Paper UC1708

Title
Out of the Box Pipeline Data Alignment (Almost)

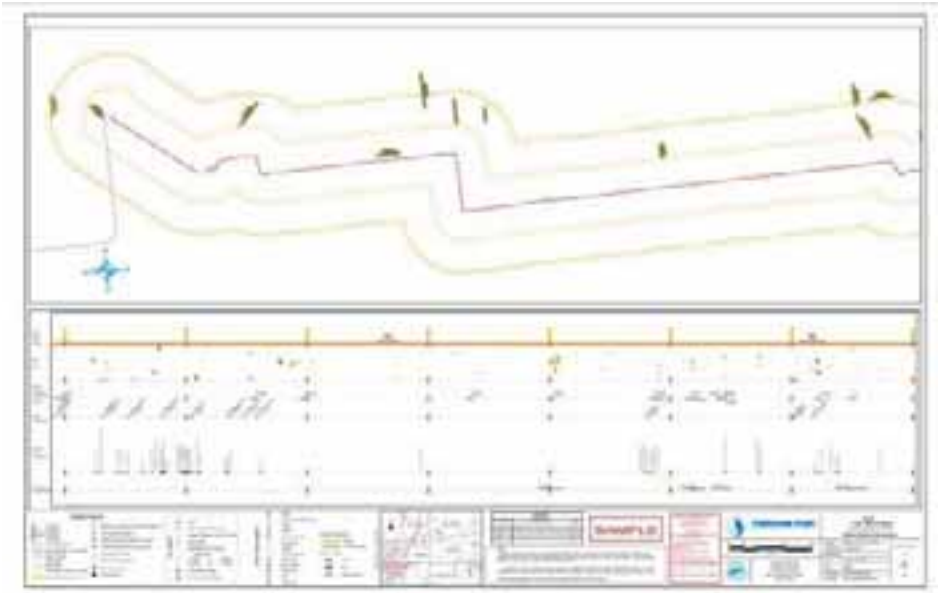
Author
David Nichter

Abstract
Pipeline operators have increasingly used GIS technologies to gather, store, and analyze data. The advent of the Pipeline Safety Act of 2002 requires that all operators gather and integrate data related to their facilities and perform more comprehensive analysis of these factors. Operators must consider data including condition of pipeline, hazards to the pipeline's safety and risks to population and environment as they interact to determine risk. This paper will focus on one operator's process for integrating all available data into a single map product that allows them the ability to examine this combination of datasets in a cohesive manner. Integration of paper records, spreadsheets, CADD drawings, GPS positions, and legacy data with data from a tabular risk modeling application was performed in an ESRI desktop environment. All work was performed with "out of the box" ArcGIS along with a few custom VBA procedures to ease data manipulation.

Body

Transmission Pipelines require specific spatial data, analysis and map layout techniques. Their specific characteristics make them different than distribution utility, natural resources and public GIS applications. While polygon overlays have their use, stationing and linear event processing are much more important. Grid mapping has limited applicability to pipelines. Pipelines and other industries such as electric transmission, communications and rail require linear map products. These specific linear mapping needs have historically been met with a document referred to as an alignment sheet.

Alignment sheets store multiple layers of data in a single hard copy document. A series of alignment sheets are created, each holding a portion of the routes length. Short pipelines may require one sheet while long interstate lines map require dozens of sheets. While only so many layers of data can be represented on a single map, alignment sheets, in addition to a conventional 2-D map window, utilize multiple parallel linear bands, one-dimensional (1-D) routes, to represent events and facilities that occurred or exist along the pipeline. Data in the multiple bands are tied together by stationing values.



Pipeline Stationing is the basis for the GIS concept of measured routes. Stationing allows multiple data sets to be stored along the pipeline centerline without creating additional geometry. This engineering concept is referred to in GIS as “linear referencing.” Stationing, or measure values, allows points to be stored along a route with two pieces of information, the name of the route and the station value. Linear events can be stored along a route with three pieces of information, the route name, the starting station value and the ending station value. Historically, all features and events along the pipeline route were captured in the field based on their station value(s).

Rec_ID	Route_ID	SECTION	FROM_M	TO_M	DIAMETER	WALL_THICK
1414	LineAB	101.01	0	16431	12.750	0.18800
1415	LineAB	102.01	16431	18326	12.750	0.21900
3679	LineAB	102.02	18326	19739	12.750	0.21900
3680	LineAB	102.03	19739	29994	12.750	0.21900
1416	LineAB	103.01	29994	30986	8.625	0.25000
3682	LineAB	105.02	30986	34475	12.750	0.25000
10795	LineAB	105.021	34475	36320	12.750	0.32200
3683	LineAB	105.03	36320	37615	12.750	0.37500
1419	LineAB	106.01	37615	37778	12.750	0.18800

Stationing is typically measured in feet. The beginning of the pipeline route is designated as zero and station values increase along the route toward the end. Stationing in the field is typically captured using a “slack-chain” technique. A flexible tape was laid along the ground and distance marked off as the surveyors “chained” the route. This captures the true 3-D length of pipe. In engineering speak; the start of the stationed route is referred to as “zero plus a pair.” This is written “0+00” which effectively shows “zero hundred feet plus zero feet.” A station of “1+25” represents a point one hundred and twenty five feet along the route. This 3-D, slope distance is always longer than the horizontal distance, unless the ground is horizontal. Capturing this slope distance in a traditional GIS environment poses some special challenges. Stationing is a straight forward way to capture location along a linear route until the pipeline requires a re-route.

Alignment sheets were initially drawn by hand. Drawings were typically performed on Mylar with ink pens. Skilled drafters used geometry and trigonometry to create highly detailed and highly accurate representations of the pipeline and its facilities. They were required to perform precise mathematical calculations based on survey measurements from the field to draw and annotate the ink and Mylar representations of the physical pipeline in the field. With the introduction of multi purpose computers and Computer Aided Drafting and Design (CADD) software the creation and update of alignment sheets became a semi-automated process. Operators who depended on hundreds of manual drafters could reduce their staffs by a factor of ten with the investment in CADD techniques.

Early CADD mapping and CADD based Automated Mapping and Facility Management (AM/FM) systems have limitations. They provide some benefits, typically supporting distribution type networks, but do not provide for all the data handling and spatial analysis requirements of transmission pipeline operators. Engineering and pipeline professionals continued to use other manual or automated tools and techniques to perform analysis based on data stored on the CADD alignment sheets. If the data in the CADD files could be used to perform analysis, then much better decisions could be made. Pipelines, with their multiple coincident datasets, have traditionally used topological rules to govern subsequent feature placement along the centerline.

GIS users benefit from superior data storage models and software capable of performing spatial analysis. Application of GIS techniques by pipeline operators provides the ability to store and use intelligent point, line and polygon data for multiple design, analysis, management and operations applications. GIS software applications provide the tools to create and maintain spatial data, perform analysis of that data and create reports based on the analysis of that intelligent spatial data. In a GIS environment, alignment sheets are reports created from the GIS database.

GIS software has made great advances since its early days but there are still some data and analysis capabilities that remain elusive. The functionality we currently have with ArcGIS 9.X provides some impressive capabilities to the pipeline GIS professional, but some industry specific needs still remain

unfulfilled. Numerous firms provide customized data and software packages to the pipeline industry. This paper focuses on the application of "Out-of-the-Box" ArcGIS functionality to create alignment sheets based on custom GIS data layers. Some custom VBA procedures were developed during this project. It should be noted that this process remains somewhat manual and therefore better suited to operators having smaller transmission pipeline systems.

The pipeline industry is, similar to other sections of the economy, frequently responding to new and challenging regulatory requirements. Interstate pipelines are regulated by both the Federal Department of Energy's (DOE) Federal Energy Regulatory Commission (FERC) and the Department Of Transportation's (DOT) Pipeline Hazardous Materials Safety Administration (PHMSA). Perhaps the regulations having the greatest impacting on the pipeline industry is the Pipeline Safety Act of 2002, A.K.A. The Pipeline GIS Proliferation Act, recorded in 49 CFR §192 Subpart O.

Section 192 subpart O "prescribes minimum requirements for an integrity management program." The subpart describes the process of determining which portion of the operators pipelines are governed by the regulation and what the operator is required to do in response to this determination. Overall, the subpart describes an entire "integrity management program" (IMP) that the operator must undertake to comply with the regulation. This "program must consist, at a minimum, of a framework that describes the process for implementing each program element, how relevant decisions will be made and by whom, a time line for completing the work to implement the program element, and how information gained from experience will be continuously incorporated into the program." The subpart describes a process of continual improvement and enhancements to the IMP.

A major component of subpart O is the requirement for operators to perform "data integration and a risk assessment." Regulations do not specify how the integration of data is to be performed, but more and more operators are turning to GIS to assist with this monumental effort. Data on a myriad of factors related to the construction, operation and maintenance of a transmission pipeline are collected and generated by operators. Much of this data was historically collected either as part of prudent operating procedures or in compliance with older safety regulations. The 2002 act specifically requires that these mountains of data be integrated and analyzed. As more operational and maintenance data is collected, it is integrated into the whole picture to enhance the operators understanding of their pipeline. Risk assessment is used to determine the segments of pipe that must be further assessed.

Risk is the combination of likelihood of failure or threat and consequence of that failure. Threats to a pipeline include factors such as third party damage, external corrosion, and incorrect operations. Consequences of failure include; injury to population, material losses and environmental damage. The proximity of high likelihood of failure and great consequence of that same failure yields the highest risk to the operator. Risk can be calculated in either relativistic or probabilistic manners. With risk scores for the various segments of pipe within their system, operators can plan preventative and mitigative measures for these segments. Riskier pipe can be replaced, inspected, and patrolled more frequently.

So, how does the raw data sitting in the filling cabinets, flat files and file servers of pipeline operators become schedules for more frequent maintenance? The answer provided by Section 192 subpart O is Data Integration. With a capitol "D" and that rhymes with "G" and that stands for Geographic Information Systems.

Linear GIS techniques provide the operator the ability to store multiple layers of data in intelligent data structures along with geoprocessing tools to perform analysis on those data layers. Location of features, events and activities is the key that allows these data to be integrated and analyzed.

The greater the volume of data acquired allows for more sophisticated analytical and display methods. This brings us back to the alignment sheets we mentioned earlier.

The GIS techniques available in ArcGIS provide for proximity and overlay analysis. Measured polylines utilize the language of stationing familiar to pipeliners. Linear events enable the overlay of multiple data layers within a single data base. GIS applications can portray these in one or multiple views. Threats and consequences can be combined based on their spatial extent to feed the algorithms used by risk specialists and engineers to rank their pipe segments. The cartographic output from the GIS disseminates the information to decision makers, contractors and field crews.

The question a GIS'er must ask is "specifically what data is required and where can it be found?" Pipeline integrity requires all nature of base map, facility and activity data which affects the pipeline and its safe operation. The characteristics of the physical pipeline, its related facilities, its related operational and maintenance activities, and the environment around the route are all necessary data for a holistic view required by pipeline integrity management. Much of this is pre-existing within the organization but it must be gathered and spatially enabled. This is no different that any other GIS project. The lion's share of the project is consumed in data collection and automation. As described earlier, the data and processes available to the IMP will evolve overtime as more and more existing data is automated and more data is captured from inspections and other activities.

The approach used during this project utilized some very basic data. Some had spatial character which existed in GPS, CADD and shape files. Some data existed only in hard copy maps and sketches. Some data was automated in spreadsheets, text files and tabular databases. Some data needed to be collected from scratch. The challenge is to make the best use of available data and focus time and resources on developing the best new datasets to maximize the operators understanding of the risk involved with their system. Perhaps the greatest effort is determining the location and attributes of "Identified Sites" as they are defined in subpart O have critical significance for consequence determination.

One critical factor to consider during the collection and data integration for an IMP is the spatial accuracy of the data. The integration of data from multiple sources involves some uncertainty and requires some care in analysis. Allowances must be made for the relative error between the various datasets. The events and features were buffered by various distances to account for the relative error between critical datasets as related to their interaction.

Data was captured for:

- Base Data
  - Elevation
  - Streams, rivers and bodies of water
  - Wetlands
  - Transportation – Streets and railroads
  - Municipal boundaries
  - Orthophotography
- Pipe characteristics
  - Pipe centerline geometry
  - Stationing
  - Diameter, Wall Thickness, Grade, Year Installed
  - Coating type
  - Depth of cover
  - Operational Data – Product, MAOP,
- Other facility data
  - Valves

        - Taps
        - Regulator and Meter Stations Locations
        - Above Ground Markers (AGM)
    - ROW information
        - Occupied Structures and Identified Sites
        - Road/Stream/Foreign Line X-ings
        - Access information
        - Pipe repairs
    - Excavation / Leak Repairs / Maintenance / Inspection History
        - Internal Line Inspection
        - Cathodic Protection (CP) surveys
        - CP Test station readings

Information from previous work was also integrated:

- Pre-segmented risk database
- Manually determined DOT Class ranges
- Manually determined High Consequence areas

The data for this project was organized in a custom data model that reflects both previous GIS work the Operator had undertaken and some components of ArcGIS Pipeline Data Model (APDM) feature classes. Much of the data began as either GPS positions or events placed by stationing along the pipeline route. In much of the system, stationing was either not being maintained or was poorly controlled. Data for facility and activities were maintained in three different datasets. Data was stored on paper by their spatial location or as a point referenced some distance and direction from a road crossing or other known point.

The second source of facility data was a risk assessment database that stored pipeline data in pre-segmented records. The operator had entered data based on earlier review of their paper records. The risk database was the single largest source of automated pipeline data within the organization. The records in this pre-segmented data was attributed with an ascending numbering scheme, line name and pipe length in footage. A challenge involved the handling of duplicate line names among different operating areas and states. Two fields were added to the database to hold the starting and ending station values. With the addition of these “From” - “To” fields and modification of some line names to eliminate duplicates, the risk data could be used for linear referencing.

The operator also maintains an Automated Mapping and Facilities Management (AM/FM) system primarily to address their distribution facilities, but also contains data for their transmission facilities. The AM/FM system uses a stick data model with limited ability to store attribution on pipe or other facilities. Drawbacks of using the AM/FM data directly include the mix of facility representations in hybrid raster/vector and poor spatial accuracy. The process of recreating stationing was a significant challenge.

Stationing values were estimated through two processes; slope stationing estimation and supplementary control points. Slope stationing was estimated by draping the pipeline routes over contour lines. Intersection points and vertices of the routes were used to create slope distances. Cumulative pipe lengths were calculated from the operators’ records. Supplementary control points were created at identifiable X,Y locations with known cumulative pipe footage. This process develops three different versions of route length that had to be reconciled.

Horizontal distance along the 2-D route, slope distance calculated with elevation data, and booked pipe footage should all generally agree with each other. Cases where these three were not reasonable close to each other were subjects of more research. These cases were sometimes the result of

human error in the data handling process but were sometimes the result of data transcription errors years or even decades earlier. Extreme cases required the operator to return to the original sketches and completion reports to verify pipe footages and characteristic. The frequency of these cases was problematic from a project management view. An acceptable error margin of 50 ft between the three legacy data sets was deemed acceptable for pipe lengths over a mile. This is an error factor of less that 1%.

After the supplemental control points were created, the slope station estimates were scaled proportionately according to the supplementary control.

Once an acceptably accurately controlled slope-stationing scheme was applied to the route, data could be located along that stationed route. Facilities and activities, which had been captured via GPS or heads up digitizing from orthophotos, were assigned M values corresponding to their location along the route. Data related to pipe character was entered using the cumulative pipe lengths. Data collected with its own measuring system, such as inline inspections (ILI) and close interval surveys (CIS), were scaled proportionately, or aligned, to the slope controlled route based on available reference markers. This data is often delivered from inspection contractors as spreadsheets or text files. Frequent, regularly located reference markers allow the data to be stationed more accurately. The scaling process assigns measures based on the stationing scheme and aligns the tabular records with the data from other sources.

Some additional data sets were created from the collected data with custom VBA procedures and geoprocessing models. Automation was created to create High Consequence Areas (HCAs) and route profiles.

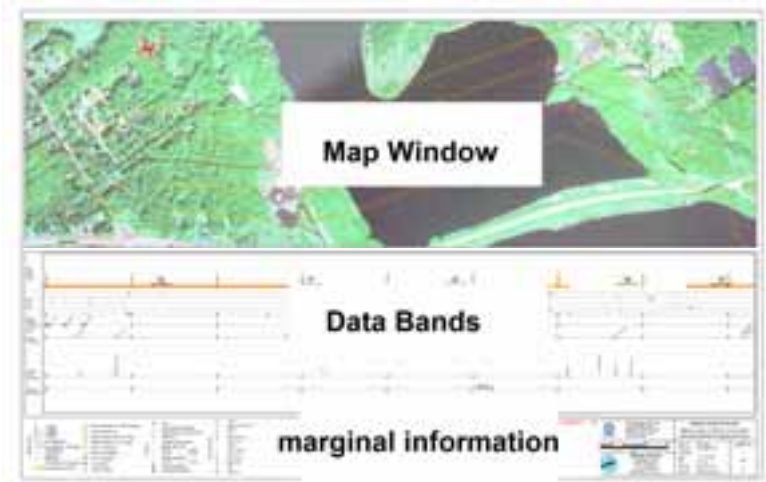
DOT Class and HCA are related but distinct ranges along the pipe centerline. For natural gas pipelines, these are delineated due to the presence of occupied structures and populated areas along the pipeline route.

For this project, occupied structures were represented as point locations. The gathering places along the route were captured as polygons. The previously determined DOT class locations were determined to be accurate. The previous HCA ranges however needed review to ensure compliance. The operator had opted to follow the DOT's option 1 which created HCA's from all DOT Class 3 locations. Additional HCA's had been created manually with the PIC concept. This project examined the use of an automated process to determine non class 3 HCA ranges based on the PIC, the pipeline route and gathering place sites. This eliminated the requirement to consider the 20 or more structures rule for HCA determination. ArcGIS model builder was used to create a tool to perform spatial analysis to create the HCA ranges. The tool buffers the gathering place polygons with the PIC for that section of pipeline. The pipe contained by the buffer is itself then buffered and designated an "HCA." This process used an error margin between 40 and 100 feet to account for spatial position and pipe attribute errors.

Profiles are required to identify low points in the pipeline. Moisture within the pipe can gather in these locations and cause interior corrosion. Another concept related to the pipe profile is critical angle. While not considered as part of this project, critical angle is the minimum pipe slope past which flow within the pipe cannot push liquids up hill. High flow rates could potentially push fluids past some relatively flat low points if their downstream pipe segments were below the critical angle.

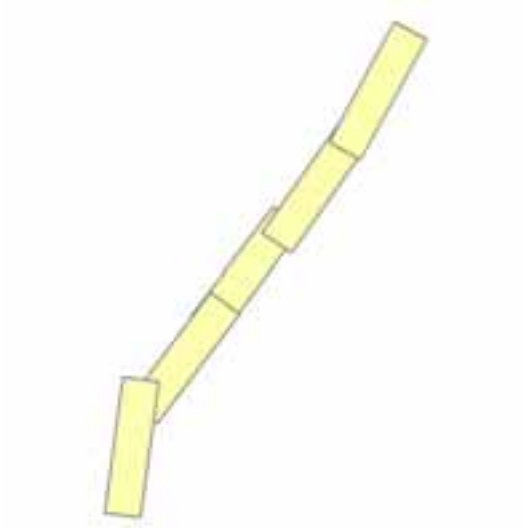
Profiles were generated from data created with the intersection of the pipe route with vector contour lines. The intersection points contained the elevation above mean sea level. The intersection points were assigned station values with the "LocateFeaturesAlongRoutes" tool.

The alignment sheet layout consists of 3 primary components. At the top is the map frame, a 2-D plan view of a section of the pipeline route. Because the rotation of the map frame is not constant, a north arrow is included here to align the user. In the center of the page are the data bands, horizontal parallel lines displaying multiple datasets. The map frame and the data bands were aligned so that features appear relatively the same vertical alignment in the two areas. At the bottom of the alignment sheet are marginal information such as; scale bar, legend, and title block.



The data bands were created with a series of horizontal line segments. Lines were created with a length scaled down 10% from the actual pipeline length to allow for differences in apparent length when the route turns in the map frame. The horizontal (1-D) lines were created as routes with measures equal to the pipeline route. Linear referencing could therefore easily symbolize events on either the 2-D route or the 1-D route.

An additional spatial dataset that had no relationship to the pipeline or its operation are sheet extent polygons. These rectangles were created to cover the pipeline and the 1-D routes from beginning to end. They overlap each other to provide enough information at each seam to allow the user to read along the series from sheet to sheet. The rectangles for the map frame were rotated to show the maximum amount of pipe on each sheet. The rectangles were sized to fit in the data frames at the desired output scale.



For each sheet, the map frame and data band frame was zoomed to the extent of the extent polygons. The Map frame was then rotated to match the rotation of the map frame extent poly. MXD bookmarks were created to store and quickly restore the extents of individual alignment sheets along the pipeline route.

Specialized data such as ILI located defects were displayed with linear referencing along the 1-D route with an attribute driven offset to show their orientation on the pipe. Profile data was displayed with a custom line chart created with the station and elevation values from the contour intersect points.

Usage of the alignment sheet is greatly affected by the choice of symbols' fonts, sizes and colors. User's attention is drawn to bright colors and dramatic symbols. Inappropriately large symbols can cloud the true meaning of the information the alignment sheet was intended to convey. Another factor in symbol size is scale. Using points to represent area features is common at small scales but becomes less meaningful at larger scales. The sizing of symbols at different scales is sometimes a difficult choice.

The initial choice of symbols was an issue during this project. End users of the alignment sheets were concerned with the size of symbols used for ILI anomalies. It was noted that the use of that symbol size could be interpreted to mean that the pipe was covered almost entirely with defects. This was a valid

criticism however the level of precision with which the ILI data could be aligned did require some allowances for potential placement errors. The final product included a marginal note describing the symbol as representing a defect located somewhere within the area covered by the symbol. This "fuzzy" description of a "fuzzy" location is difficult to process mentally, but must be considered to appropriately use this location in conjunction with other spatial datasets with their own levels of spatial accuracy.

Events displayed in the map frame and data bands were labeled and the labels converted to geodatabase annotation. Dynamic labels could not be used due to overstrikes or unattractive, illegibly placed text placements. Annotation was required to ensure that tightly bunched events could be reliably identified.

Draft versions of the alignment sheets were exported to Adobe Portable Document Format (PDF) for review by the engineering staff. The PDFs were reviewed both digitally and printed at full scale on d size sheets for review. Modifications such as symbol type and color along with sheet element layout and map scale were identified. These changes were implemented and the review process was repeated. Final production versions of the alignment sheets were also created as PDFs to allow future reprints without the need to repeat the ArcMap Layout process.

Where does this project go from the concept of an out-of-the-box alignment sheet? Considering the amount of manual work required to set up each sheet in the series, the next step would involve some level of automation. It could automate the zooming, rotating and scaling of data frames based on the geometry and attributes in the sheet extent polygons. This could build on the ESRI mapbook developer sample or could be a fresh design.

Acknowledgments

The author would like to acknowledge his colleagues who collaborated on the project. The author appreciates the knowledge, advice and assistance of Ed Northrop and Tim Driscoll at NFG and Amy Gray at ATSI who made this work and this presentation possible.

References

1. TOMLINSON, ROGER F. - Thinking About GIS, first edition: ESRI Press, 2003.
2. ESRI PIPELINE INTEREST GROUP APDM TECHNICAL COMMITTEE - APDM White Paper, Version 4.0, 2006.
3. Pipeline Safety Improvement Act of 2002. Publ. L. 107-355. 17 Dec. 2002. Stat. 116.2985.

Author Information

David Nichter
Sr. GIS Analyst
ATSI Inc.
415 Commerce Drive
Amherst, NY 14228
US
716-691-9200
dnichter@atsiinc.com