Historical Exploration Data Capture Pilot Project, Northwestern British Columbia (NTS 093L)

C.E. Kilby, Cal Data Ltd., Victoria, BC, ckilby@caldatageol.com

M.A. Fournier, MAF Geographix, Victoria, BC

Kilby, C.E. and Fournier, M.A. (2015): Historical exploration data capture pilot project, northwestern British Columbia (NTS 093L); *in* Geoscience BC Summary of Activities 2014, Geoscience BC, Report 2015-1, p. 79–84.

Introduction

There are over 32 000 mineral assessment reports available to the public in the British Columbia Assessment Report Indexing System (ARIS; BC Geological Survey, 2014a). This likely represents the largest privately funded, public geoscience databank in BC. Every exploration project beginning on newly acquired ground should start with a review of the mineral assessment reports that were written at some time in the past about that ground or the surrounding area. The new project must then present the results of some new work, which might be geological mapping, trenching, sampling or drilling. In many cases the reports will contain the geochemistry results from the analysis of samples.

However, the information provided in these reports remains locked in their analogue format. Therefore, in many cases today, preparation for the field season begins with the creation of a geographic information system (GIS) database collecting together all layers of public (and privately obtained) data. How much of the data from the assessment reports actually make it into these databases remains linked to the budget of the project and the time and abilities of the personnel employed. What if the primary data found in these assessment reports were already available as digital layers that could be loaded into a company's GIS? At the very least it could mean work performed in the past would not be duplicated, and beyond that could spur exploration work that may not have been undertaken without the insight of earlier results.

This pilot project aims to extract and convert primary analogue data from the BC assessment reports (and possibly property files and prospector's reports). Types of data being extracted include analytical chemistry information (e.g., geochemical surveys), drillhole samples, trench samples and grab samples, as well as maps displaying unique geological and geophysical information. Data inclusion in the capture process is dependent on a reasonable spatial component for each sample, and collection proceeds in order from most recent to oldest sources. This initial phase is being undertaken on the single NTS 1:250 000 map area 093L (Figures 1, 2), chosen by Geoscience BC's Minerals Technical Advisory Committee, to provide a proof-of-concept product and establish collection procedures. Techniques and protocols for undertaking this work are being developed to facilitate the continuation of the collection should the proof-of-concept products prove valuable. The converted data will be in a format that can be integrated into a GIS and web mapping systems.

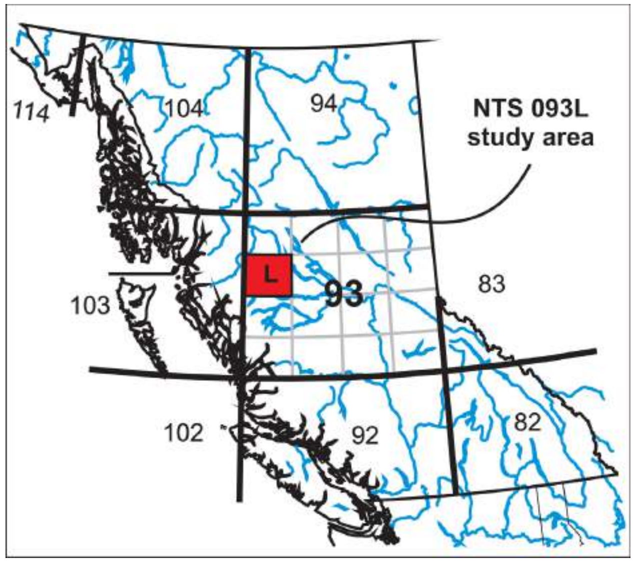

Figure 1. Location of NTS map area 093L, northwestern British Columbia.

What to Capture?

This pilot project is limited in duration as well as scope. There are over 1100 assessment reports in ARIS for over 360 MINFILE sites (BC Geological Survey, 2014b) within NTS 093L. Due to time limitations, only a portion of these reports can be included in this project. Going forward it makes sense to choose random MINFILE sites and then take advantage of common repetition in reports and maps used in consecutive years by companies on work done on a

Keywords: *ARIS, MINFILE, GIS, analogue, digital, geopositioning, geochemical data*

This publication is also available, free of charge, as colour digital files in Adobe Acrobat® PDF format from the Geoscience BC website: http://www.geosciencebc.com/s/DataReleases.asp.

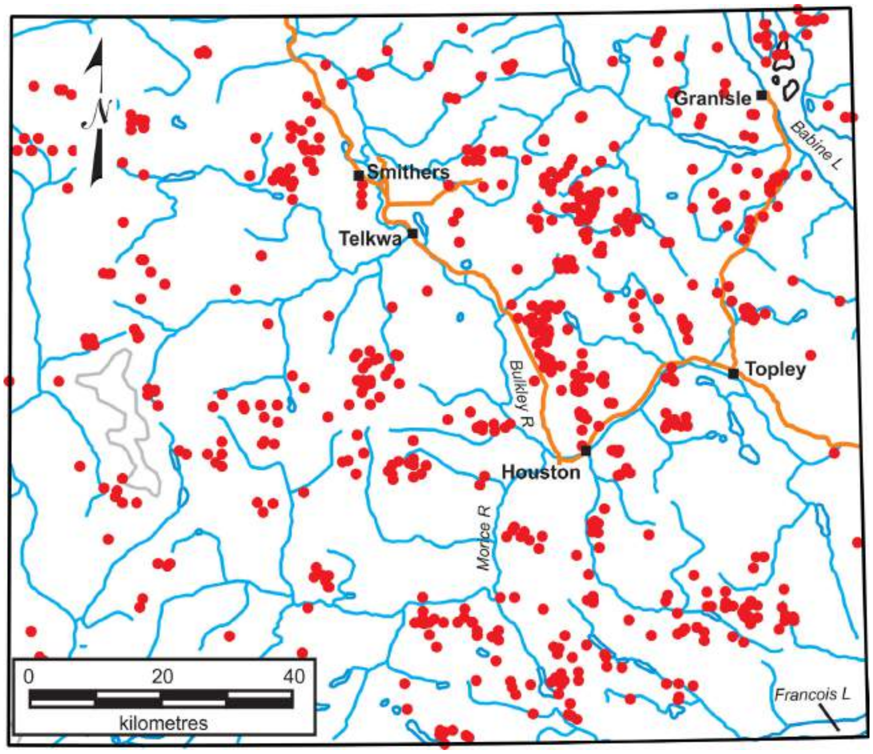

Figure 2. Red dots indicate locations of assessment reports filed in British Columbia's Assessment Reporting Index System (BC Geological Survey, 2014a) for NTS map area 093L, northwestern British Columbia.

single mineral site. Any given MINFILE site may have over 30 ARIS reports reaching back to 1947. To limit the amount of time spent on any one mineral site, only reports submitted in the last 30 years will be used. This has nothing to do with the quality of earlier reports; however, it is possible that more time would be spent geopositioning older maps. What it does allow for is the data capture to be distributed more broadly around various MINFILE sites within NTS 093L and should give a more balanced idea of the time spent to do each task required in this process. MINFILE sites with recorded production are also not included in this pilot project. This is due to the reality that after a mine has been in place some or all of the information from soil samples, trenches or drillholes may be of little value as that material may be mined out. However, the same assessment reports could likely contain data of continuing interest around (or beneath) the site; but assessing whether that is the case would be a task beyond the scope of the time available for this project. Assessment reports where map co-ordinates are incompletely provided or not provided at all, which precludes their accurate positioning, are excluded from this work; a number of reports at this point have already been excluded for this reason.

The goal is to collect primary exploration data from each report, such as soil and silt geochemistry, drillhole, trench and rock sample analyses as well as map displays, such as geology and geophysics. Each sample that is spatially locatable and has associated geochemical information is collected. Each map that can be accurately rectified and provides unique information is collected. In addition, a copy of the original data source (map, assay certificate or report table) is linked to the data as an image or PDF.

A selected report is reviewed and the decision on whether it will be used is based on an assessment of whether the primary information in that report, whatever form it might take, is located on a map that can be geopositioned. The parameters for this decision are described below in the quality control section. Alternatively, locational co-ordinates for all samples may be provided in the report text (such as in an appendix) and would eliminate the need for geopositioning a map and digitizing sample sites. For each report that meets the criteria for locational data, every sample that has associated geochemical information is collected. For a given report, the following material will be collected:

- soil samples with geochemistry,
- silt samples with geochemistry,
- rock samples with geochemistry,
- trench with samples (geochemistry),
- drillholes with samples (geochemistry),
- geology maps with unique work (new, detailed)—this is a geopositioned raster image, and
- geophysical maps (new, property-sized or smaller)—this is a geopositioned raster image.

Geology on maps is not being digitized, but images are being created and geopositioned so that the user can see the geology over a study area. Geology maps tend to be particularly unique in that they usually lean very heavily toward the singular focus of exploration and may contain layers of work by earlier workers. This is a more subjective type of information, more suited to selection (and capture) by the user.

Capture Techniques

Maps and Points

Most of the data is captured from analogue maps and tables, which are currently available in PDF format. These maps have been geopositioned against an accurate base, trying multiple projections and datums to achieve a successful fit. Then the appropriate spatial data is rectified and captured in a standard co-ordinate system (geographic projection) using North American Datum 1983 (NAD 83). Geographic co-ordinates are most easily and accurately converted to other projections by most systems. ArcMap, a GIS software application of Environmental Systems Research Institute, Inc. (Esri), is used for the data capture in this project. The ArcView level of ArcGIS provides point feature class shape files and GeoTIFF raster images as the primary products. The captured spatial data is stored in a database along with its metadata documenting the source and spatial accuracy. The rectified maps are saved in the spatial database and ultimately will be provided in a raster format, such as a GeoTIFF. Commonly, the maps contain useful information in addition to the data being captured from them. The availability of the original maps in a format that can be used in a GIS or displayed through a web mapping system is a useful byproduct of the process. Analogue data, such as tables of analyses, laboratory certificates, drillhole and trench logs, are linked to the appropriate captured data point.

Data capture is currently in progress, and it is likely that more than one accurate base may be tried before a map can be successfully geopositioned. The projection and datum used to produce many maps in assessment reports are not always included in the reports, particularly historical maps that predate the use of personal computers in preparing these reports and figures. At this point TRIM data is used to provide added precision to the co-ordinates on maps rectified to date. Once the map is considered geopositioned with a best reasonable outcome it is assigned a level of accuracy (see the discussion on quality control). The sites of interest to this work, such as grid nodes, individual sample sites, drillhole collars and so forth, are then digitized. Where a sample is from an interval, such as in a trench or drillhole, the two bounding co-ordinates will be recorded relative to an anchor co-ordinate, such as drillhole collar or end of trench. Trench and drillhole survey data (when it exists) can be used to calculate the sample positions.

Geochemical Data

The next step in the process is to capture the geochemical data and link it to each digitized point. The most accurate and cost-effective means to obtain the geochemical data is by contacting the company that completed the assessment report. When a map has been successfully rectified a letter is sent to the company requesting the geochemical data in digital format. Two options are provided for companies; they can provide the lab results from their own digital files, or they can request the lab that did the original work to provide the lab results in digital format. If the geochemical data is provided in this format, it can be added to the database. However, when the data is not provided in a digital format then it must be manually entered into the database. Manual entry is obviously more time consuming and therefore will affect the amount of material project staff are able to complete for this pilot project. Again, for the geochemical components, links to copies of the appropriate pages of the original reports will be provided to allow the user to view the original pages (such as assay certificates).

As this is a pilot project, the results will be evaluated in several ways. As procedures and protocols are created, project staff track their time on the work done. At the end of the project, this will allow the calculation of an average time to geoposition ‘x’ number of maps, digitize ‘x’ number of points, and acquire ‘x’ number of geochemical results. This will help Geoscience BC assess the cost of this work as opposed to the benefit of this work to the exploration community.

Quality Control

The project work is influenced by error from two sources: error in the original creation of the map or other item being captured, and the error that accumulates during the capture process itself. In the first instance it is only possible to know what errors may exist in the original data if possible sources of error are discussed in the assessment report itself. For example, if a sample site is only generally positioned and plotted on a map, as opposed to actually being given a measured location on site, there is no way to know this unless it is mentioned in the text. If that same sample site was located by chaining from a cut grid on the ground, then its location error would not be significant, depending on how the grid had been placed and how much of this adjacent grid had been accurately located. The sample site could be accurately located by today’s GPS equipment and its accuracy would be high, but precision would depend on the device in use. If the same sample site had been located 20 years ago with a handheld GPS the error could be large depending on the dithering, the number of satellites available, the proximity of any features, such as lakes and so forth. In other words, to know the accuracy of the data one uses in an historical assessment report one needs to consider the time and place when that report was written and consider what is discussed in the report. Make no mistake, the locational accuracy in historical reports can be very high. All of this is beyond the control of this project, but directly affects the work itself and must be kept in mind by all users. This project cannot improve the quality of locational data, but an attempt can be made to minimize the error that is added to data locations as they are captured.

Therefore, there is a significant focus in this project on quality control; essentially minimizing cumulative error in this work. There are two main areas of focus in the cumulative error: the first is in the original geopositioning of a map; and the second is during the process of digitizing the actual points associated with data. Initially, a map is brought into the GIS and an attempt is made to position the map using co-ordinates plotted on the map, as well as other unique positional information, such as lakes, streams, roads, buildings and so forth. Once a best fit is selected, an estimate of the possible error must be made. For this purpose, locational confidence categories (Table 1) are used to provide categories of error to simplify this process. At this point, one category, from A through D, is assigned to the positional accuracy of the map based on the likely error, which is estimated by scanning various sites on the map against the base map. Clearly a map in category A is the desired outcome, where locations on the map will be within 5 m of their true (plotted) location. However, this is often not possible, as many maps have introduced error from things as simple as co-ordinates on the map having been placed in error. There are many other possible sources of error. Whatever the error source may be, scale of the work also affects this assessment. However, if a map has been positioned and the error is estimated to be on the order of 100 m or greater anywhere on the map (poorer than category D), then the map is considered to be too inaccurate to make use of the associated data. There is some grey area in category D (50–100 m), depending on the scale of the map and the data associated with the plotted sites, requiring further consideration as to whether to continue with a map. Final assessment of the error margins remains to be made as the work progresses. Again, the category does not address the accuracy of the original map. However, as newer maps may contain data points located by GPS, these inherently more accurate maps (better than pre GPS) should only occur within an A or B category. It is important to note that given more time and a desire on the part of a user to ensure the best accuracy possible, a single map could be geopositioned using detailed airphotos or high resolution remote imagery and >200 points to georeference the map, if it was believed that it was worth the time and expenditure. In this project that type of focus would be unrealistic, and the gain in accuracy would be unknown until the work was complete.

Table 1. Table of locational confidence categories for historical exploration data capture pilot project.

Locational confidence category	Estimate of possible error	Data type
A	0–5 m	borehole collars, geopositioning of maps, grids
B	5–10 m	borehole collars, geopositioning of maps, grids
C	10–50 m	borehole collars (?), geopositioning of maps, grids
D	50–100 m	geopositioning of maps, grids

The creation of Table 1, and breakdown of the categories into a range of metres of possible error, attempts to set parameters for what accuracy one can expect from the data they are using. It should be noted that the table is still a work in progress. There must be a limit to the level of error a user will deem acceptable in the positioning of sample sites. Drillholes are an example of a site where the location of the collar and the potential value of the analyses from the core suggest that very limited error would be acceptable to a user. In the case of drillholes it is suggested that they should have a locational error within 50 m, or should not be digitized for this program. The final 100 m length for D is initially a random even number; however, it is not surprising at times during exploration work to begin using digital data from a particular source only to discover eventually that the sample sites you have been dealing with are 200 m from where they should be. How would the user then regard the remainder of this data? Digitizing a location can lend an air of precision to sites that are placed inaccurately to begin with. By tightening the total error margin allowed, an evaluation needs to be made regarding how many reports will be eliminated from this capture process. Finally, it is common in the georeferencing process that the sample points nearest the control points are most accurately located, those farther away are not. One way to deal with this problem is to georeference the historical maps by using control points that are closest to the data points being digitized. This might involve several separate steps but ultimately it could help reduce the overall error in locating historical data points.

The second area of cumulative error in the digitizing process is in the accuracy achieved in positioning of sites. This error can be large or small, depending on the quality of the map being used. In this project only points are being digitized, so the potential errors will include things such as the thickness of lines, where they are used to place a measured grid, and the size of the circle or other symbol used to indicate a sample site.

In the typical geopositioning procedure, a scanned map from an existing assessment report was exported to TIFF and subsequently geopositioned in ArcMap using the application software's geopositioning procedure as it relates to the position of features in TRIM. Attempts to geoposition the raster image using three different projections (UTM, BC Albers and geographic), as well as two datums (NAD 27 and 83), yielded varied results. Neither latitude nor northing capture was reasonable (Figure 3a), and longitude or easting was off by 100 m in each case (Figure 3b). This discrepancy could be attributed to a datum shift, although both NAD 27 and 83 had the same error. Another possible explanation for the discrepancy (most likely) could be that the location of longitude 127°E may not be accurately positioned on the original image. Overall, the rivers, lakes, roads and topology features in the ARIS raster image are reasonably positioned relative to those in TRIM (Figure 4). Based on an assessment of the overall fit of the geopositioned map, the grid and samples digitized from this map would each be given a locational confidence category of B to C (Table 1) for the estimate of probable locational error, which is stored with each sample in the database.

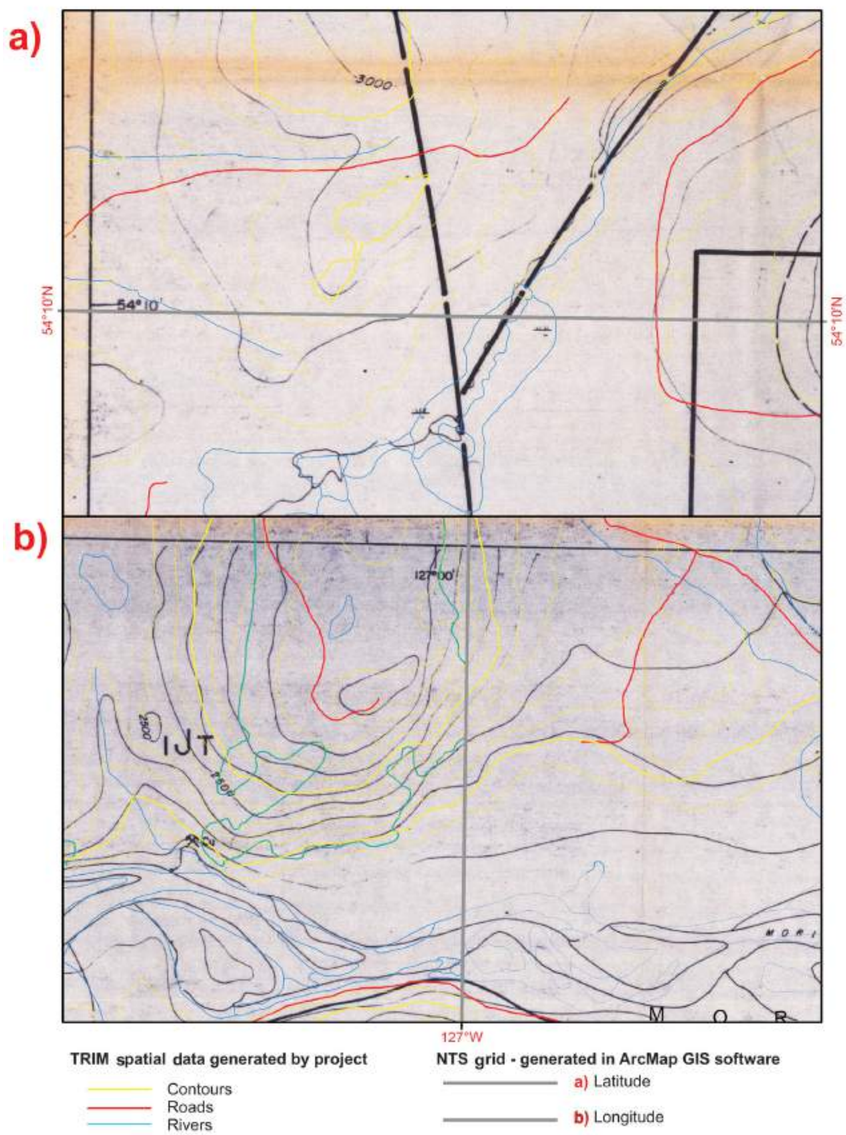

Figure 3. Views of the margins of an exploration map from an assessment report, in UTM Zone 9 projection and NAD 83 datum: **a)** the latitude or northing capture, **b)** the longitude or easting capture. A discrepancy of more than 100 m was observed in the easting for projections using both NAD 27 and 83 datums. Note: all coloured lines and red text are new data added and used to geoposition the map image. All black lines, black text and background map are part of page 36 from Assessment Report 21663 (Zastavnikovich and Bzdel, 1992).

Final Product

Presenting this historical data in a manner that makes it easily available for viewing and utilization by the exploration community is the final step in this process. The digitally captured information will be made available in downloadable and interactive formats. The downloadable option will be similar to the current format used in the Geoscience BC geochemistry releases. In the case of drillhole and trench sample data, co-ordinates for both ends of drillhole and trench data, as well as measured survey points, where available, along the length of either, and associated geochemistry will be provided in a simple, flexible format. Each unique set of data will also have the appropriate metadata attached. Map displays will be in a raster format, such as GeoTIFF. The interactive option will see all the information accessible through a web mapping interface, such as those provided by MapPlace and Geoscience BC.

Summary

Easy access to existing exploration-related information has proven to be a significant incentive in attracting exploration activity to BC. Prospective explorers typically conduct a data search to help them target areas for further investigation. All other things being equal, the jurisdiction with the best, most easily accessible, geological database will attract the most interest. Making existing information readily available for this planning process will attract exploration activity to the province. This pilot project undertakes to convert analogue assessment report primary data to digital format, which will augment and enhance the existing provincial database. This project is a proof-of-concept trial to undertake historical data capture on the single 1:250 000 NTS map area 093L. The project will ultimately develop an un-

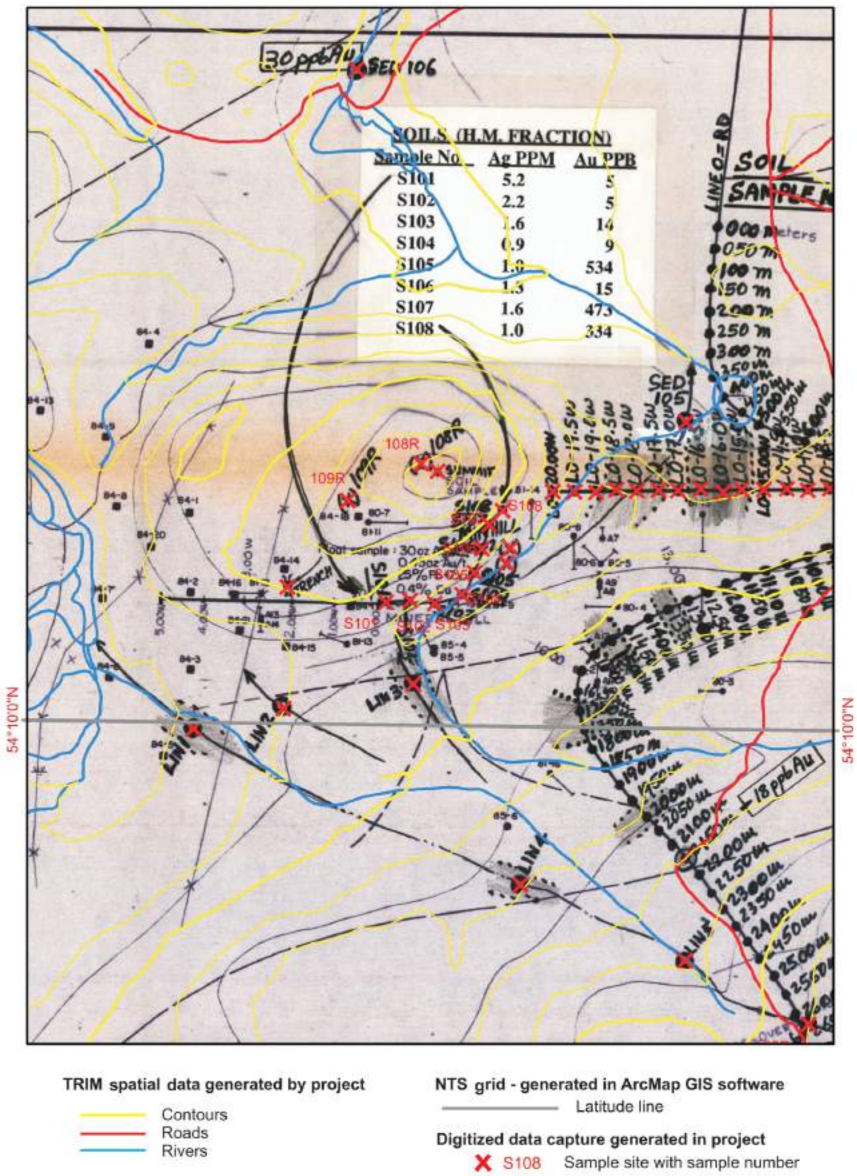

Figure 4. Portion of exploration map with the sample points (red crosses) and TRIM data (coloured lines), such as streams, roads and contours, positioned on the historical map. Note: all coloured lines and red text are new data added and used to geoposition the map image. All black lines, black text and background map are part of page 36 from Assessment Report 21663 (Zastavnikovich and Bzdel, 1992).

derstanding of what resources will be required for this work to be performed on additional areas. It will also provide an opportunity to evaluate the benefit of this type of data resurrection.

Acknowledgments

The authors thank Geoscience BC for funding this pilot project, and the BC Geological Survey, particularly T. Fuller, for help in acquiring data, and Y. Cui and L. Jones for providing insight and knowledge that assisted this project in moving forward. Many thanks to D. MacIntyre who provided a helpful review of the manuscript.

References

BC Geological Survey (2014a): Assessment Report Indexing System (ARIS); BC Ministry of Energy and Mines, BC Geological Survey, URL <http://www.empr.gov.bc.ca/MINING/GEOSCIENCE/ARIS/Pages/default.aspx/> [August 2014].

BC Geological Survey (2014b): MINFILE BC mineral deposits database; BC Ministry of Energy and Mines, BC Geological Survey, URL <http://minfile.ca/> [September 2014].

Zastavnikovich, S. and Bzdel, L.M. (1992): Geochemical and geophysical assessment report on the Fen 1-4 and Tsalit 4-8 mineral claims, Omineca Mining Division, British Columbia; BC Ministry of Energy and Mines, Assessment Report 21663, 35 p., URL <http://aris.empr.gov.bc.ca/search.asp?mode=repsum&rep_no=21663> [December 2014].