Current Research (2000) Newfoundland Department of Mines and Energy
Geological Survey, Report 2000-1, pages 69-74

GeoSymbol: UPGRADE TO ACCOMMODATE MULTIPLE UTM ZONES AND THE CONVERSION BETWEEN NAD 27 AND NAD 83

S.P. Colman-Sadd
Regional Geology Section

ABSTRACT

GeoSymbol *is a database system for managing structural symbols in a Geographic Information System (GIS). It calculates the offsets of symbols from station locations according to a choice of formats, and uses geological criteria to thin stations in order to prevent symbol overlaps. The output is a text file that can be imported directly into a GIS for plotting the symbol layer on a geological map at any desired scale.*

The original version of GeoSymbol *only worked correctly if all the UTM coordinates for input locations were within one UTM zone and referred to one geodetic datum. This upgrade allows the generation of symbol layers that straddle UTM zone boundaries, and permits input locations on a single map to refer to both NAD 27 and NAD 83. The output can use either datum.*

INTRODUCTION

GeoSymbol is a database system for managing structural symbols in a Geographic Information System (GIS). The database rationale and methodology were described previously by Colman-Sadd (1999). The main input data are station locations, expressed in UTM (Universal Transverse Mercator) coordinates, and the types and attitudes of the structures to be plotted.

GeoSymbol allows the user to decide which stations will be plotted on a map and what kinds of structures to include. It also allows three formatting options, which determine whether a symbol plots directly on a station location or is offset from it, and if the latter, how many symbols can be plotted at each station. If requested, *GeoSymbol* automatically reduces the number of stations on a map so as to include only those that can be plotted without overlap at a specified scale. It uses a set of geological criteria to determine which stations convey the least amount of information and are therefore the best to omit from the map. Once it has determined which stations and symbols are to be included, it generates a text file that specifies the UTM coordinates of the plotting location for each symbol on the map. The determination of symbol overlaps and the offset of symbols from their station locations are automatically adjusted for the scale of the map.

GeoSymbol does all spatial calculations in metres, as represented on the ground by the UTM coordinates given for its station locations. The calculations are used to identify symbol overlaps and to determine the plotting offset of symbols from station locations, and are done by coordinate geometry on the UTM grid. This method presupposes that the UTM coordinates for all the stations on a map actually refer to the same UTM grid. This assumption is invalid under two sets of circumstances:

1. Each UTM zone uses its own map projection based on its own central meridian. The UTM grid in a zone has its origin at the equator and the central meridian for that zone. It is different from the grids in zones on either side because these have their origins at different central meridians. Therefore stations on a map that straddle a zone boundary are referenced to two separate grids and it is invalid to use their UTM coordinates to calculate the distance between them.

2. North America is in the process of switching to a new datum, NAD 83 (North American Datum 1983), based on a geodetic model of the Earth derived from satellite and terrestrial data. The previous datum, NAD 27, is based wholly on terrestrial measurements dating back to 1866 (Snyder, 1987). The new datum results in a slight shift of the UTM grid within each zone and, because it also causes a shift in the lines of longitude, a slight shift in the zones themselves. The UTM grids under NAD 83 are, therefore, not coincident with the grids under NAD 27 and the distance between two stations can only be accurately derived if their sets of UTM coordinates both refer to the same datum.

The upgrade to *GeoSymbol* described in this paper adjusts the output of the database to accommodate both situations. The processing required for maps that straddle zone boundaries is entirely transparent to the user. Adjustments

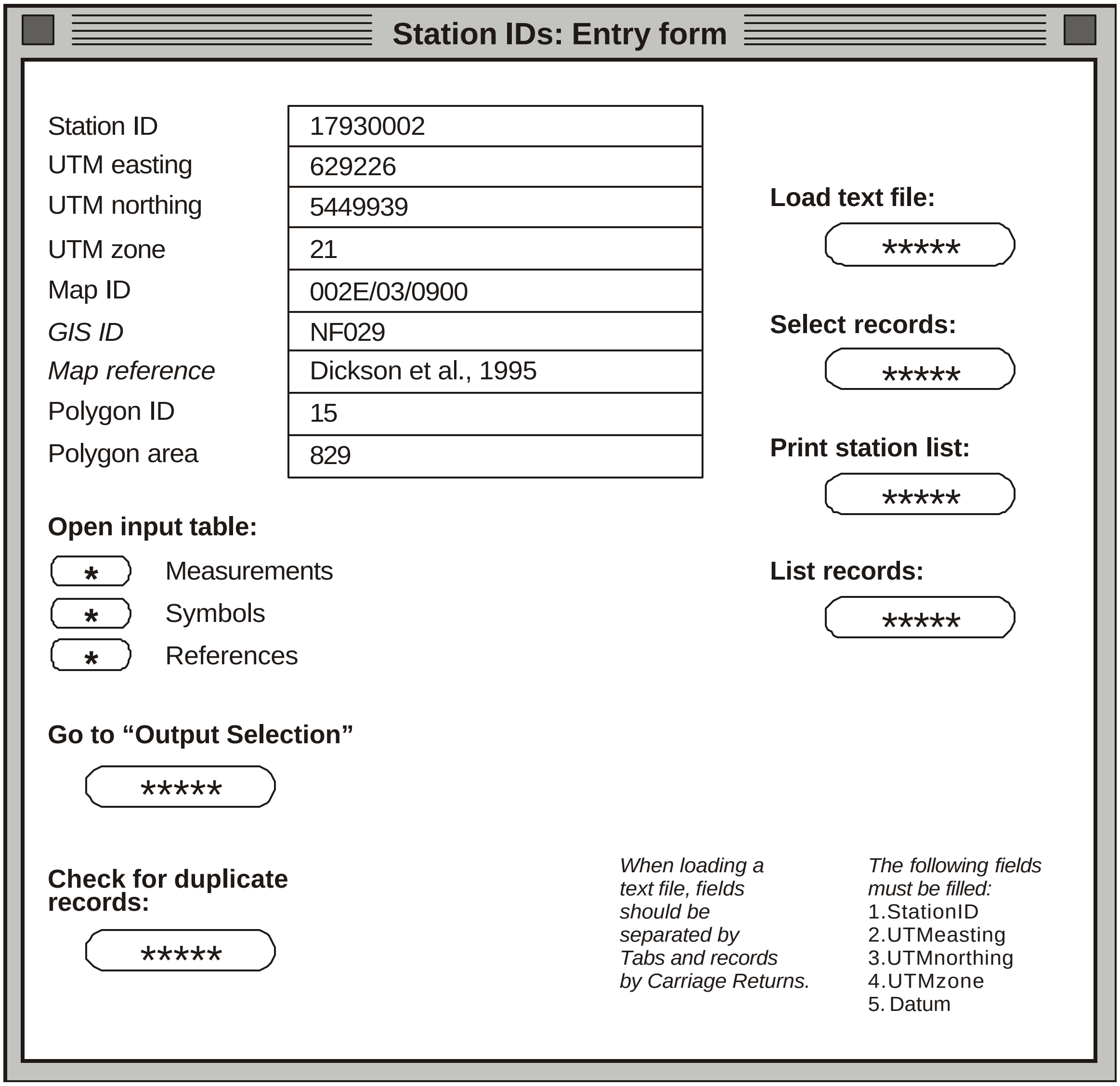

Figure 1. *Entry form for the 'Station IDs' table. The 'Datum' field has been added. Labels in Roman type indicate variables entered directly into the table. Those in italics indicate fields computed from data held in other tables. The solid triangle is a pop-up menu.*

for datum changes require an extra input field, specifying the datum used for the UTM coordinates of each station, and also that the user choose which datum should be used for the UTM coordinates of the output.

CHANGES TO DATA INPUT

STATION IDs TABLE

This table contains station location data. A record is uniquely identified by the 'Station ID' number and the location is given by 'UTM easting', 'UTM northing' and 'UTM zone'. A new variable, 'Datum', has been added after 'UTM zone' to record whether the coordinates refer to NAD 27 or NAD 83 (Figure 1). The value is picked off a pop-up menu on the entry form or is added as an extra column on a text file for batch entry. Records using either datum can be mixed together in the database and also in the subsets of data selected for making a map.

OUTPUT SELECTION FORM

The 'Output Selection' form is used to specify various spatial parameters, geological thinning criteria and other preferences for a map. Two new fields have been added to the form (Figure 2). The first is the 'Output datum', which permits the user to choose either NAD 27 or NAD 83 from a pop-up menu. All UTM coordinates are converted into the chosen datum for the output text file, no matter which datum was used for storing an individual station location in the

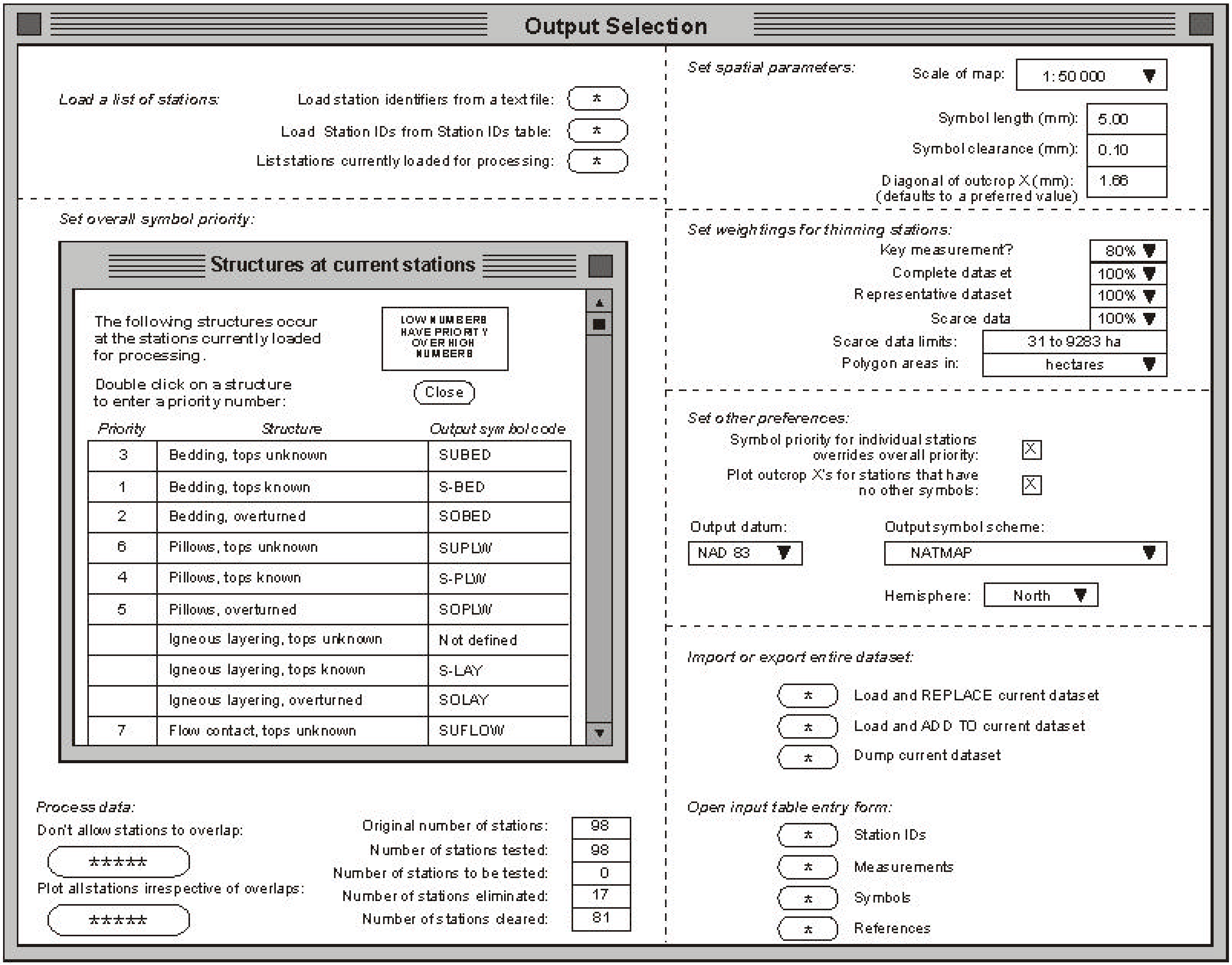

Figure 2. *'Output Selection' form containing all the controls needed to produce a customized symbol layer for a map at any scale. The 'Output datum' and 'Hemisphere' fields have been added, and there are now two choices under 'Process data'.*

'Station IDs' table. The assumption is that the output file will be plotted on a base derived from the corresponding map projection. The second field, 'Hemisphere', specifies either the 'North' or 'South' hemisphere of the Earth. The UTM coordinates in the northern hemisphere repeat in the southern hemisphere so that a location on the Earth's surface is uniquely defined by a combination of easting, northing, zone and hemisphere. Ideally, hemisphere should be defined for each record in the 'Station IDs' table, but because, from a practical viewpoint, this is only necessary for maps that straddle the equator, it is more convenient to specify it globally on the 'Output Selection' form.

Another option that has been added to the 'Output Selection' form is a button to bypass the thinning process and plot all stations selected for the map on the station location regardless of overlaps. This allows easy inspection of the data and is useful if hand selection of symbols is preferred, while at the same time benefiting from the ability to select the types of structures to be plotted and to convert between datums. Two buttons now replace the original 'Start processing' button in the 'Process data' section.

CONVERSION TABLE

A new table has been added to perform all conversions required to process maps that straddle UTM zone boundaries or have a different 'Output datum' from the datum stored for some or all of the selected stations. The table performs three basic conversions:

UTM Coordinates to Degrees

The conversion of UTM coordinates to degrees of latitude and longitude is done using equations 8-17 and 8-18 of Snyder (1987), respectively. Appropriate values for the equatorial radius and eccentricity of the ellipsoid are substituted in the equations, depending on whether the relevant datum is NAD 27 or NAD 83.

Degrees to UTM Coordinates

Degrees are converted to UTM coordinates using equations 8-9 and 8-10 of Snyder (1987) for the eastings and northings respectively, again substituting appropriate constants depending on the datum involved.

NAD 27 to NAD 83

Conversion between datums is performed on location coordinates expressed in degrees of latitude and longitude and is done by simple addition or subtraction of approximate values that vary across the province. The adjustment combines a systematic correction for the changes in the estimate of the equatorial radius and eccentricity of the ellipsoid, and a non-systematic correction for surveying errors in the NAD 27 datum as applied to Newfoundland and Labrador. The combined adjustment is derived by calculating the shift in degrees at the centre of each 1:250 000 map area, using the online demonstration of Natural Resources Canada NTv2 software at http://www.geod.nrcan.gc.ca. The shift in latitude is rounded to five decimal places and that in longitude to four decimal places, giving an accuracy that generally falls within about 5 metres of the NTv2 value for a location and is adequate for plotting a 1:50 000 map. The adjustment values are incorporated into two routines, one for latitude and one for longitude, that can easily be replaced by routines valid for other regions.

REVISED OUTPUT PROCESS

The original output process for *GeoSymbol* is described by Colman-Sadd (1999, pp. 85-93). Revisions to the process occur at two points. First, before processing, the UTM coordinates in the selected area are converted to a common grid and second, after processing, they are converted back to the separate grids required to actually plot the map. This allows processing to be done on a single UTM grid where valid calculations can be made of offsets and distances between stations.

'Station IDs Temp' and 'Conversion' Tables

The stations initially selected for a map are loaded into the 'Station IDs Temp' table, either from a text file or by selecting a subset of the records in the 'Station IDs' table. A list of these stations can be viewed by clicking 'List stations currently loaded for processing' on the 'Output Selection' form (Figure 2). The list always shows the UTM coordinates and datum originally entered in the 'Station IDs' table, even though these values may be converted to others for processing and output. The user then decides which types of structures will be plotted on the map by setting the overall priorities for structures in the 'Structures at current stations' window. Once this is done, all the data required for processing are complete and processing is started by clicking either of the buttons in the 'Process data' section of the form. The first conversion procedure has been inserted at the beginning of the sequence of actions taken by *GeoSymbol* when one of these buttons is clicked.

GeoSymbol runs two checks to select stations that need to undergo conversion and sends these records to the 'Conversion' table:

1. Any record whose datum is different from the 'Output datum', chosen on the 'Output Selection' form, is selected so that its coordinates can be converted to the corresponding values under the 'Output datum'.

2. *GeoSymbol* checks if all stations are in the same UTM zone. If they are not, it calculates the midpoint between the most easterly and the most westerly stations and determines which UTM zone this lies in. This zone, the "midpoint zone", is used for spatial calculations. Any record whose UTM zone is different from the "midpoint zone" is sent to the 'Conversion' table.

Longitude and latitude are calculated for all stations in the 'Conversion' table. Longitude and latitude for stations that have a datum different from the 'Output datum' are adjusted to this datum by addition or subtraction of the appropriate values for their location in the Province. Using the adjusted longitude and latitude, *GeoSymbol* calculates new UTM coordinates for all stations. The calculation (Snyder, 1987, equations 8-9 and 8-10) requires that a central meridian be specified for the map projection. Rather than selecting a different central meridian for stations in each of the conventional UTM zones, *GeoSymbol* calculates the coordinates using the single central meridian for the "midpoint zone". This, in effect, places all the stations into the same UTM zone, following the general procedure used by Kilfoil (1990). The new UTM coordinates are sent back to the 'Station IDs Temp' table where they are substituted for the original location data. All coordinates in the table now refer to a single grid corresponding to the 'Output datum' and the "midpoint zone". The standardized data can now go through the formatting and thinning process necessary to produce an organized symbol layer for a map with no overlaps (Colman-Sadd, 1999).

'Final Output' and 'Conversion' Tables

The 'Final output' table contains a record for every symbol that will be plotted on the map and uses UTM coordinates to specify the plotting positions. As described by Colman-Sadd (1999, pp. 79-80), there may be between one and five records for a station, each with different UTM coordinates, depending on how many symbols are to be plotted at the station.

The records that are sent to the 'Final output' table after processing have UTM coordinates that are standardized to the 'Output datum' and the "midpoint zone". The expectation is that they will be plotted on a map projection that uses the 'Output datum', which is why it would have been selected in the first place, but which also uses the conventional UTM zone system. The coordinates, therefore, have to be converted back so that they plot on the individual grids for the zones that occur within the map area.

The eastings of a UTM zone range approximately from 166018 to 833979 at the equator and from 465003 to 534996 at 84 degrees latitude, with some variation between datums. The narrowing of the range is caused by the convergence of the lines of longitude toward the poles. The UTM system is not applied at latitudes greater than 84 . Only those records whose coordinates lie outside the normal range for a UTM zone actually belong to a zone other than the "midpoint zone" and only these records need to have their coordinates converted. Records are selected for probable conversion if:

$$(E - 168{,}000)\ 31.2021 / (N + 0.01) < 1$$

or

$$(832{,}000 - E)\ 31.2021 / N + 0.01) < 1$$

where E and N are the easting and northing respectively for the record and the limits define a generous margin of safety to allow for the different datums.

The selected records are sent to the 'Conversion' table, where latitude and longitude are calculated from the UTM coordinates. All records at this stage conform to the datum selected on the 'Output Selection' form and so the calculations use the geodetic constants for this datum. The UTM coordinates are then recalculated from the latitude and longitude determinations. This time, however, the calculation uses the central meridians for the UTM zones in which the longitude values fall, instead of the central meridian for the "midpoint zone". The new UTM coordinates and zones are sent back to the 'Final output' table where the relevant records are amended as necessary. The data are then downloaded for plotting in the graphical component of the GIS.

CONCLUSION

The current upgrade allows processing of symbol layers for maps that straddle one or more UTM zones and contain location information referenced to either NAD 27 or NAD 83. The process is completely transparent to the user except for the need to specify the preferred datum for the output and the datum used for the input coordinates of each station.

The requirement for specifying the datum of input coordinates should not just apply to *GeoSymbol*. The datum is also necessary information for any precise geographic location in any database, whether expressed in UTM coordinates or degrees. NAD 27 and NAD 83 are likely to be in use together for several decades. During that time databases will be compiled containing locations referenced to both datums. Although the correct datum may be obvious to the compiler, and perhaps to all the immediate users, it will become less and less obvious as time passes and the source of the data becomes more remote. The problem is more acute because assuming the wrong datum leads to errors in position of only a few hundred metres (in Newfoundland and Labrador). The errors are too small to be obvious and easily detected. At the same time, they are big enough to be significant under most circumstances and to seriously degrade the quality and usefulness of a database.

ACKNOWLEDGMENTS

I thank Gerry Kilfoil and Larry Nolan for providing advice and, in particular for suggesting the invaluable publication by J.P. Snyder. Gerry Kilfoil also gave me a copy of his computer program for conversion between UTM and geographic coordinates. This contains the general concept for transforming UTM coordinates to a single zone, but I have used Snyder's equations for the actual programming. Gerry Kilfoil is thanked for critically reading the manuscript.

REFERENCES

Colman-Sadd, S.P.
1999: *GeoSymbol:* a database system for managing structural symbols in a Geographic Information System. *In* Current Research. Newfoundland Department of Mines and Energy, Geological Survey, Report 99-1, pages 71-94.

Kilfoil, G.J.
1990: Program UTM2.BAS. Newfoundland Department of Mines and Energy, Geological Survey Branch, unpublished computer program.

Snyder, J.P.
1987: Map projections - a working manual. United States Geological Survey, Professional Paper 1395, 383 pages.