Part 2: **Issues and Problems Relating to Cartographic Data Use, Exchange and Transfer**

A DEMONSTRATION TRANSFER OF REMOTELY SENSED DATA UTILIZING THE STANDARD FORMAT FOR THE TRANSFER OF GEOCODED POLYGON DATA

KENNETH D. KORPORAL, *Department of Geography, Carleton University, Ottawa, Canada*

INTRODUCTION

THE IMPETUS for attempting a transfer of remotely sensed data was spawned during the course of a graduate research tutorial which was examining relationships between computer cartography and remote sensing as they existed in 1982 (Korporal and Manore, 1982). It was known that the Standard Data Transfer Committee (Goodenough et al, 1979) had developed a format which was accepted by all member agencies as well as other federal and provincial organizations. Although the Standard Data Transfer Format (SDTF) was accepted as a standard, no agency to our knowledge, had successfully completed a data exchange using the SDTF. During the course of this graduate tutorial program it was realized that remotely sensed data in the SDTF could be used as an input into GIMMS (Waugh, 1982), a geographic information and mapping system.

GIMMS is a very large, user-friendly, integrated geographic information mapping and manipulation system. The program has been developed over a number of years and presently consists of a number of inter-connected applications including mapping, graphics, data manipulation, and interactive design components. The mapping and data manipulation applications are perhaps the best developed of the system and are those which have the most immediate relevance to the data transfer. Being device independent, GIMMS can produce medium to high quality output on a number of line drawing devices, utilizing black and white or colour shading patterns, point symbols, dot patterns, and/or solid colour if available.

In the context of the tutorial project, an experimental transfer was designed utilizing the facilities at the Canada Centre for Remote Sensing (CCRS), Department of Energy, Mines and Resources, and the Geocartographics Subdivision (GCG) of Statistics Canada. Both agencies were very supportive of the project, donating computer and other machine time, program support, and valuable assistance. Following the first set of experimental transfers, improvements were made both to the programs residing on the GCG VAX 11/780 and to those resident at CCRS, thus providing an opportunity to test new and larger data files. This paper will describe the first transfer attempts, subsequent developments and demonstrations, and examine planned and potential future enhancements for an automated transfer and geographic information/mapping system.

THE FIRST TRANSFERS

The first transfer attempts successfully achieved the objective of transferring remotely sensed data from one agency to another and subsequently inputting the data into a geographic information and mapping system. In order to meet the primary objective of mapping remotely sensed data within the strict time-frames imposed in the academic environment, program development had to be kept to a minimum. By taking advantage of utility programs resident on DEC computers (PDP 11/70 at CCRS, VAX 11/780 at GCG) the first series of transfers required only minimal program development for data receipt and display. Program development was required to convert the data from mixed ASCII and BINARY to ASCII files formatted for input into our mapping system, GIMMS. Many problems were encountered at all phases of the process but most were rectified as they were discovered. The most serious problem existed in the data output from the raster to vector conversions at CCRS where it was found that topological errors were created which prevented the formation of closed polygons as required for thematic map production. Nonetheless, the problems were resolved sufficiently to permit a simple thematic map to be produced and fulfill the primary course objective (Figure 1). Subsequent developments and demonstrations have differed largely in one respect: the SDT format has been used instead of the DEC dependent formats of the first transfers.

STANDARD DATA TRANSFER FORMAT

The Standard Data Transfer Format (SDTF) uses polygon chains as the basic structure around which the format was developed. The SDTF also contains general geographic, spatial and descriptive information while conforming to standards developed for geocoded Computer Compatible Tape (CCT). Design of the SDTF incorporates several important properties essential for present and future applications including machine and language independence, generality, expandability, and, self-definition (Goodenough et al, 1979).

At present, spatial data in the SDTF are composed of three files: the Volume Directory File, the Master Header File and the Geographic Detail File which includes the detailed polygon chain information. For demonstration transfer purposes this third file is the most important because in addition to the actual X and Y coordinates, this file contains the chain numbers, left and right polygon descriptors, left and right polygon numbers, number of points in the chain, and left and right polygon areas. This is the essential geocoded information required for input into GIMMS.

PROCESSING AT CCRS

The remotely sensed data used for SDTF transfer purposes require mimimal special processing other than the final formatting and transfer to tape. Although

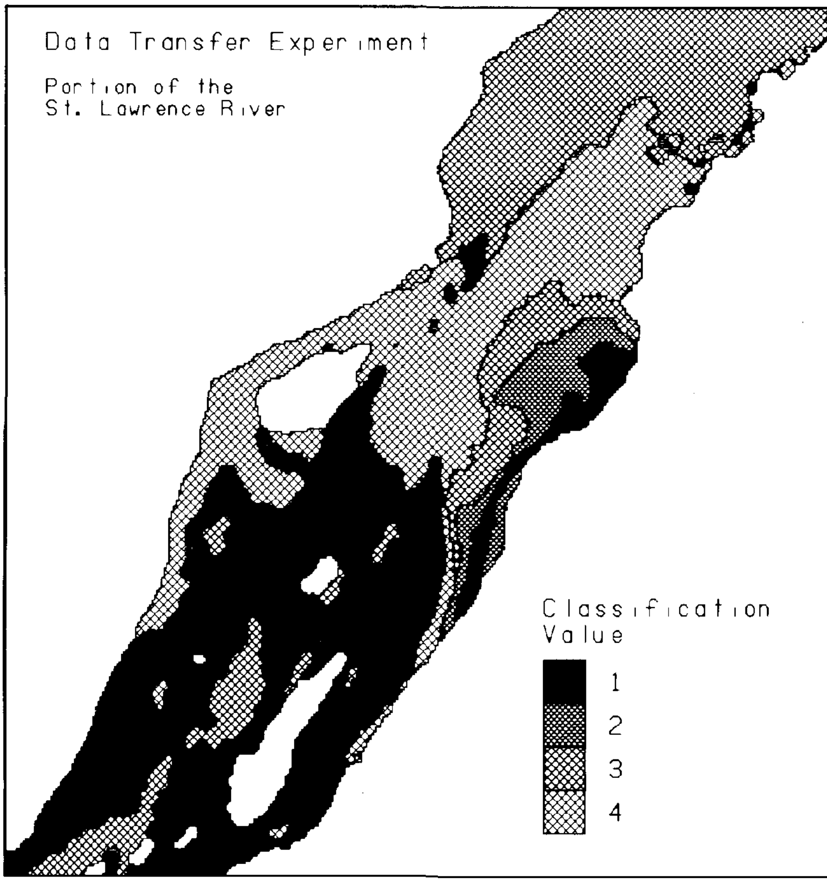

FIGURE 1. *One of the first thematic maps produced from transfer data.*

any classified digital data can be used, LANDSAT Multispectral Scanner (MSS) data were used for most of these demonstration transfers. Standard image processing and classification procedures are utilized by CCRS for creating a thematic raster image.

For demonstration transfer purposes, some theme manipulation was required to remove single or small groups of scattered pixels causing 'speckle' in the image classifications. Two reasons for this are a cartographic generalization for thematic mapping, and b the need to avoid excessive data storage and computing requirements. Generalization of thematic data into contiguous classes may improve map readability and perhaps better represent real conditions. If, however, patterns produced by 'spurious' pixels are important, then such theme manipulation may not be beneficial.

After completion of the theme manipulation and processing, the thematic data are vectorized using raster to polygon software implemented for SDTF purposes.

The final step at CCRS is writing the vectorized data to tape using software which converts the data to the SDTF description. A more detailed outline of the processes and the SDTF software is given in a companion paper (Goodenough et al, 1983).

PROCESSING AT STATISTICS CANADA

Before processing of the remotely sensed data at Statistics Canada could be undertaken, some software development was required to read and format the data from the SDTF tape. A VAX 11/780 mini-computer is the base for the software, including the thematic mapping program and the new software written for this demonstration transfer.

The GCG processing can be divided into four major operations:

a reading of the SDTF tape;
b converting and formatting the digital files;
c line and polygon file creation;
d thematic mapping and graphics.

Figure 2 illustrates these major processing steps at Statistics Canada.

SDTF Tape Reading

Reading of the tape is a relatively straightforward matter using standard system VAX foreign tape mount procedures. Normally, the tape will have three files giving volume, header and segment information. Details essential for accurate reading such as record lengths, number of records, and some geographic information including projection, scale, etc. are on the first two files. A program called SDTREAD was developed to convert the files from mixed ASCII and BINARY to ASCII files. The third file, Geographic Detail File, is processed separately because it relies on information contained in the first two files.

File Conversion and Formatting

Reading, converting to ASCII, and formatting of the Geographic Detail File presented the most difficult and time consuming developments for the transfer. The program created to do this task is called CONVERT. Much of the difficulty in defining the program was due to the uncertainties of the quality of each step in the transfer process. For example, when spurious lines would appear in a file which was in the thematic mapping stage, it was difficult to determine if the problem was related to GIMMS, to our conversion program CONVERT, or to incorrect information on the SDTF tape data files. Therefore, considerable time was spent in tracing problems to the source.

CONVERT, in its present form, creates two output files. The first is a formatted segment file for GIMMS input with segment labels, coordinate pairs and unique GIMMS command identifiers. The second output file is a formatted and sorted data file of classification values and polygon labels, also in GIMMS input format specifications. The first of the above files is used to create a line segment file then a closed polygon file, while the second file is used only at the thematic mapping phase.

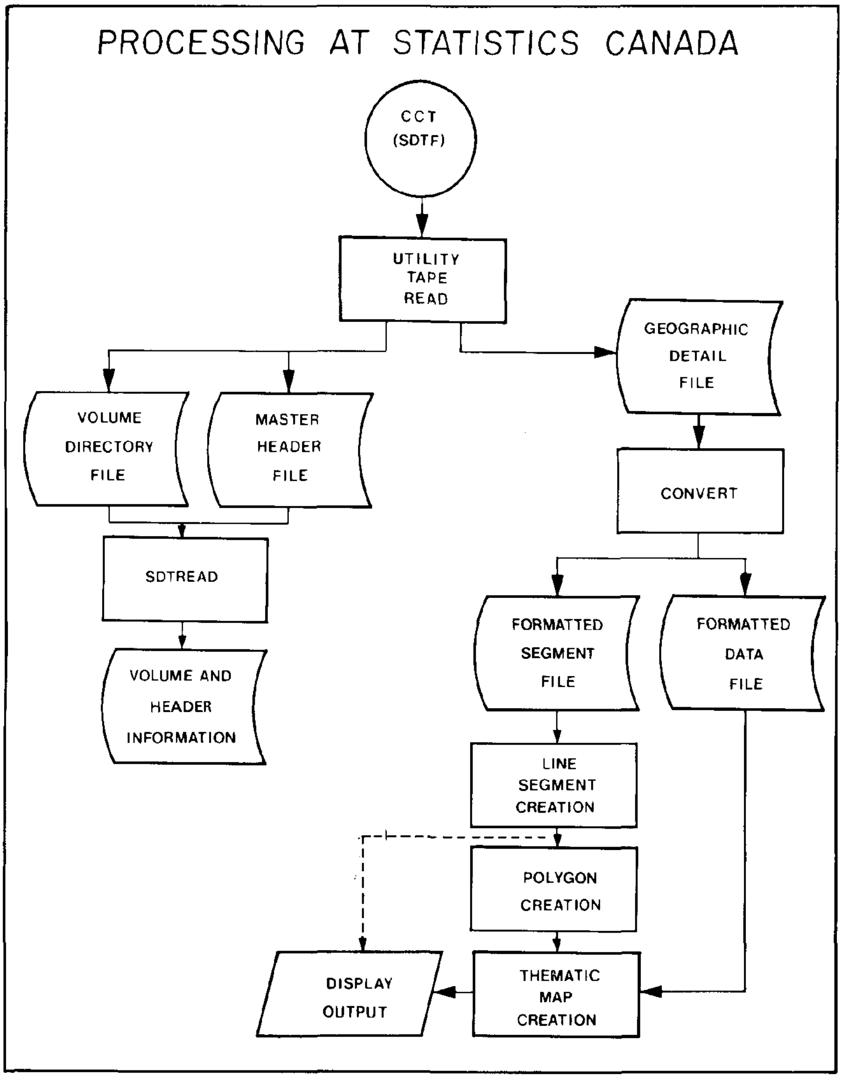

FIGURE 2. *Processing at Statistics Canada.*

Line and Polygon File Creation

Upon completion of CONVERT processing, the segment data file is ready for GIMMS processing to create two more output files and, later, merge with the data classification file to produce thematic maps.

Using the CONVERT generated segment data file, a segment line file is derived which, when drawn (Figure 3), maps the segments as encountered in the raw data file. In the context of thematic mapping, the main purpose of this file is as the input for creation of a polygon descriptor file. Polygon or area type files differ

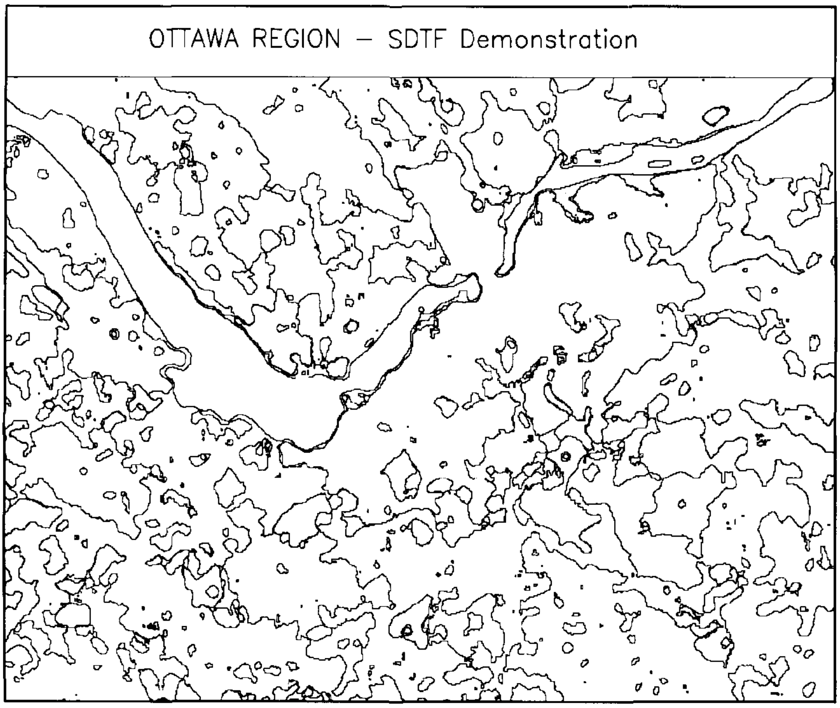

FIGURE 3. *Plot of a* GIMMS *segment file. Scale is 1:250,000.*

from the segment file in that a complete 'closed' boundary description is required where the start and end points for each polygon are identical. The polygon file is required for most thematic mapping purposes where shading or dot patterns are wanted. The very large size of the data files generated from remotely sensed data, combined with the many small areas described by small groups of pixels, puts great demands on GIMMS whose usual mapping application is for much larger areas such as those of political or census type boundaries. GIMMS has the capability to allow the easy upgrading of 'normal' program limits for unusual or unique applications such as those encountered in this demonstration transfer.

An important feature of GIMMS is that a check can be made on the segment data to test for errors such as polygon closure, at much lower processing costs than required for creating the polygon data files. This check is recommended if errors are anticipated in the segment file description. In addition, a generalization function can be introduced into the polygon file creation process to reduce the number of points and generalize the boundaries, thus reducing the processing costs associated with the thematic mapping component.

Thematic Mapping

The GIMMS mapping module is capable of producing medium to high quality maps with supplemental graphics at virtually any scale or size, limited mainly by the capabilities of the output device. Cartographic enhancements are an integral

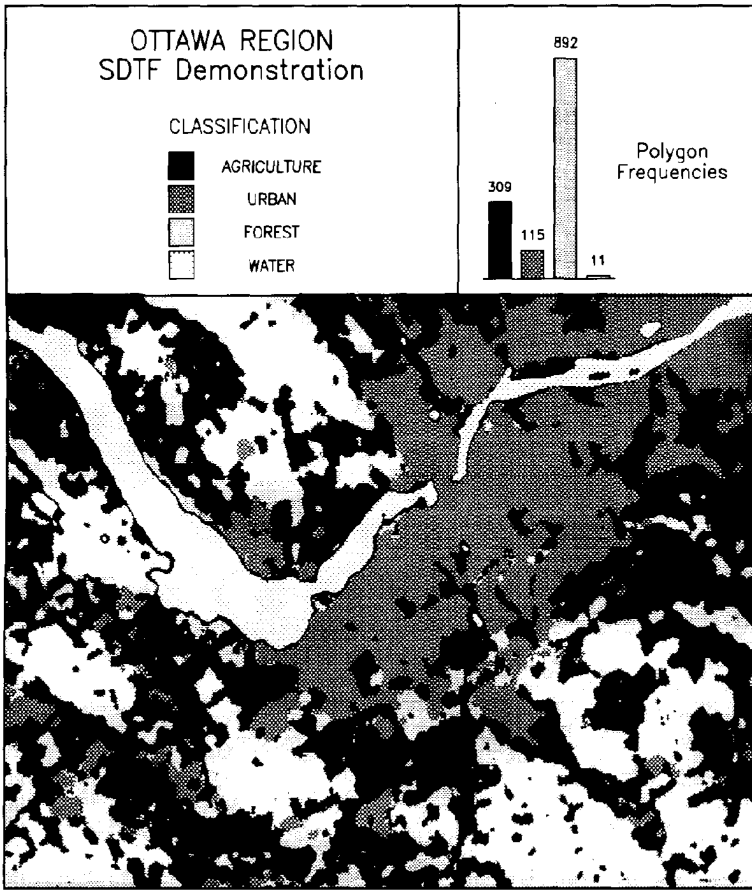

FIGURE 4. *GIMMS thematic map and frequency histogram produced from CONVERT formatted remotely sensed data in SDTF.*

part of the program enabling the addition of titles and other textual features, legends, north arrows, shielding, overlays, offsets, many text fonts, multi-map sheets, etc. User specified black and white or colour shading and dot patterns are used for the effective presentation of data.

The classification values contained in the classified data file, the second file output from CONVERT, are required at the thematic mapping phase. This is the thematic data which will be represented by the selected symbolism and on which the graphics and user specified manipulations are based. An example of a GIMMS thematic map produced from classified digital data is illustrated in Figure 4.

FUTURE DIRECTIONS

As a demonstration system, the developments at GCG are adequate and suitable for the intended purpose: to demonstrate the transferability of data using the

SDTF description. A number of system enhancements have been identified and are presently in various stages of development.

System Automation

An automated production system is being designed to permit the transfer of any data in SDTF description for use in GIMMS. The system will be 'fronted' by a user-friendly module which will guide the user through the relevant steps: tape read, SDTREAD, CONVERT, segment creation and polygon creation. However, the user will still require a knowledge of GIMMS to draw the polygon files or to produce thematic maps.

Two-Way Transfers

To be a truly useful system there is a requirement for two-way data transfers. A number of steps are necessary, including:

1 conversion of polygon data to segment data;
2 conversion of segment data to SDTF description;
3 conversion of SDTF data to raster image.

Each of these components is available in either prototype or production systems in the two departments. PLUSX (Goodchild et al, 1982) can convert polygon data to a segment data format. Features within GIMMS provide the basis for conversion of segment data to SDTF based files. Programs have already been developed at CCRS for reading SDTF files and converting them to raster format. This task, therefore, is one mainly of moulding various independent programs into a modified format to create an automated system.

CONCLUSION

The transfer and mapping of digital remotely sensed data has demonstrated that the Standard Data Transfer Format is a viable working format. The very large file sizes inherent in digital data, such as classified digital images, provide severe testing of system abilities and limitations. While much development is required to create an automated production type system, this demonstration has shown that the advances to date are significant and that upgrading and improving on these developments could be implemented at other agencies. With the ability to quickly and easily exchange data between CCRS and Statistics Canada, and others adopting the SDTF, practical applications will become feasible without the acquisition of expensive and unique hardware.

ACKNOWLEDGMENTS

The author wishes to acknowledge the important contribution of Michael J. Manore as initial collaborator involved in the conception and the first set of experimental transfers. Special thanks to Dr. Gordon F. Deecker, Statistics Canada, for his programming expertise and valuable assistance. Thanks also to C. Shadbolt, M. Lemay and L. Brunet for cartographic and compilation assistance. And finally, to the many persons at CCRS who have been involved, including Dr. D.G. Goodenough, J. Palimaka, G. Plunket and K. Dickinson.

REFERENCES

GOODCHILD, M.F. 1982. *PLUSX Documentation*. Department of Geography, University of Western Ontario.

GOODENOUGH, D.G., K.J. O'NEILL, L.A. GORDON, J. YAN, T. FISHER, C.L. MACDONALD, and A. DESROCHERS, 1979. Standard format for the transfer of geocoded information in spatial data polygon files. *CCRS Research Report*, Ottawa.

GOODENOUGH, D.G., G.W. PLUNKETT, and J.J. PALIMAKA, 1983. On the transfer of remote sensing classifications into polygon geocoded data bases in Canada. *Auto-Carto VI Proceedings*, Ottawa.

KORPORAL, K.D. and MANORE, M.J. 1982. An experimental transfer of remotely sensed data. Unpublished graduate tutorial, Department of Geography, Carleton University, Ottawa.

WAUGH, T. 1982. *GIMMS reference manual*. GIMMS LTD., Edinburgh, Scotland.