

WEST VIRGINIA SCHOOLS

Updating the Geographic Names Information System

By: Laura Cathers
West Virginia GIS Technical Center
West Virginia University
April 2005

Purpose

The Geographic Names Information System (GNIS) schools' data had not been collectively updated in up to 20 or 30 years. Therefore, the majority of the 1,675 school records in the GNIS were out of date; closed, consolidated, or changed names. In addition, the coordinates originally stored for each record were collected using topographic maps. The three goals of this project were to update the overall list of all schools in the GNIS data base, add a subtype field for all school records, and update the coordinates.

Data

In 2004 the West Virginia National Guard (NG) collected GPS coordinates for almost all grade schools in West Virginia, as part of the Counter Drug Task Force. The West Virginia GIS Technical Center requested use of the data to update the GNIS and statewide schools data sets.

Countywide lists of schools and associated data were received from the NG on the 16th of February 2005, in Excel spreadsheets. Each spreadsheet contained the following fields:

- School Name
- School Code
- Address
- County
- Telephone Number
- Fax Number
- School Type
- Grades
- GPS Coordinates

Randall Kirk, of the WV Education Information System office, West Virginia Department of Education (WVDE), provided current lists of open schools and of schools which have recently closed. These lists, along with the WVDE online Education Directory (http://wvde.state.wv.us/ed_directory/), were used to validate the names and lists provided by the NG.

Methods

The countywide spreadsheets were compiled into one statewide spreadsheet, and many fields needed to be parsed in order to separate the information in cells that contained

multiple fields of data. The address field was parsed into five fields; Address 1, Address 2, City, State, and Zip Code. The School Type field contained both the School Type Code and the descriptive School Type, which were split into separate fields. The GPS Coordinates field, which contained coordinate pairs in one cell, was parsed into two fields; Latitude and Longitude. The coordinates were provided in degree-minute-second (DMS) format, but needed to be in decimal degrees for verifying spatial locations, generating map books, and submitting to the GNIS. Therefore, the coordinates were converted from DMS using the equation: **(degrees + (minutes/60) + (seconds/3600)) = decimal degrees.**

The DMS coordinates and the address fields are not currently recorded in the GNIS, though were used in the creation of the maps of each school which are being compiled into a statewide schools map book.

The coordinates and School Type (subtypes) data from the National Guard lists were used to update old data and determine what schools are new and which schools in the GNIS no longer exist. Another main goal was to add a Subtype field for which each record could be categorized more specifically. The Subtypes were derived from the School Type field determined by the National Guard. The NG assigned each school a School Type and School Type code, from a list of seven categories:

- 1 – Public Primary School
- 2 – Public Middle School
- 3 – Public High School
- 4 – Alternative
- 5 – Vocational / Technical
- 6 – Private School
- 7 – Other

Previously in the GNIS, all schools were grouped together in one main Feature Type, regardless of level or classification. The importance of subtypes is mainly to provide more comprehensive querying abilities. For instance, to separate out public schools from private schools, or grade schools from higher education institutions. However, the School Types provided by the NG were unfavorable for use in the GNIS due to the frequency in which records fitted the definition for multiple School Type categories, and were therefore vaguely classified as Other. This would potentially skew the outcome of a query if the data user did not think to skim the Other category for records that also fit their desired classifications. To resolve this issue, the NG School Types were generalized into more user friendly subtypes:

- Public
- Private
- Alternative
- Vocational/Technical
- Early Childhood
- College
- University

The Public subtype combined all three levels of public schools: NG School Type Codes 1, 2 and 3. Some of the records in the Other School Type fit the Early Childhood subtype; such as Day Care and Head Start centers, otherwise, most schools in the Other category were public schools which overlapped multiple NG School Type classifications; such as consolidated elementary/middle schools.

Subtypes:	Count
Public	724
Private	116
Alternative	12
Vocational/Technical	39
Early Childhood	4
TOTAL	895

Microsoft Access reports were generated for the NG schools data, grouped by counties then sorted by subtypes. Then, all school features for West Virginia were queried by county in the GNIS and printed. The data from each source was compared, county-by-county, to determine what edits needed to be made. It was determined that it would be necessary to sort the schools into four categories based on the type of updates; Closed Schools, Current Records, Name Edits, and New Records.

Closed Schools (Historical)

Using the “Closed Schools” list from the WVDE to compare to the GNIS records, I went through and highlighted all closed schools in blue, which were not already ‘historical’ in the GNIS. These were then compiled into a spreadsheet, *WV_Schools_GNIS_Make_Historical_042005.xls*, to submit to GNIS for updating these records to ‘historical.’ Then another spreadsheet was compiled for the schools that were not on the closed list, but were not included on any of the other lists indicating that they were still open. Therefore, these were assumed to have closed, with the exception of private schools.

Once all validation was complete and I was confident that the lists were as accurate as currently possible, the second list was combined with the main “Closed Schools” spreadsheet. A new field was added to the *WV_Schools_GNIS_Make_Historical_042005* spreadsheet in order to distinguish the two sources of reference. The new field is a check column, *WVDE_list*, in which "Y" represents schools that have been officially recognized as closed by the WV Department of Education. All these records should be updated in the GNIS to reflect “Historical” status.

Current Schools – Update Coordinates

The GNIS county lists were compared to the current list of schools provided by the NG, in the Access reports layout. All school names that matched those which were in the GNIS were highlighted in purple. A separate spreadsheet was generated with only these schools, *WV_Schools_GNIS_CoordinateUpdates_042005.xls*. These records need their coordinates updated and the Subtype field added in the GNIS data base.

Names Edits

All schools whose names had been altered or changed were highlighted in pink, and were compiled into a new spreadsheet, *WV_Schools_GNIS_Name_Edits_042005.xls*. These records need their name and coordinates updated, and the Subtype field added in the GNIS data base as well.

New Schools

All schools in the NG list that were not in the GNIS already were separated into a new spreadsheet, *WV_Schools_GNIS_NewRecords_042005.xls*. These schools need to be entered into the GNIS, with all appropriate data.

Validation

Once the NG list was organized into separate spreadsheets for the different types of updates, the WVDE Education Directory was used for validation of the overall list and subtypes. The Education Directory provided complete countywide lists of all public grade schools, and the select private schools which use the WVDE resources. The purpose of the comparison was to validate the NG data with more official lists from the WVDE. A few inconsistencies were discovered within the data; some which required corrections to specific records and some records were found in one source but not included in the other's list. Based on these inconsistencies I verified the level of completeness for which each source intended their data to meet. The National Guard had not yet decided whether to include private schools in their data collection, and therefore had not completed the collection of those schools. However, they did start to GPS private schools, so I received updated coordinates and names for some. To my knowledge, there is no list of private schools that is 100% comprehensive, due to the fact that they are private entities and not required to report to any governing authority. The WVDE list only contains those private schools which request the use of WVDE resources, and even at that, it is possible that schools on their list may have closed and simply not informed the WVDE. Due to these issues with data collection, I was very conservative in what changes I made to the private schools subtype. The only Private Schools that were made historical are those on the list of schools that have definitely closed. As far as Public schools, that subtype should be complete and the most accurate of all the subtypes.

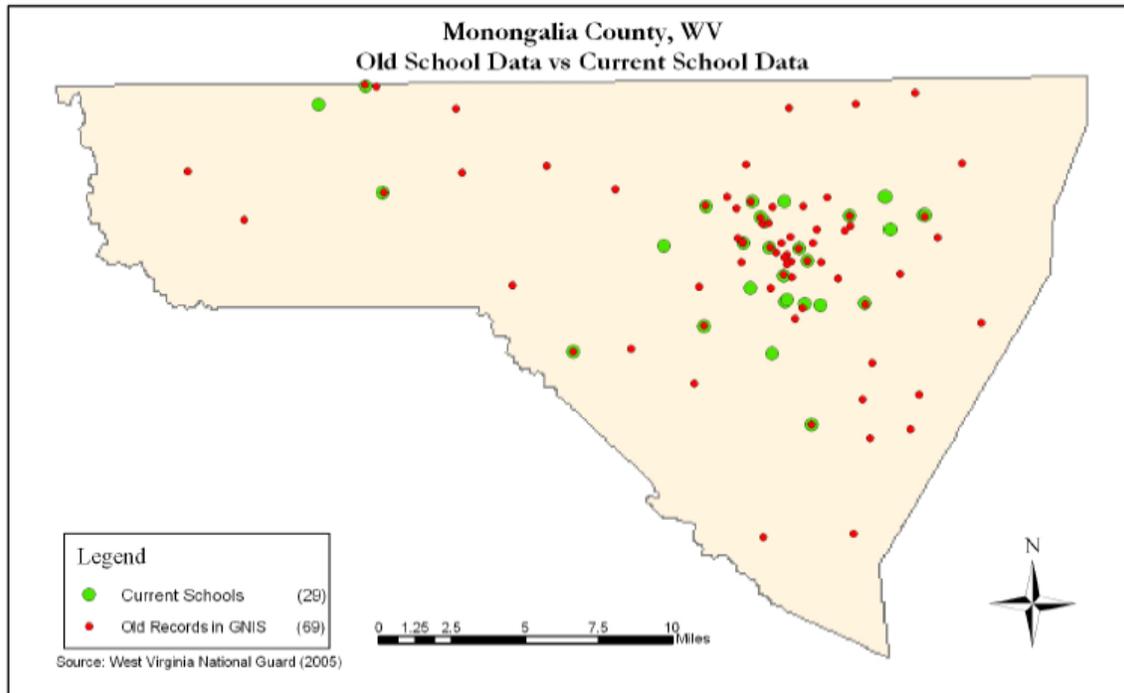
There are 13 records that were in the WVDE lists, but not collected by the National Guard, that required editing or entry into the GNIS; five required editing the name, and eight were new records to add to the database. The current schools which required name edits simply kept their coordinates already stored in the GNIS, and the new records were geocoded using Tele Atlas' Eagle geocoder. Only three of the eight new records returned a good match when geocoded, either a block face match or near match. The remaining five returned a zip code or zip+2 match, which were decidedly not precise enough for entry into the GNIS.

UPDATES - Final Count:		
Total Current Schools (New and Edits):	<u>895</u>	
New Records:	<u>311</u>	
National Guard data:	<u>303</u>	
WVDE only (not in NG):	<u>8</u>	
Coordinate Updates:	<u>408</u>	
Name Edits:	<u>176</u>	
National Guard data:	<u>171</u>	
WVDE only (not in NG):	<u>5</u>	
Make Historical:	<u>1,006</u>	
WVDE Closed list:	<u>114</u>	(matched w/ GNIS)
Assumed closed:	<u>892</u>	

GNIS (pre-updates):		
Total School Features Initially in GNIS:	<u>4,153</u>	
Historical:	<u>2,481</u>	
Not-Historical:	<u>1,672</u>	

Spatial Data Verification

ArcMap was used to verify the new coordinates. Once the NG data was received and compiled into a usable statewide spreadsheet, it was saved as a DBF, and imported into an empty ArcMap project, where the coordinate data was exported as a shapefile. The shapefile was then opened in ArcMap, with only the county boundary file for West Virginia. All the records in the attribute table for the schools shapefile were selected one county at a time and viewed on the statewide map to ensure that each record was at least located in the correct county. Then the old GNIS schools shapefile was added to the map, in order to visualize the difference in the number of records in the old data versus the current data, and the difference in location of the GPS coordinates versus the old coordinate locations for the schools which still remain.



Comparison of the out of date GNIS school records vs. the updated current schools data (April 2005).

Submission to GNIS

Typically there is a form to fill out and submit online for each feature in the GNIS individually. There are three different forms depending on the size of the feature (point or polygon) and whether it is a new record or an edit. However, since there were over 1,900 schools which required updating, a batch submission was most practical. Therefore, all the data was organized into the spreadsheets mentioned above, in which all the fields necessary for submission to the GNIS were added. The necessary data includes:

- Feature name
- Feature ID (edits only)
- Name Reference Code (using bibliographic codes predefined by the USGS)
- Feature Type [School]
- Country [US]
- State [WV]
- County
- Primary Latitude (decimal degrees, NAD83 or WGS84 is acceptable)
- Primary Longitude (decimal degrees, NAD83 or WGS84 is acceptable)

For the GNIS files the address, phone number, fax number, and grades; which were included in the original files from the NG, were excluded in order not to crowd the data that is necessary. However, all data was combined for use in creating the map books. Additionally, in the future it would be of great help if the GNIS would begin storing the address for cultural features.

Recommendations for Future GNIS Entries

Storing the address for cultural features in the GNIS would significantly simplify the task of updating large sums of features, and help make future updates more accurate by erasing the guess-work that can occur when all the other data associated with a feature has changed or been updated. The reason for this request is that during this process, multiple records within close proximity were found, of whom also shared similar names. When one of their names changes, it is difficult to decide which is the correct feature to edit, especially during an update such as this when edits have been made to the attributes which would otherwise be used to link the new name to the existing feature. With both the coordinates and name having changed, and no records keeping track of the specific changes, there is no data directly linking the new information to a specific feature. If addresses were stored for these features there would be no question as far as which feature the new data was to replace.