Nathan B. Piekielek

Best practices for georeferencing large scale historical fire insurance maps of the USA

Keywords: fire insurance maps, Sanborn Map Company, georeferencing, accuracy assessment, best practices

Summary: There is increasing interest in transforming historical print maps from passive archival documents into value-added digital geospatial data through the process of digitization and georeferencing. Central to this activity is the quantitative evaluation and documentation of the quality of the original map and resulting data product(s) produced. Best practices have been proposed for small scale (i.e. large area) historical print maps, while a similar set of recommendations for large scale (i.e. small area) maps are lacking. The current study presents methods for and the results of a project to digitize and georeference historic fire insurance maps produced by the Sanborn™ map company during the late 19th and early 20th centuries in the state of Pennsylvania, USA. A random sample of 202 out of approximately 29,000 individual map sheets was evaluated for their horizontal positional accuracy against contemporary reference data. Results were summarized into a set of proposed best practices for georeferencing large scale historic fire insurance maps of the USA.

Introduction

Historical maps are treasured for the cultural, artistic and cartographic contributions that they have made to past and modern human societies. As some of the most important scientific documents of their time, they also offer reliable information about the past in ways not afforded by other historical sources of data and information (Rumsey and Williams 2002). Georeferencing historical maps has become a common practice in a variety of academic and applied contexts. Most georeferencing to date has been done with a single specific application in mind. As such, researchers (i.e. those doing the georeferencing), have typically georeferenced maps for small geographic areas to an accuracy standard that met their needs, but rarely have they shared their methods of evaluation, georeferencing results, and raw datasets 1. More recently, the motivations for, and people involved in georeferencing historical maps are transitioning from a focus on individual research projects, to the production-scale creation of quantitative datasets across large geographic areas (e.g. Liknes et al. 2011), with the primary intent being that the data produced are broadly useful to others. This shift brings with it a necessary change in the workflow and methods being employed to georeference (Balletti 2006, Fuse 1998). The production-scale transformation of historical maps into useful geospatial data requires as a first step, the development and adoption of data standards and commonly agreed upon “best practices” for georeferencing. Favretto (2012) suggested that the first step in georeferencing small scale (large area) maps of ancient origin should be to assess the geometric accuracy of the spatial information contained on historical maps; and to document that information in metadata so that future users of the data can evaluate for

---

* University Libraries, Pennsylvania State University [nbp104@psu.edu]
1 This is also true of many non-georeferencing projects that produce research data.
what applications they are appropriate? The method recognized that small scale maps of ancient origin were not based on inconsistent information that was available about different parts of the mapped area and therefore, the quality of information presented on the map likely varied geographically. Similarly, Grosso (2010) identified the primary challenges to georeferencing small scale maps of ancient origin as matching map data to a contemporary projection and coordinate system; and choosing a georeferencing transformation that balances accuracy versus overfitting.

More recently than the ancient past, mapping techniques became more sophisticated and larger scale maps began to be produced for specialized applications. One such application was fire insurance mapping for USA cities and towns during the nineteenth and twentieth centuries (Ristow 1981). The production of fire insurance maps typically included carefully carried out field survey measurements using steel tape measures (and other instruments). Based on field measurements, detailed diagrams were created and transferred to a drafter and cartographer for map production, and later, publication (Oswald 1997). The Sanborn™ Map Company was established in 1866 and became the single largest producer of USA fire insurance maps, creating more than 1.5 million individual map sheets and updates (i.e. remapping) for more than 12,000 USA cities and towns, typically at 1:50 and 1:100 map scale (Oswald 1997). Sanborn™ maps contain information on building footprints, heights and construction materials, land uses and tenet names, street names and locations, the location of water lines and fire hydrants, among other information. Today, Sanborn™ fire insurance maps are used extensively by architects, planners and environmental engineers for applications including historic preservation, zoning, genealogy, industrial archeology and environmental assessments among other uses (Berry 2003), (Keister 1993), (Leonard and Spellane 2013), (Thrall and Knetsch 2014), (Wright 1983), (Wrigley 1949). As such, Sanborn™ maps have been a popular historical map to georeference across small areas (i.e. individual cities and towns) for specific research projects. Academic libraries in the USA often own regional collections of Sanborn™ maps that are among their most highly used map collections. The Library of Congress houses a complete set of Sanborn™ fire insurance maps and the entire collection was microfilmed in 1977 (Oswald 1997).

Fire insurance maps likely present a different set of challenges to georeferencing than small scale maps of ancient origin for a number of reasons. First, because they are based on field measurements that were collected in a consistent manner their geometric accuracy should not vary across the mapped area. Second, because they cover such small geographic areas (i.e. very large scale) they will not suffer significant distortion when placed into projected map space, therefore higher order transformations are not likely necessary and consequently overfitting should not be a concern. Third, given their large scale and detailed field measurements, finding contemporary reference base map data that is of demonstrably greater positional accuracy (a requirement of reference base maps, FGDC 1998), likely presents a challenge for fire insurance map georeferencing efforts that span large spatial domains.

With these potential challenges in mind, the present study developed and applied a georeferencing procedure to a set of digitized Sanborn™ fire insurance maps of Pennsylvania (PA), USA with the following study objectives, to:

i. Quantify the mean geometric accuracy of PA Sanborns™ maps.

ii. Compare the mean geometric accuracy of PA Sanborn™ maps to the positional accuracy of the best available reference base maps.
iii. Summarize the lessons learned from the above results into a set of proposed best practices for georeferencing large scale historical fire insurance maps of the USA.

**Methods**

The area and time period of study for the present project included Sanborn™ fire insurance maps produced for the state of Pennsylvania (PA), USA that are owned by The Pennsylvania State University Libraries. The Pennsylvania State University Libraries owns a nearly complete set of Sanborn™ maps for PA, 29,498 of which have been digitized as TIF files at 600 dots per square inch (dpi) (complete digitization specifications are available upon request to the author). Pennsylvania is a state in the mid-Atlantic region of the eastern United States (USA) (latitude 40.882, longitude -77.803) (Figure 1). Pennsylvania was an early colonial settlement in USA history and its cities and towns were commonly the subject of fire insurance mapping that was typical of the nineteenth and twentieth centuries. In the years that have followed fire insurance mapping, many PA cities and towns have undergone a land use transition from industrial to other uses.

A random selection of 202 Sanborn™ map sheets was drawn for inclusion in the present study. Production years for the entire Sanborn map set ranged from 1884 to 1947 and for the sample set 1884 to 1933. Visual inspection of the sample set indicated that it was geographically dispersed across the state including map sheets drawn from the major cities of Philadelphia, Pittsburgh and Scranton, but excluding any sheets from within the city boundaries of the state capital, Harrisburg (Figure 1).

![Figure 1: Study area and location of selected Sanborn map sheets.
*Note that some sheet locations are obscured by one another at this map scale](image)
Quantifying the geometric accuracy of historical maps requires a contemporary reference base map of known positional accuracy. There was no single suitable reference base map of statewide coverage available, so two similar base maps were used that together provided statewide coverage. Orthoimages were compressed as LizardTech MrSID files and mosaicked by the data producer for each county (Table 1). Orthoimagery mosaics were of one foot pixel resolution and at least 4.5 (PAMAP\(^2\)) and 5 (DVRPC\(^3\)) feet mean positional accuracy at a 95 percent confidence upper bound respectively, which is consistent with the Federal Geospatial Positioning Accuracy National Standard for the USA (FGDC 1998). Both sources of orthoimagery dated from the mid-2000s (PAMAP 2003-2006, DVRPC 2005). DVRPC imagery covered the southeastern PA counties of Bucks, Chester, Delaware, Montgomery, and Philadelphia and PAMAP imagery covered the remaining 62 counties in the state. Countywide orthoimagery mosaics were downloaded in the North American Datum 1983 StatePlane Pennsylvania North FIPS 3701 (US Feet) and South FIPS 3702 (US Feet) (ESRI:102728 and ESRI:102729; www.spatialreference.org), with division of north and south dictated by the data providers. All georeferencing steps that follow were performed in Environmental Systems Research Institute’s ArcGIS10.3 in the projection native to countywide orthoimagery mosaics.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Data Producer</th>
<th>Online linkage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digitized Sanborn™ Maps</td>
<td>A</td>
<td><a href="https://www.libraries.psu.edu/psul/digital/sanborn.html">https://www.libraries.psu.edu/psul/digital/sanborn.html</a></td>
</tr>
<tr>
<td>PAMAP cycle 1 orthoimagery</td>
<td>B</td>
<td>ftp://pamap.pasda.psu.edu/pamap_imagery/cycle1</td>
</tr>
<tr>
<td>DVRPC orthoimagery</td>
<td>C</td>
<td>ftp://www.pasda.psu.edu/pub/pasda/dvrpc/</td>
</tr>
</tbody>
</table>

A – The Pennsylvania State University Libraries  
B – Pennsylvania Department of Conservation and Natural Resources Bureau of Topographic and Geologic Survey  
C – Delaware Valley Regional Planning Commission

| Table 1: Data Sources. |

There were two primary steps in quantifying the geometric accuracy of PA Sanborn™ maps. First, in addition to a random sample of map sheets, a random sample of ground control points (i.e. places that link reference base map locations to historical map locations), was required in order to make inference across the entire mapped area. Second, because random samples of geographic data are often spatially clustered and the goal was to make inference to the total mapped area, a spatially-balanced random sample of ground control points was employed (Theobald et al. 2007). The following procedure was used to georeference and assess the geometric accuracy of Sanborn™ maps (hereafter referred to individually as ‘map sheets’).

Map sheets were overlaid in their approximate position relative to reference base maps. To identify the geographic space across which the spatially-balanced random sample was to be drawn; four ground control points were rapidly identified, placing the map sheet and rotating it appropriately in the projected space of the geographic information system (GIS) environment. A spatially-balanced random sample of 100 potential ground control points was then drawn from the geographic extent of the roughly placed map sheet. The four initial ground control points were then deleted and georeferencing commenced with the spatially-balanced random sample of potential ground control points.

---

\(^2\) Digital Base Map of Pennsylvania  
\(^3\) Delaware Valley Regional Planning Commission
Potential ground control points were evaluated based on whether they intersected an identifiable feature that did not change location through time on both the map sheet and reference base map. Fine-scale adjustment of the placement of suitable ground control points followed. Features used for ground control were most often the right-angle intersections of buildings, fencerows or other artifacts of the human built environment. Ground control points were less often the centerpoints of road intersections. Road and sidewalk corners were not used because roads in the study area were often widened over time, but typically were widened uniformly on all sides. Similarly, stream and river banks were not used as ground control points because they are dynamic through time (Micheli and Kirchner 2002). Georeferencing continued until the 100 spatially-balanced potential ground control points were exhausted, or 25 acceptable ground control points were identified. A goal of 25 ground control points was chosen based on USA federal guidelines (FGDC 1998), and following the lead of other proposed methods (Favretto 2012). The first-order (i.e. affine) root mean squared error (RMSE) was recorded for each map sheet that was successfully georeferenced, along with the number of ground control points that were used. A polynomial transformation was not used even through it would have resulted in a superior historical map RMSE because the objective of the study was to quantify the geometric accuracy of the mapped information in its native form (i.e. without stretching or warping). An affine transformation places the historic map in projected geographic space, rotates it and the remainder of ground control points quantify the degree to which the location of mapped objects differ from the same objects in a reference base map that itself has some positional error associated with it. This procedure was applied to all 202 randomly selected Sanborn™ map sheets. Some of the areas represented on map sheets had changed too extensively for ground control points to be identified and this was recorded as “Not enough GCPs” (GCP meaning ‘ground control point’). The mean of all mapsheet RMSEs was then compared to the reported mean positional accuracy of reference base maps.

The PAMAP program quantified the positional accuracy of their countywide orthoimagery mosaics by field visiting identifiable locations in their imagery, logging survey-grade global positioning system (GPS) coordinates and calculating the difference between field-surveyed coordinates and mapped orthoimagery coordinates. PAMAP makes all of their positional accuracy data publically available whereas DVRPC reported only the 95% confidence interval upper bound (5 feet positional accuracy). Access to the detailed PAMAP accuracy reports allowed a more detailed comparison to historical mapsheet RMSE than was possible statewide. To make a more detailed comparison across the PAMAP domain, base map orthoimagery mosaics and their positional accuracy reports were subsetted to those that intersected mapsheets that returned a successful RMSE. Similarly, statewide mapsheet RMSEs were subsetted to those that intersected with PAMAP base map orthoimagery. Mapsheet RMSEs and and reference base map positional accuracies for the PAMAP domain were compared so that if they overlapped in their 95 percent confidence upper or lower bounds, their means were interpreted to be statistically indistinguishable from each another, thereby, not meeting the criteria that reference base maps be of distinguishably higher positional accuracy than the historical maps that they were being used to georeference.

Finally, because we learned that mean RMSE was different for map sheets of cities and towns that were either not laid out in a right-angle grid, or had numerous streets with irregular angles between them; we also subsetted RMSE to map sheets of gridded cities and towns and reported those results separately.

[88]
Results

One hundred thirty-eight out of 202 randomly selected map sheets were successfully georeferenced (Table 2). Of the 64 map sheets that could not be georeferenced, extensive land use conversion (e.g. from industrial to other uses) was the most common reason that ground control points could not be identified. In few cases where a map sheet could not be georeferenced, was it because there were acceptable ground control points available, but they did not intersect with the randomly selected potential ground control points. Successfully georeferenced map sheets used between seven and 25 ground control points (mean = 19) and for 20 map sheets a complete set of 25 ground control points were identified. Statewide map sheet RMSE was 7.85 feet at its 95 percent confidence upper bound and 6.27 feet at its lower bound compared to reference orthoimagery positional accuracy that was better than a mean of 5 feet at its 95% confidence interval upper bound. Comparing mean RMSEs and positional accuracies for only the PAMAP domain, the upper bound of the mean reference orthoimagery positional accuracy did not overlap with the lower bound of Sanborn™ RMSE (3.04 compared to 6.06). Mean RMSE for map sheets that depicted cities and towns that were laid out on a right-angle grid was somewhat lower than for those that depicted cities and towns that were laid out with irregular angles (Figure 2) although these two distributions of RMSEs were statistically indistinguishable (i.e. their 95% confidence intervals of the mean did overlap). However, the lower confidence interval bound of gridded map sheets nearly overlapped with the maximum mean positional accuracy allowed by reference orthoimagery (5.1 feet compared to a maximum of 5 feet).

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of Map Sheets</th>
<th>Mean RMSE +/- 95% C.I. [ft]</th>
<th>Reference Imagery Positional Accuracy +/- 95% C.I. [ft]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sample</td>
<td>202</td>
<td>7.06 +/- 0.79 [2.01 – 38.47]</td>
<td>&lt; 5.00 @ 95%</td>
</tr>
<tr>
<td>Not enough GCPs</td>
<td>64</td>
<td>6.20 +/- 1.10 [2.01 – 14.62]</td>
<td>&lt; 5.00 @ 95%</td>
</tr>
<tr>
<td>Total Sanborn™ RMSE</td>
<td>138</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gridded Sanborn™ Domain</td>
<td>126</td>
<td>7.44 +/- 1.38 [2.50 – 38.47]</td>
<td>2.33 +/- 0.71* [0.51 – 4.96]</td>
</tr>
<tr>
<td>PAMAP Domain</td>
<td>115</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*PAMAP positional accuracy is for 43 county orthoimagery mosaics

Table 2: Quantification of map sheet geometric accuracy and comparison to reference base map positional accuracy.

Discussion

Georeferencing large scale historical fire insurance maps of the USA that were based on high-quality field survey information presents a set of challenges that are unique from many other historic map georeferencing projects. Challenges include finding reference base maps of suitable positional accuracy, quantifying the geometric accuracy of historical maps with inference to the entire map set of interest, and documenting accuracy with enough detail that future users can evaluate the quality of

---

4 Confidence interval as computed by: 2*(mean(rmse)/sqrt(n))
georeferenced maps relative to their research data needs. To meet these challenges, the lessons learned from the present project were summarized into a workflow and proposed set of best practices for georeferencing large scale historical maps that are based on high-quality field survey measurements.

Figure 2: Example of a map sheet depicting a town not laid out on a right-angle grid. Map sheet is of Larksville, PA Sheet 6, 1911; rmse was measured at 29.16 feet.

*All works published in the USA prior to 1923 are in the public domain as a result of expired copyright terms; 17 U.S.C. § 304(b).

Sanborn™ fire insurance maps of PA, USA were found to be highly accurate, even relative to the positional accuracy of commonly available contemporary geospatial data layers. This was true for the entire historical map study period as a qualitative evaluation (i.e. visual inspection of a scatterplot graph), of the relationship between map sheet publication date and RMSE did not show a trend of increasing or decreasing accuracy through time. In practice, highly accurate historical map data creates challenges for identifying suitable reference base maps as convention is that base maps are of distinguishably higher accuracy than the maps that they are being used to georeference (FGDC 1998). Orthoimagery positional accuracy is independent of spatial resolution. The 1-foot spatial resolution of the reference base maps used in the present study has been superseded by orthoimagery of finer spatial resolution, but positional accuracy standards remain unchanged in the USA. Current USA standards generally refer to positional accuracies of less than 1-foot for map scales like those common of

---

5 Orthoimagery of higher spatial resolution than what was used in the present study was available for portions of the study area, but lacked positional accuracy reports.
fire insurance maps (FGDC 1998). Local-scale reference data like these are rarely available and are not practical to assemble for large areas like those required of a production-scale fire insurance map georeferencing project. Further complicating the challenge of identifying suitable reference base maps is that many contemporary geospatial data sources either are not evaluated for their positional accuracy, or their positional accuracies are not reported in metadata documentation or elsewhere. Without the known positional accuracy of reference base maps, the quantification of RMSE for georeferenced historical maps lacks important context that give it meaning.

The procedure of randomly identifying map sheets and ground control points did not prove to be an overly time consuming. The task took untrained university undergraduates following a step-by-step protocol approximately 30-45 minutes per map sheet to complete. This included the automation of as many steps in an ArcGIS ModelBuilder tool as was possible. The successful georeferencing of 138 map sheets was sufficient to identify mean RMSE to within a 95% confidence interval tolerance that was less than the spatial resolution of base map orthoimagery (0.79 feet compared to 1 foot). The 95% confidence interval is a product of the variability of observed RMSEs across map sheets, and the number of map sheets considered (see footnote ³). Assuming that georeferencing technicians cannot locate ground control points with precision less than the spatial resolution of orthoimagery base maps, this means that fewer map sheets could have been used to adequately quantify mean geometric accuracy (i.e. RMSE) for this map set. Means and confidence intervals remain the standard metrics for reporting RMSE and positional accuracy for orthoimagery data and are also highly applicable to georeferenced historical map data (Favretto 2012).

The present study applied a spatially-balanced approach to identifying potential ground control points rather than the grid approach suggested by Favretto (2012). This was because for large scale fire insurance maps there was no expectation that the quality of maps varied across the mapped area like is common of historical maps of ancient origin. Also, the goal of the present study was to make accurate inference to the entire mapped area, which requires a random sample of potential ground control points rather than a user selected sample that could introduce systematic bias. There are likely mistakes on fire insurance maps introduced by human error in the map production process, but errors do not likely exhibit coherent spatial patterning like in maps of ancient origin. By the same reasoning, the present study did not report RMSEs as deviations in both the x and y coordinate directions, although this could be done to be consistent with standards for orthoimagery positional accuracy assessment (FGDC 1998), and the suggestion of georeferencing standards for small scale historical maps of ancient origin (Favretto 2012).

An outstanding question for georeferencing large scale historical maps that are based on high-quality field survey information is, “how many ground control points are necessary to quantify the geometric accuracy of each map sheet?” The present study used a goal of 25 points per map sheet, with most sheets achieving a result that was closer to 20 ground control points. Using an affine georeferencing transformation (i.e. no warping) each successive ground control point should first place the historic map in projected space, then rotate it appropriately, and then merely provide more observations of a relatively consistent RMSE. If as this project has suggested, the quality of mapped information is consistent across the mapped area, then the variability of observed RMSEs should stabilize at some low level after the selection of three to four ground control points. Using polynomial transformations, stabilization of map sheet RMSE would require more ground control points, but should stabilize at a somewhat lower RMSE. The implication of RMSEs that stabilize after relatively few ground control
points is that far fewer than 20 ground control points may adequately quantify the geometric accuracy of large scale historical map sheets. Fewer ground control points required per map sheet would reduce the amount of time required to georeference the remaining sheets in a map set following the quantification of geometric accuracy. In the case of the map set considered for the present study, any time savings as a result of a reduced number of ground control points required per map sheet would cascade over approximately 29,000 additional map sheets to georeference. The number of ground control points per map sheet required to stabilize RMSE could be derived empirically for each map set considered, or potentially even for large scale historical maps that are based on high-quality field survey in general? Coordinated investigation of this topic by researchers who are using large scale maps of different origins would likely be fruitful in further developing best practices for georeferencing.

The lessons learned from the present study can be summarized into a set of recommendations for best practices of georeferencing large scale historical maps that are based on high-quality field survey information. This workflow (Figure 3), is scalable from a single map sheet up to a large map set like the one used in the present study. Best practices start with identifying a reference base map of known positional accuracy for the study area, preferably one that is expected to be at least as highly accurate as the historical map that is being georeferenced. The next two steps estimate the geometric accuracy of historical maps in a way that should have statistical inference to the entire mapped area being considered. If working with a map set, quantifying mean RMSE to within a 95% confidence interval tolerance that is smaller than what georeferencing technicians can reasonably identify in base map data seems like a good goal. Next, using all of the reference base map positional accuracy information available in comparison to historical map geometric accuracy results provides the foundation for understanding the appropriateness of georeferenced historical map data for other applications. If working with a map set, this is also the step where one could investigate the number of ground control points required for each map sheet to achieve a stable RMSE. Once geometric accuracy has been adequately quantified, the remainder of the maps can be georeferenced using a minimum convenient user selected ground control points. As has already been emphasized by others (Favretto 2012), reporting the results of accuracy assessment in metadata documentation is the final and perhaps most important step in transforming historical map information into useful digital geospatial data. This workflow is likely applicable to other large scale fire insurance maps of the USA during the nineteenth and twentieth centuries, including those produced by Hexamer for southeastern PA, by Perris and Browne among others for the state of New Jersey, the Rascher Map Company for areas throughout the mid-western states of the USA, The Dakin Map Company for the state of California, USA, the Charles E. Goad and Manufacturers Mutual Fire Insurance Company for the USA and Canada, and other producers of fire insurance maps in the USA and internationally.
Conclusion

Large scale historical maps that are based on high quality field measurements offer a wealth of information that is currently largely inaccessible in its analog format. Transforming the information on these maps into useful quantitative geospatial data has the potential to enable researchers and others to rapidly ask and answer questions that they have not been able to previously. As interest in transforming historical maps into geospatial data increases to the point of production-scale efforts involving many people and large map sets, the importance of georeferencing standards and best practices also increases. In the absence of a central governing body for historical maps that will create georeferencing standards, it seems more likely that best practices will come from researchers and practitioners who are familiar with historical maps and georeferencing techniques. Rather than being a final statement on the matter, the present study and proposed workflow is meant to be the start of a conversation, the result of which could encode best practices on project webpages, in conference proceedings and in the peer-reviewed literature. Failure to develop georeferencing standards and best practices may jeopardize broad appreciation for large scale historical maps and the truly remarkable information they contain that remains relevant today.

References


