Kelly Jean Fergusson, P. E., Ph. D.
Center for Integrated Facility Engineering
Civil Engineering Department
Stanford, CA 94305-4020
(415) 855-9770

Chip Eitzel
Geodesy
8 California, Suite 400
San Francisco, CA 94111
(415) 677-8748

QUALITY CONTROL METHODOLOGY FOR INFRASTRUCTURE DATA CONVERSION: A MUNICIPAL APPROACH

Abstract: The City of Palo Alto, California, has developed methods to ensure that data delivered by its conversion vendor meet both contractual requirements and user needs. The complexity of the utility infrastructure systems and a high accuracy standard required the development of a comprehensive quality control (QC) program integrated with the conversion process. This paper describes the basemapping and conversion process, proposes a typology of errors, and presents manually- and computer-based QC methods to identify errors. Computer based methods include automated checks, computer-assisted checks, and a red-lining process.

INTRODUCTION

Why do so many conversion projects fail? One contributing factor is that the conversion process contains many opportunities to *add* error beyond that which is inherent in source documents. Quality control procedures are an essential, yet often underutilized way to identify and correct errors introduced by the conversion process. A well designed quality control process improves the potential for a successful conversion project, especially if the conversion is complex and demands high accuracy standards.

This paper addresses quality control in the context of Palo Alto's base mapping and infrastructure conversion effort. First, an overview of Palo Alto's Geographic Information System (PAGIS) is given so the reader understands the strategic context of the project. Second, the conversion steps in which errors may be introduced are described. Third, a typology and description of errors are proposed. Fourth, methods used to identify errors at pertinent steps in the conversion process are presented. The paper concludes with a set of recommendations for municipalities and utility companies that are considering or planning a conversion project.

PALO ALTO GIS OVERVIEW

A primary goal of Palo Alto's Geographic Information System (PAGIS) is to serve as an infrastructure management tool for this city of 55,000 residents. Utility services including water, gas, wastewater, and electrical distribution are provided by the city's Utilities Department. The Public Works Department is responsible for the roads, the stormwater collection system, and the parcel records. Manual records for each of these data themes are being converted to digital format during Palo Alto's $1.2 million conversion project. When completed, the data themes will be accessed by software applications that will assist with evaluating the existing infrastructure, identifying risk areas, prioritizing capital improvement projects for city council approval, and designing such projects once they have been approved. To provide context for these themes, digital orthophotos of the city have also been produced. As of April 1995, 25% of the conversion work has been completed.

Objectives of Palo Alto's GIS implementation effort include spatially accurate, application-ready data, distributed data access, and assured data protection and maintenance. To achieve these objectives, an extensive and rigorous quality control methodology has been developed which will be described in detail in the body of this paper.

The need for high-accuracy spatial data is driven by infrastructure replacement projects designed by in-house engineers. The design documents are based on the digital basemap including the utilities, curb lines, and parcel lines. As-builts of the completed projects are used to update the base data. City management views the integration of CAD and GIS as a strategic priority for managing the city's infrastructure. Improvements beyond existing traditional infrastructure management tools are particularly relevant given the potential profitability of utilities and the ever-increasing competition in the utilities marketplace.

The conversion process must yield application-ready data, meaning that the spatial data have clean graphic characteristics, and that the tabular attribute content is at least as reliable as on the source documents. Producing data of this quality level provides two major benefits: streamlining the work of application developers, thus reducing application development and maintenance costs; and assuring engineering and operations personnel's acceptance of the converted data.

PALO ALTO'S APPROACH TO DATA CONVERSION

Palo Alto's data conversion effort is led by a strategically oriented steering committee composed of representatives from the Planning, Utilities, Public Works, and Information Systems Departments. The steering committee has overseen hardware and software acquisition and the procurement of the conversion contract. It identifies and prioritizes GIS applications for current and future development. For each of these major steps in GIS implementation, the committee has prepared and defended funding requests that have been approved by the city council. In addition to drawing on their own knowledge, committee members solicit technical input from outside consultants, including the two authors as well as others, particularly Bruce Joffe of GIS Consultants, Oakland, CA, and Jim Cannistra of PlanGraphics, Silver Spring, MD.

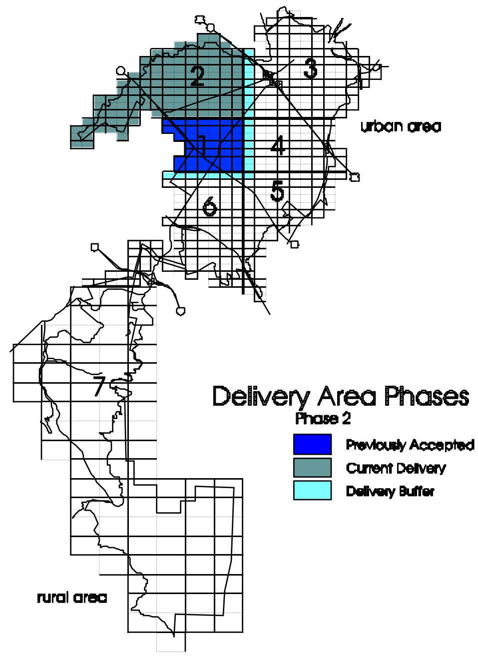
In support of the physical database design, a data dictionary was developed, an excerpt of which is shown in Figure 1. The data dictionary identifies the components of all data themes, the

FIGURE 1

organization responsible for each component, the data source, line styles and character styles, the relationships between components, and component attributes.

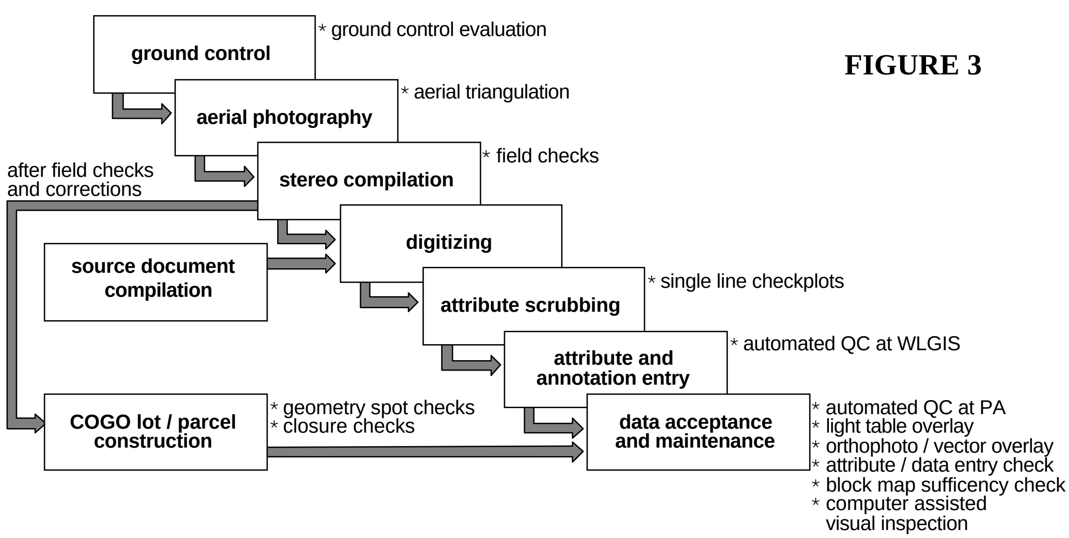
Westinghouse Landmark GIS (WLGIS) of Cary, North Carolina was selected as the vendor to perform the conversion. The conversion process was developed jointly by the steering committee, city staff, consultants representing Palo Alto, WLGIS, and their subcontractors Hammon, Jensen, Wallen and Associates (HJW) of Oakland, California, and Bestor Engineers of Monterey, CA. A first step was to agree upon an overall delivery strategy. Delivery of the converted data is proceeding in an incremental fashion as shown in Figure 2. Delivery starts with a small prototype area in which the conversion method and data dictionary are tested and refined, followed by Area 1, Area 2, and so on through Area 7. When a data delivery is received, Palo Alto personnel have 30 days to evaluate it. If quality control checks indicate that the data delivery is of satisfactory quality, it is formally accepted and the conversion vendor is paid for that portion of its work.

FIGURE 2



There are nine main steps to the conversion process: ground control, aerial photography, stereocompilation, source document compilation, digitizing, attribute scrubbing, attribute data entry, coordinate geometry input, and data maintenance. These steps are flowcharted in Figure 3, with quality control methods shown to the right of the conversion step to which they correspond. (Conversion steps for electrical themes are somewhat different, and are not addressed by this paper.) Because of the interdependence of the quality control methodology and conversion, the steps in the conversion process are described in this section.

FIGURE 3



1) Ground Control

With the assistance of Global Positioning System (GPS) technology, horizontal and vertical ground control were established using both existing and new monumentation. Least squares adjustments were made to unify the new and existing points into a coherent set of coordinated points. The monumentation was distributed throughout the relatively flat, urban, northern part of the city as well as the hilly, rural, southern part of the city. NAD83 datum standard was used.

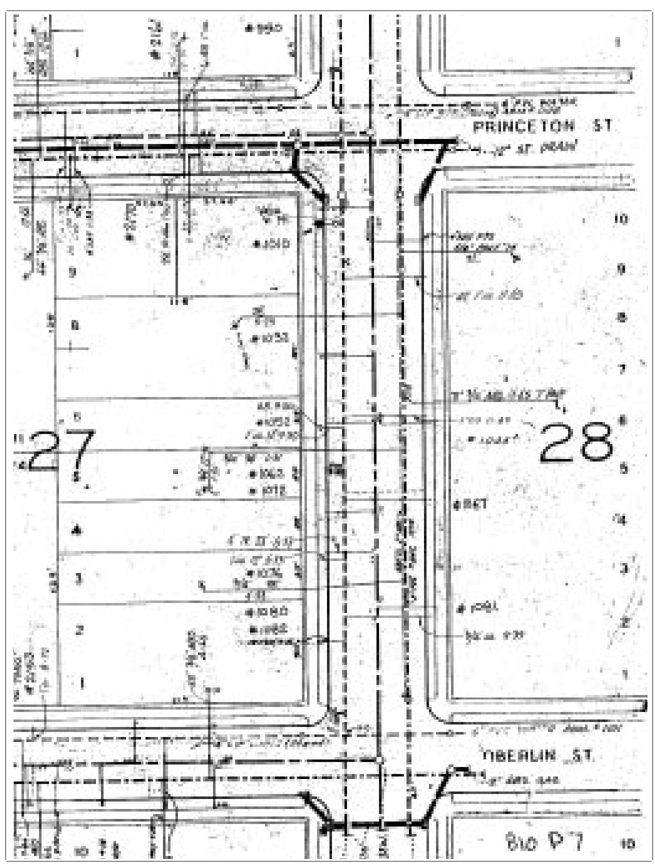
2) Aerial photography and terrain modeling

HJW performed the aerial photography, terrain modeling, and digital orthophoto production. The three variables of low flight elevation (2100 feet), high overlap of the flight paths, and the precision and density of the ground control enable the development of the terrain model to be well within horizontal and vertical tolerances relative to the true ground position in the urban area.

3) Stereocompilation of planimetric basemap

WLGIS operators stereocompiled several digital planimetric themes. The most important features to subsequent steps in the conversion process are curb lines, valves, and manholes. These features are expected to possess absolute positional horizontal accuracy of ± 0.4 foot wherever they are visible on the photographs, i.e. not obscured by tree foliage, cars, or other view obstructions. Operators distinguish between visible and obscured curb lines by using different line styles for each when performing the stereodigitizing. Curb-to-curb dimension data from block books and the pavement management manual are used to refine stereodigitized curb locations where required.

FIGURE 4



4) Compilation and preparation of source documents

Palo Alto's infrastructure records reside on several different sets of source documents: block maps (1" = 40' scale), distribution maps (1" = 600' scale), and stormwater and wastewater handbooks. Block maps are the primary source for the water, gas, and wastewater data, and a secondary source for the stormwater system. There are approximately 600 block maps to cover the city, each typically containing (as the name indicates) utilities data for one or two blocks. A block map excerpt is shown in Figure 4. The master set of Mylar™ block maps was shipped to WLGIS on an incremental basis for photographic reproduction. As WLGIS returns the block maps to Palo Alto, any subsequent updates to them (e.g. relocated gas lines) are made in red ink so that the edits are readily distinguishable for input purposes once the digital data is delivered and accepted. Palo Alto provided paper copies of the distribution maps and handbooks to WLGIS.

5) Digitizing the block maps

First, stereodigitized valves, manholes, and curb lines are translated into Microstation. Then, valves and manholes are used to register the block maps and the stormwater handbook for digitizing of the mains. Next, curb lines are used for registration on a curb-block-face by curb-block-face basis for digitizing of the service laterals. Resulting from the digitizing step are four single-line check plots (one for each of water, gas, wastewater, and stormwater themes) for each of the 561 urban-area project grid cells that were shown in Figure 2.

6) Attribute scrubbing

WLGIS personnel then transfer attribute information from the source documents to the single-line check plots by visually matching features on the two documents, then manually writing the information shown on the source onto the check plot. For example, a length of pipe on the source document is typically shown with a diameter, material type, and install date. This information is written in a specific order onto the single-line check plot in a data-input-ready format, with an arrow from the attribute data to the feature. The manually attributed single-line plots are then sent to Palo Alto for review.

7) Translation, attribute entry, and annotation

Once the single-line plots for each discipline have been reviewed by city personnel, they are returned to WLGIS. Graphic revisions are made in Microstation, and then the files are translated into GDS, which is Palo Alto's GIS/CAD software platform. After translation, data entry of the manually written attributes residing on the single-line plots is performed. These attributes are associated with each feature and are stored non-graphically in GDS's built-in attribute data store. Then, wherever text is shown on the original block maps, corresponding digital annotation is added in a semi-automated fashion.

8) Coordinate geometry (COGO) input of parcel lines

Digital parcel lines were a required theme in creating a finished map product that would be accepted by and useful to the engineering and operations personnel. The steering committee decided COGO input of parcel lines could be accomplished effectively in-house. The source documents are records of surveys and subdivision plats. Because the majority of subdivisions could not be tied back to original monumentation, these internally precise groups of parcels are "floating," with no precise coordinates to locate them on the earth's surface. Since Palo Alto will not be budgeting for a thorough land survey for several more years, a compromise solution was to position the "floating" sets of parcels according to a best fit within the stereodigitized curb lines within a ± 0.4 foot horizontal accuracy. This solution satisfies Palo Alto's use requirements for the parcels, which include accurate and legally defensible area calculations, and buffer search notification. The solution also achieves a reasonable correspondence between the COGO parcel lines and those shown on the block maps, enabling the creation of a final map product that satisfies the current block map users throughout the city.

9) Acceptance and maintenance

If the data meet acceptance criteria, Palo Alto accepts the delivery. Attribute data are ported from GDS's hierarchical attribute data repository in the drawing files to a relational data base. It is crucial to have software for data maintenance developed, tested, and loaded on the system at this time, with trained users ready to begin maintaining the digital files. Without software and trained users for maintenance, the substantial investment of money and energy in such a detailed and accurate system will quickly erode in value as the digital data becomes out-of-date.

TYPES OF ERRORS

Some types of errors originate in either the source documents or the conversion process. Other types are introduced exclusively by the conversion. At the very least, a conversion project should avoid introducing additional error, beyond that contained in the source documents, into an organization's information base. The advantages gained by having continuous maps linked to a relational database should not be subverted by degraded data.

Degradation in data quality due to conversion is not a new concept; it has been occurring periodically throughout the development of infrastructure records. Each technical innovation for improving the storage of infrastructure records (e.g. linen to vellum, vellum to Mylar™, Mylar™ to digital) potentially pays a price in terms of data quality.

The authors propose, in general order of severity, six broad error categories: sufficiency, classification, position, attribute value, data structure, and cartographic representation errors. Several subcategories further classify the error types within each main category.

1) Sufficiency errors concern an inappropriate amount of information. They exist when the quantity of map features does not match the quantity of features in the real world. Data should be complete, non redundant, and free of superfluous information.

Completeness errors exist in the source documents, or occur when the source document is not completely scrubbed: graphic or text features are inadvertently omitted from the conversion effort. This is a particularly difficult problem for Palo Alto's conversion because the conversion project grid (shown in Figure 2) does not match the block map grid. The block map grid is laid out in correspondence to the northwest-southeast orientation of Palo Alto's street grid. This means that parts of six or seven block maps typically need to be referenced (scrubbed) to completely transfer information to the four single-line plots (water, gas, wastewater, and stormwater) for each grid rectangle. Clearly, this situation could easily lead to errors of omission.

Superfluous errors are "made up" graphic or attribute data, and can be thought of as the opposite of completeness errors. For example, a block map may not have been manually updated when a service lateral was removed in the field, creating a graphic which does not represent a real feature.

Redundancy errors are repeated data. For example, symbols placed directly one on top of the other, which can occur if a translation program is executed twice with the same target file.

2) Classification errors exist when features or feature attributes are incorrectly classified at one of three levels: theme, feature, or attribute. The primary cause of misclassification is misinterpretation of the source at any stage in the conversion.

Theme classification errors occur when features actually belonging to one system are represented as belonging to another. For example, a water service lateral on the source document might be incorrectly input as a wastewater lateral. Another example the authors experienced was caused by the translation table, which placed all hydrants in the wastewater theme.

Feature classification errors occur in two forms. A feature on the source document can be misinterpreted and classified, for example, as "main text" instead of "lateral text." Alternatively, a name could be used for the feature that does not exactly match a name in the data dictionary.

Attribute classification errors are the assignment of values to the wrong attribute. For example, the assignment of an install-date value to the abandoned-date attribute is an attribute classification error. This type of error occurs due to misinterpretation of source documents or data entry mistakes.

3) Position errors exist when the positions of map features do not reflect their real world positions. Engineering drawings, construction crews, and Underground Service Alert (USA) markings depend on positional accuracy. Incorrect positions can have disastrous consequences, such as accidentally digging up a live gas main. Less serious are the costly construction change orders necessary when field conditions do not match those in a design. In addition to position errors in the source documents, stereodigitizing that exceeds tolerances, incorrect source document registration in the digitizing process, and digitizing that exceeds tolerances are three of many sources of positional errors.

Absolute position errors are those in which mapped features are not located within expected tolerances, or have incorrect geometric relations. For example, a gas valve may be represented on the map 10 feet from its actual field location.

Relative position errors are those where the spatial relationship between map features is incorrect. For example, a water service lateral is mapped to the north of a gas service lateral, when their actual position in the field is reversed. This is also fondly referred to as the "hopscotch" problem. Relative position errors are not uncommon in block map source documents, particularly if features are mapped using different reference points, e.g. wastewater service laterals are measured from the upstream manhole, but water service laterals are measured from the property line. Relative position errors can also be introduced when data from a high positional accuracy source, such as COGO parcel lines, replace data from a moderate relative accuracy source, such as block map parcel lines. In this situation, service laterals can hopscotch property lines.

4) Attribute value errors are those in which annotation text and/or attribute values misrepresent actual field conditions. The source of these errors can be in the source document or any step in the conversion process from stereodigitizing onward, particularly attribute scrubbing. Scrubbing is very challenging for vendor personnel for two reasons. First, they are not experienced at reading engineering records, and may miss the subtleties such as implied association of annotation with features. Secondly, the source documents may be physically very difficult to read. Different lettering styles, small lettering, and the cumulative effect of many erasures and revisions over the years can increase the opportunity for misinterpretation. Attribute value errors can also be caused by scrubbing from the wrong source document, or data entry mistakes.

Wrong value errors are the use of legal, yet incorrect values. For example, ABS is shown and/or stored as the apparent material value, but the actual material is polyethylene. Another example is a misspelling, e.g. "Hamlton" instead of "Hamilton" for a street name.

Value not in list errors are made when an assigned value is not a member of a legal list of values.

Missing mandatory value errors are made when the data dictionary specifies the attribute field to be "not null," and a value was not entered for the field.

Non-unique key value errors are made when the data dictionary specifies the field must contain unique values, and redundant values exist. For example, a theme may contain two manholes with the identification designation "MH856".

5) Data structure errors can be classified into five types, one of which (GDS data type errors) is specific to GDS. All five types concern how data are stored, and how data elements relate to each other.

Spatial data type errors concern the topology of points, lines, and polygons. They exist when graphic elements that are specified to be related in a certain way do not, as shown in Figure 5. Pipes in the network of a gravity system such as wastewater should have a unified direction of flow among adjacent elements. If a pipe segment (P13) is represented as flowing west and is located between two pipe segments (P12 and P14) that are flowing east, then an incorrect relationship exists. Another example is shown in Figure 6. In this case, the patterned tape representing a sewer line came loose on the original Mylar™ block map, and happened to reconnect itself to a storm water system manhole. As a final example, Figure 7 shows line breaks in a pipe crossing a project grid

FIGURE 5

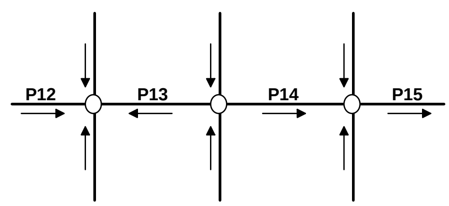
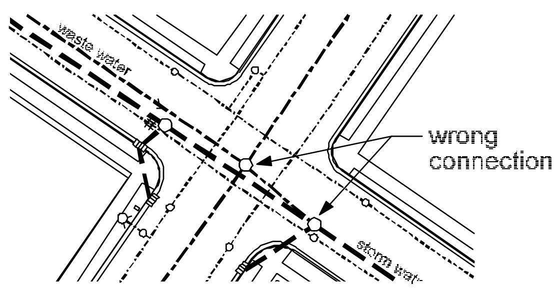
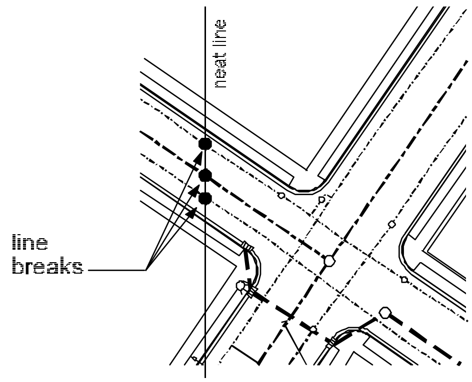


FIGURE 6



neat line which introduces artificial nodes into infrastructure networks. Conversely, a line passing through a manhole should break; a continuous line in this case is an error.

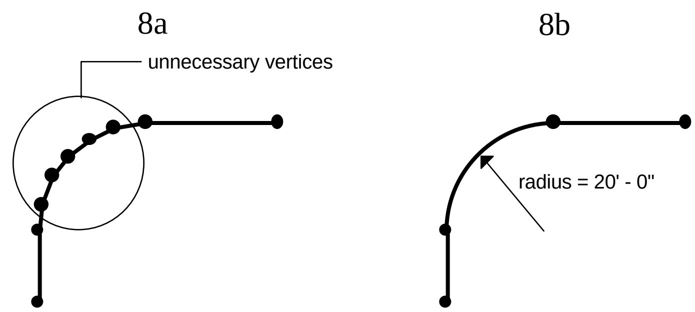
FIGURE 7



GDS data type errors occur when the GDS data type used to represent a linear feature does not match that specified in the data dictionary. (GDS has four data types that can represent a linear feature.) For example, in GDS, features can be represented as *objects*, which include a scale, a rotation, and an origin, or *items*, with no such overhead. A feature which is represented as an object, but is specified as an item in the data dictionary, has an incorrect data structure. Feature representation using the correct data type is important to orderly application development.

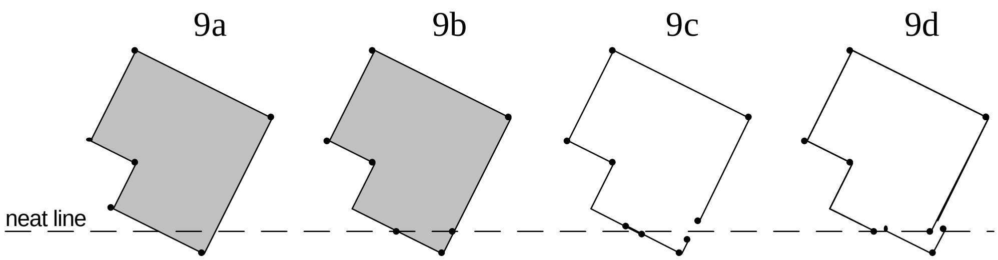
Association errors exist when expected logical grouping links between features are incorrect or missing. An example would be a symbol, e.g. the elevation text of a manhole, that is not associated with that manhole, when the data dictionary specifies that the association should exist. As another example, laterals that lack association with a main cannot be used to "tap" to notification information (owner, address) for the properties they service, should the main need to shut off.

FIGURE 8



Unnecessary vertices are vertices that do not contribute to the shape of a feature. They usually occur as a result of fill points in splines or arcs, and frequently appear in road and cadastral features. When the stereodigitizing or regular digitizing platforms do not support true curves, the only way to capture the geometry is to insert shape points, as shown in Figure 8a. In a system which supports true curves, the extra line segments unnecessarily increase data storage requirements, and should be removed in an editing process to achieve the true curve shown in Figure 8b. Unnecessary vertices often occur at neat lines. The graphics crossing the neat line should be

FIGURE 9



continuous lines, without the vertex, as shown in Figure 9a. Figure 9b shows extra vertices.

Misplaced vertices errors are characterized by untidy linework such as overshoots, undershoots, and offsets, as portrayed in Figure 9c and 9d. On infrastructure themes, unconnected lines such as these cause difficulties in building continuous networks of system components for analysis purposes. Misplaced vertices can distort the shape of a feature. Overshoots can cause extraneous components in the network structure. Such errors are commonly caused by poor edge matching, incorrect use of hit codes, or poor registration of source documents.

Feature justification errors concern symbols and text which do not scale or rotate in a cartographically acceptable fashion. Figures 10 and 11 show correct justification examples.

FIGURE 10

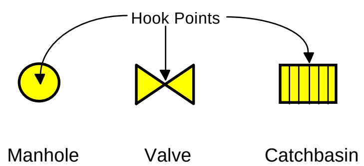
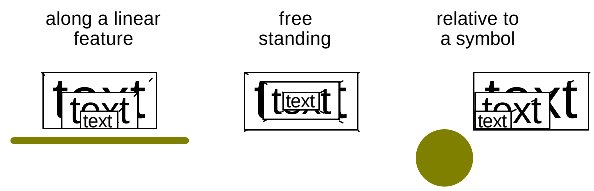


FIGURE 11



6) Cartographic Representation errors are those in which graphics are incorrectly shown on the map, or where graphic standards have not been followed. They can cause a map to be misinterpreted.

Symbol errors are those instances in which the wrong symbol is used. For example, a valve symbol is used instead of reducer, or an earlier version of a symbol is used by a translator instead of the current version.

Style errors are instances in which the wrong character style or line style is used. For example, a line style representing an active water main is used rather than a line style representing an abandoned water main. This kind of error is characterized by a mismatch between the style specified for a feature in the data dictionary and the style of the feature in the delivered product. Incorrect font size and incorrect line thickness also fall into this subcategory.

Annotation placement errors exist when rules of placement are violated, causing difficulty in map legibility and interpretation. Annotation placement rules concern orientation, proximity, alignment, use of leader lines, and order of precedence for feature obstruction, e.g. text should not overlay pipes or curbs.

Congestion errors occur when annotation and features are too dense to read clearly. Annotation blocks may appear to run together.

Scale errors occur when the scale of a symbol does not match that specified in the data dictionary. Scale errors generally are introduced in translation process.

Rotation errors exist if features are not aligned correctly. For example, a catch basin symbol should be rotated to align to a curb.

QUALITY CONTROL METHODOLOGY

The objective of Palo Alto's quality control (QC) process is to identify errors in a timely fashion to ensure a final product that meets the project specifications and intent. Each step in the QC process employs one or more methods of identifying errors. As shown to the right of the conversion steps in Figure 3, it is critical to implement the individual QC steps at the appropriate time in the conversion process, so that errors can be identified and corrected before they adversely affect downstream conversion steps.

Ground Control Evaluation

In order to ensure the positional accuracy of its monumentation, Bestor Engineering performed its own internal quality check by running a five mile level loop traverse which closed to within 0.07 foot. Palo Alto's City Surveyor used standard surveying methods to field check several ground control points for positional errors. He also evaluated the residuals from the least squares adjustments on the control points to ensure that they were well within tolerance specifications.

Aerial triangulation

This technique is used to link all the aerial photography into a unified photogrammetric block of the project area (Montgomery, 1993). First, ground control points are identified on the photographs, and a mathematical modeling tool is used to determine if any of the ground control points fit the model poorly. Ground control points that are incorrectly located on the photographs are identified by this process, so that their positions can be corrected before stereodigitizing and orthophoto production commence.

Prototype evaluation

The purpose of the prototype is to put specifications to the test of implementation on a small subset of the first delivery area, and to flush out any fundamental problems or misunderstandings in the conversion methodology. The prototype evaluation is an opportunity to test quality control methods, and to train users in how to follow the methods. It is an opportunity to understand quality control dependencies so that check methods can be worked in parallel, and moved as far upstream in the conversion process as possible. Prototype evaluation requires several rapid iterations. Palo Alto provides quality control feedback, and in response, WLGIS reworks the existing data and delivers the reworked data plus a new grid cell area. It is particularly important to trace back and find the sources of systematic errors during the prototype evaluation, because fixing them may require adjustments to the conversion process itself. Once the prototype area is satisfactorily converted, authorization is given to the conversion vendor to proceed with the conversion of the rest of the project.

Field check of curb locations

After completing stereodigitizing of a grid area, WLGIS provides "raw" digital graphics including curb lines and other stereodigitized features. Palo Alto personnel use a simple GIS dimensioning application to add street width dimensions, aligning the dimensions with building edges for easy location identification in the field. The curb lines and dimensions are plotted and field checked by the City Surveyor and assistants. Sample measurements with out-of-tolerance discrepancies between the plot and the field measurement are identified, and the information is communicated back to WLGIS. Positional accuracy of curbs at this point in the process is necessary so that the digital curbs lines can be used as a stable basis for registering the block maps on the digitizing table, and for positioning COGO parcel blocks.

Single-line check plot review

After attribute scrubbing, WLGIS sends the single-line check plots to Palo Alto for review. Civil engineering student interns and staff check for attribute value correctness and sufficiency, as well as for feature classification errors, for example, a water lateral on the wastewater theme. Utilities technicians and engineers are available for consultation by the interns. As they review each graphic or attribute datum on the check plot, the corresponding feature is highlighted in yellow on a blueprint copy of the block map, so omissions can later be identified.

Block map sufficiency check

After the single-line check plot review has been completed for a substantial continuous geographic area, the blueprints can be reviewed to identify any non-highlighted features that were omitted by the digitizing and attribute scrubbing conversion steps. These omitted features will be input by Palo Alto personnel once the digital data for the area is accepted.

Automated check at WLGIS and Palo Alto

After WLGIS personnel input the attribute data and text annotation, they execute an automated quality control application given to WLGIS by Palo Alto. This program runs extensive checks on the digital data, comparing them against the data dictionary specifications. The program checks for theme classification, GDS data type, manditory attribute value, illegal attribute value, style, and feature justification errors. The program places error symbols on an overlaid QC red-line drawing and generates an error report listing the location and type of each error encountered. This approach was particularly beneficial during the prototype evaluation to identify systematic errors. This automated check is also performed by Palo Alto upon receipt of a data delivery.

Light table overlay check

When a delivery is received, plots are made corresponding to all the block map extents for the current delivery area. These plots often cross one or more neat lines. The themes on the plots include curb lines, water, gas, wastewater, and storm. The original Mylar™ block map is placed on a light table and overlaid by the corresponding plot. Curb edges are used to register the two documents together, mimicking the registration process used by WLGIS in the digitizing step. The plot is examined for positional errors indicated by lateral location discrepancies between the two documents. In addition, Palo Alto checks its COGO parcel lines by identifying any instances in which the parcel lines in the two documents differ by more than 1/10". Finally, the documents are examined for any "hopscotch" problems caused by the change in parcel line source. These are highlighted for review by senior engineers and/or field checking.

Comparison of single-line check plots to attribute plots

Once Palo Alto receives a delivery, a digital single-line check plot, the "attribute plot," is created for each of the water, gas, wastewater and stormwater themes. The plot is created by a program which generates text annotation containing attribute values, and places the annotation on each graphic feature from which the attributes were derived. The student interns compare the attribute plots to the single-line check plots containing the manually written attributes, to ensure that no data entry mistakes were made. They also check to ensure that all the edits they made to the single-line check plot were faithfully corrected.

FIGURE 12

Sufficiency and source attribute value checks with orthophotos

On the computer screen, the infrastructure vector data is overlaid on the orthophoto, as shown in Figure 12. The orthophoto is examined to ensure that features, particularly valves and manholes, that can be seen on the raster image have been captured stereographically. Sufficiency errors for these features are serious, since the mains are located according to valve and manhole positions. In addition, curbs are examined to check the correct use of "visible" versus "obscured" attribute values.

Computer assisted visual inspection checks and red-lining

In addition to the automated and computer assisted checks already described, several more computer assisted checks help interns and technicians inspect the data and discover errors. They mark up errors on a digital quality control drawing which is overlaid on the on-line view. Plots are made which show the data in black and the markups in red.

The *feature classification check* identifies incorrectly named features. On a theme by theme basis, all of the graphics corresponding to a selected legal name are displayed in yellow, and all other features appear in gray. For example, all manholes are displayed in yellow; everything else is gray If a yellow feature is *not* a manhole, then it has been incorrectly named. This routine is especially helpful for finding incorrectly named annotation features.

The *direction topology* check places arrow heads at the end of a line to indicate the implied direction of flow for gravity systems as shown in Figure 5. Incongruous segments can be identified in this way.

The *closed block check* forces a closed block of line segments to a solid fill line style as shown in Figures 9a and 9b. Non closed block are easy to identify visually, as demonstrated by Figures 9c and 9d. This check is used primarily on building outlines.

The *unnecessary vertices* check places a dot at every vertex so that visually it is easy to identify instances of stream digitizing, as shown in Figure 8a, or extra vertices at neat lines, as shown in Figure 9b.

Visual inspection of digital plots

Finally, an automated routine makes digital plots for each grid area corresponding to the graphic look and content of the original block map source. These plots are examined for cartographic representation errors.

RECOMMENDATIONS

There are a few recommendations the authors can make regarding developing a quality control process for a conversion project.

1. Research the history of the engineering records. One method we found helpful was tapping into the old-timers' network. Jim Abler, a retired Palo Alto engineer, was able to describe the history and development of the block map source documents. This history included three media types, with each media change distorting or degrading the maps in some way. Knowing the maps' history and reliability, and knowing the manual methods used to add and delete graphics, assisted us in proposing sensible process and QC choices.

Be diplomatic by avoiding criticism of the source documents since many of the engineers have a strong emotional investment in them.

2. When deciding on a visual look for the new digital map products, consider maintaining the look of the source document that is currently being used in the organization, unless there is a compelling reason to change. This will aid in early cultural acceptance of the digital data.

3. Understand the consequences of reconciling data from different sources. For example, when COGO parcel lines replace the block map parcel lines in the digital map product, service laterals drawn relative to the less accurate block map parcel lines can hopscotch over into an adjacent parcel. The replacement of parcel lines requires another step in the QC process to identify discrepancies outside a certain tolerance.

4. When it comes time to accept or reject a delivery, the deciding factor will often be whether the remaining errors are easier to fix in-house than to describe the problem to the conversion vendor.

5. Identify and provide authority to a Quality Control champion. Managing the QC process should be this person’s primary responsibility.

CONCLUSIONS

It is apparent from the authors' experience and this discussion that quality control is an integral part of the conversion process throughout its course. In order for an infrastructure conversion project to be successful, quality control planning must be done before the project commences. Do not think of quality control as something that is done only when the final product of the conversion process is delivered. Instead, think of it as an ongoing effort that requires thoughtful planning and adequate budget and schedule resources.

The methodology used for quality control will vary from project to project to suit each project's unique conversion process. The authors hope that the error typology and quality control methods presented here provide useful guidelines to fellow pilgrims on the road to higher data quality.

REFERENCES

American Society of Photogrammetry (ASP), Manual of Remote Sensing, Volumes I and II, ASP, Virginia, 1983.

Bamberger, W. L., T. J. Richard, and J. L. Trowbridge, *"Data Conversion Lessons Learned in the Trenches,"* GIS World, August 1994.

Joffe, B., and D. Matson, *"Better and Cheaper: Technical Innovation Builds Palo Alto's High-Quality Map Base,"* URISA Conference Proceedings, Vol. 1, Milwaukee, WI, August 1994, pages 400 - 410.

Lillesand, T. M., and R. W. Kiefer, Remote Sensing and Image Interpretation, Wiley and Sons, New York, 1979.

Montgomery, G. E., and H. C. Schuch, GIS Conversion Handbook, GIS World Inc., Fort Collins, CO, 1993.

ACKNOWLEDGEMENTS

We thank Glenn Loo and Dave Matson of the City of Palo Alto for their guidance of this effort. Bruce Joffe of GIS Consultants was an early advocate of the need for QC in the project. Becky Tumicki of the City of Palo Alto has been of great assistance in keeping QC alive in its darkest hour. Kevin Crachian, also of Palo Alto, helped with image capture for figures. Jean-Paul Lavoie of Geodesy contributed as an editor extraordinaire. Finally, we thank Jerry Ray and the rest of the folks at WLGIS for their perseverance, and their efforts to help us understand the origin of many of the errors described in this paper: "that which does not destroy us makes us stronger."