Step-by-step procedures
to migrate
National Guard Bureau data
to a
SDSFIE Personal Geodatabase

There are eight steps in the migration process:

1. Browse the SDSFIE Data Model for NGB data
2. Create a correlation matrix between NGB data and SDSFIE
3. Create a SDSFIE Filter of the NGB features in correlation matrix
4. Determine the Spatial Reference of the NGB data
5. Create a SDSFIE Geodatabase, using the Filter and the Spatial Reference
6. Migrate NGB data into the SDSFIE Geodatabase
7. Quality check the SDSFIE Geodatabase to ensure loss-less data migration
8. Add metadata information to the SDSFIE Geodatabase records

Each step will be described in detail, along with any necessary procedures.

1. Browse the SDSFIE Data Model for NGB data

The Spatial Data Standards for Facilities, Infrastructure, and Environment (SDSFIE) are large scale/high resolution geospatial data content and classification standards for GIS. They are an integrated model of multi-thematic data content standards and an example of a Federal Geographic Data Committee data content standards implementation. The SDSFIE Data model consists of five parts:

1. Entity Sets - Broad grouping for data management
2. Entity Classes - Grouping of data within each Entity Set
3. Entity Types - Grouping of Items that appear graphically on a map or drawing
4. Attribute Tables - A relational database table containing non-graphic information, or attribute data
5. Domain Tables - Contains lists of “valid” or “permissible” values for specific attributes in an Attribute Table.

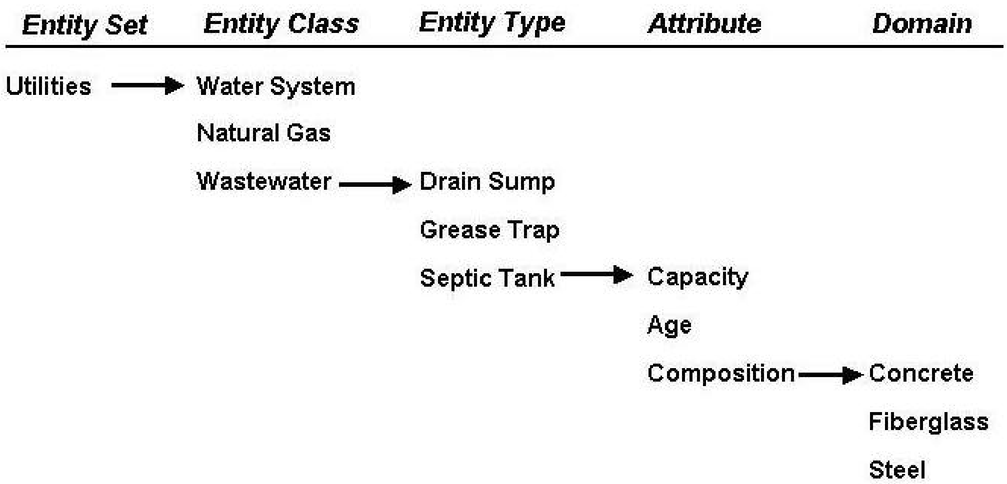
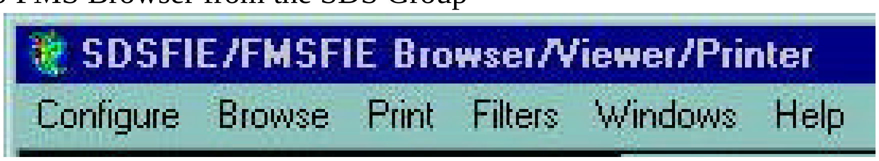


Figure 1 is a graphic representation of the SDSFIE.

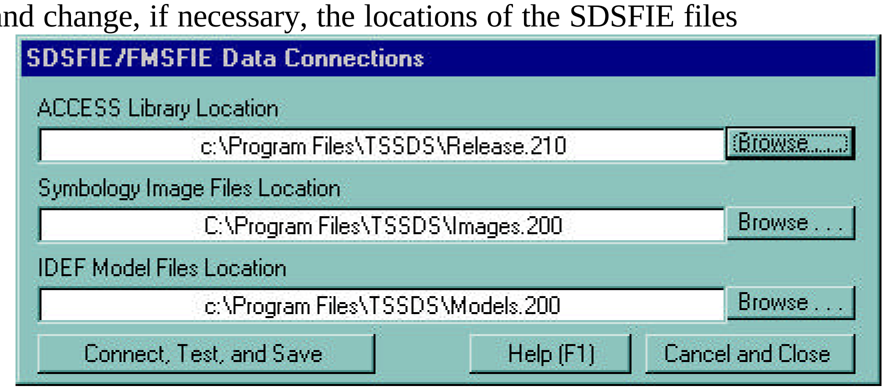
The SDSFIE Browser allows user to browse the contents of the SDSFIE Library or Release.

Configure the SDS FMS Browser

1. Select SDS FMS Browser from the SDS Group

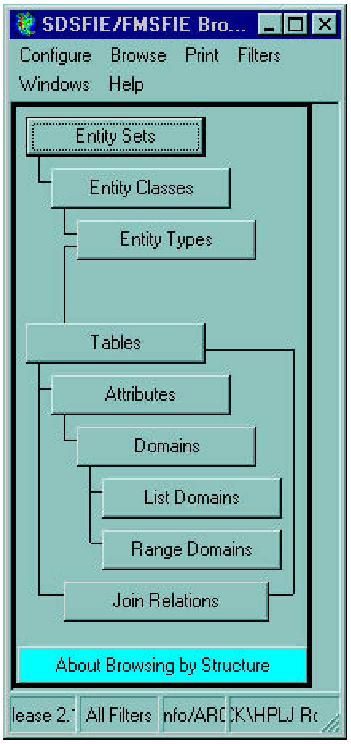


2. Configure the Browser Data Connections
   a. Select Configure → Connect
   b. Verify and change, if necessary, the locations of the SDSFIE files

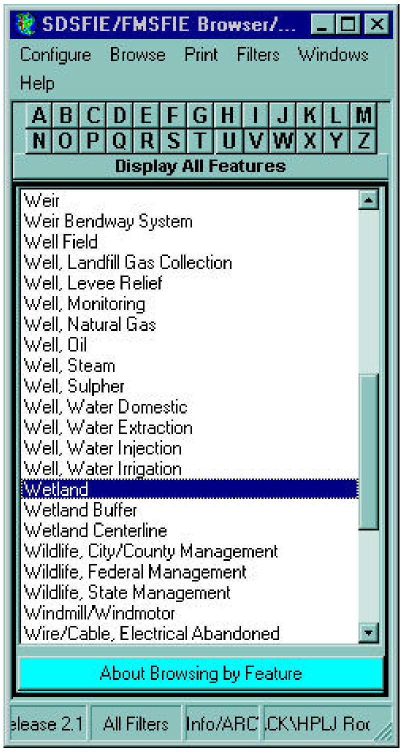


3. Configure the SDSFIE Options
   a. Select Configure → Options
   b. Select the desired GIS Software application and select OK or Apply.
   c. Select the appropriate special features and output file path and select OK or Apply.

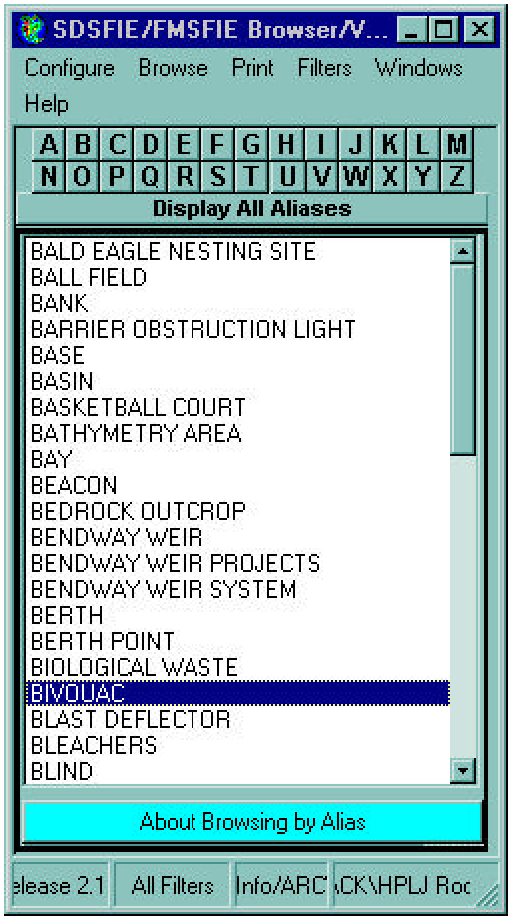
4. There are five different methods available to browse the SDS:
   a. Select Browse → By Structure (default)



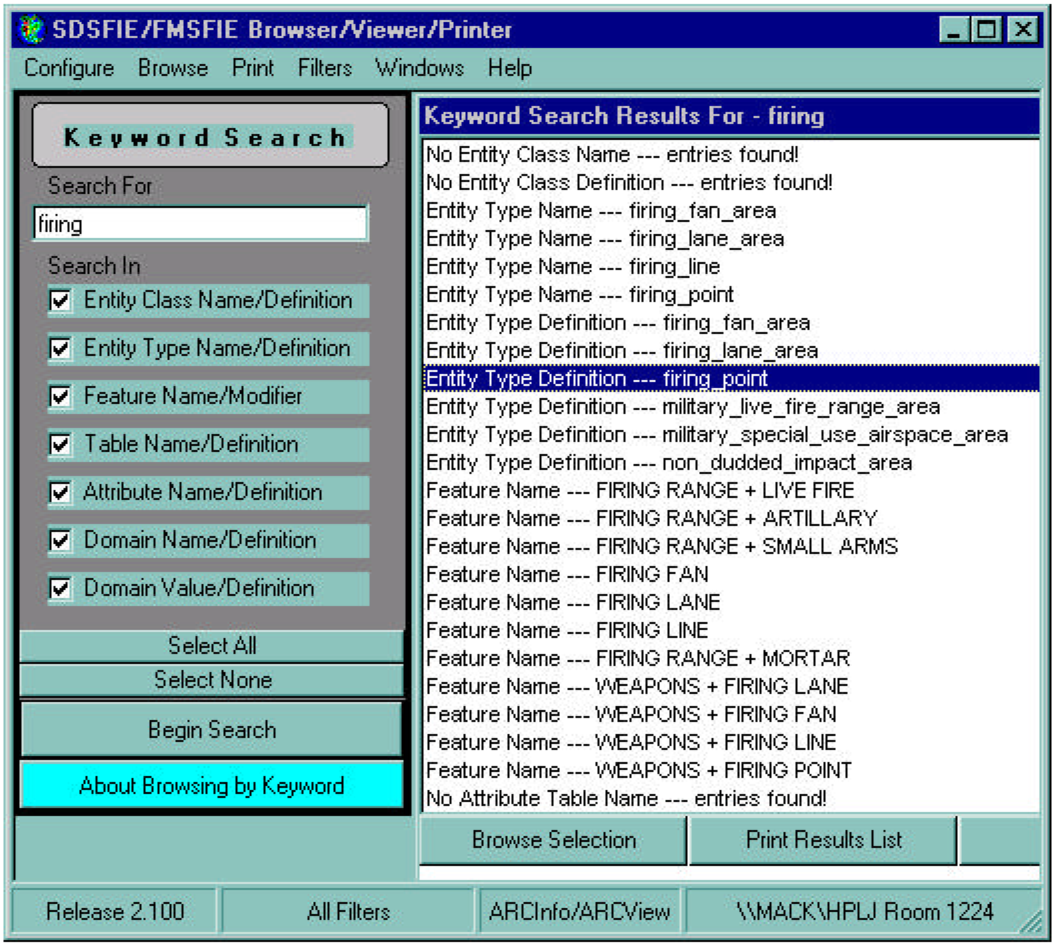
   b. Select Browse → By Feature



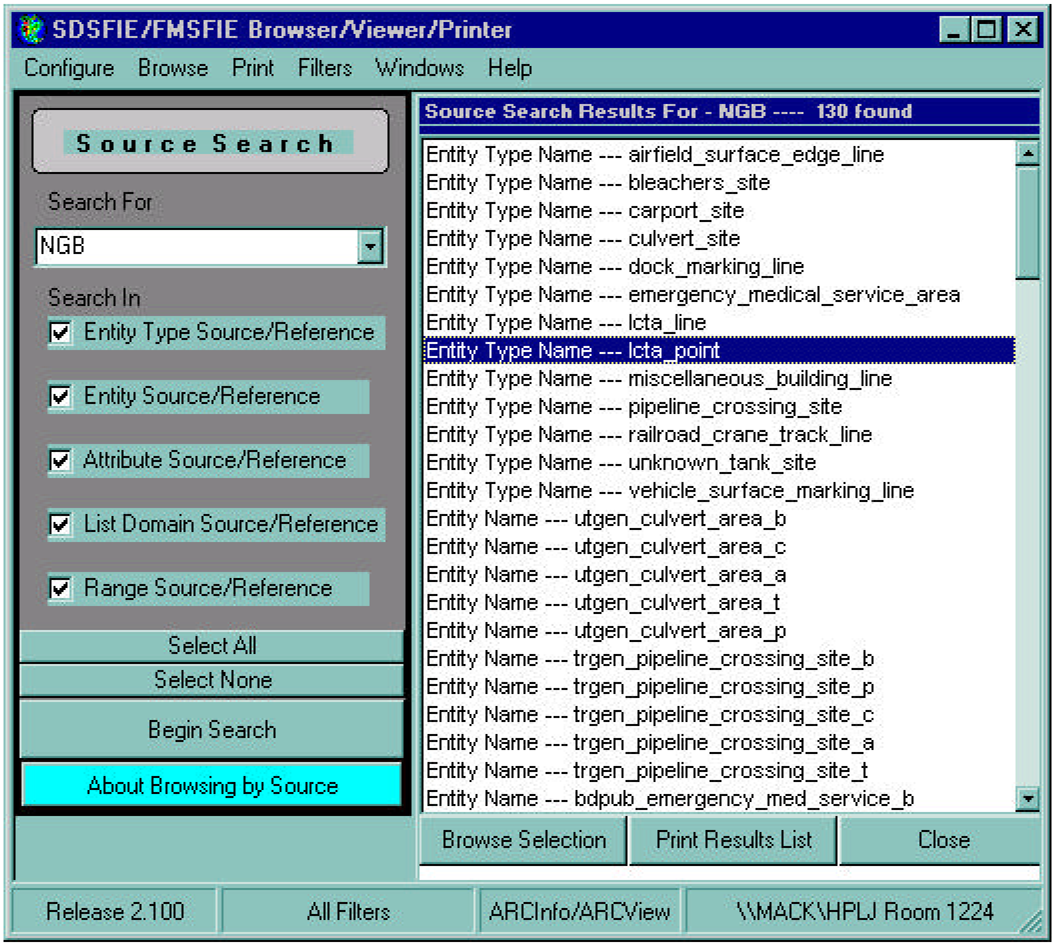
c. Select Browse → By Alias



d. Select Browse → By Keyword



e. Select Browse → By Data Source

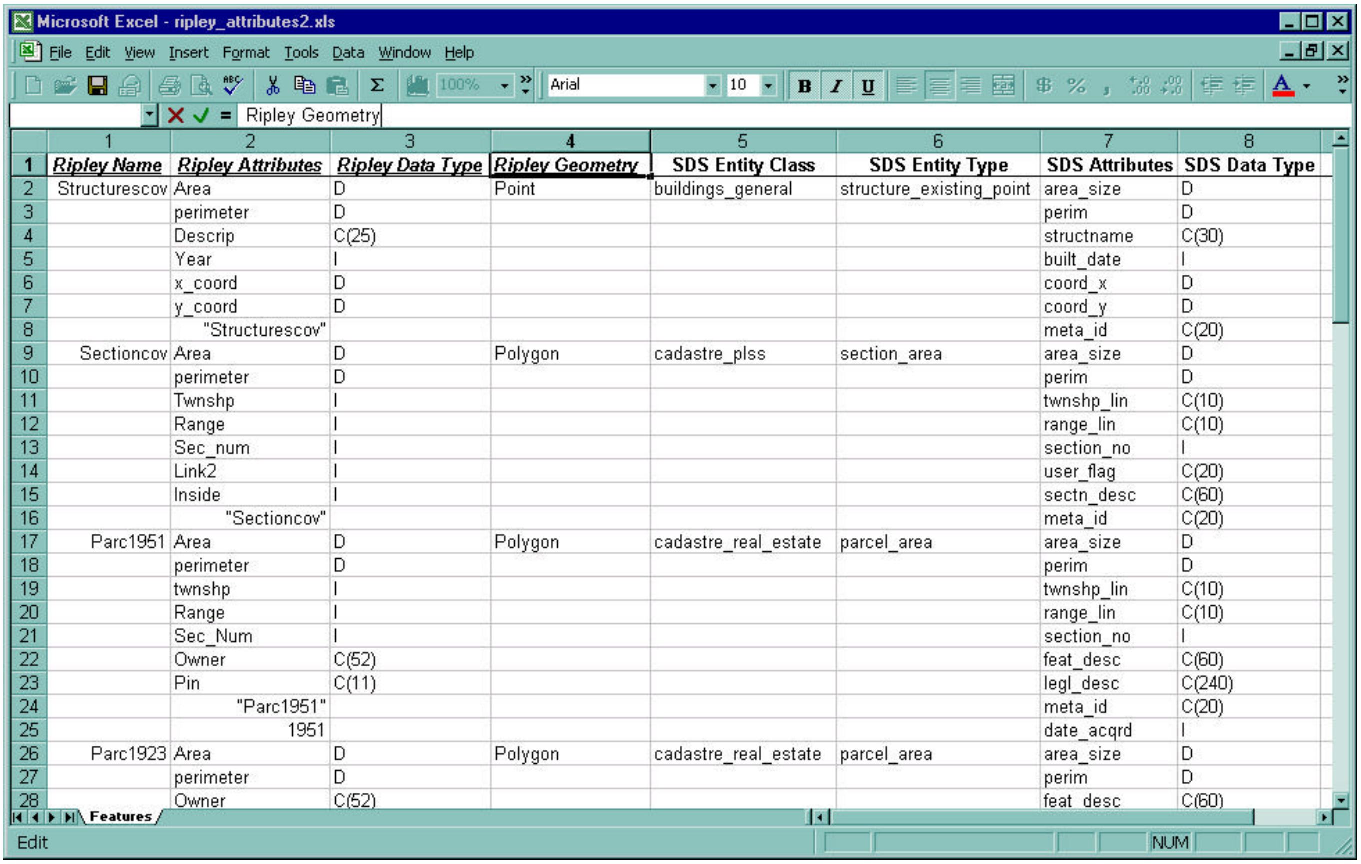


2. Create a correlation matrix between NGB data and SDSFIE

The correlation matrix will be used to correlate existing source data with the Spatial Data Standard (SDSFIE) features and attributes.

1. Enter all the *Source* coverages and shapefiles attributes into a spreadsheet including the *Source* Name (e.g. Ripley Name), *Source* Attributes (e.g. Ripley Attributes), *Source* Data Types, and *Source* Geometry Type (e.g. Ripley Geometry).
2. Create columns for the SDSFIE Entity Class, SDSFIE Entity Type, SDSFIE Table, SDSFIE Attributes, and SDSFIE Data Types. The SDSFIE Entity Class is equivalent to a Feature Dataset in a geodatabase and the SDSFIE Entity Type is equivalent to a Feature Class in a geodatabase.
3. Correlate the *Source* Name to the SDSFIE Entity Class and Entity Type. The correlation matrix lists the *Installation* Name, *Installation* Attributes, *Installation* Attributes' Data Types, and *Installation* Geometry that is currently being used. For example, one *Installation* Name for Camp Ripley is Structurescov, which is a name of a NGB ESRI ArcInfo coverage. The correlation matrix lists the SDSFIE Entity Class, which is the equivalent of a Feature Dataset in a geodatabase. Structurecov correlates to the SDSFIE Entity Class buildings_general and SDSFIE Entity Type structure_existing_point. The SDSFIE Entity Type is the equivalent of a Feature Class in a geodatabase.
4. Correlate the *Source* Attributes to the SDSFIE Table and Attributes. Every SDSFIE Entity Type/Feature Class has many attributes or fields. These attributes are correlated to the Installation attributes. An example is: Year correlates to built_date in the SDSFIE Feature Class structure_existing_point. The attribute media_id should be populated with the Installation's original source data (e.g. Structurecov) to identify which coverage or shapefile that data came from in the original dataset. This will be further explained in Step 8 Add metadata information to SDSFIE Geodatabase records.
5. Correlate the *Source* Data Types to the SDSFIE Data Types. If correlating to a domain attribute (_d), ensure the *Source* attribute's values are contained in the SDSFIE domain table)

The NGB_all_features.xls spreadsheet (Appendix A) is the correlation matrix of all NGB ESRI coverages and shapefiles to SDSFIE features and attributes. This correlation can be used as a guide for the correlation.

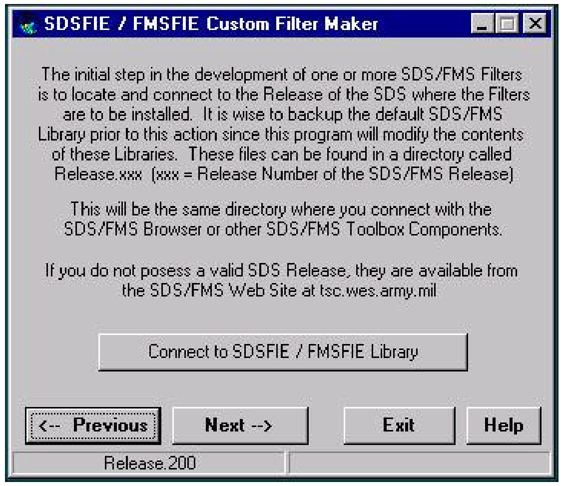


	1	2	3	4	5	6	7	8
1	Ripley Name	Ripley Attributes	Ripley Data Type	Ripley Geometry	SDS Entity Class	SDS Entity Type	SDS Attributes	SDS Data Type
2	Structurescov	Area	D	Point	buildings_general	structure_existing_point	area_size	D
3		perimeter	D				perim	D
4		Descrip	C(25)				structname	C(30)
5		Year	I				built_date	I
6		x_coord	D				coord_x	D
7		y_coord	D				coord_y	D
8		"Structurescov"					meta_id	C(20)
9	Sectioncov	Area	D	Polygon	cadastre_plss	section_area	area_size	D
10		perimeter	D				perim	D
11		Twnshp	I				twnshp_lin	C(10)
12		Range	I				range_lin	C(10)
13		Sec_num	I				section_no	I
14		Link2	I				user_flag	C(20)
15		Inside	I				sectn_desc	C(60)
16		"Sectioncov"					meta_id	C(20)
17	Parc1951	Area	D	Polygon	cadastre_real_estate	parcel_area	area_size	D
18		perimeter	D				perim	D
19		twnshp	I				twnshp_lin	C(10)
20		Range	I				range_lin	C(10)
21		Sec_Num	I				section_no	I
22		Owner	C(52)				feat_desc	C(60)
23		Pin	C(11)				legl_desc	C(240)
24		"Parc1951"					meta_id	C(20)
25		1951					date_acqrd	I
26	Parc1923	Area	D	Polygon	cadastre_real_estate	parcel_area	area_size	D
27		perimeter	D				perim	D
28		Owner	C(52)				feat_desc	C(60)

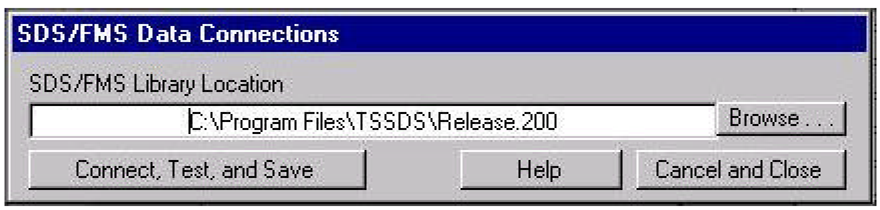
3. Create a SDSFIE Filter of the NGB features

The SDSFIE Filter Maker allows users to create filters, which serve the purpose of choosing the required components of a the SDSFIE for a specific project. Once a filter is created, it can be used in all SDSFIE Toolbox applications. Using the SDSFIE Filter Maker tool, create a Source Filter, which matches the correlation matrix (e.g. Core NGB Filter). This filter can be used to create a SDSFIE personal geodatabase.

1. Select Filter Maker from the Spatial Data Standards Group



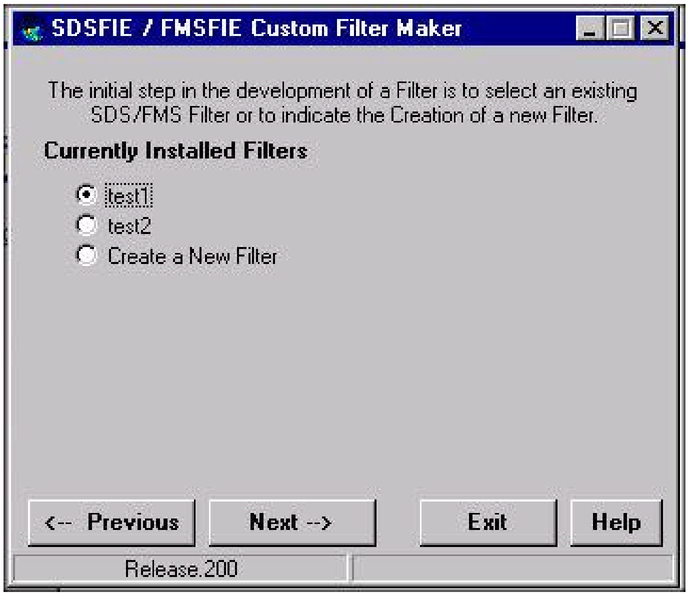
2. Click On "Connect to SDSFIE/FMSFIE Library"
   a. Verify and change, if necessary, the location of the SDSFIE/FMSFIE Library



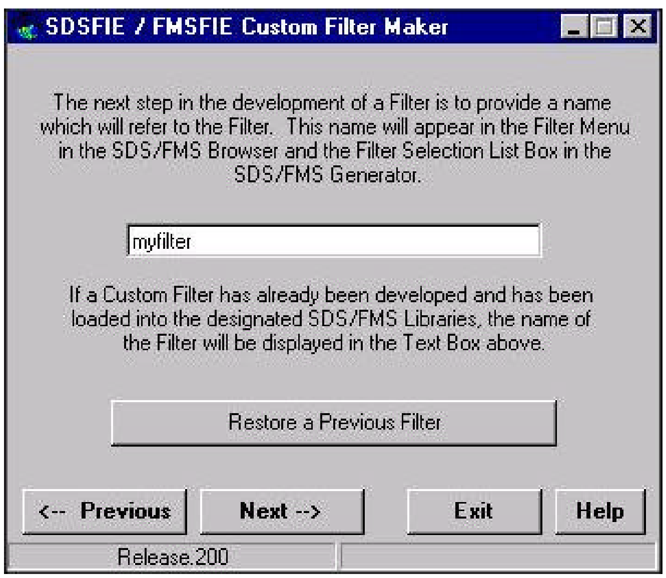
   b. Click "Connect, Test, and Save" to test and validate the SDSFIE Connection

3. Click On "NEXT →" to move to Restoring or Creating a Filter
   If you have previously created custom filters, the names of them will appear in this window. To restore an existing filter, click in the radio button and then click 'Next'.

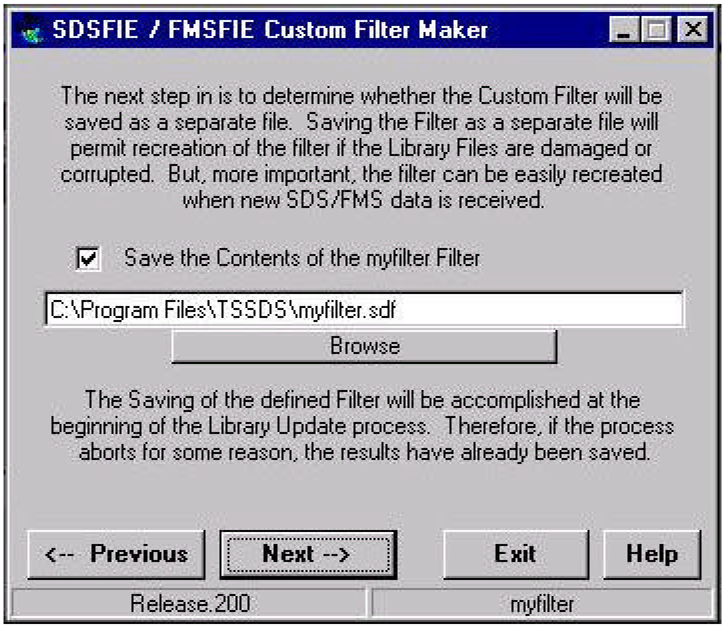
4. To create a new filter, click in the radio button and then click 'Next'.



5. Enter the Name of the Filter you wish to create (Maximum 20 Characters) (Spaces permitted but no special characters)



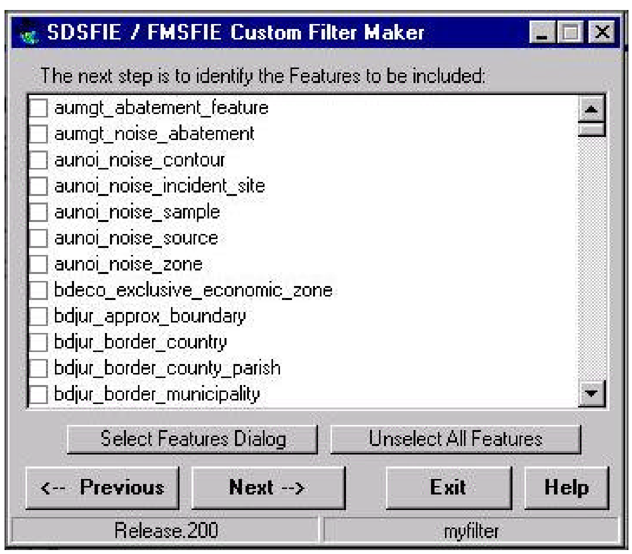
6. Click On "NEXT →" to move to screen four (Saving a Filter). Ensure that the lower right status bar panel displays the name of the SDSFIE/FMSFIE Filter



7. Click "Save the Contents" Check Box to save the filter.

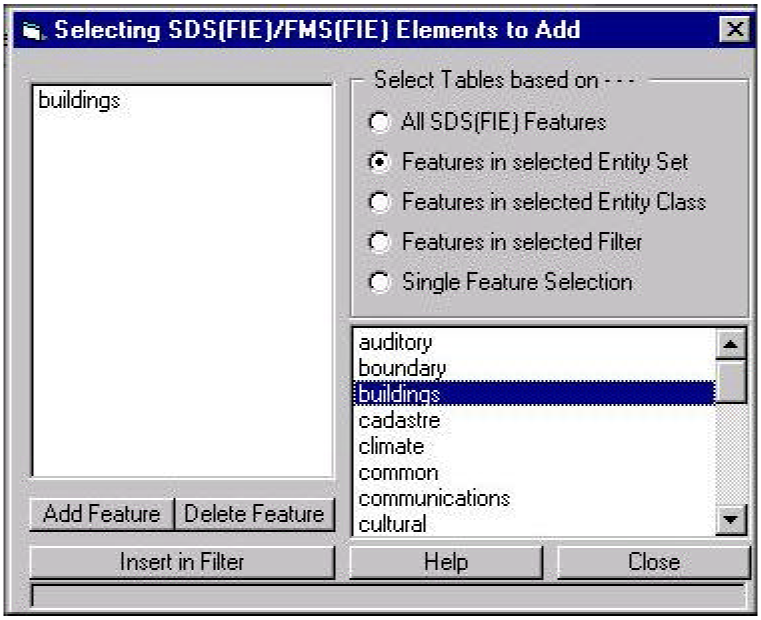
8. Click On "Save Filter File Name Assignment". Select and confirm a File Name and Path to save the filter file [*.sdf]

9. Click On "NEXT →" to move to Selecting Features. At this point, the software will automatically read the SDSFIE/FMSFIE Library and populate the List of Features included. The software will display a progress bar. When complete, the List of Features will be displayed.

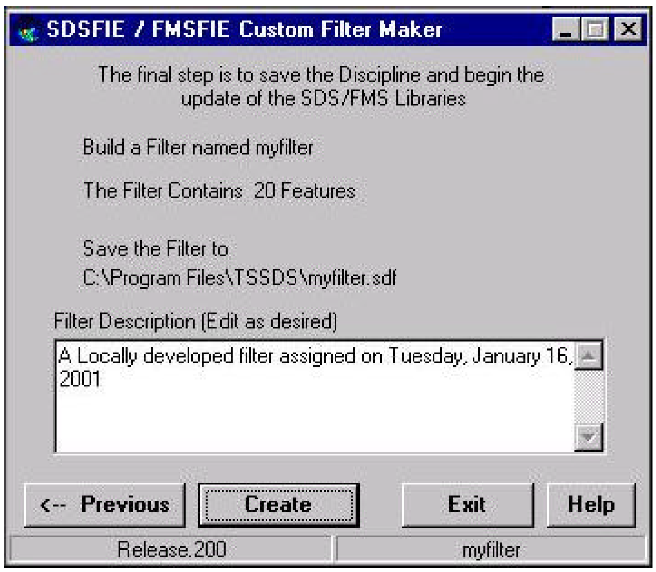


10. Check those Features to be included (Click on the Check Box or Double Click on the Feature Text).

11. Or, click on "Select Features Dialog" to select features in one of five ways: all features, an entire entity set, an entire entity class, an existing filter, or single features.



12. Then click 'Insert in Filter' to insert the selected features in the filter.



13. Confirm the information displayed with respect to the Filter desired including Filter Name, number of Features included, and whether to save the Filter File (*.sdf) and where. You may also edit the filter description (which is saved in the Access library).

14. Click On "Create" to begin the Filter Creation process. The first step will always be the saving of the Filter information, if saving has been configured.

4. Determine the Spatial Reference of the NGB data

The coordinate system needed for the geodatabase must be defined. All feature classes in a feature dataset use the same coordinate system, and all coordinates in all features in all feature classes must fall within the coordinate domains. Geodatabase feature coordinates must fall within the feature class x and y domain extent and are set when the feature dataset or stand-alone feature class is created. It cannot be modified after creation. The spatial domain is best described as the allowable coordinate range for x and y coordinates, m- (measure) values, and z-values. The precision describes the number of system units per one unit of measure. A spatial reference with a precision of 1 will store integer values, while a precision of 1,000 will store three decimal places. Once the spatial reference for a feature dataset has been set, only the coordinate system can be modified - the spatial domain is fixed.

The spatial or X/Y Domain of the dataset is dependent on the minimum x & y, and the precision values. The precision is the number of system units per unit of measure, and therefore specifies the degree of resolution. Likewise, the elevation or Z Domain is dependent on the minimum and maximum values and the precision. The precision for the z values is also the number of system units per unit of measure. Therefore, it also specifies the degree of resolution for elevations just as it does for the x & y. Coordinate precision refers to the mathematical exactness of a coordinate and is based on the possible number of significant digits that can be stored for each coordinate. Note, however, that mathematical precision does not, by itself, define accuracy, but it can be a major factor.

The Geodatabase Grid

The geodatabase stores coordinates as positive 4 byte integers that have a maximum value of 2,147,483,648. By default, the geodatabase requires the dataset to allow for the entire 2.14 billion by 2.14 billion size grid even if your data only covers a small portion of the grid. (Max - Min values) * precision = 2,147,483,648.

Precision

Precision is the number of storage units in one map unit. The dataset precision is determined by dividing map units by storage units. For example, if the map units are in meters and the desired storage units are centimeters, the precision would be 1000 since there are 1000 centimeters in a millimeter. Since ArcGIS does not allow the user to change spatial domain values in a dataset, precision should account for the future accuracy levels that data may be collected.

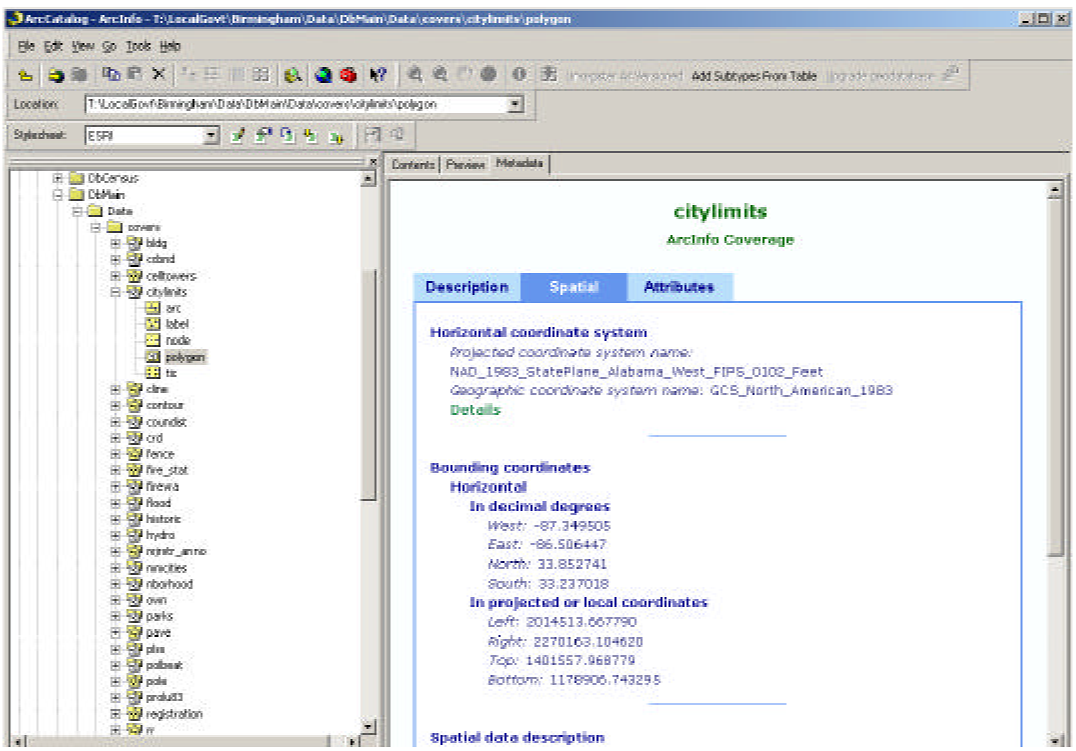
When determining dataset precision, it is important to know the current precision of existing data in order to prevent the geodatabase from modifying data when it is loaded into the geodatabase. A current double precision ArcInfo coverage allows for sub millimeter accuracy, so setting a geodatabase precision to allow only centimeter accurate data may cause some slight modification of the data as it is loaded into the geodatabase.

If the map units are meters, a precision of 3000 would allow for data accurate to approximately 1/3 of a millimeter. If the map units are feet, a precision of 1000 would allow for data accurate to 1/80th of an inch to be loaded into the geodatabase without modification.

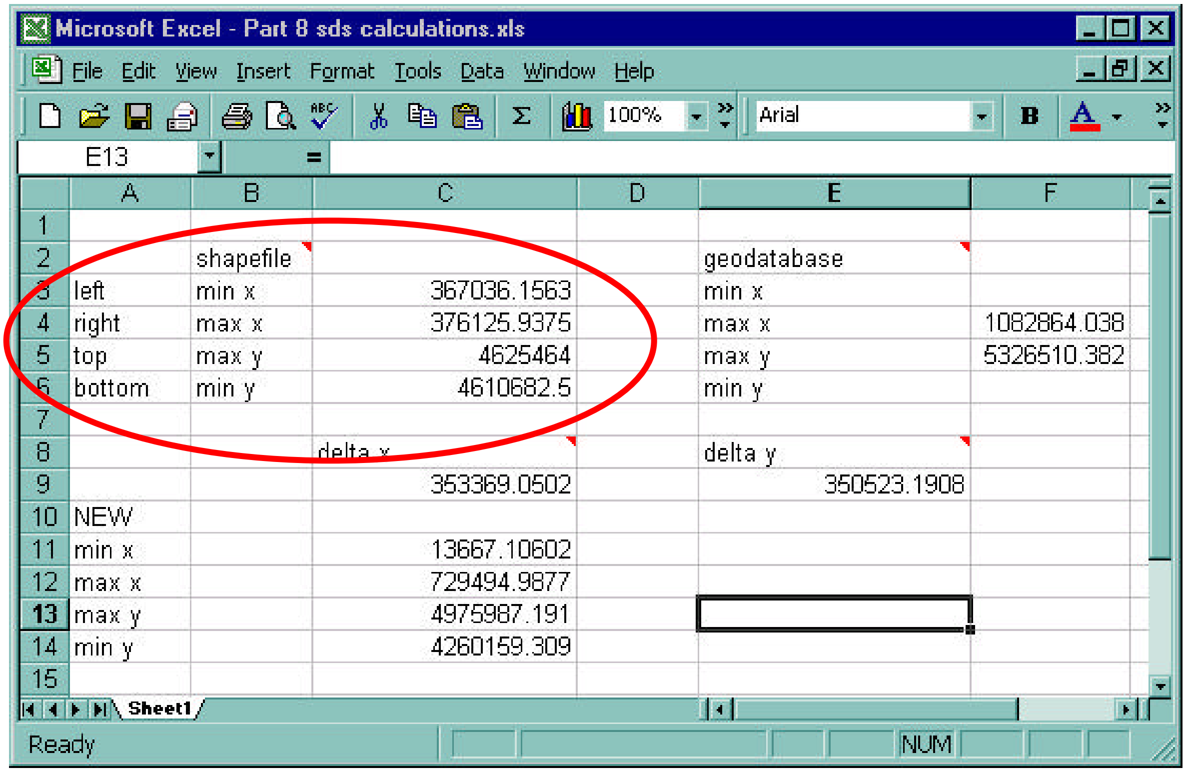
Spatial Domains

The Spatial Domains are specified at the Feature Dataset level in the geodatabase. Setting these values correctly is very important in preserving the accuracy of the data and allowing for future growth.

1. In ArcCatalog: Use the Metadata tab to locate the actual extents of the largest dataset to be loaded.
    a. Select the largest dataset in the Table of Contents Screen.
    b. Select the **Metadata tab** on the right-hand side of the ArcCatalog Window.
    c. In the Metadata display, select the **Spatial Tab**.
    d. Write down the **bounding coordinates** located under the **In projected or local coordinates** header. Left = X Min, Right = X Max, Top = Y Max and Bottom = Y Min.



2. Using Microsoft Excel, enter these values as the left, right, top, and bottom.
3. The remainder of the calculations will be performed in the following step.



	A	B	C	D	E	F
1						
2		shapefile			geodatabase	
3	left	min x	367036.1563		min x	
4	right	max x	376125.9375		max x	1082864.038
5	top	max y	4625464		max y	5326510.382
6	bottom	min y	4610682.5		min y	
7						
8			delta x		delta y	
9			353369.0502		350523.1908	
10	NEW					
11	min x		13667.10602			
12	max x		729494.9877			
13	max y		4975987.191			
14	min y		4260159.309			
15						

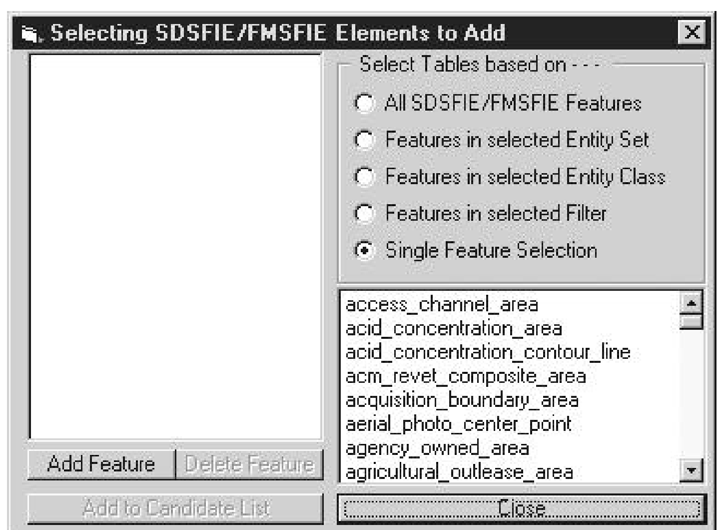
5. Create a SDSFIE Geodatabase, using the Filter and the spatial reference

1. Select SDS ARCBuild from the Spatial Data Standards Group
   Using menu item Open → Library Connection

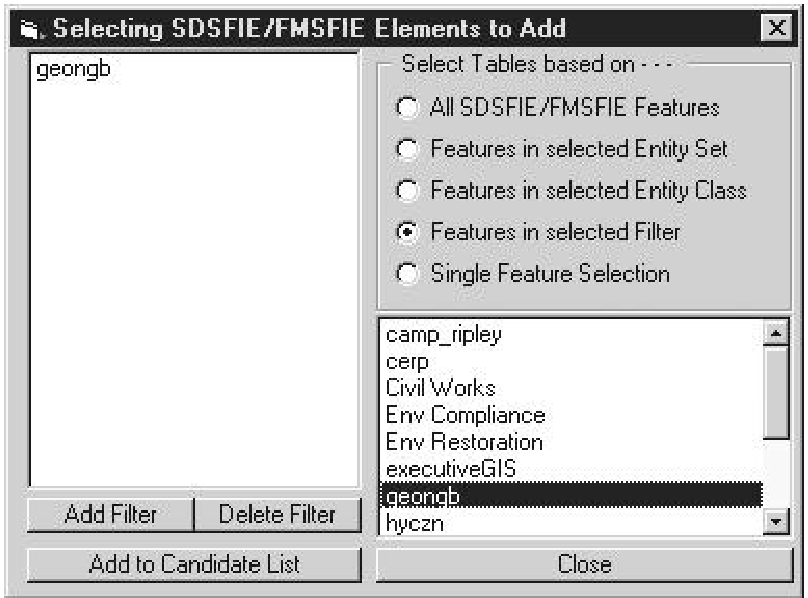
Click **Connect, Test, and Save** to Connect to Release 2.0 or later

2. Select Open → Create Geodatabase from the main menu
   a. Specify the path and filename of the Geodatabase (e.g. Ext Lights)
   b. Verify the connection to the new Geodatabase.

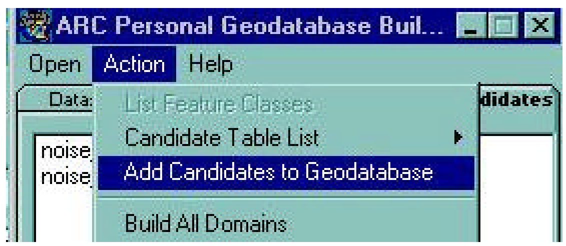
   c. The Builder will display the Element Selection Dialog as shown.



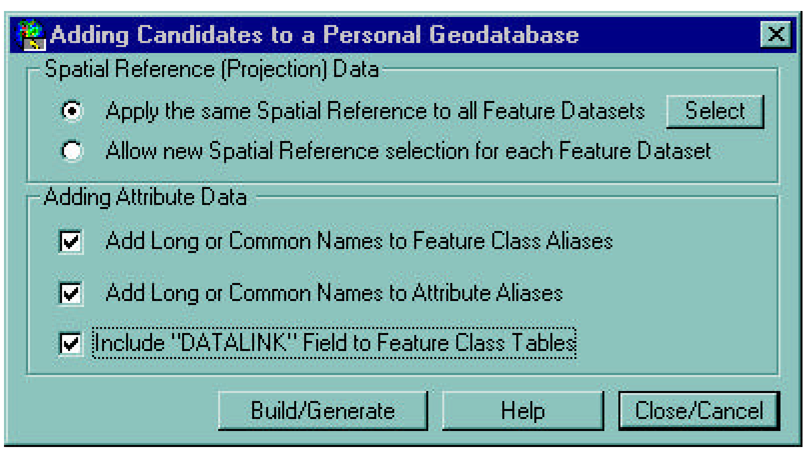
   d. Select **Features in selected Filter** from the available options
   e. Double click on the filter created in the previous step.
   f. Select **Add to Candidate List**.



g. Select **Action** → Add Candidates to Geodatabase to add the candidates to the geodatabase.

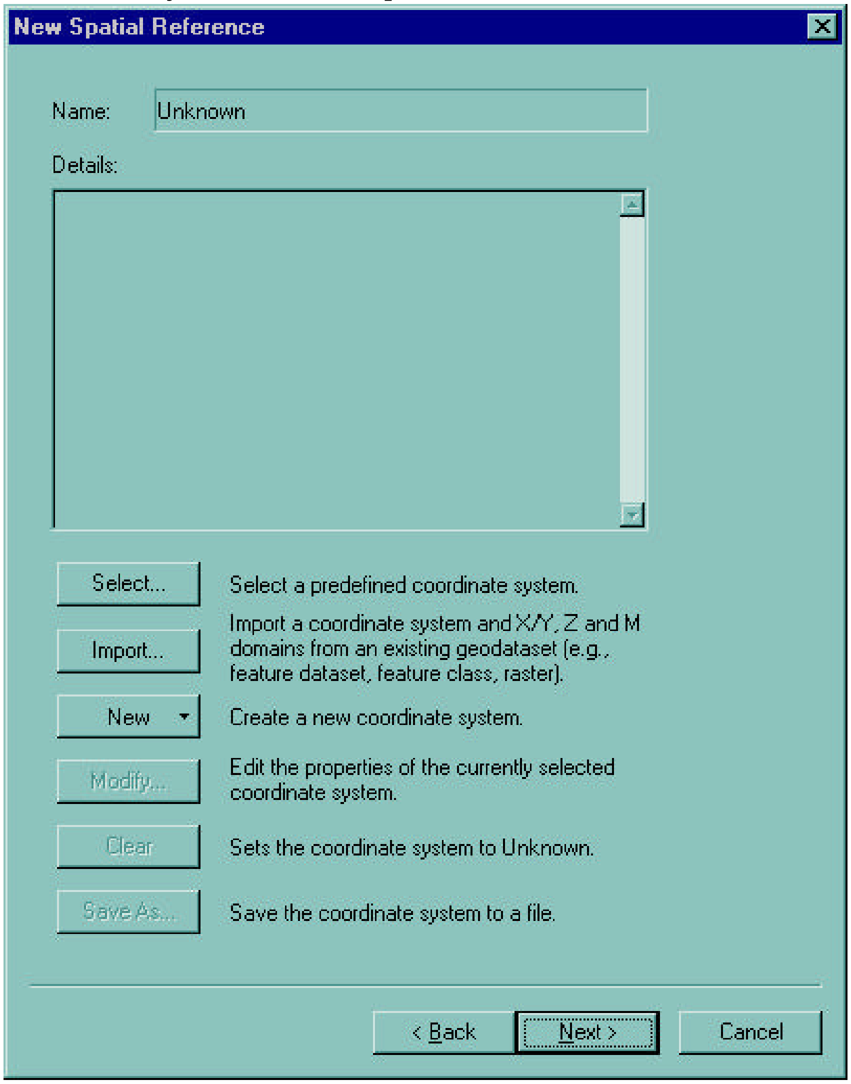


h. The ARC Build Tool will display the "Adding Candidates" Dialog.

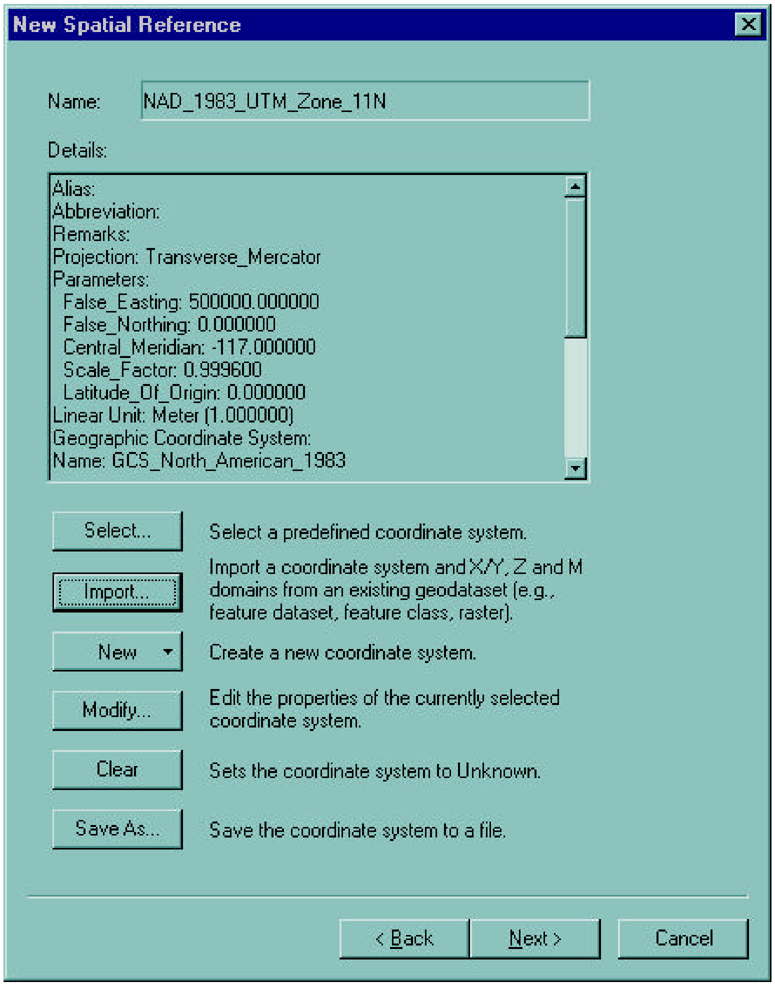


If the entire geodatabase will have the same spatial reference, select "Apply the same Spatial Reference to all Feature Datasets". If each feature dataset in the geodatabase will have different spatial references, select "Allow new Spatial Reference selection for each Feature Dataset". When a feature dataset is created by ArcBuild, the spatial reference will need to be defined.

i. From the New Spatial Reference window you are given the opportunity to select a predefined coordinate system, import a coordinate system from an existing dataset, or create a new coordinate system. If you have existing data already projected into the proper coordinate system use the Import tool.

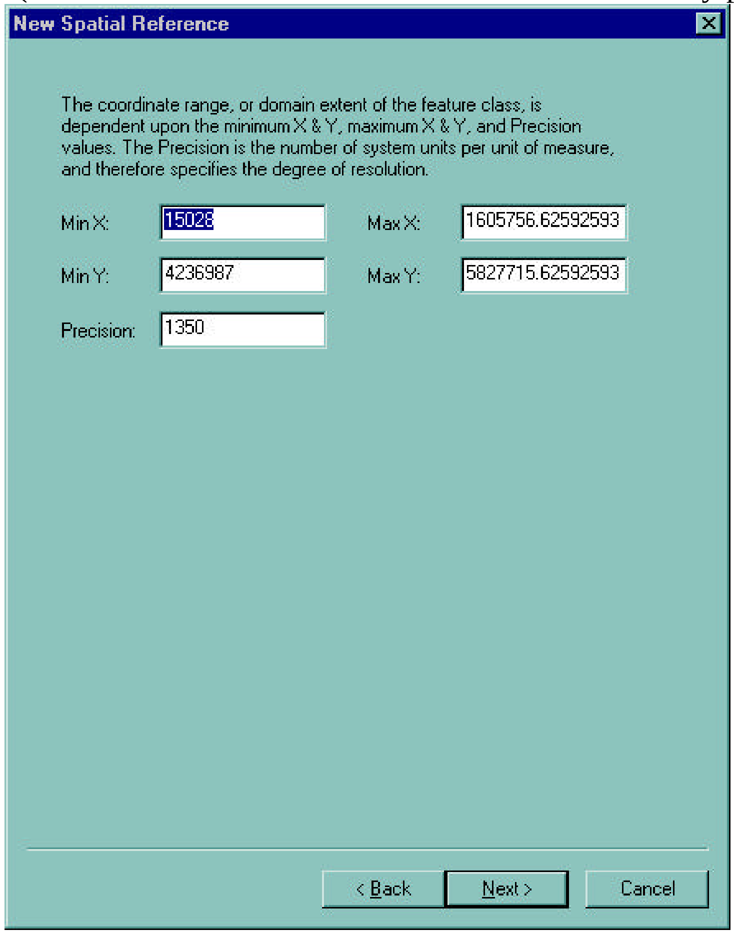


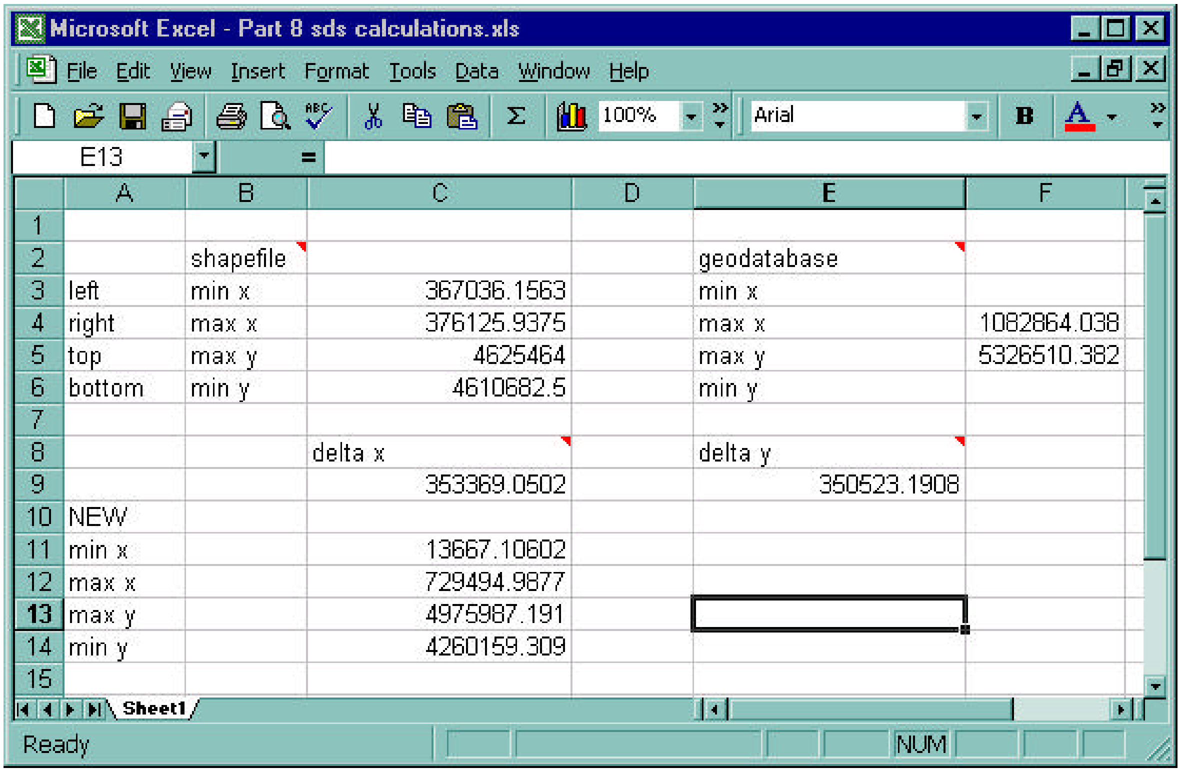
1. Select **Import**
2. Browse to the location of your existing data with the largest dataset extent
3. Select a dataset and select **Add** to import the Coordinate System and Extents into your Geodatabase.
4. Click Next



f. X/Y Domains and Precision

a. Enter the **X and Y Min values** and dataset **Precision** from the largest dataset. (Note: Max X and Max Y values will be automatically populated)





	A	B	C	D	E	F
1						
2		shapefile			geodatabase	
3	left	min x	367036.1563		min x	
4	right	max x	376125.9375		max x	1082864.038
5	top	max y	4625464		max y	5326510.382
6	bottom	min y	4610682.5		min y	
7						
8			delta x		delta y	
9			353369.0502		350523.1908	
10	NEW					
11	min x		13667.10602			
12	max x		729494.9877			
13	max y		4975987.191			
14	min y		4260159.309			
15						

g. Centering the Data

Attachment B is a spreadsheet with calculations preset. After each step, the cell of this spreadsheet will be referenced.

a. After entering the Min X, Min Y and Precision, the system automatically calculates the max values to complete the entire 2.14 billion by 2.14 billion grid. The data needs to be centered in the grid so that the data boundaries can expand equally in all directions in the future and allow for an equal size buffering to occur. Enter these values into the spreadsheet with Max X in Cell F4 and Max Y in Cell F5.

b. Calculate Delta X: **Subtract** the **Max X value** determined by viewing the metadata from the **largest dataset** from the **Max X value** that was **calculated by the system** after you entered the Min values and dataset precision. **Divide** the result **by 2** to get **Delta X**. (Cell C9)

c. Calculate Delta Y: **Subtract** the **Max Y value** determined by viewing the metadata from the **largest dataset** from the **Max Y value** that was **calculated by the system** after you entered the Min values and dataset precision. **Divide** the result **by 2** to get **Delta Y**. (Cell E9)

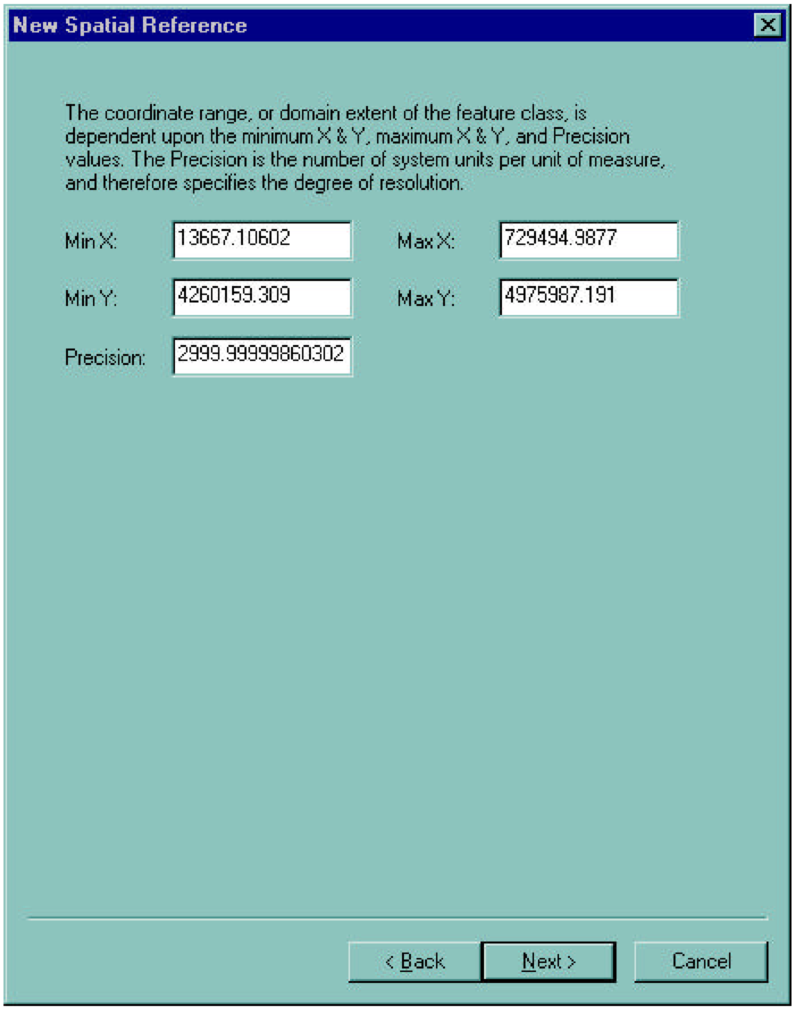
d. For new Min X: **Subtract Delta X** from the **Min X value** from your largest dataset. (Cell C11)

e. For new Min Y: **Subtract Delta Y** from the **Min Y value** from your largest dataset. (Cell C13)

f. For new Max X: **Add Delta X** to the **Max X value** of your largest dataset. (Cell C12)

g. For new Max Y: **Add Delta Y** to the **Max Y value** of your largest dataset. (Cell C14)

h. Enter these values into the dataset Min X/Y and Max X/Y Domains.



i. Check your precision value. It should be exact or very close to the original value you entered.

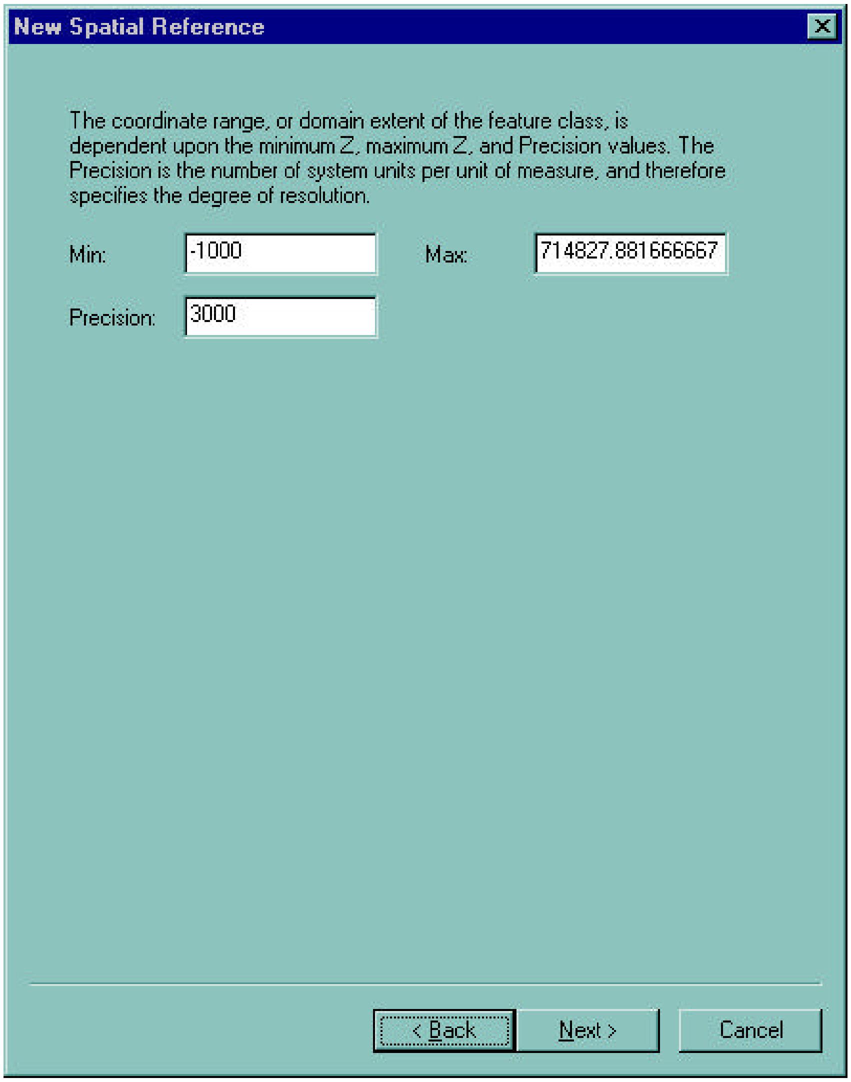
j. If necessary, change the precision to the desired value; x and y max values will change slightly.

Setting the Z Domain

k. The Z extent is also limited to 2.14 billion units. The precision value should be set to the same value for Z extents as it was for the X/Y.

l. Enter the **Min value** and **Precision** and allow the system to calculate the Max value.

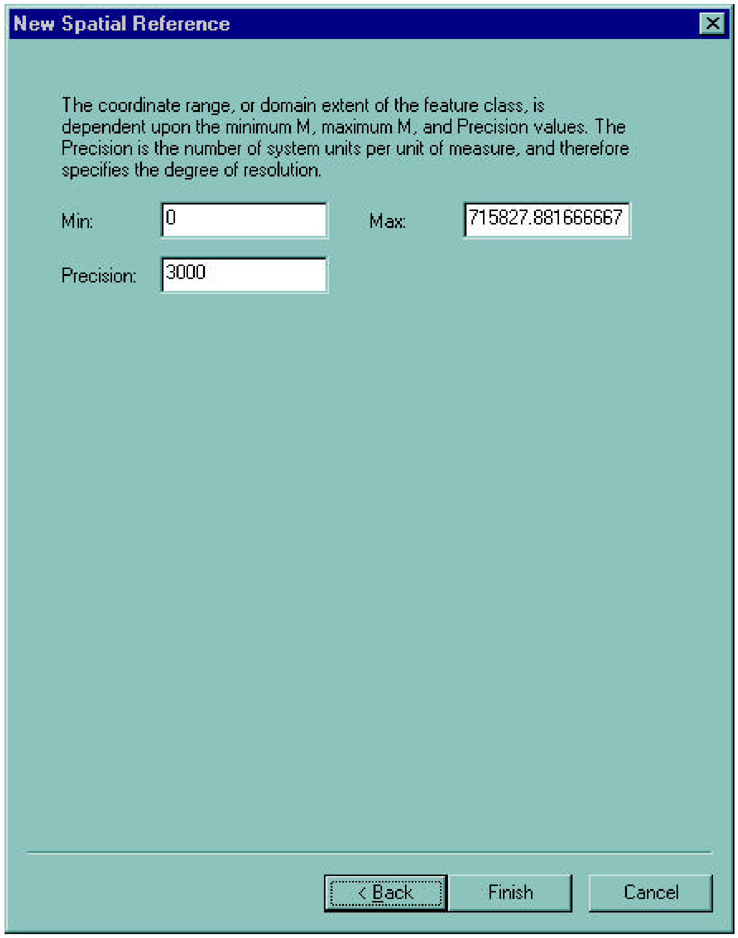
m. For the minimum value, it is important to account for underground features such as buildings, utilities and possibly bathymetry.


New Spatial Reference
The coordinate range, or domain extent of the feature class, is dependent upon the minimum Z, maximum Z, and Precision values. The Precision is the number of system units per unit of measure, and therefore specifies the degree of resolution.
Min:
-1000
Max:
714827.881666667
Precision:
3000
< Back
Next >
Cancel

Setting the M Domain

n. The M extent is also limited to 2.14 billion units. The precision value should be set to the same value for M extents as it was for the X/Y.

o. Enter the **Min value** and **Precision** and allow the system to calculate the Max value.



o. Select "Build/Generate" to generate the Geodatabase.

p. The ARC Build Tool will:

- Create the Spatial Reference (Coordinate System, X/Y, M, and Z Domains)
- Create the Feature Datasets
- Create the Feature Classes
- Mark the Candidate with "Done" when completed.
- Display the Progress of the Generation.
- Display "Compacting Database"

6. Migrate NGB data shapefiles and coverages into SDSFIE Geodatabase

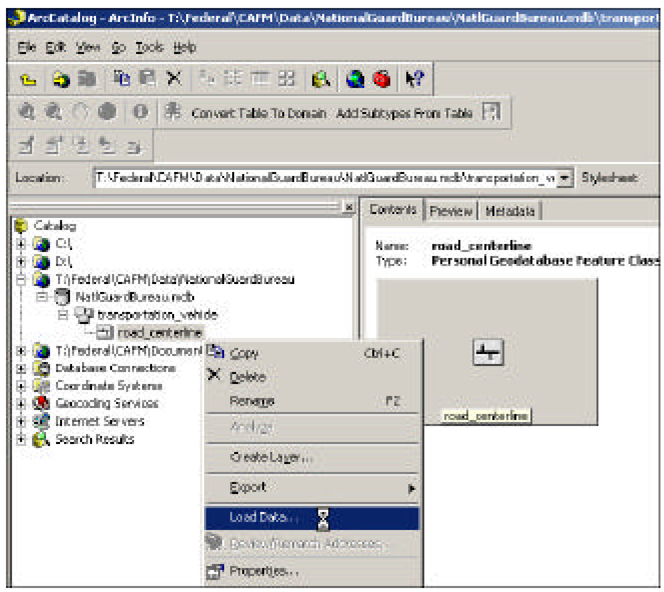
There are two ways to migrate data into a NGB SDSFIE Personal Geodatabase:

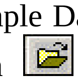
- Simple Data Loader (single migration)
- Automated Scripts (batch migration)

A. Simple Data Loader

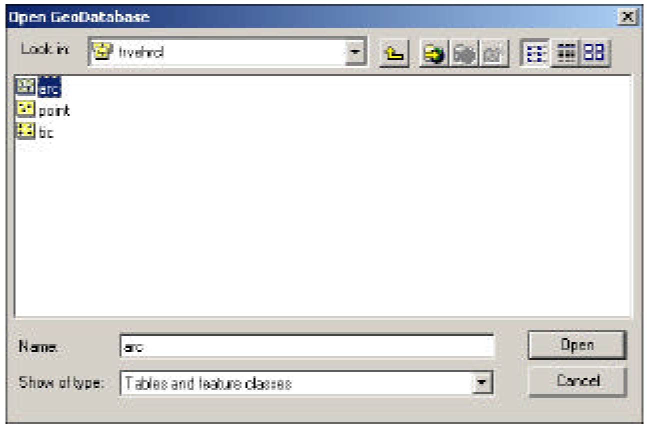
1. ArcInfo Coverages

Use the correlation matrix to determine the original coverage and its target feature dataset and feature class. In the ArcCatalog browser, right-click on the feature class and select Load Data to load ArcInfo coverage data into an existing feature dataset and feature class in a SDSFIE personal geodatabase.

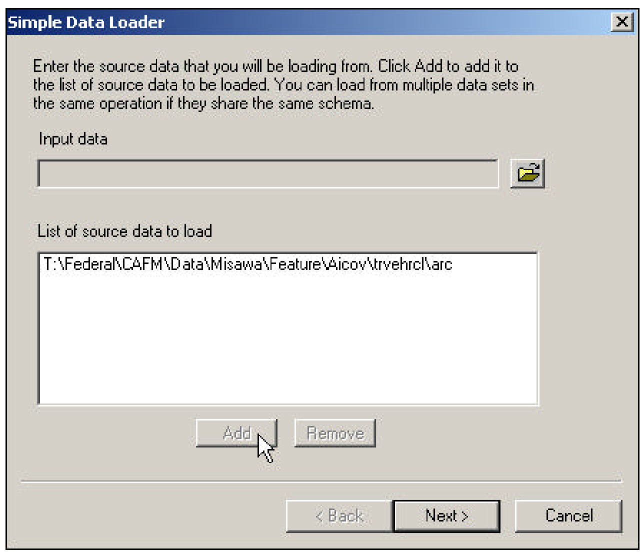


a. When the Simple Data Loader dialog box appears, click on the **Folder** button  to browse to the directory containing the Coverage file to be migrated. Double click on the coverage file (e.g. the **trvehrcl** coverage).

b. The coverage geometry type selected must match the feature class geometry type in the correlation matrix. Click on the **Open** button to select the data.

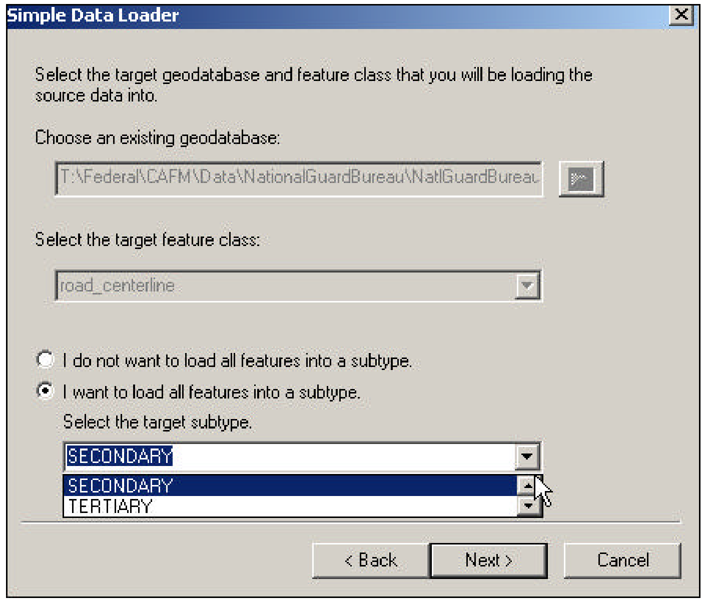


c. Click on the **Add** button to move the data source from the 'Input data' window to the 'List of source data to load' window.

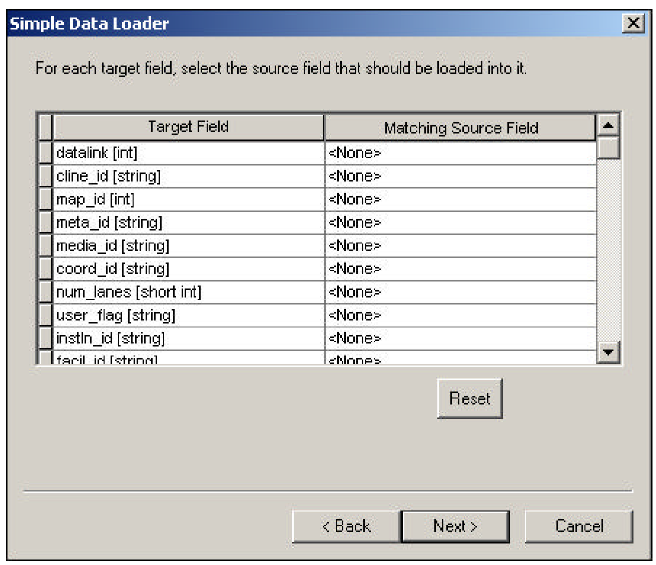


d. Click on the **Next** button to go to the next screen.

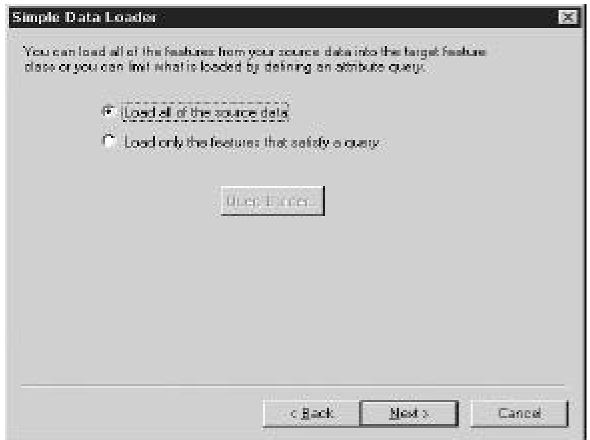
e. If the feature class has subtypes, select the subtype in which the data will be loaded. Otherwise, leave the 'I do not want to load all features into a subtype' radio button selected. Click the **Next** button to go to the next screen. In this case, road_centerline feature class data will be loaded into the SECONDARY Subtype.



f. Associate source data attribute fields with existing fields in the feature class, using the correlation matrix (Step 2) to determine which target field is associated with which matching source field. Click on the **Next** button to go to the next screen.

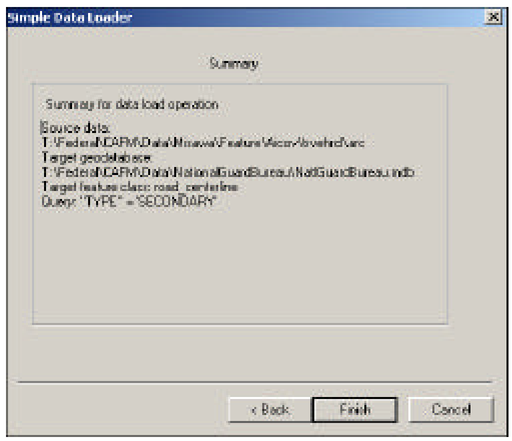


g. Since the entire source data is to be migrated, leave the **Load all of the source data** button selected.



h. Click on the **Next** button to go to the next screen.

i. A summary is displayed of all the parameters used to load the data. Review and select the **Finish** button to start the loading process.



2. ArcView Shapefiles

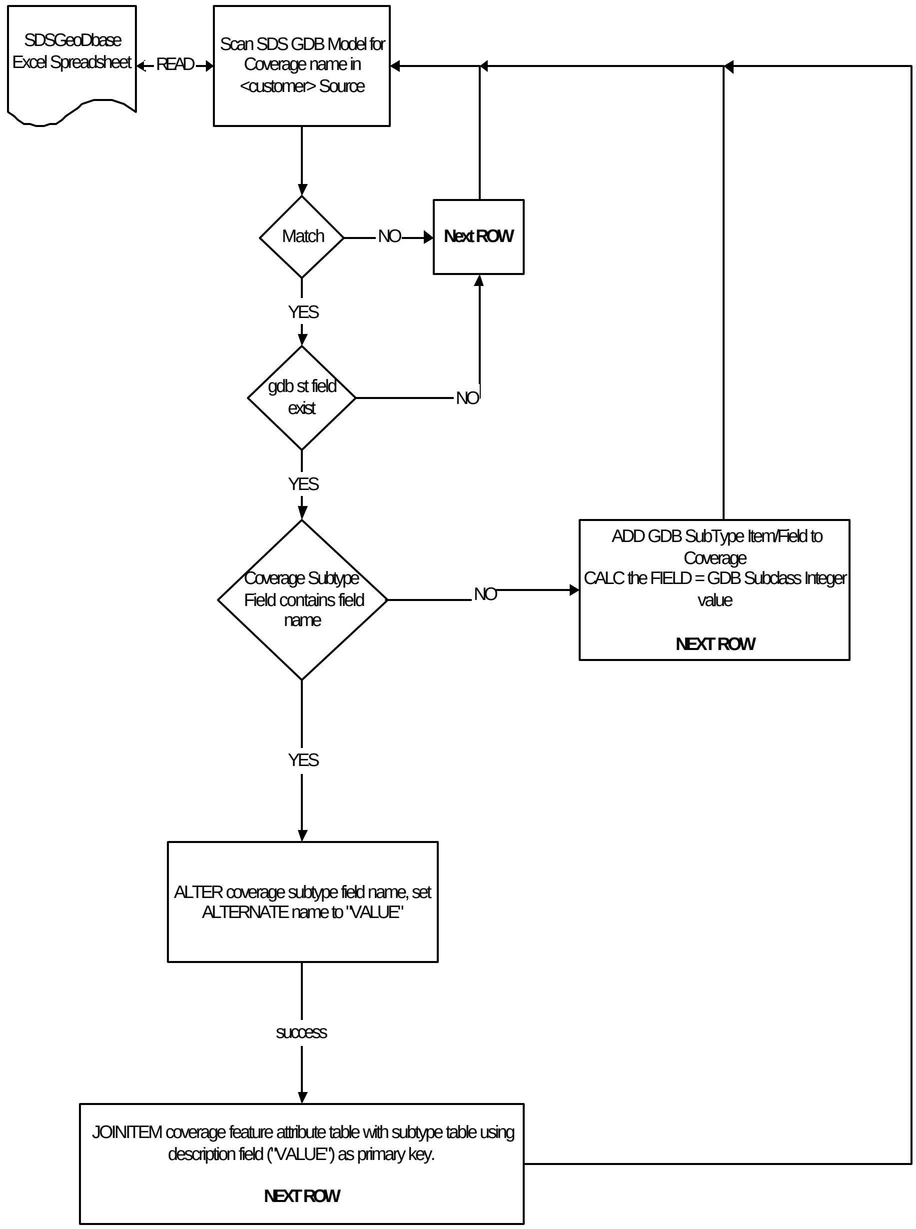
ArcView Shapefiles are loaded the same way as ArcInfo Coverages. The only exception is choosing file selection instead of feature type.

B. Automated Scripts for Batch Migration

1. Pre-Process

a. PREPROCESS.AML – This AML calculates the USER_FLAG field to be the value of the coverage or shapefile name.

b. ADDNONSDS.SQL – SQL script, which adds non-SDSFIE fields to various tables in Geodatabase (not necessary)



c. SUBTYPE.AML – AML Script that adds the subtype field to coverages or shapefiles (if not present) and calculates the subtype value (integer). Refer to the Microsoft Visio diagram below.

2. Data Migration

a. LOAD.AML – This script reads through the spreadsheet and creates "cov2sde" commands for the data load. It tests for attribute mapping files (text files that match the attributes in the data with SDSFIE compliant attributes in the geodatabase) in an attribute maps (am) directory and includes those if present. It also creates a truncate script, which empties the geodatabase before running a load.

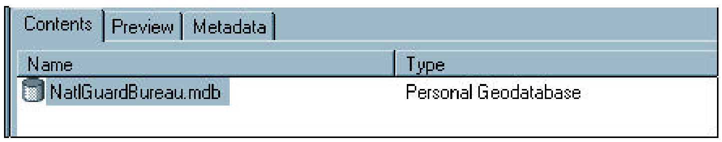
3. Post-Processing (QA/QC)

a. Currently each script creates a log file using the conventional <script_name>.log. Review these log files and make notes where problems occurred.

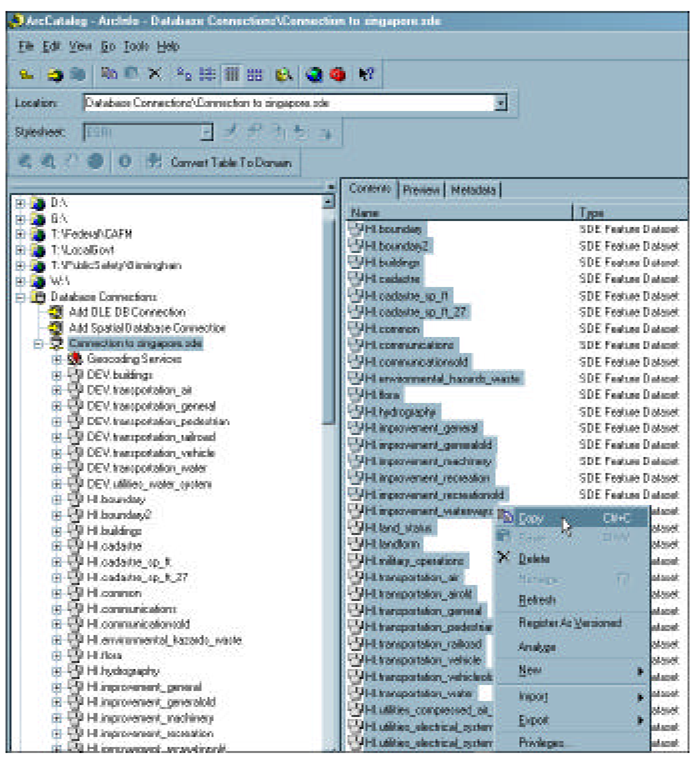
b. Run an ArcCatalog Visual Basic script, which creates a list with feature counts per layer. This will allow a quick scan on data layers to determine where problems might occur.

C. Migrating SDE/Oracle Geodatabase to Personal Geodatabase

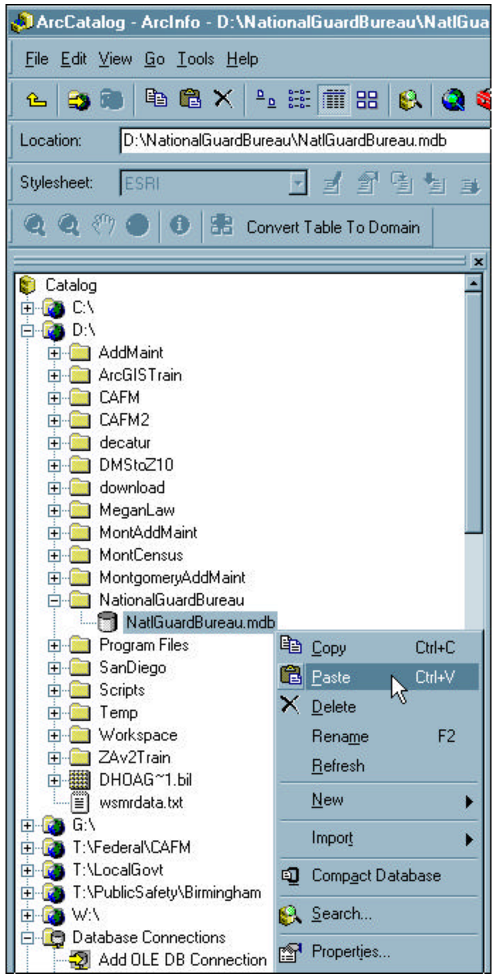
1. Type in the name of the Personal Geodatabase



2. Open the SDE/Oracle Geodatabase in ArcCatalog. Select all of the feature datasets, “right-click” anywhere on the highlighted feature datasets and select **Copy**.



3. *"Right-Click"* on the Personal Geodatabase and select **Paste**.



7. Quality check SDSFIE Geodatabase to ensure loss-less data migration

This is a VERY important step! Compare every attribute in every record of every shapefile and coverage that was migrated to the geodatabase attributes and records. Each piece of data should be accommodated for in the SDSFIE geodatabase to ensure loss-less data migration.

8. Add metadata information to SDSFIE Geodatabase records

The attribute meta_id should be populated with the Source Name (e.g. Structurescov) to identify the original data source. To populate this attribute, edit the geodatabase using Access 2000. Create an update query to insert the Source Name in the meta_id attribute (e.g. UPDATE structure_existing_area SET structure_existing_area.meta_id = "STRUCTURESCOV"). This should be done for every shapefile and coverage that is migrated.