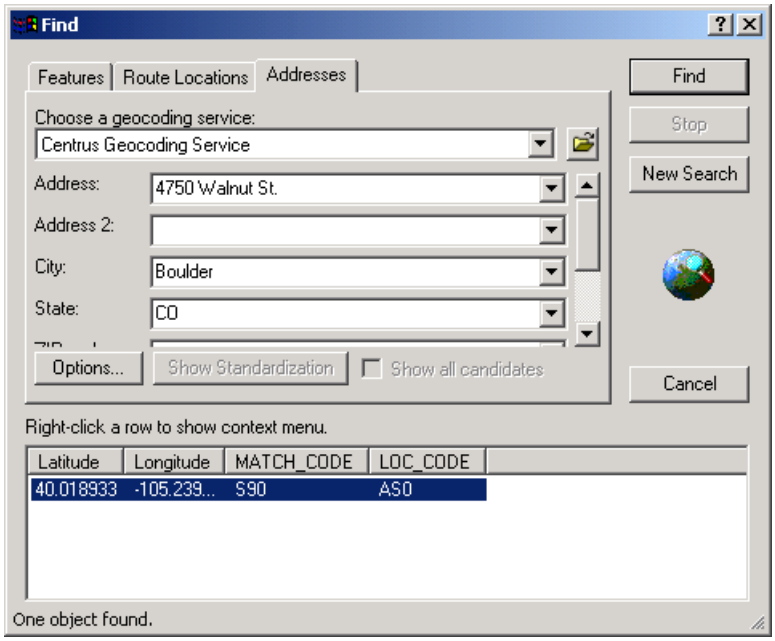
2. Select the *Centrus Geocoding Service* in the **Choose a geocoding service** drop down list. If the service is not listed, press the folder button to the right of the list and browse for the *Centrus Geocoding Service* specification file located in the Bin subdirectory of the ArcGIS program folder.



3. Click **Options…** to define the desired match options and output fields. See the sections "Output Fields" on page 13 and "Processing Options" on page 15 for more information.

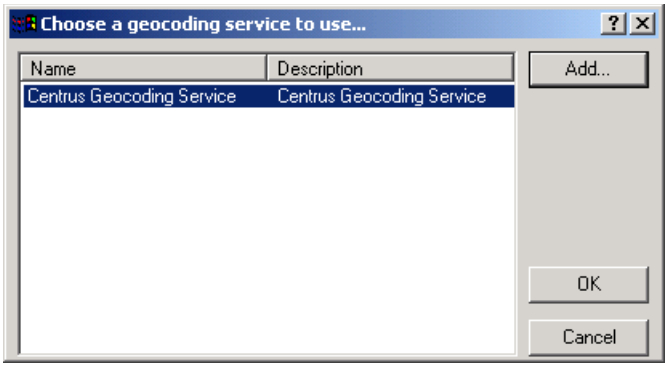
4. Enter the input address information in the correct input fields, and click **Find**.

   Note: Modifying your match options affect the results of the current input address.



Geocoding a Table of Addresses

1. Select **Tools > Geocoding > Geocode Addresses…** from the ArcMap window.
2. Select the *Centrus Geocoding Service* from the list of available geocoding services. If the service is not listed, click **Add…** to the right of the list and browse for the *Centrus Geocoding Service* specification file located in the Bin subdirectory of the ArcGIS program folder.



3. Click **OK**.
4. Browse for the input Address Table you want to geocode.
5. Define the input fields to be used for geocoding. See the section "Input Fields" on page 12 for more information.
6. Click **Geocoding Options…** to define the desired match options and output fields. See the sections "Output Fields" on page 13 and "Processing Options" on page 15 for more information.
7. Specify the table which is to receive the results of the geocoding operation.
8. Click **OK** to begin processing. See the section "Interactive Address Processing Options" on page 19 for more information.

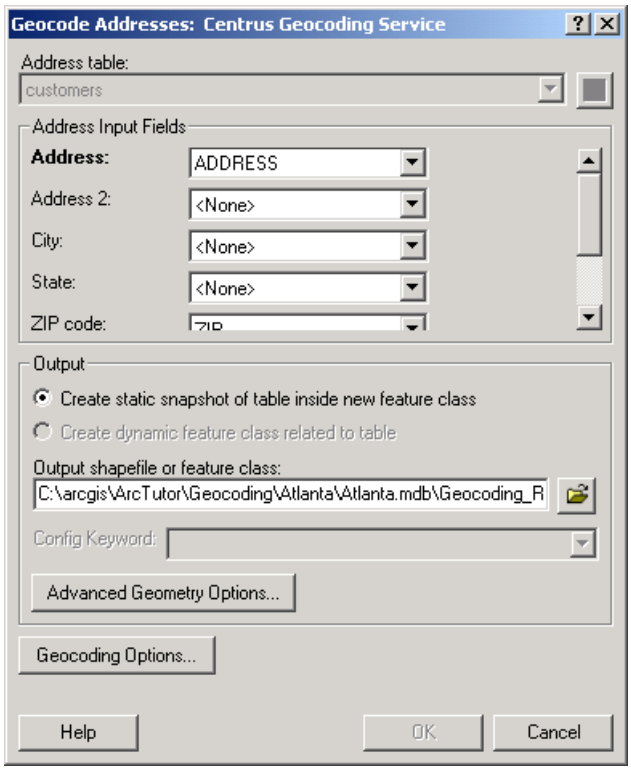
ArcCatalog

With ArcCatalog, you can geocode a single address or a table of addresses.

Geocoding a Single Address

1. Locate the address table you wish to geocode using the ArcCatalog browser.
2. Right click on the desired table and select **Geocode Addresses….**
3. Select the *Centrus Geocoding Service* from the list of available geocoding services. If the service is not listed, click **Browse for service…** to the right of the list and browse for the *Centrus Geocoding Service* specification file located in the Bin subdirectory of the ArcGIS program folder.

4. Click OK.
5. Define the input fields to be used for geocoding. See the section "Input Fields" on page 12 for more information.



6. Specify the table which is to receive the results of the geocoding operation.
7. Click **Geocoding Options…** to define the desired match options and output fields. See the sections "Output Fields" on page 13 and "Processing Options" on page 15 for more information.
8. Click **OK** to begin processing. See the section "Interactive Address Processing Options" on page 19 for more information.

Chapter 4

Using GeoCoder for ArcGIS

This chapter provides a comprehensive reference for the GeoCoder for ArcGIS tabs and options.

Topics include:

- Input Fields
- Output Fields
- Processing Options
- Interactive Address Processing Options

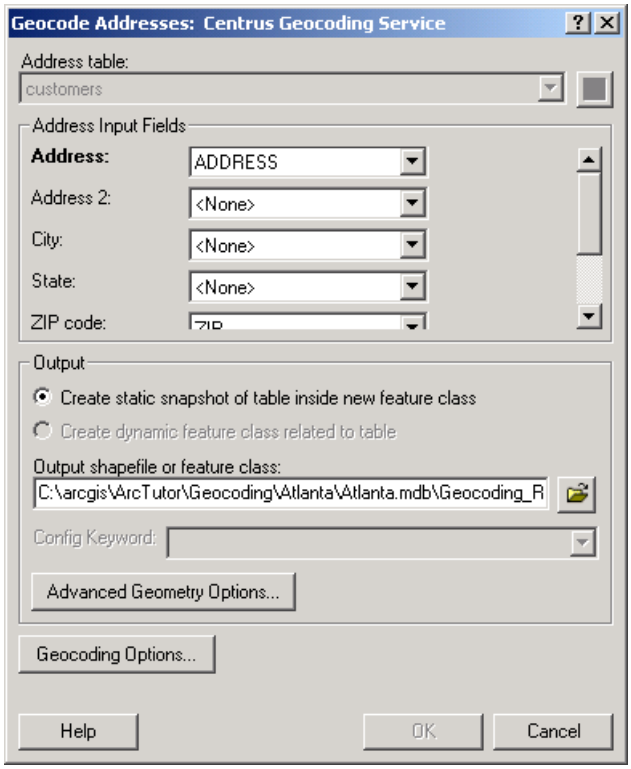
Input Fields

You may assign input fields from the source table that are to be validated against the street network data.

Assigning Input Fields

To assign input fields:

➢ Choose the appropriate field name from each drop-down list box as desired.



Output Fields

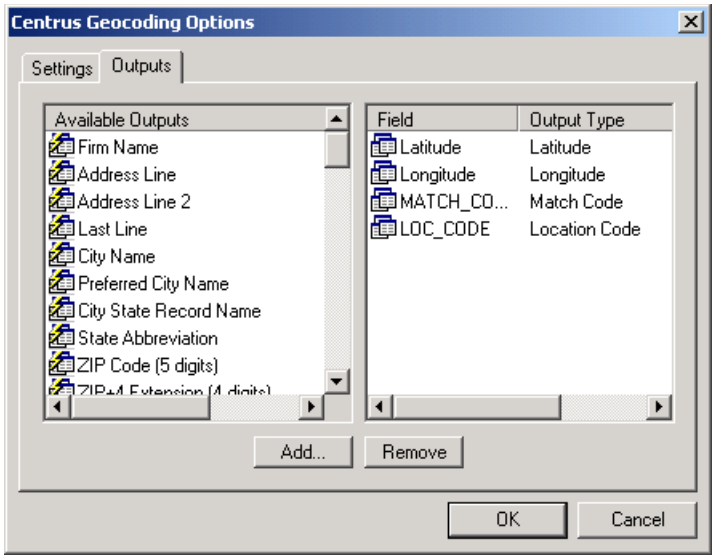
Use the **Output Fields** tab to assign fields that will be output with available output types. You may only output information to new fields; you may not replace existing field values. When you click the **Output Fields** tab, all of the fields that are already defined are listed in the Field column. The Available Outputs column lists all the available address standardization and geocoding output fields.

Note: According to ESRI standards, these output fields are always provided as columns filled with data whether or not they are requested: ARC_Firm, ARC_Address, ARC_Addr2, ARC_City, ARC_State, ARC_ZIP, and ARC_ZIP4.

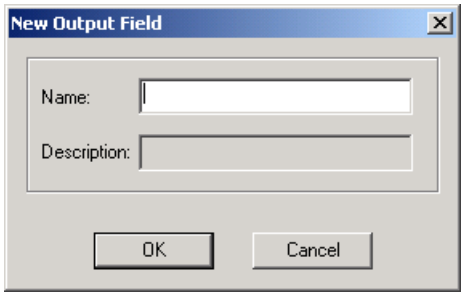
Adding Output Fields

To add an available output field:

1. From the Available Outputs list, select an Available Output type.



2. Click **Add...** The **New Output Field** dialog displays.



3. Type a field name and click **OK**, or accept the default field name by clicking **OK**.

Note: You can also open the **New Output Field** dialog by double-clicking an Available Output type.

The new field name appears in the Field column with the output type.

Removing Output Fields

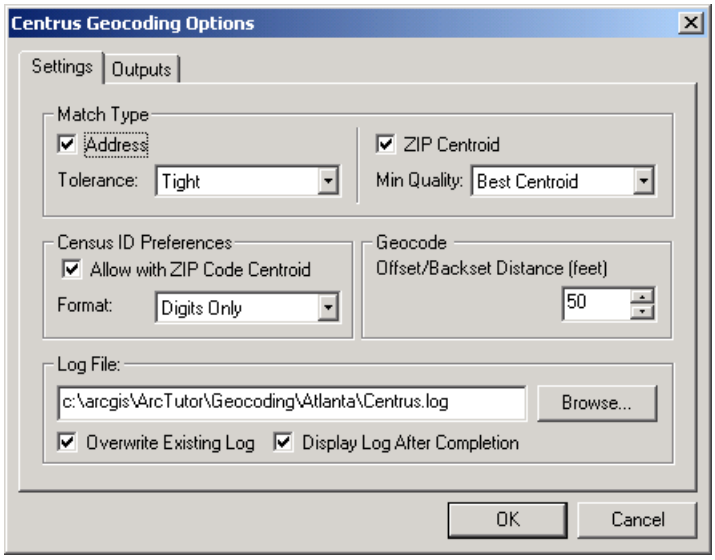
To remove an output field:

1. Select a field from the Field column at the right.
2. Click **Remove**. This removes the selected field from output.

Note: You can also unassign an output field by double-clicking it.

Processing Options

The Centrus GeoCoder for ArcGIS processing options appear on the Centrus Geocoding Options dialog.



Address Matching

When the **Address** matching check box is selected, Centrus GeoCoder for ArcGIS performs address standardization and address geocoding. The setting indicates the type of match that is required. The possible settings for address matching **Tolerance** are:

- **Tight** requires that addresses are virtually an exact match. For the fastest processing, use Tight only when the addresses have already been standardized to USPS standards.
- **Close** is the recommended setting for most scenarios. It allows minor misspellings, as well as incorrect or missing directionals or street types. Close is the default setting.
- **Extended** is similar to Close, except that the street name search considers all streets that have the same first letter, rather than using pattern matching. This setting also performs the widest search possible, using the USPS Finance Area. Group 1 Software recommends that you review all matches made by this setting in which the street name was modified. You can identify these matches via the match code. Extended is the slowest setting.

ZIP Centroid Matching

When the **ZIP Centroid** matching check box is selected, Centrus GeoCoder provides a ZIP+4, ZIP+2 or ZIP Code centroid geocode when an address geocode is not available. The **Minimum Quality** setting indicates which type of centroids are used. The best possible centroid is always used if multiple centroids are available. The possible settings are:

- **9 Digit Best Location** uses the most positionally accurate 9-digit centroids. These centroids are almost as accurate as an address geocode, in that they are accurate to a single block face in most cases. Census accuracy varies, but most centroids in this class are accurate to the Block Group level. This is the default value.
- **9 Digit Good Location** uses 9-digit centroids with very good positional accuracy. These centroids are accurate to a single block face in most cases. Census accuracy varies, but most centroids in this class are accurate to the Block Group or Census Tract level.
- **9 or 7 Digit Location** uses all 9- or 7-digit centroids. A 7-digit centroid is positionally accurate to within several blocks in most areas. Census accuracy varies. This setting yields all but 5-digit centroids.
- **Block Group Accuracy** uses centroids that are accurate to the Block Group level. Positional accuracy varies, but is normally accurate to the city block on which the address is located.
- **Census Tract Accuracy** uses centroids that are accurate to the Census Tract level. Positional accuracy varies, but is normally accurate to within several city blocks of where the address is located.
- **5 Digit Only** uses only 5-digit centroids. This is the least accurate geocode, but it is the fastest to assign.
- **Best Centroid** uses the best centroid available.

Census ID Preferences

The Census ID **Format** option determines the format of the Census FIPS Code. The default setting is Census. The possible formats are listed below with an example of how a Block Group is displayed. When selected, the **Allow with 5 digit ZIP Code Centroid** check box permits the assignment of the Census FIPS Code even when only 5-digit ZIP Code centroids are available. This option is not selected as the default setting.

- **Delimited** 08 013 0002.01-1
- **Census** 80130002.011
- **Digits Only** 080130002011

Offset/Backset

If an offset distance is used, that distance is calculated perpendicular to the portion of the street segment in which it lands. This distance calculation yields the best visual representation for any mapping package, as well as be the most accurate location able to be imputed from the geographic data. Block data is returned for the side of the segment on which a match was found. This option allows you to override the default offset distance of 50 feet. The backset value is the same as the offset value or half the length of the segment, whichever is smaller. The offset/backset value must be an integer value between 0 and 999.

Log File

The Log report returns a complete list of the settings used to process the table. It also contains statistics regarding how clean the file was before being processed and how many records were processed by Centrus GeoCoder for ArcGIS.

Centrus.log - Notepad

File Edit Format Help

```
Centrus Address Coding Report
-------------------------------------------------------------------------------
  Output Table........................... Geocoding_Result
  Process Completed...................... 04/16/2003 02:31:58 PM

Options:
  Search Path for Address Geocoding...... C:\Program Files\Centrus\Datasets
  Address Matching....................... YES
    Matching Tolerance................... Tight
  ZIP Centroid Matching.................. YES
    Minimum Quality...................... Best Centroid
  Census ID Format....................... Digits Only

ADDRESS STATISTICS:
Records Processed.......................     50
  Addresses Matched.....................     49 (98.00%)
    Standardized........................     45
    Intersections.......................      1
    Non-USPS............................      3

  Unmatched.............................      1 (2.00%)
  Valid Unmatched Zips..................      0
    E001: Low level error...............      0
    E002: State file not found..........      0
    E003: Invalid file..................      0
    E004: File is out of date...........      0
    E010: Unrecognized last line........      0
    E011: ZIP not found.................      0
    E012: City not found................      0
    E013: City not unique...............      0
    E014: Out of licensed area..........      0
    E020: No matching streets...........      1
    E021: No cross streets..............      0
    E022: No matching ranges............      0
    E023: Match is unresolved...........      0
    E025: Too many cross streets........      0
    E026: Unable to find Address input...     0
    E027: Invalid Directional attempted..     0
    E028: Also matched to EWS data......      0

  Standardized Address Quality:
    Original Address Unchanged..........     49 (100.00%)
    Original Last Line Unchanged........      0 (0.00%)
    Corrected prefix direction..........      0
    Corrected street name...............      0
    Corrected street suffix.............      0
    Corrected postfix direction.........      0
    Corrected city name.................     48
    Corrected state abbreviation........     48
    Corrected ZIP Code..................      4
    Corrected ZIP+4 add on..............     45

  Geocodes Assigned.....................     50 (100.00%)
    Address...........................     48 (96.00%)
    ZIP Centroids.....................      2 (4.00%)
       Location Accuracy:
          ZIP+4.......................      1 (2.00%)
          ZIP+2.......................      0 (0.00%)
          ZIP Code....................      1 (2.00%)
       Census Accuracy:
          Block Group.................      1
          Census Tract................      0
          County......................      1

Note: All percentages are based on records processed.
```

Note: By default, any previous log file information is truncated after each job. If instead you wish to append log information to the end of a previous log file, select the check box labeled **Overwrite Existing Log** to clear it.

You may display this report in Windows Notepad (if Notepad is available) by selecting the **Display Log After Completion** check box. An example report is displayed below.

To specify a log report:

➢ Click **Browse** next to the **Log File** text box, and specify the path and file name in the **File Selection** dialog box.

Interactive Address Processing Options

The **Input Address** section of the **Process** dialog displays the record as it exists in the input table. Inputs are directly editable. The processed information is shown in the **Results** section of the dialog box. The data is presented as a hierarchical tree. You can double-click an item to expand or contract the information display. The **Records** section contains the processing statistics. It shows the total records processed and the number of records standardized and geocoded, as well as the estimated completion time. Within the **Process** dialog box, you can perform several different functions using the buttons on the right side of the dialog box. A description of these functions follows.

Find

Click **Find** to process the current address as it appears in the **Inputs** section. The results of the process are displayed in the **Results** section of the dialog box. The address may be edited and geocoded as often as necessary in order to receive a proper match.

Query

If you are unable to produce a match by editing an address and using **Find**, click **Query** to display the **Query** dialog box. See the section "About the Query Dialog Box" on page 31 for more information.

Reload

Click **Reload** to load the original data from the address file into the **Inputs** and **Results** sections. To geocode the record, click **Find**. This button is useful if you have made a number of edits to an address in an attempt to receive a geocode, and would like to return to the original address in the address file.

Next

Click **Next** to save the current information in the **Results** section to the output file and load the next address for processing. This function is useful for interactive processing of a file, or if you wish to verify the results of every match.

Next Error

Click **Next Error** to process the address file until the next unmatched record is encountered. The unmatched record displays in the **Inputs** and **Results** sections, along with the reason that no match could be made. The address can be edited to correct any errors or other flaws in the address. Click **Next**, **Next Error,** or **Batch** to continue processing. This feature is useful if you wish to maximize the number of matches that can be received.

Goto

Click **Goto** to jump to an absolute record number.

Note: You may only jump forward in the file.

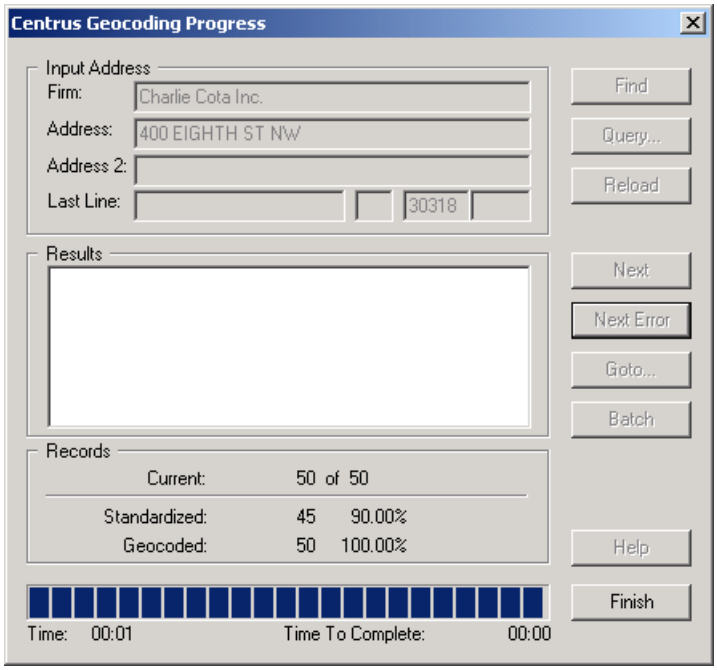
Batch

Click **Batch** to begin processing records without interruption. Centrus GeoCoder for ArcGIS processes the entire file, displaying statistics for the number of records processed. The **Inputs** and **Results** areas of the dialog box do not update during batch processing. Once batch processing begins, all buttons except **Stop** appear dimmed.

Click **Stop** to interrupt processing on the current record and display the same options that were available before batch processing began. The current record is displayed in the **Inputs** and **Results** sections.

Close

Click **Cancel** to abort the current geocoding process. The process stops, and records that have been processed remain coded.



Finish

Once processing is complete, click **Finish** to close the **Process** dialog and optionally display the processing log file.

Appendix A

System Messages and Codes

Centrus GeoCoder for ArcGIS uses match codes and location codes to provide detailed information about the matches retrieved. Match codes provide information about the match itself, and location codes provide information about the assigned geocode.

The following topics describe the two types of codes and explain how to interpret them:

- Match Codes
- Location Codes

Match Codes

When you use Centrus GeoCoder for ArcGIS to perform address standardization, a match code is returned in the `MatchCode` output field. The match code is an alpha-numeric code that encapsulates information about the address standardization process—including whether or not a match was found, information about the type of match found (when applicable), and information about why no match was found (when applicable). The following sections provide the information you need to interpret these codes:

- Match Codes for the United States

Match Codes for the United States

The following table contains the match code values.

Note: You can find a description of the hex digits for the different match codes in the table following the match code table.

Code	Description
Ahh	Same as Shh, but indicates match to an alias name record or an alternate record.
D00	Match is a small town with P.O. Box and/or General Delivery only.
Gxx	Match is to an auxiliary file.
Nxx	Match is to the nearest address. Used with reverse geocoding. The following are the only values for N: n NSO – Nearest street center match (nearest street segment interpolated) n NS1 – Nearest unranged street segment n NP0 – Nearest point address n NX0 – Nearest intersection
Qhh	Match to USPS range records with unique ZIP Codes. This match code was introduced for clarity because CASS rules prohibit altering an input ZIP if it matches a unique ZIP Code value.
Shh	Match found in USPS data. This is considered the best address match, because it matched directly against the USPS list of addresses.
Thh	Match to a street segment record. Street segment records do not contain ZIP Code information. If you enter a ZIP Code, the the application returns the ZIP Code you entered. If the input city and state has only one ZIP Code, the application returns that ZIP Code.
Uhh	Match found in USPS data but cannot resolve the ZIP+4 code without the firm name or other information. CASS mode returns an E023 (multiple match) error code.
Xhhh	Match found is for an intersection of two streets, for example, "Clay St & Michigan Ave." The first hex digit refers to the last line information, the second hex digit refers to the first street in the intersection, and the third hex digit refers to the second street in the intersection. **NOTE:** The USPS does not allow intersections as a valid deliverable address.
Yhhh	Same as Xhhh, but indicates the application used an alias name record for one or both streets. The first hex digit refers to the last line information, the second hex digit refers to the first street in the intersection, and the third hex digit refers to the second street in the intersection.
A	Address matches to an alias record.
Z	No address given, but the application verified the validity of the given ZIP Code.

The following table contains the description of the hex digits for the match code values.

Code	In first hex position means:	In second and third hex position means:
0	No change in last line.	No change in address line.
1	ZIP Code changed.	Street type changed.
2	City changed.	Pre-directional changed.
3	City and ZIP Code changed.	Street type and Pre-directional changed.
4	State changed.	Post-directional changed.
5	State and ZIP Code changed.	Street type and Post-directional changed.

Code	In first hex position means:	In second and third hex position means:
6	State and City changed.	Pre-directional and Post-directional changed.
7	State, City, and ZIP Code changed.	Street type, Pre-directional, and Post-directional changed.
8	ZIP+4 changed.	Street name changed.
9	ZIP and ZIP+4 changed.	Street name and Street type changed.
A	City and ZIP+4 changed.	Street name and Pre-directional changed.
B	City, ZIP, and ZIP+4 changed.	Street name, Street type, and Pre-directional changed.
C	State and ZIP+4 changed.	Street name and Post-directional changed.
D	State, ZIP, and ZIP+4 changed.	Street name, Street type, and Post-directional changed.
E	State, City, and ZIP+4 changed.	Street name, Pre-directional, and Post-directional changed.
F	State, City, ZIP, and ZIP+4 changed.	Street name, Street type, Pre-directional, and Post-directional changed.

The following table describes the values returned when the application cannot find a match code.

Code		Description
Enn n		Indicates an error, or no match. This can occur when the address entered does not exist in the GeoStan directory, or the address is badly formed and the application cannot parse the address correctly. The last three digits of an error code indicate which parts of an address the application could not match to the GeoStan Directory.
	`nnn = 000`	No address found.
	`nnn = 001`	Low level error.
	`nnn = 002`	Could not find GSD file.
	`nnn = 003`	Incorrect GSD file signature or version ID.
	`nnn = 004`	GSD file out of date. Only occurs when CASS Mode is on.
	`nnn = 010`	No city+state or ZIP Code found.
	`nnn = 011`	Input ZIP not in the directory.
	`nnn = 012`	Input city not in the directory.
	`nnn = 013`	Input city not unique in the directory.
	`nnn = 014`	Out of licensed area. Only occurs if using Group 1 licensing technology.
	`nnn = 015`	Record count is depleted and license has expired.
	`nnn = 020`	No matching streets found in directory.
	`nnn = 021`	No matching cross streets for an intersection match.
	`nnn = 022`	No matching ranges.
	`nnn = 023`	Unresolved match.
	`nnn = 024`	No matching ranges. (Same as 022.)
	`nnn = 025`	Too many possible cross streets for intersection matching.

Code	Description
`nnn = 026`	No address found when attempting a multi-line match.
`nnn = 027`	Invalid directional attempted.
`nnn = 028`	Record also matched EWS data, therefore the application denied the match.

Location Codes

Location codes indicate the accuracy of the assigned geocode. There are two types of geocodes—*Address* and *ZIP+4* centroids.

Address geocodes are simple to interpret because they indicate a geocode made directly to a street segment (or two segments in the case of an intersection). *ZIP+4* centroids, however, have a range of "confidence" depending upon how the ZIP+4 centroid was determined. An "E" code indicates that no geocode was possible. The following sections describe the location codes:

- Address Location Codes
- ZIP Centroid Matches

Address Location Codes

Address location codes detail the known qualities about the geocode.

An address location code has the following characters.

1st character Always an A indicating an address location.

2nd character May be one of the following

- *G* Auxiliary file data location
- *I* Application infers the correct segment from the candidate records
- *P* Point-level data location
- *S* Location on a street range
- *X* Location on an intersection of two streets

3rd character Digit indicating other qualities about the location.

The following table contains the address codes.

Code	Description
AGn	Indicates an auxiliary file for a geocode match where n is one of the following values:
n = 0	Geocode is a point geocode
n = 1	Geocode is an interpolated address along a segment
n = 2	Geocode is an interpolated address along a segment, and GeoStan cannot determine the side of the street from the data provided in the auxiliary file record
n = 3	Geocode is the midpoint of the segment
APnn	Indicates a point-level geocode match where nn is one of the following values:

Code	Description
nn = 02	Parcel centroid Indicates the center of an accessor's parcel (tract or lot) polygon. When the center of an irregularly shaped parcel falls outside of its polygon, the centroid is manually repositioned to fall inside the polygon as closely as possible to the actual center.
nn = 04	Address points Represents field-collected GPS points with field-collected address data
nn = 05	Structure centroid Indicates the center of an addressable building footprint polygon. An addressable structure is typically a structure that receives mail or has telephone service. Usually a residential address consists of a single building. For houses with outbuildings (detached garages, shed, barns, etc.), only the residences have a structure point. Condominiums and duplexes have multiple points for each building. Larger buildings, such as apartment complexes, typically receive mail at one address for each building and therefore individual apartments are not represented as discrete structure points. Shopping malls, industrial complexes, and academic or medical center campuses where one building accepts mail for the entire complex are represented as one point. When addresses are assigned to multiple bu8ildings within one complex, each addressed structure is represented by a point. If the center of a structure falls outside of its polygon, the center is manually repositioned to fall inside the polygon.
nn = 07	Manually placed Address points are manually placed to coincide with the midpoint of an accessor's parcel's street frontage at a distance from the center line.
AIn	Application infers the correct segment from the candidate records at match time.
ASn	House range address geocode. This is the most accurate geocode available.
Both AIn and ASn share the same qualities for n as follows:	
n = 0	Best location.
n = 1	Street side is unknown. The application assigned the Census FIPS Block ID from the left side; however, there is no assigned offset and the application placed the point directly on the street.
n = 2	May indicate one or both of the following: n Application interpolated the address onto a TIGER segment that did not initially contain address ranges initially. n The application changed the original segment name to match the US Postal Service spelling. This specifically refers to street type, predir, and postdir. **NOTE:** Because the application only completes segment range interpolation for TIGER data, only the second case is valid for non-TIGER data.
n = 3	Both 1 and 2.
n = 7	Placeholder. Used when starting and ending points of segments contain the same value and shape data is not available.
AXn	Intersection geocode. The digit at the end indicates the following:
n = 3	Standard single-point intersection computed from the center lines of street segments.
n = 8	Interpolated (divided-road) intersection geocode. The application will attempt to return a centroid for the intersection.

ZIP+4 Centroid Location Codes

ZIP+4 centroid location codes indicate the quality of two location attributes: Census ID accuracy and positional accuracy.

A ZIP+4 centroid location code has the following characters.

1st character
Always Z indicating a location derived from a ZIP centroid.

2nd character
Census ID accuracy.

3rd character
Location type.

4th character
How the application defined the location and Census ID. Provided for completeness, and may not be useful for most applications. If this is the case, you can set the width of the location code field to three so that the application does not return this identifier.

The following table contains the values and descriptions for the location codes.

Character Position	Code	Description
2nd Character		
	B	Block Group accuracy (most accurate).
	T	Census Tract accuracy.
	C	Unclassified Census accuracy. Normally accurate to at least the County level.
3rd Character		
	5	Location of the Post Office that delivers mail to that address, a 5-digit ZIP Code centroid, or a location based upon locale (city). See the 4th character for a precise indication of locational accuracy.
	7	Location based upon a ZIP+2 centroid. These locations can represent a multiple block area in urban locations, or a slightly larger area in rural settings.
	9	Location based upon a ZIP+4 centroid. These are the most accurate centroids and normally place the location on the correct block face. For a small number of records, the location may be the middle of the entire street on which the ZIP+4 falls. See the 4th character for a precise indication of locational accuracy.
4th Character		
	A	Address matched to a single segment. Location assigned in the middle of the matched street segment, offset to the proper side of the street.
	a	Address matched to a single segment, but the correct side of the street is unknown. Location assigned in the middle of the matched street segment, offset to the left side of the street, as address ranges increase.
	B	Address matched to multiple segments, all segments have the same Block Group. Location assigned to the middle of the matched street segment with the most house number ranges within this ZIP+4. Location offset to the proper side of the street.

Character Position	Code	Description
	b	Same as methodology B except the correct side of the street is unknown. Location assigned in the middle of the matched street segment, offset to the left side of the street, as address ranges increase.
	C	Address matched to multiple segments, with all segments having the same Census Tract. Returns the Block Group representing the most households in this ZIP+4. Location assigned to the middle of the matched street segment with the most house number ranges within this ZIP+4. Location offset to the proper side of the street.
	c	Same as methodology C except the correct side of the street is unknown. Location assigned in the middle of the matched street segment, offset to the left side of the street, as address ranges increase.
	D	Address matched to multiple segments, with all segments having the same County. Returns the Block Group representing the most households in this ZIP+4. Location assigned to the middle of the matched street segment with the most house number ranges within this ZIP+4. Location offset to the proper side of the street.
	d	Same as methodology D except the correct side of the street is unknown. Location assigned in the middle of the matched street segment, offset to the left side of the street, as address ranges increase.
	E	Street name matched; no house ranges available. All matched segments have the same Block Group. Location placed on the segment closest to the center of the matched segments. In most cases, this is on the mid-point of the entire street.
	F	Street name matched; no house ranges available. All matched segments have the same Census Tract. Location placed on the segment closest to the center of the matched segments. In most cases, this is on the mid-point of the entire street.
	G	Street name matched (no house ranges available). All matched segments have the same County. Location placed on the segment closest to the center of the matched segments. In most cases, this is on the mid-point of the entire street.
	H	Same as methodology G, but some segments are not in the same County. Used for less than .05 % of the centroids.
	I	Created ZIP+2 cluster centroid as defined by methodologies A, a, B, and b. All centroids in this ZIP+2 cluster have the same Block Group. Location assigned to the ZIP+2 centroid.
	J	Created ZIP+2 cluster centroid as defined by methodologies A, a, B, b, C, and c. All centroids in this ZIP+2 cluster have the same Census Tract. Location assigned to the ZIP+2 centroid.
	K	Created ZIP+2 cluster centroid as defined by methodologies A, a, B, b, C, c, D, and d. Location assigned to the ZIP+2 centroid.
	L	Created ZIP+2 cluster centroid as defined by methodology E. All centroids in this ZIP+2 cluster have the same Block Group. Location assigned to the ZIP+2 centroid.
	M	Created ZIP+2 cluster centroid as defined by methodology E and F. All centroids in this ZIP+2 cluster have the same Census Tract. Location assigned to the ZIP+2 centroid.
	N	Created ZIP+2 cluster centroid as defined by methodology E, F, G, and H. Location assigned to the ZIP+2 centroid.
	V	Over 95 % of addresses in this ZIP Code are in a single Census Tract. Location assigned to the ZIP Code centroid.
	W	Over 80 % of addresses in this ZIP Code are in a single Census Tract. Reasonable Census Tract accuracy. Location assigned to the ZIP Code centroid.
	X	Less than 80 % of addresses in this ZIP Code are in a single Census Tract. Census ID is uncertain. Location assigned to the ZIP Code centroid.

Character Position	Code	Description
	Y	Rural or sparsely populated area. Census code is uncertain. Location based upon the USGS places file.
	Z	P.O. Box or General Delivery addresses. Census code is uncertain. Location based upon the Post Office location that delivers the mail to that address.

ZIP Centroid Matches

In the **Options** dialog box, there are seven choices for **Minimum Quality** when selecting ZIP Centroid Matching. The following indicates which methodologies (the fourth character in the location code) are used for each **Quality** selection:

9-Digit Best Location	A, a, B, b, C, c, D, d
9-Digit Good Location	A, a, B, b, C, c, D, d, E, F, G, H
9- or 7-Digit Location	A, a, B, b, C, c, D, d, E, F, G, H, I, J, K, L, M, N
Block Group Accuracy	A, a, B, b, E, I, L
Census Tract Accuracy	A, a, B, b, E, I, L, C, c, F, J, M, V
5-Digit Only	V, W, X, Y, Z
All Centroids	Can return all Methodology codes.

Appendix A

Reference

This chapter describes the sections to the **Resolve the Address** dialog, also referred to as the **Query** dialog. It includes the following information:

- About the Query Dialog Box
- Specifying a Search Filter
- The Query Tree

About the Query Dialog Box

When you click **Query** in the **Process** dialog box, the current address record is loaded into the **Query** dialog box and displayed. Query processes the input record and expands as many levels as possible in the **Records** section of the dialog box.

The **Query** dialog box consists of three different sections. The top section lists the **Input Address**. This is a fixed, non-editable area that displays the record as entered in the **Input Address** section of the **Process** dialog box. It is displayed as a reference to the original address.

The middle section of the **Query** dialog box is the **Search Filter**. This section consists of fields for House, Street, Unit, City, State, and ZIP. Only street records matching all of these fields display in the **Records** section.

The bottom section of the **Query** dialog box is the **Records** section. This section displays the information from the address standardization and geocoding database. It is hierarchical in nature: it displays only the requested level of detail. The level of detail can be increased or decreased.

Specifying a Search Filter

A query can display all streets that start with “K” in Washington, but queries can also be much more refined. For example, you could have a query that specified “show all streets in ZIP Code 80302 that begin with APPLE and have a house number range that includes 4300.”

The **Search Filter** section of the **Query** dialog box is used to specify the limits of the query. In general, the more information entered in the search filter, the more refined the search results. This means that fewer possible matches display in the **Records** section.

The search filter requires the user to enter a city and state, or a ZIP Code, or both. If the **City** and **State** entry are for one state but the **ZIP** is for a different state, the two states are displayed in the **Records** section. You must choose which state to search. Centrus GeoCoder for ArcGIS does not list more than two states in the **Records** section.

The **City** entry needs only the first letter of the cities that are to be displayed. Centrus GeoCoder for ArcGIS displays all the cities within the given state which match the letters in the **City** entry. The state abbreviation, however, must be a recognized state abbreviation or the entries for both **City** and **State** are ignored.

The **ZIP** entry is used solely to generate the list of possible cities. It does not determine which entries are displayed in the **Records** section. The **ZIP** entry must be a valid five-digit or three-digit ZIP Code. Three-digit ZIP Codes are commonly referred to as *Sectional Centers*. A Sectional Center is comprised of all five-digit ZIP Codes that begin with those three digits. The **Street** entry may have zero or more characters entered. If no street data is entered all streets are displayed. If the **Street** entry was “APPLEW”, streets with the name of “Applewood” or “Appleworm” would both match, but a street with the name “Apple” would not match and would not be displayed.

Note: The **Street** entry should be a base street name only; it should not contain pre-directionals, post-directionals, or street types. For example, “Vine” would be a valid entry for Boulder, CO, but “Vine PL” would be invalid because of the additional street type (“PL”).

The **House** entry specifies a house number. The **House** entry is used to constrain the search to show only those streets that have a block on which that house number would fall. If the entry for House was “1000”, a house range of “200 to 300” would not match and would not be displayed. This entry may be blank, in which case all ranges would match and the search would not be constrained by house number.

The **Unit** entry is very similar to the **House** entry, except that it contains the unit number, such as “12” or “E”. House ranges that do not contain the unit number entered do not display.

Note: The U. S. Postal Service does not list separate unit numbers for all buildings. Due to this limitation, house range entries that do not have any unit numbers match any unit number entered. This is consistent with USPS CASS requirements.

The Query Tree

Queries are displayed in a hierarchical nature in the **Records** section of the **Query** dialog box. Each level of the hierarchical tree is indented below the rest. All entries are sorted alphabetically, including Street Block and House Range. The levels of the tree are:

Country

State

City

Street Name

Street Block

House Range

Country

Always displays USA. Centrus GeoCoder for ArcGIS's area is the 50 states and Washington, D.C., plus Puerto Rico and other U.S. Protectorates.

State

Displays the two-letter state abbreviations that match the search filter criteria. If the ZIP Code is for a different state, two state abbreviations display.

City

Displays the city names that match the search filter. If the City entry in the filter is "C," then all city names beginning with "C" are displayed. Some city names are displayed with a different name in parentheses. The name in parentheses is the city name that should be used when mailing to that location. The name outside of the parentheses is a name that the USPS recognizes and may not be a city name at all. These names might be buildings, military installations, or even large corporations. For example, a search for "READERS DIGEST" in the city field and "NY" in the state field results in "READERS DIGEST (PLEASANTVILLE)". When a house range record is chosen, the proper USPS City name is displayed in the bottom left corner of the dialog box.

Street Name

Contains the full street name including directionals and street type. For example, there are separate entries for "N Main St" and "S Main St". The entries are sorted by street name, then street type, then predirectional, then postdirectional. A street may also appear with *(Alias)* next to the street name, which indicates a different name was used to match street name with the search filter.

Street Block

Contains one “block” of house ranges. A Street Block often is the same as an actual city block, but not always. In some instances, a Street Block may represent a partial city block. Less frequently, a Street Block represents several city blocks. Street blocks list a range of all house numbers that might appear on that block. In most instances, all house numbers within the range shown are valid. On some blocks, however, there may be gaps within the range where a house number is invalid. For example, the Street Block might indicate “100-199 Main St” but inspection of the House Range records shows records for “100-120 Main St” and for “150-199 Main St”. In this instance, a house number of 140 would be invalid.

House Range

Displays the actual range information contained in the USPS files. An example would be “1000-1098 Kearney St NE, 20017-4526 (S,E)”. The first part of this line is the house range and full street name. The next entry is the nine-digit ZIP Code for that house range. In parentheses are displayed the Record Type and Record Parity, described as follows:

Record Type

Indicators are as follows:

- **S** for Street
- **R** for Rural Route/highway contract
- **F** for Firm
- **H** for High-rise, building, or apartment
- **P** for P.O. Boxes and General Delivery
- **G** for General Delivery records

A Street record consists of a house number range and can include any unit ranges given. A Firm record has a firm name associated with it. This record should only be chosen if the firm name is a correct part of the address being searched for. A High-rise record usually denotes an apartment or office building and has a range of unit numbers. A High-rise record should only be chosen if the unit number of the input address is within the unit range of the High-rise record. P.O. Box records simply list all available box numbers or the General Delivery record. When a Firm record is chosen, the Firm name is displayed in the bottom right corner of the dialog box. Should there not be a match on firm name or unit number, both Firm and High-rise records should always have a “default” record that is either for the building as a whole, or for the entire block. Assigning a non-matching firm or high-rise record may result in incorrect ZIP+4 or carrier route information being assigned.

Range Parity

Indicates whether the House Range contains (O)dd, (E)ven or (B)oth types of house numbers. If the parity of the record does not match the input record, it should not be chosen, as incorrect ZIP+4 and carrier route information may be assigned.

Navigating the Query Tree

When a query result is first displayed, the **Records** section displays Country and State. Double-click on State to expand the query tree. Double- clicking on any entry expands or contracts the tree, unless the selected entry is at the House Range level of the tree. If the selected entry is at the House Range level, double-clicking the entry accepts it as the correct record and the information is pasted into the **Results** section of the **Process** dialog box. If the selected entry is the last entry in the tree, and it is not a House Range entry, the tree is expanded, showing the next lowest level. If the selected entry is in the middle of the tree, the tree is collapsed to that point.

Selecting a Match Using Query

You select a match within the **Query** dialog box by selecting a House Range entry from the query tree. If there are entries in the House and Unit section of the Search Filter, then the selected entry's information is simply pasted to the **Process** dialog box in the **Results** section.

If either the House entry was blank, or if the Unit entry was blank but the House Range entry chosen contained unit numbers, Centrus GeoCoder for ArcGIS displays a dialog box allowing you to select House and Unit information before returning to the **Process** dialog box.

Glossary

Glossary

A

Address standardization

Address standardization is the process of taking an address and verifying that each component meets USPS guidelines for addresses. For example, “123 Main Avenue” should be abbreviated as “123 Main Ave.” During standardization, minor misspellings, dropped components, and abbreviations are all corrected. The correct city, state, and ZIP are also provided.

B

Buffer

An area drawn around an object. Buffers can be used to test whether points are “near” other objects, such as points, lines, or polygons.

C

CASS

USPS Coding Accuracy Support System. A service offered to mailers, service bureaus, and software vendors that improves the accuracy of delivery point codes, ZIP+4 Codes, 5-digit ZIP Codes, and carrier route information on mail. CASS provides a common platform to measure the quality of address matching software and useful diagnostics to correct software problems. CASS Certified mailings qualify for substantial postage discounts.

Note: This product may not currently be used to produce CASS-certified mailing lists for postage discounts.

Census ID

Also called a Census FIPS Code. A Census ID uniquely identifies each piece of Census geography.

Centroid

The calculated center of an area. The coordinates that define a centroid are the average of the sets of coordinates that describe the area.

D

Datum

A mathematical model of the Earth used to calculate the coordinates on any map, chart, or survey system. Surveyors take an ellipsoid model of the Earth and fix it to a base point. The North American Datum (NAD) is the official reference ellipsoid used for the primary geodetic network in North America.

E

eLOT

The Enhanced Line of Travel (eLOT) Product was developed to provide mailers the ability to sort their mailings in approximate carrier-casing sequence. To aid in mail sorting, eLOT contains an eLOT sequence number field and an ascending/descending code. The eLOT sequence number indicates the first occurrence of delivery made to the add-on range within the carrier route, and the ascending/descending code indicates the approximate delivery order within the sequence number. Mailers can use eLOT processing to qualify for enhanced carrier route presort discounts.

F

Finance Area

A USPS Finance Area is an area defined by the Postal Service to collect cost and statistical data. It is frequently used for area searches, since it covers some or all of the ZIP Code areas in a city or town.

FIPS code

A Federal Information Processing Standards (FIPS) Code, also called a Census ID, uniquely identifies each piece of Census geography. Centrus GeoCoder for ArcGIS assigns County, Census Tract, and Block Group, all of which are Census geographies. The syntax of the FIPS code is as follows:

ssccctttt.ttgbbb

where

ss is the two-digit State FIPS Code

ccc is the three-digit County FIPS Code

tttt.tt is the six-digit Census Tract FIPS Code

g is the single-digit Block Group FIPS Code

bbb is the Block FIPS Code

G

Geocode

Geocoding is the process of assigning data based upon location information. Centrus GeoCoder for ArcGIS uses the address or ZIP to assign latitude, longitude, and Census FIPS information. Geocoding, coupled with demographic assignment or mapping software, can yield new insights regarding the address information.

GIS

Geographic Information System. A computer-based tool for enhancing geographic data by analyzing both the physical location in space and the set of characteristics associated with a location.

L

Layers

GIS applications separate the different types of information into data “layers.” For example, store locations might be defined in on one layer and main roads on another. This allows for separate display and processing when necessary but does not prevent cross referencing between data layers during query and analysis. Layers are referenced to a common spatial domain so that they can be scaled and overlain in such a way that any given reference point can be located on any of the layers and the data value extracted.

Location code

Location codes indicate the accuracy of the assigned geocode.

P

Point-in-Polygon analysis Point-in-Polygon analysis determines which polygon or polygons a point lies within.

Z

ZIP Code Zone Improvement Plan Code. Established in 1963, the system of 5-digit codes that identifies the individual post office or metropolitan area delivery station associated with an address. ZIP+4 is an enhanced code consisting of the 5-digit ZIP Code and four additional digits that identify a specific range of delivery addresses.

Index

S

T

U

Z