Going From Greenline Mylar to Digital Geologic Map

By Thomas G. Whitfield, P.G.

Pennsylvania Bureau of Topographic and Geologic Survey
3240 Schoolhouse Road
Middletown, PA 17057
Telephone: (717) 702-2017
Fax: (717) 702-2065
e-mail: twhitfield@state.pa.us

ABSTRACT

Land development pressures in glaciated northeastern Pennsylvania and the Poconos have resulted in a great demand for information about the surficial deposits of the area. Surficial deposit mapping of this area has been an on-going Statemap project for many years (Statemap is a component of the USGS National Cooperative Geologic Mapping Program). Two or three 7.5-minute USGS quadrangles are usually mapped each year. Until recently, the finished (but not finalized) project consisted of a text report and one or more clear or greenline mylar quadrangle maps. A finished project is one in which the author has completed his or her fieldwork, maps, and documentation, and has had a minimum level of review. A finalized project is one that has had a more formal review and has met all the standards necessary for formal publication. Geologic contact lines, isochors, bedrock outcrop ledges, etc. were drafted directly onto the mylar maps or on mylar overlays. Other features were hand drafted or rub-on transferred to the mylar sheets.

The initial intent was to release these maps as formal publications at a later date, but given the demand for the data, they were released in the open file series. Each open-file report consisted of large, at-scale photocopies of the mylar maps and various overlay combinations, in widely varying quality, and a copy of the report.

When GIS and digital map data began to be widely used in the 1990's, users began to request these maps in a digital format, preferably as a georeferenced GIS file. Early attempts to convert the mylar maps to digital were problematic. Many of the greenline mylars had lines drafted directly on them. Scanning these maps and separating the drafted line from the background was very difficult and time consuming. Some semi-automatic digitizing was tried, but most of the digitizing had to be done by hand.

The problem was studied and, after many trials, a solution emerged. New and improved scanning techniques and software solved the problem of digitizing lines that had been drafted directly on the greenline mylar. The drafted lines could be separated at the scanning station and saved as a separate binary image. A binary image is a raster image with just two values. Each pixel is either a one (1) or a zero (0). Much improved auto-tracing software that allowed interactive image editing was also a great step forward. Now, we include many different georeferenced and attributed data layers in the digital map products released in our open-file series.

THE PROCESS

Heads-up digitizing of a scanned image is generally a straightforward process. Automated or semi-automated digitizing speeds the process up considerably. For successful tracing, however, most automated digitizing programs require a binary or black and white image. The tracer will follow ones or zeros, but not the number ranges associated with color designations. Producing such an image from a greenline mylar can be a difficult task.

During the scanning operation, a threshold setting determines the sensitivity of the scanner. The threshold sets the values used for dividing tonal ranges into black and white output. Setting the threshold high enough to drop out the greenline background and noise in one area may cause the black object lines to be dropped out in another area. Setting it too low will increase noise (speckling) and will pick up unwanted background lines. A "happy medium" can be elusive.

Many of the new scanning interfaces have an automatic thresholding feature. During the scanning process, the scanner will analyze small sections of the scanned object map and determine the optimal threshold setting for that section within a variability range that is set at the interface. By varying the threshold for each section of the object map, noisy areas are cleaned up and light or faded object lines are more reliably detected. Although this process was designed for maps such as blue-line ozalids and older maps that tend to degrade to a yellowish color, it worked very well for us in dropping the greenline background from our mylars and keeping the black contact lines.

We used a Vidar Titan II scanner. It is a color scanner capable of scanning maps up to 40-inches wide and (assumed) unlimited length. It has a dual roller feed, three cameras, and an optical resolution of 400 dpi. The dpi can be increased in the software, but anything above the optical resolution of the cameras is done through software interpolation. The scanning software we used is Vidar TruInfo v1.4.6, which was supplied with the scanner.

Obtaining a good scan also depends on an effective contrast setting. The scanner illuminates the mylar being scanned but, because the mylar is translucent, random background noise is picked up from the light bouncing back from the paper hold-down bar of the scanner. To correct this, we made a sheath by folding a large piece of clear acetate in half. We then placed a clean sheet of plotter paper behind the mylar and placed both inside the acetate sheath. This gave us a very clean, white background behind the mylar, and protected the mylar during scanning. This method is also very useful when scanning delicate maps and papers. The acetate sheath holds loose or tattered pieces in place and protects them from damage that the scanner rollers may cause.

Georeferencing the scanned image is an essential step in this process. Before scanning, we had to ensure that the reference tic marks were clearly visible on the mylars. The reference tic marks on the greenline mylar also were green. We had to change them to black by either hand drafting or by rub-on transfer so they would be detected by the scanner.

Once the scanning was completed, we brought the images into ArcMap 8.3 and georeferenced them to a 2.5-minute georeferenced point and line grid. Once the images were georeferenced, we prepared them for vectorization by converting them into grids.

Vectorization of most of the linework was accomplished using the ArcScan module extension of ArcMap 8.3. ArcScan allowed us to do interactive raster editing and clean-up while previewing how ArcScan was going to vectorize the lines. Raster line intersections have always been one of the hardest things for ArcScan to interpret. "T" intersections would commonly have a deep "V" in them where the tracer would move to the center of the pixel cluster in the middle of the "T" before continuing down the pixel line. Also, lines intersecting at low angles often have pixel in-fills between the lines as they approach the actual line intersection. The tracer often interprets the line intersection short of its actual location and at a larger angle than intended. Interactive raster editing allows you to remove the in-fill pixels between the low angle lines and clean up a "T" intersection, while observing how ArcScan intends to interpret and vectorize the intersection. This allows you to obtain a vector trace that is truer to the actual intersection, and avoid time-consuming clean-up of the vector lines.

Once the raster editing was done, we used the ArcScan automatic vectorization feature to vectorize the entire scan. The results were very good, but some final clean-up was necessary. Discontinuous lines, points, and other features were digitized by hand. Line and point placement were checked for accuracy and corrected where necessary. Individual data layers were attributed where appropriate and checked for accuracy. Author review of the vector files in some cases resulted in changes and clarifications. Editing was easily accomplished on the digital files.

The overall goal of this project was to quickly create digital data and release it to the public without the delays involved in creating a formal publication. Caveats apply until formal publication, but the data is quickly available. Nevertheless, many in the user community prefer hard copy paper maps. For this reason, a generic map template was created in which any of the thirty open-filed quads can be placed, and with minimal editing and adjustment, printed within several hours. The map document is saved as a PDF file and can be reprinted at any time.

For those using the digital data in ArcGIS, we include the ArcMap MXD (ArcMap document) file in our open-file data release. We also include a PMF (ArcPublisher created) file for use with ArcReader. ArcReader, a free download from ESRI, is a limited version of ArcGIS that allows the user to view the data, do some limited data manipulation, and print the map in the same template that was used to create it in ArcGIS.

A text write-up of formation descriptions, stratigraphy, structure, geologic settings, etc. also is available as a separate document for each quad. By not including much text in the marginalia, map production time is greatly reduced.

A 1:100,000-scale composite map template also was constructed for regional studies. As digital conversion of the 1:24,000-scale quads is completed, their digital files are referenced in the 1:100,000-scale map MXD template. When printed, the 1:100,000-scale map appears to be a composite, but is actually a mosaic of the many digital 1:24,000-scale quadrangles. Printing this map as a mosaic instead of a composite adds to the processing overhead. In a composite map, polygons of the same formation along quad boundaries are merged together. In a mosaic, like polygons along the quad boundaries are not merged, and remain separate, thereby increasing the amount of data to be processed for a print file.

The composite map is also self-correcting. Because the 1:24,000-scale quadrangle data sets are referenced by the 1:100,000-scale composite map MXD file, edits and corrections are reflected in the composite map as they are completed on the 1:24,000-scale data. Because this map is a preliminary, open-file release, caveats are included in the map template margin and in the metadata. These caveats warn of possible edgematching problems along quad borders and the changeable nature of the data.

CONCLUSION

One of the uses of the Open File Report series has been to disseminate maps and data that are incomplete or are finished and waiting for formal publication. Usual distribution methods included photocopied maps, map composites, notes, etc., where the user could obtain the data, with various caveats. The increasing demand for digital data, however, made distribution of analog paper reports and maps less desirable.

Using some simple techniques, we were able to vectorize various map layers into a digital format. Development of a generic map template allows us to "make a map" by dropping the various map layers into the template.

In northeastern Pennsylvania, we have more than 30 USGS 7.5-minute quadrangles already mapped in analog form. Each quad has from 3 to 5 mylar overlays along with a greenline quad base. Using the techniques described above, we were able to complete a digital rendering of a quad area in about 3 to 5 man-days.

These maps are simple but functional. The digital data is now available for those who don't want to wait for formal publication, with the typical open file caveat that the data is subject to change. Because it is digital, the data now has more utility than it had when it was available only as composite photocopies.