Lake Champlain
Basin Program

PUBLICATION SERIES

Technical Report No. 3B

Handbook of GIS Standards and Procedures for the Lake Champlain Basin Program

Prepared by
Vermont Center for Geographic Information, Inc.

for
Lake Champlain Management Conference

March 1993

THIS PROGRAM IS SPONSORED BY U.S.E.P.A. AND THE STATES OF NEW YORK AND VERMONT.

This technical report is the third in a series of reports prepared under the Lake Champlain Basin Program. Those in print are listed below.

Lake Champlain Basin Program Technical Reports

1. *A Research and Monitoring Agenda for Lake Champlain.* Proceedings of a Workshop, December 17-19, 1991, Burlington, VT. Lake Champlain Research Consortium. May, 1992.

2. *Design and Initial Implementation of a Comprehensive Agricultural Monitoring and Evaluation Network for the Lake Champlain Basin.* NY-VT Strategic Core Group. February, 1993.

3. (A) *GIS Management Plan for the Lake Champlain Basin Program.* Vermont Center for Geographic Information, Inc., and Associates in Rural Development. March, 1993.

   (B) *Handbook of GIS Standards and Procedures for the Lake Champlain Basin Program.* Vermont Center for Geographic Information, Inc. March, 1993.

   (C) *GIS Data Inventory for the Lake Champlain Basin Program.* Vermont Center for Geographic Information, Inc. March, 1993.

This report was funded and prepared under the authority of the Lake Champlain Special Designation Act of 1990, P.L. 101-596, through the U.S. Environmental Protection Agency (EPA grant #EPA X 001840-01). Publication of this report does not signify that the contents necessarily reflect the views of the States of New York and Vermont, the Lake Champlain Basin Program, or the U.S. Environmental Protection Agency.

March 30, 1993

Lake Champlain Basin Program

Handbook of GIS Standards and Procedures

PART 1 - STANDARDS

PART 2 - PROCEDURES

APPENDICES

November, 1993

PART 1 : STANDARDS - SECTION A

DIGITAL DATA CONVERSION

I. PURPOSE

The purpose of this document is to provide guidelines covering the digital conversion of existing manually drawn data into GIS format. The goal is to obtain the best possible digital map representing an original analog map. As such, it is assumed that a suitable source map exists. While the ultimate quality of the database is dependent on the quality of the source materials and the care used in automation, the procedures for original mapping are not discussed. It must be remembered that precise automation procedures will not improve an inferior base map.

Digital database standards have been proposed to provide:

1. Agencies and vendors contracted to automate spatial information with guidelines for digital data to be input into the LCBP GIS with respect to format, quality control and data documentation.
2. Users of the database with information on the technical specification of existing data to aid in determining the viability of those data for specific applications.

GIS Standards are necessary to assure the quality of GIS data, ensure compatibility with existing data in terms of format and attribute coding, and facilitate the most efficient organization, manipulation and analysis of these data.

II. SOURCE MAP MEDIA

- Stable-base mylar

Whenever possible, the most scale-stable media should be used in the automation process. Most preferable would be the original mylars upon which the map data were drafted. If originals are not available for direct encoding, then positive contact prints (i.e., mylar original to mylar copy) would be desirable. The following represents, from most to lease desirable, the media to be used as the source material for automation:

1. Mylar original
2. Mylar contact reproduction from mylar original
3. Non-stable base paper from mylar original
4. Non-stable base paper

III. GEOGRAPHIC REGISTRATION

- *Use corner tics for orthophoto base maps or USGS topographical maps*

A minimum of four registration points (tics) are required for each automated map. Tics are generally located in the four corners of the manuscript map.

In cases where new tic marks are to be established, they should never be created through digitization. New tic locations should be established through keyboard entry of exact geodetic coordinates.

IV. DIGITAL TOLERANCES

- *RMS not to exceed 0.005*
- *Fuzzy tolerance equivalent to 0.002 inches at the scale of the manuscript map*
- *Dangle distance approximately 0.10 inches*

There are several digital tolerances which affect the accuracy and resolution of a digital map that can be explicitly defined during the map automation process. The Fuzzy Tolerance defines the minimum distance separating arc coordinates. The Dangle Distance defines the minimum allowable length for a dangling arc. The Root Mean Square Error Tolerance (RMS) defines the error incurred when predefined tics are used to register a map on the digitizer.

The RMS error is a very important tolerance to consider because it indicates, in part, the accuracy with which the digitizing technician captures the locations of the tic marks for a map in the X and Y directions. In order to preserve the spatial integrity of map features during the automation process, it is extremely important to keep the RMS error as low as possible when a map is registered on the digitizing table. The recommended maximum RMS is 0.005. LCBP GIS requires that records of the RMS be kept during the automation process and that these records be delivered along with the digital data.

V. DATA CAPTURE

- *Use a minimum number of coordinates to define a line.*
- *Eliminate unnecessary pseudo nodes.*

In setting the standards for quality control, it is important to consider the resources for storing basin-wide digital data. In order to facilitate the most efficient use of computer storage, it is necessary to encode only the minimum number of vertices needed to capture the essence of a cartographic feature within the 0.010 inch accuracy limit is discussed below in the section on Cartographic Accuracy.

A pseudo node separates two arcs from one another. These should only used when an attribute change occurs along an arc. Unnecessary pseudo nodes increase the number of arcs and the size of the attribute tables, and degrade the quality of plotted line symbols.

VI. COINCIDENT FEATURES

- *Do not double digitize map features*

Coincident features are those which are in common between two or more data layers. For example, if a political boundary is formed by a river bank, the river bank and political boundary are coincident features.

Coincident features must be digitized only once. Regardless of the care taken in digitizing, slight differences are inevitable if a feature is digitized more than once. These differences lead to problems with small "slivers" if the layers are topologically joined.

Coincident features, digitized once, can be placed into master template coverages from which they can be retrieved when needed for incorporation into another layer, or they can be taken directly from the original coverage. In the above example, if the river was digitized first, the arcs necessary for the political boundary can be selected and placed in the new coverage.

VII. ATTRIBUTE CODING SCHEMA

- Use LCBG GIS approved coding schemes

All attributes should be coded according to the coding standards adopted by LCBP GIS. These are detailed in "Attribute Definitions and Codes", *LCBP GIS Handbook*, Part 2, Guidelines, Section A.

Typically, the adopted coding schemes reflect those commonly in use by organizations responsible for collecting and maintaining the particular type of data. For instance, coding of a soils database will follow the conventions of the USDA Soil Conservation Service. Nationally recognized formats also exist for coding of data layers such as wetlands and land use.

In cases where a coding scheme has not been adopted, it is crucial that whatever system is used be fully documented. Also, when devising a new scheme, it is important to allow for as many different foreseeable uses of the data as possible. For example, in a land use coding scheme, it is preferable to have individual codes for different types of agricultural use (e.g., cropland, pasture, barnyard) than just a single code specifying agricultural use. The individual codes may be aggregated, but the single code cannot be broken into its components, thereby limiting its use.

VIII. CARTOGRAPHIC ACCURACY

- *Digitize map features to within the equivalent of 0.01 inch of the original manuscript, at the scale of the manuscript.*

Digital map quality warrants special consideration. It is very important that the quality of the source maps be preserved as carefully as possible throughout the automation process. Careful, consistent and systematic digitizing, plus thorough verification are essential. Source map(s) should meet or exceed National Map Accuracy Standards (NMAS). (See Appendix A)

The LCBP GIS standard for feature accuracy are:

- *90% of the planimetric features on the digital map must be within 1/100 inch (0.010) of the centerline of that feature on the manuscript map when plotted at the original scale.*
- *100% of all features must be within 0.020 inch.*

The 0.010 inch interval is equivalent to a standard 0.010 plotter pen width. When a proof hardcopy plot of the digital map is overlayed on the original base map, discrepancies will be seen as an open space between the plotted feature and the original manuscript.

IX. SPATIAL TOPOLOGY

- *No overshoots*
- *No undershoots*

- *No slivers*
- *No open polygons*
- *No unlabeled polygons*
- *No unresolved line segment intersections*

Data submitted to the LCBP GIS shall be topologically clean and free of errors. The ARC/INFO commands CLEAN and BUILD are used to create topology, while the LABELERRORS and NODEERRORS can be used to detect certain topology problems.

X. EDGE MATCHING

- *Adjacent coverages must be edge matched*
- *Map features are not to extend beyond coverage boundaries*

Line segments (arcs) which intersect the boundaries of a coverage must be accurately edge matched with the corresponding line segment in the adjacent coverages. Computer edge matching techniques ensure an exact match. In lieu of an exact match, arcs must be matched to within 0.010 inch, centerline to centerline.

Arcs must not extend beyond (overshoot) or fall short of (undershoot) the coverage boundary.

XI. DOCUMENTATION

- *All coverages shall be documented*

The producing agency is responsible for providing a standard list of descriptive documentation for each digital map submitted to LCBP GIS. Standard documentation methods are described in "Data Layer Documentation", *LCBP GIS Handbook,* Part 1, Standards, Section B. The reasons for including documentation with each map are as follows:

- Documentation provides a history of each map.
- Documentation provides users with the necessary information to determine the utility of a map for a specific purpose.

XII. DELIVERY OF DIGITAL DATA

Several data formats for delivery of data include:

- *ARC/INFO Coverage*
- *ARC/INFO IMPORT/EXPORT interchange file format*
- *1600 BPI CCT - ARC/INFO TAPEWRITE utility - blocksize 8000, LRECL 80 (preferred method)*
- *DOS diskettes - DOS BACKUP utility*

Digital spatial information shall be delivered to LCBP GIS in a format agreed upon in advance by the suppliers and receivers of data. Files delivered on computer compatible tape (CCT) should be written at 1600 BPI with a block size of 8000 and a logical record size of 80. The ARC/INFO TAPEWRITE utility is useful for transferring an EXPORT interchange file to tape.

If PC ARC/INFO is used in automation, small data sets can be delivered on diskettes by first converting to the EXPORT interchange format. Large data sets should be transferred to computer compatible tape using the specifications described in the previous paragraph.

XIII. PROOF PLOTS

- *Mylar proof plots*
- *Pen width - 0.010 inch*

Acceptable methods for producing proof plots should be agreed upon by the data supplier and receiver. Preferably, for each digital map submitted, a mylar proof plot of that map must be delivered along with the corresponding source maps. The proof plot will be of the same scale as the manuscript. All lines on the proof shall use a linewidth of 0.010 inch or less and be drawn with liquid ink pens.

XIV. LOG FILES

The ARC/INFO LOG file or comparable file describing the steps used in the automation of each map shall be included in the delivery of each digital map.

XV. PRODUCTION HISTORY AND RMS TRACKING

Because the ARC/INFO LOG file does not record the history of the RMS error during automation, a separate written record must be kept of the RMS. The details for recording the production history and RMS tracking are provided in "Data Layer Documentation", *LCBP GIS Handbook*, Part 1, Standards, Section B.

XVI. ACCURACY ASSESSMENT

- *90% of all features within 0.010 inch*
- *100% of all features within 0.020 inch*
- *Topology complete and accurate*
- *100 % of all attributes coded correctly*

It is the responsibility of the producing agency to verify that the original data have been encoded within the accuracy limits set by LCBP GIS. Ninety percent of the cartographic features on a map must be digitized within 0.010 inch measured from the center line or center point of a feature. 100% of all cartographic features must be digitized within 0.020 inch.

The topology of all maps must be free from overshoots, undershoots, slivers and open polygons.

APPENDIX A

UNITED STATES NATIONAL MAP ACCURACY STANDARDS

With a view to the utmost economy and expedition in producing maps which fulfill not only the broad needs for standard or principal maps, but also the reasonable particular needs of individual agencies, standards of accuracy for published maps are defined as follows:

1. **Horizontal Accuracy.** For maps on publication scales larger than 1:20,000, not more than 10 percent of the points tested shall be in error by more than 1/30 inch, measured on the publication scale; for maps on publication scales of 1:20,000 or smaller, 1/50 inch. These limits of accuracy shall apply in all cases to positions of well-defined points only. Well-defined points are those that are easily visible or recoverable on the ground, such as the following: monuments or markers, such as bench marks, property boundary monuments; intersections of roads, railroads, etc.; corners of large buildings or structures (or center points of small buildings); etc. In general what is well defined will also be determined by what is plottable on the scale of the map within 1/100 inch. Thus while the intersection of two road or property lines meeting at right angles would come within a sensible interpretation, identification of the intersection of such lines meeting at an acute angle would obviously not be practicable within 1/100 inch. Similarly, features not identifiable upon the ground within close limits are not to be considered as test points within the limits quoted, even though their positions may be scaled closely upon the map. In this class would come timber lines, soil boundaries, etc.

2. **Vertical accuracy,** as applied to contour maps on all publication scales, shall be such that not more than 10 percent of the elevations tested shall be in error more than one-half the contour interval. In checking elevations taken from the map, the apparent vertical error may be decreased by assuming a horizontal displacement within the permissible horizontal error for a map of that scale.

3. **The accuracy of any map may be tested** by comparing the positions of points whose locations or elevations are shown upon it with corresponding positions as determined by surveys of a higher accuracy. Tests shall be made by the producing agency, which shall also determine which of its maps are to be tested, and the extent of such testing.

4. **Published maps meeting these accuracy requirements** shall note this fact on their legends, as follows: "This map complies with National Map Accuracy Standards."

5. **Published maps whose errors exceed those aforestated** shall omit from their legends all mention of standard accuracy.

6. **When a published map is a considerable enlargement** of a map drawing (manuscript) or of a published map, that fact shall be stated in the legend. For example, "This map is an enlargement of a 1:20,000-scale map drawing," or "This map is an enlargement of a 1:24,000-scale published map."

7. **To facilitate ready interchange and use of basic information for map construction** among all Federal mapmaking agencies, manuscript maps and published maps, wherever economically feasible and consistent with the uses to which the map is to be put, shall conform to latitude and longitude boundaries, being 15 minutes of latitude and longitude, or 7.5 minutes, or 3-3/4 minutes in size.

U.S. BUREAU OF THE BUDGET

Issued June 10, 1941
Revised April 26, 1943
Revised June 17, 1947

DATA LAYER DOCUMENTATION

I. PURPOSE

This document is primarily for use by GIS data developers, managers and technicians. The purpose is to provide standard methods and formats for documenting Lake Champlain Basin Program GIS (LCBP GIS) data layers. Comprehensive documentation must fully describe the data layer, its attributes, its development and its updates over time. Data layer documentation is an essential component of the LCBP GIS Quality Assurance/Quality Control program.

II. DOCUMENTATION FILE FORMATS

Data layer documentation may be maintained on paper or in a database management system. The documentation standards described here assume the use of an xBase relational database management system (RDBMS), which is compatible with pc ARC/INFO and most other desktop GIS software. The standard can be used, with minor modifications, with virtually any RDBMS, and is therefore compatible with a wide variety of GIS software on workstations, minicomputers and mainframes.

An example of the complete documentation for a data layer is given in Appendix B (page 12).

III. THE GISDOC PROGRAM

The Vermont Center for Geographic Information (VCGI) has developed a user interface program to make data layer documentation easier. This program, called GISDOC, can be used to develop or update LCBP GIS standard documentation files in an XBASE format. The program runs on any MS-DOS compatible personal computer; no database management software is required for its use. The GISDOC program is listed in the *VGIS Data Catalog* as data product \GISDOC, and is available upon request by contacting them.

IV. TILED DATA LAYERS

A data layer may be broken up geographically into numbered tiles, in which case each digit of the tile number is represented by an "n". Each tile would be a separate ARC/INFO coverage. The data layer's name thus incorporates the tile identifier, as in "RDSnn" for data layer "RDS" with tile number "nn". Individual coverages for different geographic areas would thus be named RDS11, RDS12, etc.

For tiled data layers, it is not practical to maintain digital documentation files for each tile's coverage. In this case, a separate directory should be created at the same level as the layer's coverages with the data layer's root name (without the tile number, for example, "RDS"). The layer's documentation files can be stored under this directory, along with other files used globally by the data layer (such as lookup tables).

Tiled data will be distributed by LCBP Data Clearinghouse with this directory structure (see Figure 1). Users who combine or manipulate the data for a specific project should copy the documentation files to the resulting coverage and record the changes made as described in the section below on **How to Document a Data Layer Update.**

DATA LAYER DIRECTORY

RDS

Coverage subdirectories (tiles)

Subdirectory for global documentation files, lookup tables, etc.

RDS11	RDS12	RDS13	. . .	RDS
LOG	LOG	LOG		DOC.DBF
ARC	.	.		AAD.DBF
AAT.DBF	.	.		UPD.DBF
HST.DBF	.	.		LUT1.DBF
.				.
.				.
.				.

Figure 1. Example of a directory structure for a tiled data layer.

V. SUMMARY OF THE DIGITAL DOCUMENTATION FILES

The following eight digital files are used for data layer documentation. They are available from LCBP Data Clearinghouse in DBF (Xbase) and INFO formats, as described in the *VGIS Data Catalog* under data product \DOCFILES. VCGI maintains a master documentation database for all data layers, which is used to generate the *VGIS Data Catalog* and to track data distribution.

Note that a DBF file referred to here as "<layer>.DOC" by ARC/INFO naming conventions will actually be named DOC.DBF by DOS under directory <layer>.

The documentation DBF files are summarized here, followed by file structures and field descriptions.

<layer>.DOC

Comprehensive documentation for the data layer. Each data layer has a single record. Information recorded for a data layer falls under several headings, including:

- Data layer summary
- Data manager
- Geographic area and tile structure
- Data source
- Data format
- Accuracy and tolerances
- Data development and automation
- Detailed narrative descriptions

The unique identifier for each record is the coverage name combined with its "path", which can be used to distinguish separate coverages with the same name. For example, roads data developed specifically for Stowe might be uniquely identified as coverage "RDS" with path "\STOWE".

Several fields of <layer>.DOC require lengthy descriptions. In DBF format, these are memo fields (variable length). In INFO format, these are 16 character fields holding "See <layer>.DCT", referencing a separate documentation text file where the more detailed information can be found.

<layer>.PAD — **Polygon or Point Attribute Documentation.** Attribute, value and code descriptions for the PAT items. There is one record for each PAT field. The ARC-defined attributes (AREA, PERIMETER, <layer>_ and <layer>_ID) don't need to be documented, except for <layer>_ID if it's used as a code.

<layer>.AAD — **Arc Attribute Documentation.** Attribute, value and code descriptions for the AAT items. There is one record for each AAT field. The ARC-defined attributes don't need to be documented, except for <layer>_ID if it's used as a code.

<layer>.TDD — **Tabular Data Documentation.** Tabular data files related to and stored with a coverage are described here. Tabular data bases (in .DBF or INFO format) are documented in essentially the same way as the PAT and AAT attribute tables. Simple tabular data bases, such as lookup tables, are more easily described under the ASSOC_FILE field of the .DOC file. Note that the GISDOC program does not handle .TDD files as of version 1.06.

<layer>.AND — **Annotation Documentation.** A table describing the data layer annotations by annotation level. There is at least one record for each annotation level.

<layer>.UPD — **Update file for recording any changes made to the features and attributes of the data layer.** This file is used to record major updates to the data, and to establish new versions of the data for distribution. The update is used to keep track of additions made, errors corrected, and who needs to receive updated data (as for specific tiles). Routine modifications made to the data during data development can be recorded in the history file (.HST), described below.

For a data layer tiled geographically into separate coverages, only one update (.UPD) file should be used for the complete data layer.

<layer>.HST — **History file for recording data layer development.** This file can be used to record specific tasks of data layer development, and can be used as the RMS log. The contents of the history file supplement the contents of the ARC/INFO LOG file. Its use is optional.

<layer>.DCT — **Documentation Text File (ASCII).** This file is for use with documentation maintained in INFO format, which is not recommended but which may be required for sites not able to use the GISDOC program MS-DOS personal computers. The .DCT file takes the place of the memo fields used in the xBase format documentation files. Note that the .DCT file, and any other non-ARC/INFO ASCII files, will not be EXPORTed with a coverage when that command is used.

VI. HOW TO DOCUMENT A DATA LAYER UPDATE

Data layers are typically updated over time with new features or attributes, or as errors are corrected. For the purpose of data distribution, the *version* of a data layer is the date of its last update. These steps should be taken to document a data layer update:

1. First, archive the previous version if you might want it at a later date. All data layers distributed through the LCBP GIS Clearinghouse are archived at the LCBP GIS Clearinghouse before being updated.

Archived versions are available by listing their update date on the LCBP GIS Data Request Form.

2. Add a record to the .UPD file describing the changes made since the last update, and recording the date of this update.
3. Update the LASTUPDATE field of the .DOC file to be the same as the record added to the .UPD file.
4. Update the ENTRY_DATE field of the .DOC file to agree with the date of last update. (The entry date may be changed by itself if the documentation is updated, but not the data).
5. Update any of the other documentation fields as need be to reflect the changes made to the data (including any changes to attributes).

VII. THE .DOC FILE

The .DOC file contains a comprehensive description of a data layer. Individual field contents are described here.

FILE STRUCTURE FOR DOC.DBF (<layer>.DOC):

Field	Field Name	Type	Width	Dec	Description
					DATA LAYER SUMMARY:
1	ENTRY_DATE	Date	8		Date of this entry; updated when any changes are made to the documentation or to the data. (The data layer *version* is given by the LASTUPDATE field.)
2	LAYERNAME	Char	20		Maximum of 8 characters for pcarc/INFO; tile numbers are referred to with an "n" for each digit.
3	DATALAYER	Char	60		A one line description of the data layer.
4	FEATURE	Char	50		Valid coverage feature types include point, line, poly, link and network. Link = line & point; network = line & poly.
5	STATUS	Num	2	0	1: Available for distribution at LCBP Data Clearinghouse; meets any LCBP GIS standards described in the *Catalog*'s detailed description. 2: Not available for distribution due to some restrictions (check with data manager listed in the documentation for more information. 4: In development. Data conversion to GIS format is in process. 5: These data a planned but data conversion has not yet begun. 6: Available for distribution in a pre-release version from LCBP Data Clearinghouse. More thorough data checking or other revisions are expected with the final release.
DATA MANAGER:					
6	DATAMGR	Char	40		Name and Title of the person responsible for data layer development; this person should be contacted if there are questions about the source or use of the data.
7	ORGANIZATN	Char	80		Organization of Data Manager
8	STREET_ADD	Char	50		Street address of organization
9	TOWN_ST_ZP	Char	40		Town/state/zip of organization
10	PHONE	Char	16		Phone number of organization

FILE STRUCTURE FOR DOC.DBF (<layer>.DOC) (...continued)

Field	Field Name	Type	Width	Dec	Description
GEOGRAPHIC AREA AND TILE STRUCTURE:					
11	GEOG_AREA	Char	60		Area covered by the coverage; use DESCRIPTN if more than 60 characters are needed, e.g., New York, Franklin County.
12	TILE	Char	40		Description of how the coverage is divided up geographical units, e.g. Basin, Sub-Basin, Vermont, New York, County, Town, Ortho (VT), QUAD (VT & NY). This will be state defined.See the *VGIS Data Catalog* contain a complete list and description of standard tile structures.
13	AVG_TILE	Num	6	1	Average size of the tiles, in kilobytes (1000 bytes).
DATA SOURCE:					
14	SRC_MAP	Char	60		Source of coordinate information.
15	SRC_DATE	Char	30		Published date(s) of the source maps--original compilation and latest revision.
16	SRC_SCALE	Char	30		Scale(s) of all source material; more than one scale may be given if the areas covered by each scale are clear. If need be, give details in SRC_DESCRP.
17	SRC_MEDIA	Char	40		Media of the source maps - paper, mylar, etc.
18	SRC_DESCRP	Memo	10		Type of data, dates collected, where stored - a full description of the source material. Details on how the data are used and its limitations should be given with the detailed descriptions (fields 40-43).
DATA FORMAT:					
19	DATAFORMAT	Char	20		Format used to store these data (e.g. ARC/INFO).
20	COORDSYS	Char	30		Coordinate system these data are stored in (e.g., UTM, State Plane, etc.)
21	PROJ_UNITS	Char	30		Vermont or New York State Plane Meters, UTM or other.
22	PJ_PRMTR	Memo	10		Description of projection name and any parameters needed to transform from lat/long to this projection e.g. zones, shifts, NAD27 to NAD83 conversion method used, etc.
23	H_DATUM	Char	10		Horizontal control datum (NAD 1927 or other).
ACCURACY AND TOLERANCES:					
24	FEATACC	Memo	10		Statement of the likely maximum differences between the feature coordinates and true ground location. (specify meters or feet). See Appendix A (p. 10) for help in assessing feature accuracy. Attribute accuracy is given in the PAD and AAD files.
DATA AUTOMATION:					
25	DIG_MGR	Char	40		Name/title of a single digitizer or of the digitizing manager.
26	DIG_ORG	Char	50		Organization responsible for data automation.
27	DIG_ADDR	Char	50		Street address.
28	TWN_ST_ZIP	Char	40		Town/State/Zip.
29	DIG_DATES	Char	30		Period over which features were digitized.
30	DIG_DEVICE	Char	40		Model of digitizing tablet, scanner, or other.
31	DIG_SOFTWR	Char	30		ARC/INFO or other; include software version(s).

FILE STRUCTURE FOR DOC.DBF (<layer>.DOC) (...continued)

Field	Field Name	Type	Width	Dec	Description
32	MAX_RMS	Char	25		Maximum RMS error allowed; generally .005. Note if there are exceptions, or use the .HST file as an RMS log.
33	QC_POSACC	Memo	10		Description of method(s) used (or not used) to assess the positional accuracy of data layer features; for example, "All linework checked by mylar proof plots", or "attribute code values not verified". This field can be added to over time, and used as a record of the QC checks used.
34	QC_ATTACC	Memo	10		Description of method(s) used (or not used) to assess the encoding of attributes associated of data layer features; for example, "All attributes had two independent checks" or "attribute code values not verified". This field can be added to over time, and used as a record of the QC checks used.
35	UPDATE_SCH	Char	50		Planned updates for distribution.
36	LASTUPDATE	Date	8		Date of last update of the data for distribution. This must agree with the .UPD file (described below).
DETAILED DESCRIPTIONS:					
37	DESCRIPTN	Memo	10		Feature definitions, methodology, data collection methods, etc. This should enable the proper use of the data, and include any pertinent information not found elsewhere in the documentation.
38	ASSOC_FILE	Memo	10		Describes files associated with the coverage, such as lookup tables, symbol sets or related attribute tables.
39	LIMITATION	Memo	10		Limits on data use, interpretation and use with other layers. Red flags should go here.
40	DISCLAIMER	Memo	10		Prepared by the data originator or manager.
** Total **			1175		Bytes per record

VIII. .PAD AND .AAD FILES

These two files are for documenting individual attributes (point, arc or polygon). The .PAD is for Polygon or Point Attribute Documentation, while the .AAD is for Arc Attribute Documentation. Both files use the same fields.

STRUCTURE FOR FILES PAD.DBF AND AAD.DBF (<layer>.PAD and <layer>.AAD):

Field	Field Name	Type	Width	Dec	Description
1	ENTRY_DATE	Date	8		Date of this documentation entry; updated when any changes are made to this item.
2	COVERNAME	Char	16		Identical to the .DOC file's COVERNAME.
3	PATH	Char	18		Identical to the .DOC file's PATH.
4	ITEMNAME	Char	16		From the PAT or AAT file; this is unique for each coverage.
5	ATTRIBUTE	Char	60		Short description of the attribute.
6	ATT_SOURCE	Char	80		The source of the attribute values.
7	TYPE	Char	1		From the PAT or AAT file.
8	WIDTH	Num	3	0	From the PAT or AAT file.
9	DEC	Num	2	0	From the PAT or AAT file.
10	CONTENT	Memo	10		Detailed description of the attribute, including any codes used and the accuracy of the item values.
** Total **			214 bytes per record		

IX. .TDD FILE

This file is for documenting the items (i.e., the fields) of a tabular data base associated with a coverage. Any tabular data bases should be noted in the memo field ASSOC_FILE (associated files) of the .DOC file. Simple tabular data bases are more easily described under the ASSOC_FILE field than in the .TDD file. They are usually in .DBF or INFO format.

A .TDD file can also be used to document tabular data bases not associated with any particular coverage. Note: the GISDOC program does not handle .TDD files as of version 1.06.

STRUCTURE FOR FILE TDD.DBF (<layer>.TDD):

Field	Field Name	Type	Width	Dec	Description
1	ENTRY_DATE	Date	8		Date of this documentation entry; updated when any changes are made to a field.
2	COVERNAME	Char	16		Identical to the .DOC file's COVERNAME. This may be a directory name if there is no single associated coverage.
3	PATH	Char	18		Identical to the .DOC file's PATH.
4	TABLENAME	Char	16		Name of the data table (8 character limit for DOS).
5	FIELDNAME	Char	16		The field (or item) name from the tabular database file (10 character limit for xBase).
6	DESCRIPTN	Char	60		Short description of the field and its values.
7	SOURCE	Char	80		The source of the field values.
8	TYPE	Char	1		From the structure of the tabular data file.
9	WIDTH	Num	3	0	From the structure of the tabular data file.
10	DEC	Num	2	0	From the structure of the tabular data file.
11	CONTENT	Memo	10		Detailed description of the field, including any codes used and the accuracy of the field values.
** Total **			230 bytes per record		

X. .AND FILE

The .AND file is for documenting coverage annotations.

STRUCTURE FOR FILE AND.DBF (<layer>.AND):

Field	Field Name	Type	Width	Dec	Description
1	COVERNAME	Char	16		Identical to the .DOC file's coverage name.
2	PATH	Char	18		Identical to the .DOC file's path.
3	LEVEL	Num	3	0	Unique for this coverage.
4	SYMBOL	Char	30		Often one symbol per level, but may be many; assumes PLOTTER.TXT symbols unless otherwise noted.
5	SIZE	Char	30		Often one or two sizes per level, but may be many.
6	TYPE	Char	10		The ANNOTYPE (POINT1, POINT2 or LINE).
7	DESCRIPT	Memo	10		Description of the annotations in this level.
** Total **			117 bytes per record		

XI. .UPD FILE

The .UPD file is for documenting coverage updates for the purposes of data distribution, archiving and tracking. This file is to be maintained by the designated data manager. Each data layer should have only one .UPD file, even if the data layer is tiled into separate coverages.

Every data layer available for distribution must have a .UPD file, in which case the first record of the file establishes the original date of completed data compilation.

Fields for describing any source maps used, RMS error or a tile number are provided for convenience. If not used, these fields can be left blank.

STRUCTURE FOR UPD.DBF (<layer>.UPD):

Field	Field Name	Type	Width	Dec	Description
1	DATE	Date	8		Date of this entry to the .UPD file, when the update occurred.
2	ACTION	Memo	10		A description of the changes or additions made.

** Total ** 18 bytes per record

XII. .HST FILE

The .HST file is for documenting coverage development prior to data distribution, or during data revisions. Each coverage should have an .HST file, even for tiled data layers. Its use is optional.

STRUCTURE FOR FILE HST.DBF (<cover>.HST):

Field	Field Name	Type	Width	Dec	Description
1	DATE	Date	8		Date of this entry to the .HST file.
2	OPERATOR	Char	20		Digitizer, editor or other responsible person.
3	SRC_MAP	Char	40		Source of digitized coordinate information.
4	SRC_DATE	Char	25		Published date of the source map
5	SRC_SCALE	Char	25		Scale of the source map.
6	SRC_MEDIA	Char	25		Media of the source maps - paper, mylar, Cronaflex orthophoto, etc.
7	RMS	Char	10		Tic registration RMS error
8	NOTES	Memo		10	

A description of the changes or additions made.
** Total ** 163 bytes per record

APPENDIX A
ASSESSING FEATURE ACCURACY

A. FEATURE ACCURACY

Primary GIS features generally include points, lines and polygons. Documenting feature accuracy is complex due to the many sources of error occurring throughout the process of data development. The discussion here refers to the conversion of maps to GIS format with a digitizing tablet. Many aspects of the discussion can be applied to other methods of digitizing (such as scanning and GPS).

An assessment of feature accuracy should describe how close a digitized feature is to its true location on the ground (in feet or meters). For most data layers, there are two general sources of error for geographic features: source map error, and data conversion error. All sources of error in the feature locations should be described in the Feature Accuracy field (FEATACC) of the .DOC file.

B. SOURCE MAP ERROR

Source map compilation often involves preparing source maps upon a chosen base map. The base maps most used in Vermont are the 1:5000 orthophotos and the 7.5' USGS topographic maps.

The *process* by which source maps were developed should be described in the Source Description field (SRC_DESCRP) of the .DOC file. The *error* in feature representation resulting from source map compilation should be described under the Feature Accuracy field (FEATACC) of the .DOC file.

Source map error arises from:

- Error inherent in the base maps. Vermont's orthophotos and USGS topographic maps claim to meet National Map Accuracy Standards. For horizontal accuracy, this means that 90% of the features fall within 0.02 inches of their true locations at the map scale.

- Source map distortion, particularly for paper maps which have stretched out, or which shrink or swell with humidity.

- Difficulty in locating features on the map (e.g., the location of an underground storage tank or of a stream on a 1:5000 orthophoto). Features are particularly difficult to locate when they are far from roads, rivers or other base map reference features.

- Difficulty in feature interpretation. Examples include the interpretation of land use/cover classes, and estimating the location of soil unit boundaries.

- The approximate nature of interpolated features (as with interpolated contour lines or slope maps generated from digital elevation models).

Except for the factor of base maps meeting National Map Accuracy Standards, these aspects of source map error are often difficult to express in terms of "meters on the ground". Often, a good knowledge of the data allows one to judge the expected error. In other cases, it will be appropriate to test the accuracy of the data by comparing it against ground truth.

C. DATA CONVERSION ERROR

The process of converting source maps to digital format introduces several sources of error:

- error in map registration while digitizing (i.e., the RMS error),
- digitizer error (e.g., the accuracy of the digitizing tablet), from the manufacturer's specifications,
- map shifting (for source maps which cannot be taped smoothly to a digitizing tablet),
- digitizing error (e.g., the precision with which features were traced with the digitizing puck), which can be minimized when proof plots are thoroughly checked and errors corrected.

D. ESTIMATING ERROR IN FEATURE LOCATIONS: AN EXAMPLE

Source map error must be estimated from the data developer's knowledge of the data, or by testing the data against a source of known accuracy. Often it's enough to say how good a mapped location is expected to be. For example, if well locations are mapped on USGS topographic maps, the source map preparer might be able to specify that:

> "Most wells are within about 0.1 inch of their true location, except when they were far from a road in which case they may be off by 0.4 inches."

For 1:24,000 scale maps, 0.1 inch translates to 200 feet on the ground.

Other factors might be easier to quantify, as shown in the Table A1 below.

Table A1

Source of error	Map inches	Feet on the ground (1:24,000 map scale)
Putting wells on source maps	0.1" to 0.3"	200 ft (800 worst case)
Source map distortion	negligible[1]	
National Map Accuracy Std.	0.02"	40 ft
Map registration error (RMS)	0.005"	10 ft
Digitizer error	0.01"	20 ft
Map shifting	negligible[1]	
Digitizing error	0.02"	40 ft
Totals	0.155" to 0.355"	310 to 810 ft

[1]Source map distortion and map shifting can be considered negligible for source maps in excellent condition.

These totals constitute a worst case scenario. It is likely that the smaller errors due to data conversion will cancel each other out to a degree. As a rule of thumb, the likely error can be considered to be about two thirds of the total error. For this example, it would be appropriate to say:

> "Most well coordinates are expected to fall within about 200 ft of their true ground location, except where wells were distant from roads, in which cases the errors may be as great as 800 ft. The primary source of error is the difficulty in locating wells on the source maps."

Examples of NMAS for horizontal accuracy are contained in Table A2 below.

Table A2

Horizontal Accuracy Examples

Scale	NMAS[1] Feet / Meters
1:1,200	+/- 3.33 / 1.01
1:2,400	+/- 6.67 / 2.03
1:4,800	+/- 13.33 / 4.06
1:9,600	+/- 26.67 / 8.13
1:10,000	+/- 27.78 / 8.47
1:12,000	+/- 33.33 / 10.16
1:24,000	+/- 40.00 / 12.19
1:63,500	+/- 105.60 / 32.19
1:100,000	+/- 166.67 / 50.80

[1] National Map Accuracy Standards define the requirements for meeting horizontal accuracy as 90 percent of all measurable points must be within 1/30th of an inch for maps at a scale of 1:20,000 or larger, and 1/50th of an inch for maps at scales smaller than 1:20,000.

E. OTHER FACTORS

Many data layers are susceptible to change over time. For example, rivers and streams slowly change their course over time. The documented feature accuracy is only expected to reflect the accuracy at the date of source map compilation. When appropriate, the lack of currency of the data should also be noted.

F. ATTRIBUTE ACCURACY

Attribute accuracy should be described individually for each attribute in the .PAD and .AAD files. If appropriate, global statements on attribute accuracy can be made under the QUAL_CNTRL or DESCRIPTN fields.

APPENDIX B
SAMPLE OF DATA LAYER DOCUMENTATION

COVERAGE: TB250
DETAILED DOCUMENTATION
Digital file TB250.DOC
Entry date: 12/30/91

DATA LAYER SUMMARY:
Coverage: **TB250**
Data layer: Town Boundaries at 1:250,000
Feature type(s): Polygon and Line
Status: 1: Available from VCGI

DATA MANAGER:
GIS Database Administrator
VCGI
206 Morrill Hall, UVM
Burlington VT 05405-0106
(802) 656-4277

GEOGRAPHIC AREA AND TILE STRUCTURE:
Area: State of Vermont
Tile structure: STATE
Average tile size: 260.0 kilobytes

DATA SOURCE:
Map(s): USGS GIRAS data
Map date(s): Pre-1980
Scale: 1:250,000
Media: Digital
Description: This coverage was originally obtained by the UVM School of Natural Resources and Distributed to various VGIS users.

DATA FORMAT:
Data format: ARC/INFO
Projection/units: State Plane Meters
Datum: NAD 1927

ACCURACY AND TOLERANCES:
Feature Accuracy: The line accuracy of this coverage is questionable due to its uncertain history and modifications. The boundaries are very approximate and do not line up well with the Digital Line Graph water features (1:100,000 scale). Errors of up to 200 meters are common, and errors greater than 300 meters have been observed.
Fuzzy tol: 12.700 map units
Dangle: 317.50 map units

DATA DEVELOPMENT AND AUTOMATION:
Manager: None
Digitizing -
Dates: Unknown
Quality control: Line accuracy has not been checked. Attribute values have been spot checked by VCGI technicians, but some attribute errors are possible.
Update schedule: None. Revised attribute tables when available.
Last update: December 30, 1991

DETAILED DESCRIPTIONS:
Description: Vermont town boundaries at 1:250,000 scale. TB250 town boundaries are appropriate for thematic mapping, very small scale (1:250,000) maps, and test or demonstration plots. Most appropriately used for thematic maps of any size, maps with approximate town boundaries, or maps at 1:250,000 or smaller scale. A

1:250,000 map of Vermont would be about 40" high.

The Lake Champlain boundary is used as the town boundary for towns bordering the lake, so that the lake can be separately shaded. This differs from coverage TB24, in which the town boundaries extend into the lake.

Note that more accurate town boundaries are available for Addison, Orange, Rutland and Windham counties in coverage TWNBND, but these are not necessarily closed polygons.

Associated files: Arc lookup tables (line symbols in parentheses):

Lookup table	Town line	County line	State Line
TB250.LU1	thin (1)	thicker (5)	thicker yet (9)
TB250.LU2	dashed(77)	solid (5)	solid (5)
TB250.LU3	thin (1)	thick (9)	thick (9)

Limitations: These boundaries do not line up well with coincident river features from the Digital Line Graph data (1:100,000). Producing maps at 1:50,000 or larger scale should be done with caution to avoid obvious data accuracy problems.

Disclaimer: VCGI makes no representations of any kind, including but not limited to the warranties of merchantability or fitness for a particular use, nor are any such warranties to be implied with respect to the data.

Coverage: TB250
Polygon Attributes

Digital file TB250.PADItem Name:FIPS6 **Type:** N**Width:** 6 **Decimals:** 0 **Date of entry:** 12/30/91

Attribute: Unique code for each Vermont town
Source: Federal Information Processing Standard and Census codes
Content: Town FIPS6 codes are made up of 6 digits or "cccttt". The first three digits ("ccc") are a county FIPS code (001 to 027 by odd numbers, alphabetical by county). The last three digits ("ttt") are a Census Place code (in multiples of 005, alphabetical within county). See the Geographic Area Codes standard in the LCBP handbook for a description and list of geographic area codes used in New York and Vermont.

Item Name: TOWNNAME **Type:** C **Width:** 20 **Decimals:** 0 **Date of entry:** 06/01/91
Attribute: Name of town, upper case letters
Source: Unknown
Content: Town names, in upper case letters.

Item Name: RPC **Type:** C **Width:** 2 **Decimals:** 0 **Date of entry:** 06/01/91
Attribute: Two letter code for the town's RPC
Source: Assigned by VCGI
Content: AC = Addison County RPC
BC = Bennington County PC
CC = Chittenden County RPC
CV = Central Vermont RPC
FG = Franklin-Grand Isle RPDC
LC = Lamoille County PC
NV = Northeastern Vermont Development Association
RR = Rutland RPC
SW = Southern Windsor County RPDC
TR = Two Rivers-Ottauquechee RC

UV = Upper Valley-Lake Sunapee Council
WR = Windham Regional Commission

Item Name: PATHCODE **Type:** C **Width:** 8 **Decimals:** 0 **Date of entry:** 12/30/91
Attribute: A standard 8-character abbreviation (can be a DOS pathname)
Source: VGIS Handbook: Geographic Area Codes standard
Content: A town abbreviation of 8 or fewer letters, used to organize computer information by town. The PATHCODE can be used as the 8-character DOS directory name for town information.

Item Name: TOWNANGLE **Type:** N **Width:** 3 **Decimals:** 0 **Date of entry:** 06/01/91
Attribute: Angle the town can be drawn for best fit in a square
Source: Determined by VCGI
Content: This attribute can be used to rotate a town relative to north for automatic town map production. Used in conjunction with MAPZOOM, TOWNANGLE can be used to automatically plot towns for best fit within square map limits.

Item Name: MAPZOOM **Type:** N **Width:** 5 **Decimals:** 3 **Date of entry:** 06/01/91
Attribute: Zoom factor for a town once it's been rotated by TOWNANGLE
Source: Determined by VCGI
Content: A town map can be zoomed out to fit a given square plotting area by this factor once it's been rotated by the TOWNANGLE.

Item Name: LABELTEXT **Type:** C **Width:** 1 **Decimals:** 0 **Date of entry:** 06/01/91
Attribute: Flag for plotting text at the label point.
Source: Determined by VCGI
Content: Some towns which include islands in Lake Champlain are made up of more than one polygon. This flag can be used to suppress plotting of a value at more than one label point for a town.
LABELTEXT = 'Y' should have text at the label points,
LABELTEXT = 'N' should not.
Use "resel tb250 polys labeltext = 'Y'" before using the LABELTEXT command for proper placement the label text. Note also that level 2 annotation is suitably placed for use with label text.

Coverage: TB250
Arc Attributes
Digital file TB250.AAD

Item Name: RPC **Type:** N **Width:** 1 **Decimals:** 0 **Date of entry:** 06/01/91
Attribute: RPC boundaries
Source:
Content: RPC boundary code:
1 = State boundary
2 = RPC boundary (not state)
3 = Town boundary (not state or RPC)

Item Name: TBLINE **Type:** N **Width:** 1 **Decimals:** 0 **Date of entry:** 06/01/91
Attribute: Town boundary line code
Source: VCGI
Content: Town boundary code:
1 = State boundary
2 = County boundary (not state)
3 = Town boundary (not county or state)

Coverage: TB250
Annotation
Digital file TB250.AND

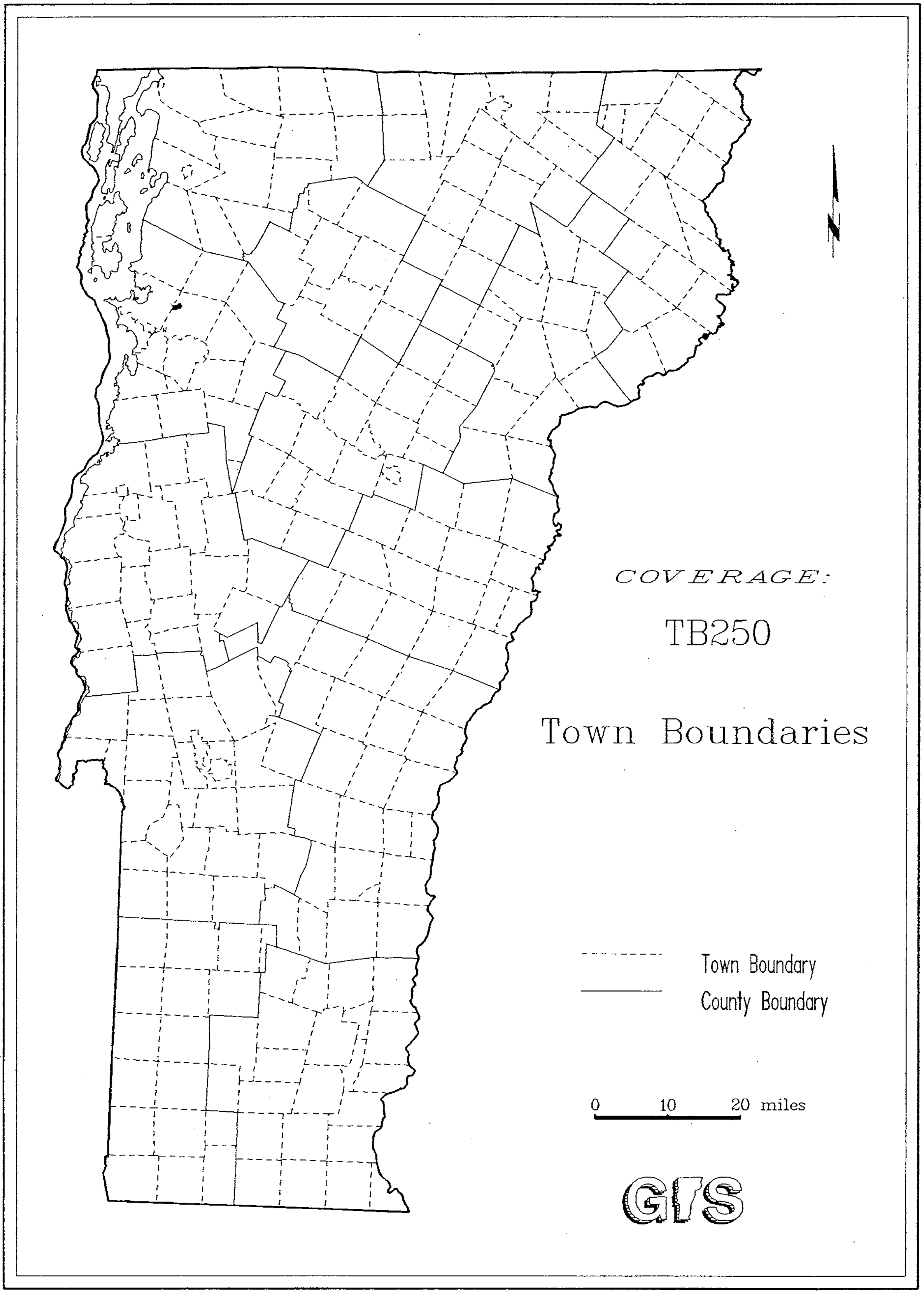
Annotation level: 1
Symbol: 1
Anno Size: 900
Anno Type: POINT2
Description: Town names in upper case letters. City names are in lower case (capitalized). These annotations are placed to be suitable when only the town name appears; see level 2 annotation.

Annotation level: 2
Symbol: 1
Anno Size: 900
Anno Type: POINT2
Description: Town names in upper case letters. City names are in lower case (capitalized). These annotations are placed to be suitable when the town name is used along with text placed at the label point. In this case, the town name may be away from the middle of town to allow room for the other text. See also the LABELTEXT polygon attribute.

Coverage: TB250
Data Updates
Digital file TB250.UPD

Date of Update: 12/02/91
Operator: Steve Bower,VCGI
Tile(s): STATE
Action: TB250 has undergone numerous changes which were not well recorded over the last couple of years. It has previously been known as TOWN and TOWN250. Item FIPS6, the unique identifier for each community, has been renamed from FIPS to meet VGIS standards.

Date of Update: 12/30/91
Operator: Steve Bower, VCGI
Tile(s): STATE
Action: Modifications made to level 1 annotations. Level 2 added for use when attribute information is output at the label point. Spelling of "BUELS GORE" corrected.


COVERAGE:
TB250
Town Boundaries
Town Boundary
County Boundary
0 10 20 miles
GIS

MAP COORDINATE SYSTEM

I. PURPOSE

Standards for a map coordinate system are critical to retaining the integrity of the many layers of the LCBP Geographic Information System (LCBP GIS). This means that sources of information coming from other coordinate systems require conversion for use in the LCBP GIS.

II. COORDINATE SYSTEM STANDARDS

The **standard** map projections used in the LCBP GIS are both the New York and Vermont State Plane Coordinate System (SPCS) of 1927 and the Universal Transverse Mercator grid system. The **standard** for storing coordinates is meters.

These are the two plane coordinate systems commonly used to express positions. *Universal Transverse Mercator (UTM)* coordinates (northings and eastings) are derived from a rectangular grid based on a transverse Mercator projection. The UTM system consists of 60 north-south zones, each 6° wide in longitude. The zones are numbered consecutively, starting with zone 1 between 180° and 174° and 180° E. longitude.

The UTM system is designed to be used for latitudes between 80° S. and 84° N. and does not include the polar regions. Unlike state plane coordinates, UTM coordinates are available for the whole world (minus the poles). UTM coordinate values are given in meters.

The *State Plane Coordinate Systems (SPCS)* is a system of rectangular (X and Y) grid coordinates derived from one of two projections over a state or part of a state (zone). A Lambert conformal conic projection with two standard parallels is used for states having a large east-west extent. A transverse Mercator projection is used for states having a large north-south extent. Most states have more than one zone to minimize the distortions inherent in the projections, and the zone boundaries follow county lines. Computed state plane coordinates are available for horizontal geodetic stations established by Federal agencies.

The State Plane Coordinate System of 1927 is now out-of-date. However, it is used rather than the new SPCS of 1983 because most of the commonly used base maps (topographic maps from USGS and Vermont orthophoto maps) use SPCS 1927. Nearly all existing computerized map data, such as the digital line graph (DLG) files from USGS, use NAD 1927 as well. The use of NAD 1983 for the LCBP GIS database will be accomplished by use of a conversion procedure. Conversion of all data to NAD 1983 will most likely occur during the time frame of this project. Upon completion of this conversion, this standard will be revised.

In addition, NAD 83 is the official civilian horizontal datum for U.S. Surveying and mapping activities performed or financed by the Federal government. To the extent practicable, legally allowable and feasible, all Federal agencies using or producing coordinate information should provide for an orderly transition from NAD 27 to NAD 83.

In certain cases coordinates may be entered in units of U.S. survey feet rather than meters. To convert from feet to meters, the following relationship is used:

1 meter = 3937/1200 (or 3.280833333) U.S. survey feet.

1 foot = .30480061 meters

III. DATUM CONVERSION PROCEDURES

Procedures exist to convert from NAD 1927 to NAD 1983 and vice versa. All Global Positioning Satellite (GPS) data uses NAD 1983. Maps utilizing NAD 1983 are also available. Methods for conversion shall be stated in the data coverages documentation (see LCBP GIS Data Documentation). NADCON or CORPSCON are federally adopted software programs that do the conversion. Any data collected using should be clearly labeled so that the data does not get merged incorrectly with NAD 1927 based datalayers. (NOTE: Vermont Statutes require all new projects by the year 2000 to use NAD83 (Ch17;Sec 679).)

The differences in the NAD27 and NAD83 datum are significant! These differences or errors amount to about a 7 meter shift in northing or latitude and a 35 meter shift in easting or longitude for the champlain basin. In addition to the base error there are local differences on the magnitude of +/- 2 meters in northing and/or easting. These differences tend to be gradual and spread over the entire basin.

IV. MAP PROJECTION CONVERSION PROCEDURES

Occasionally, it may be necessary to automate a source map which is based on USGS topographical maps. These maps must be digitized in the Universal Transverse Mercator (UTM) projection. Coordinates are in meters. The UTM zone is determined as follows: If the USGS quad is west of 72 degrees west longitude and east of 78° west longitude, zone 18 is used. The Lake Champlain Basin is in zone 18.

Note that a "yshift" of -4000000 meters is applied to all UTM northing measurements. This is to preserve the precision of coordinates which are stored with approximately 7 significant digits. Since all northings in the Lake Champlain Basin are in the 4000000 range, the millions digit is not required.

It is important to note that zone 18 is distinct and separate map coordinate system. UTM and SPCS are different. Before data based on different map projections can be used together such as NY SPCS 1927 and VT SPCS 1927, it is necessary to convert the coordinates to a common projection, using the ARC/INFO PROJECT command or equivalent. Appendix A contains the PROJECT commands needed to convert from UTM to state plane.

V. REFERENCES

1) NOAA manual NOS NGS 5
"State Plane Coordinate System of 1983"
James Stem
January 1989
2) NOAA technical memorandum NOS NGS-50
NADCON
"The Application of Minimum corvature derived surfaces in the

transformation of positional data from the north american datum of 1927 to the north american datum of 1963"
Warren T. Dewhurst
January 1990

3) USGS Professional Paper 1395

"Map Projections a Working Manual"
by John P. Snyder
1987

4) NOAA Professional Paper nos 2

"North American Datum of 1983"
Charles Schwarz
Editor
Dec 1989

APPENDIX A
Method for Converting a Map Projection

ARC/INFO PROJECT commands to convert a digitized map from the UTM projection to SPCS of 1927:

```
INPUT
PROJECTION UTM
UNITS METERS
YSHIFT 4000000
ZONE 18
OUTPUT
PROJECTION STATE
UNITS METERS
ZONE 5526
PARAMETERS
NEED
```

Note the YSHIFT 4000000 command in the UTM projection specification. This causes the northing coordinate to be restored to its true value. The northing values are stored with the millions digit removed to prevent computer round-off error.

LAND USE/LAND COVER CODES

I. PURPOSE

Maps of land use and land cover are fundamental tools for natural resource management and planning. These land use/cover codes provide a standard system by which units of land can be categorized. Land use/cover maps serve to:

- assess the area and distribution of individual land use/cover categories (such as cropland and pasture, code 21)
- show the diversity of the landscape
- contrast the land diversity of different regions
- monitor changes in land use and/or cover over time

The LCBP GIS Land Use/Land Cover Codes have been developed to:

- provide compatibility with existing USGS codes,
- provide codes tailored to Lake Champlain Basin's landscape,
- permit land coding at a various resolutions (local, regional and state),
- avoid and clarify redundant classifications which depend on interpretation.

II. LAND USE AND LAND COVER

GIS provides new tools for mapping the landscape. Although land use and land cover are quite different, they are frequently lumped together on a single map. Land use and land cover, however, are very different ways of looking at the land:

Land Cover:
- The vegetative, barren, liquid or artificial (human-made) material covering the land.
- Land cover can be interpreted from aerial photographs or satellite imagery with varying degrees of accuracy.

Land Use:
- The actual use of the land by its human inhabitants.
- Land use may at times be based on property boundaries (as in a municipal park), but the actual use of the land will not necessarily correspond to land ownership.

In a GIS it is feasible to view land use and land cover in a combined data layer, or as separate data layers. Some individual land categories will be developed as GIS data layers completely separate from land use/cover mapping. For example, surface waters and wetlands, both components of a land cover map, are being developed as separate data layers. A land use/cover mapping project should make use of existing data layers when possible.

Land use is the actual human activity for which the land is used, which may differ from the promoted use of the land. For example, zoning is generally not considered a land use since an area with "three acre zoning" may actually have much less than one unit per three acres. The delineation of residential areas may be particularly tricky. Mappers should be systematic in their methods, which should be well documented to enable proper interpretation of their maps in the future. Documentation might include language such as: "Land use and cover were photo-interpreted from 1:5000 orthophotos. Residential areas were delineated at the boundaries between homesteads (including yards or gardens) and the surrounding woods and fields. The density of residential single family/duplex areas (code 113) was determined by assigning a number of homes to residential area polygons, and computing the number of units per acre."

III. THE CODING HIERARCHY

LCBP Land Use/Land Cover Codes are an expansion of USGS land categories developed by Anderson and others (1976) for use with remotely sensed images. The codes are organized into four hierarchical levels. At the top of the hierarchy, first level includes:

1. Urban and built-up land
2. Agricultural land
3. Brush or transitional between open and forest
4. Forest
5. Water
6. Wetlands
7. Barren land
8. Tundra
9. Permanent snow and ice (no known examples in the Basin)

Each sub-level contains at most 9 categories (1 through 9). Level one has the lowest resolution, or detail of the landscape, while level four has the highest resolution. For example, a permanent pasture would be coded as follows at each of the four levels:

Level	Code	Category	Resolution
1	2	Agricultural land	Low
2	21	Cropland and Pasture	|
3	212	Hay/rotation/permanent pasture	|
4	2123	Permanent pasture	High

The USGS level 1 and level 2 codes were used for the LCBP GIS codes with these changes:

- Rangeland (3) was changed to Brush or Transitional Between Open and Forested, since there are no real rangelands in the basin,
- a Mixed Herbaceous and Shrub/Brush code (33) was added,
- level 2 codes were added for Tundra (8).

IV. RESOLVING CONFLICTS BETWEEN CODES

For some land units, more than one code will be possible. *The recommended rule for resolving conflicts is to use the code that comes first in the numbering sequence.* This rule is meant to be used for a combined land use/land cover map. For example, a grove of maples used for producing maple syrup may be considered either a Sugarbush (code 225) or as Northern Hardwoods, Sugar Maple (code 4121). In this case, the lower code (225) would be used. Likewise, a campground with mixed woods would be coded as Campground (1936) instead of as a Mixed Coniferous-Broadleaf Forest (43), because 1936 precedes 43 in the numbering sequence.

Only levels 1 and 2 (Urban/built-up and Agricultural land) emphasize land use, while levels 3 to 9 are almost exclusively land cover categories. If the rule above is used to resolve coding conflicts, the resulting map will emphasize land use over land cover. Land cover can be emphasized by selecting the code that is *higher* in the numbering sequence when conflicts occur. Whichever method is used, it should by followed systematically and clearly documented.

V. USE OF THE CODES

Mappers will choose how to use the codes to best suit their needs. Maps can be made in which:

- only land use is coded
- only land cover is coded
- a single code is used to for mixed land use/land cover
- separate land use and land cover codes are assigned to each unit of land (thus to each GIS polygon)

It will not always be feasible to code land for land use only, because remote areas often will not have appropriate land use codes. Land use maps will generally be restricted to developed areas when using these codes. More extensive land use coding schemes exist, however, and can be provided by LCBP GIS Clearinghouse.

If land use and land cover are separately coded, it is likely that some boundaries will be shared between land use and land cover maps. The following ARC/Info polygon attribute table (PAT) items are recommended for a combined land use/land cover GIS map:

Item	Type	Width	For
LUSEyy	I	4	Land use only
LCOVERyy	I	4	Land cover only
LULCyy	I	4	Combined land use/land cover

"yy" refers to the year of interpretation, so that LCOVER90 would be the land cover interpreted for 1990. All item types are I (integer).

VI. LAND USE / LAND COVER CODES

1 URBAN AND BUILT-UP LAND

11 RESIDENTIAL

111 Multi-family, medium to high rise apartments & condominiums (3 or more stories)
- 1111 more than 50 units/acre
- 1112 36 to 50 units/acre
- 1113 21 to 35 units/acre
- 1114 10 to 21 units/acre
- 1115 1 to 10 units/acre

112 Multi-family, low rise apartments and town houses, but not duplexes; (1 or 2 stories)
- 1121 more than 18 units/acre
- 1122 12 to 18 units/acre
- 1123 8 to 12 units/acre
- 1124 less than 8 units/acre

113 Single family/duplex
- 1131 more than 8 units/acre
- 1132 5 to 8 units/acre
- 1133 3 to 4 units/acre
- 1134 1 to 2 units/acre
- 1135 less than 1 unit/acre

114 Mobile home parks
- 1141 more than 9 units/acre (legal Vermont maximum of 9 units/acre)
- 1142 6 to 9 units/acre
- 1143 2 to 5 units/acre
- 1144 less than 2 units/acre

115 Group and transient quarters
- 1151 Rooming and boarding houses
- 1152 Shelters
- 1153 Residence halls and dormitories
- 1154 Retirement homes and orphanages
- 1155 Religious quarters
- 1156 Residential hotels
- 1157 Fraternities/sororities
- 1159 Other group & transient quarters

119 Other residential

12 COMMERCIAL, SERVICES AND INSTITUTIONAL

121 Commercial retail
- 1211 Building materials, hardware and farm equipment
- 1212 General merchandise/department store
- 1213 Food/groceries
- 1214 Automotive, marine craft, aircraft and accessories/dealers
- 1215 Furniture, home furnishings and equipment
- 1216 Eating and drinking (restaurants)
- 1217 Gasoline/petroleum
- 1218 Farm stand
- 1219 Other (Not elsewhere recorded)

VI. LAND USE / LAND COVER CODES (continued)

1 URBAN AND BUILT-UP LAND (continued)

12 COMMERCIAL, SERVICES AND INSTITUTIONAL (continued)

122 Commercial wholesale
- 1221 Food/sundries/beverages
- 1222 Agricultural products/supplies
- 1223 Lumber/hardware/building supplies/paper
- 1224 Industrial product/chemical/petroleum
- 1225 Motor vehicles/parts/supplies
- 1229 Other

123 Services
- 1231 Banking
- 1232 Personal services
- 1233 Professional services
- 1234 Self-storage
- 1235 Day care
- 1236 Contract construction services
- 1237 Automotive services/car washes
- 1238 Agricultural sales/services
- 1239 Other services

124 Lodging
- 1241 Hotel/motel
- 1242 Bed and breakfast
- 1249 Other lodging

125 Government
- 1251 Correctional
- 1252 Military
- 1253 Courthouse
- 1254 Postal service
- 1255 Administrative offices
- 1256 Emergency services
- 1257 Public works
- 1259 Other government

126 Institutional
- 1261 Religious
- 1262 Hospital
- 1263 Medical clinics
- 1264 Nursing
- 1265 Social organizations/associations
- 1269 Other institutional

127 Educational
- 1271 Preschool
- 1272 Elementary
- 1273 Middle
- 1274 High school
- 1275 College/university
- 1276 Vocational/trade
- 1279 Other educational

128 Indoor cultural/public assembly
- 1281 Museum
- 1282 Theatre
- 1283 Amusement
- 1284 Sport

VI. LAND USE / LAND COVER CODES (continued)

1 URBAN AND BUILT-UP LAND (continued)

12 COMMERCIAL, SERVICES AND INSTITUTIONAL (continued)

128 Indoor cultural/public assembly (continued)

1289 Other indoor cultural assembly

129 Other commercial, services and institutional

13 INDUSTRIAL

131 Primary metal production
132 Petrochemicals
133 Primary wood processing and paper mills
134 Stone, clay, glass
135 Metal & non metal fabrication
136 Food processing
137 Mining
1371 Hard rock
1372 Quarry
1373 Sand/gravel
1379 Other mining
138 Home/cottage industries
139 Other industrial

14 TRANSPORTATION, COMMUNICATION AND UTILITIES

141 Air transportation
1411 Commercial Airports
1412 Non-commercial Airports
1413 Heliports
1419 Other air transportation
142 Rail transportation
1421 Rail yard
1422 Station or terminal
1423 Active rail line
1424 Inactive rail line
1429 Other rail transportation
143 Water transportation
1431 Port facilities
1439 Other water transportation
144 Road transportation
1441 Limited & controlled highway
1442 Road
1443 Truck terminal
1444 Bus terminal
1445 Park/ride lot
1446 Parking structure/lot
1449 Other road transportation
145 Communication
1451 Radio station
1452 Telecommunications facilities
1453 Television station
1454 Transmitters/towers
1459 Other communication

VI. LAND USE / LAND COVER CODES (continued)

1 URBAN AND BUILT-UP LAND (continued)

14 TRANSPORTATION, COMMUNICATION AND UTILITIES (continued)

146 Electric, gas and other utilities
- 1461 Electric transmission lines
- 1462 Electric power generation facilities
- 1463 Gas or oil pipeline (natural, propane or other)
- 1464 Gas or oil storage facilities
- 1469 Other utilities

147 Water & wastewater utilities
- 1471 Water treatment plant
- 1472 Aqueduct
- 1473 Water storage tanks
- 1474 Wastewater treatment
- 1475 Interceptors
- 1479 Other water/wastewater utilities

148 Solid waste utilities
- 1481 Active landfills
- 1482 Inactive dump sites
- 1483 Recycling/transfer stations
- 1484 Incinerator
- 1485 Composting facility
- 1486 Sludge processing facility
- 1487 Hazardous waste collection/processing facility
- 1488 Junkyards
- 1489 Other solid waste facilities

149 Other communication, utilities

15 INDUSTRIAL AND COMMERCIAL COMPLEXES

151 Industrial park
152 Office park
153 Shopping center/mall
159 Other industrial/commercial complexes

16 MIXED (areas where level 2 uses are mixed with no one use predominating)

161 Multiple stories, residential in upper stories only
- 1621 Commercial lower level
- 1622 Services lower level
- 1623 Mixed commercial/services lower level
- 1629 Other mixed uses in lower level

169 Other mixed level 2 uses

17 OUTDOOR AND OTHER URBAN AND BUILT-UP LAND

171 Outdoor cultural
- 1711 Botanical gardens and arboretums
- 1712 Zoos
- 1719 Other outdoor cultural

VI. LAND USE / LAND COVER CODES (continued)

1 URBAN AND BUILT-UP LAND (continued)

17 OUTDOOR AND OTHER URBAN AND BUILT-UP LAND (continued)

172 Outdoor public assembly
- 1721 Amphitheaters
- 1722 Drive-in movies
- 1723 Stadiums
- 1724 Racetracks
- 1725 Fairgrounds
- 1726 Amusement parks
- 1729 Other outdoor public assembly

173 Outdoor recreation
- 1731 Playground or play field
- 1732 Sports areas
- 1733 Water recreation/marina/launch
- 1734 Ski area
- 1735 Golf course
- 1736 Campground
- 1737 Parks and recreation
- 1739 Other outdoor recreation

174 Cemeteries

179 Other outdoor urban, built-up and developed

2 AGRICULTURAL LAND

21 CROPLAND AND PASTURE

211 Row crops (not including orchards and berries, code 22)
- 2111 Corn (land cover includes silage corn)
- 2112 Silage (corn or grasses)
- 2113 Potatoes
- 2114 Soybeans
- 2115 Alfalfa
- 2118 Other root crops
- 2119 Other non-root crops

212 Hay/rotation/permanent pasture
- 2121 Hay
- 2122 Rotation pasture
- 2123 Permanent pasture
- 2124 Idle agricultural land; use level 3 for land cover
- 2129 Other pasture

213 Grains (not including row crops)
- 2131 Wheat
- 2132 Oats
- 2133 Sorghum
- 2134 Barley
- 2139 Other grains (not including row crops)

VI. LAND USE / LAND COVER CODES (continued)

2 AGRICULTURAL LAND (continued)

22 ORCHARDS, BUSH FRUITS, VINEYARDS AND ORNAMENTAL HORTICULTURE

- 221 Tree fruits, regularly spaced (planted)
  - 2211 Apples
  - 2212 Pears
  - 2219 Other tree fruits
- 222 Bush and vine fruits (in rows)
  - 2221 Strawberries
  - 2222 Raspberries
  - 2223 Blueberries
  - 2224 Grapes
  - 2229 Other bush and vine fruits
- 223 Ornamental horticulture & nurseries (floriculture, etc.)
- 224 Tree farms
  - 2241 Christmas tree farms
  - 2242 Other coniferous trees
  - 2243 Deciduous trees
  - 2249 Other tree farms
- 225 Sugarbush
- 229 Other orchards, etc.

23 CONFINED FEEDING OPERATIONS

- 231 Livestock
  - 2311 Beef
  - 2312 Dairy
  - 2313 Swine
  - 2314 Lamb
  - 2319 Other livestock
- 232 Poultry
  - 2321 Chickens
  - 2322 Turkeys
  - 2329 Other
- 233 Pisciculture
  - 2331 Aquaculture
  - 2332 Fish hatchery
- 239 Other confined feeding operations

24 OTHER AGRICULTURAL LAND

- 241 Farmsteads
- 242 Greenhouses and mushroom houses
- 243 Stables & racetracks (non-commercial training areas)
- 244 Sod
- 249 Other

3 BRUSH OR TRANSITIONAL BETWEEN OPEN AND FORESTED

- 31 HERBACEOUS, NON-WOODY VEGETATION
- 32 SHRUB/BRUSH
- 33 MIXED HERBACEOUS AND SHRUB/BRUSH

VI. LAND USE /

4 FOREST ([nn] = Society of American Foresters type number)

LAND COVER CODES (continued)

41 BROADLEAF FOREST (generally deciduous)

411 Boreal hardwoods
- 4111 Aspen [16]
- 4112 Pin cherry [17]
- 4113 Paper birch [18]
- 4119 Other boreal hardwoods

412 Northern hardwoods
- 4121 Sugar maple [27]
- 4122 Sugar maple/Beech/Birch [25]
- 4123 Sugar maple/Basswood [26]
- 4124 Black cherry/Maple [28]
- 4125 Beech/Sugar maple [60]
- 4126 Red maple [28]
- 4129 Other northern hardwoods

413 Upland oaks
- 4131 Chestnut oak [44]
- 4132 White oak/Black oak/Northern red oak [52]
- 4133 White oak [53]
- 4134 Black oak [110]
- 4135 Northern red oak [55]
- 4136 Northern pin oak [14]
- 4139 Other upland oaks

414 Other hardwoods
- 4141 Gray birch/Red maple [19]
- 4142 Black ash/Elm/Red maple [39]
- 4143 Hawthorn [104]

42 CONIFEROUS FOREST (generally evergreen)

421 Boreal conifers
- 4211 Jack pine [1]
- 4212 Balsam fir [5]
- 4213 Black spruce [12]
- 4214 Black spruce/Tamarack [13]
- 4215 White spruce [107]
- 4216 Tamarack [38]
- 4219 Other boreal conifers

422 Spruce-fir types
- 4221 Red spruce [32]
- 4222 Red spruce/Balsam fir [33]
- 4223 Red spruce/Fraser fir [34]
- 4224 Northern white ceder [37]
- 4229 Other spruce-fir types

423 Pine and hemlock
- 4231 Red pine [15]
- 4232 White pine [21]
- 4233 White pine/Hemlock [22]
- 4234 Eastern hemlock [23]
- 4235 Eastern red ceder [46]
- 4239 Other pine and hemlock

VI. LAND USE / LAND COVER

4 FOREST (continued)

CODES (continued)

43 MIXED CONIFEROUS-BROADLEAF FOREST

431 Mixed spruce-fir types
- 4311 Red spruce/Yellow birch [30]
- 4312 Red spruce/Sugar maple/Beech [31]
- 4313 Paper birch/Red spruce/Balsam fir [35]

432 Mixed pine/hemlock types
- 4321 White pine/Red oak/Red maple [20]
- 4322 White pine/Chestnut oak [51]
- 4323 Hemlock/Yellow birch [24]

5 **WATER** (see 143 for transportation uses and 233 for agricultural uses)

51 RIVERS, CANALS AND OTHER WATERWAYS

511 Canal
512 River or Stream
519 Other waterways

52 LAKES AND PONDS

53 RESERVOIRS (and other artificial water surfaces)

54 BAYS/ESTUARIES

6 **WETLANDS** (not including open water)

61 FORESTED WETLAND

651 Broad-leaved deciduous
652 Needle-leaved deciduous
653 Broad-leaved evergreen
654 Needle-leaved evergreen
655 Dead

62 NON-FORESTED WETLAND

621 Aquatic bed
- 6211 Algal
- 6212 Aquatic moss
- 6213 Rooted vascular
- 6214 Floating

622 Moss-lichen
- 6221 Moss dominant (bogs and fens)
- 6222 Lichens dominant

623 Emergent wetland
- 6231 Wet meadow
- 6232 Shallow marsh
- 6233 Deep marsh

6 **WETLANDS (continued)**
62 NON-FORESTED WETLAND (continued)

CODES (continued)

624 Scrub-shrub wetland
- 6241 Broad-leaved deciduous
- 6242 Needle-leaved deciduous
- 6243 Broad-leaved evergreen
- 6244 Needle-leaved evergreen
- 6245 Dead

7 BARREN LAND

71 SALT FLATS

72 BEACHES AND RIVER BANKS

- 721 Sand beach
- 722 Gravel beach
- 723 River banks
- 729 Other beaches/riverbanks

73 SANDY AREAS (NON-BEACHES)

- 731 Sand dunes
- 732 Sand pits
- 739 Other sandy areas

74 BARE/EXPOSED ROCK

- 741 Rock knobs
- 742 Escarpments/Cliffs
- 743 Shoreline rock outcrop
- 744 Riverbank rock outcrop
- 749 Other bare/exposed rock

75 STRIP MINE/QUARRY OR GRAVEL PIT

79 OTHER BARREN LANDS

8 TUNDRA

81 SHRUB AND BRUSH

82 HERBACEOUS

83 BARE GROUND

84 WET

85 MIXED

9 PERMANENT SNOW & ICE (no known examples in the basin)

ATTRIBUTE DEFINITIONS AND CODES

I. PURPOSE

GIS attributes contain information associated with geographic features on a map. GIS users, present and future, will most likely need to develop attribute tables for new layers of map information. Existing attribute tables will need to be expanded as well. The LCBP GIS effort encourages Lake Champlain Basin users of geographic information to use the same base attributes and codes so that data can be more easily exchanged and understood by all users.

If you wish to develop a GIS data layer from scratch, please contact LCBP GIS to see what data already exists and for general guidance.

If you are developing new attributes for geographic data, please:

- contact the LCBP GIS Clearinghouse to see if standard attributes have already been defined,
- follow the guidelines below for developing new attributes,
- submit your attributes and codes to LCBP GIS Clearinghouse so that others may use them and for general review.

Individual users can expand on existing attributes as needed. If new attributes are defined and used, please inform LCBP GIS Clearinghouse so that attributes and codes remain consistent.

II. DEFINITIONS

Data Layer: A separate map layer with a specific theme (e.g., public wells), generally corresponding to an ARC/INFO "coverage".

Coverage: An ARC/INFO or other GIS data layer.

Map Feature: For these guidelines map features are represented as points, lines and areas. Areas are also called polygons. Examples are public wells (points), roads (lines) and wetlands (areas). Each individual feature (or element) corresponds to one *record* in a database.

Attribute: Information associated with each feature in a data layer. For example, depth (feet) is an attribute of public wells. An attribute corresponds to one *field* in a database record.

III. GENERAL GUIDELINES

Several factors should be considered when defining attributes and codes:

The Data Layer: The general purpose and feature types of the data layer should be well defined. What are the potential applications of the data layer? Has the same data layer been used elsewhere? Contact the LCBP GIS Service Center for help if you are just starting out.

Unique ID: Each geographic feature in a database generally has a unique identifier that singles it out from all other features of the same type. For example, each public well would have its own number. The unique identifier is usually the first attribute defined. It may be defined in part by the feature's location, such as by drainage basin. The unique ID should generally be different from the <cover>-ID, described next.

<cover>-ID: ARC/INFO automatically generates a <cover>-ID for each data layer. The <cover>-ID must be unique for each record of information; if not, other attribute values will be lost when **BUILD**ing or **CLEAN**ing the coverage. If you are not using ARC/INFO directly, you do not need to be concerned with the <cover>-ID.

Field width: Each attribute has a width, or number of characters (or bytes) for each entity (or record). For large data sets the attributes should be kept small. In any case, attributes should only be as wide as is needed for the largest possible value. For example, the longest county name in Vermont requires 10 characters, so a width of 10 would be used.

Description: Attributes must be explicitly defined to avoid confusion by other users.

Units: The units must always be given for measured values (e.g., feet or meters), either in the field name or in its description.

Codes: Codes can be used in place of full words or concepts. When possible use a mnemonic code, such as one or two letters which stand for something. Dropping vowels is a common practice. Codes can save considerable storage space in the database.

Possible Values: Many attributes have a specific set of allowable values, such as "1 to 99", or "three characters made up of a capital letter followed by two digits". The range of acceptable values should be explicitely given.

Complexity: Coding schemes should be easily understood by all users, even when more complex schemes would be computationally more efficient.

IV. DATA LAYER DESCRIPTION

The data layer should be described as fully as needed to develop the attributes (an example is given in Section VI). For GIS layers the description should at least include:

Coverage Name: The computer (or GIS) name for the data layer, limited to 8 characters by DOS. Letters, digits, underscores and dashes may be used; no spaces.

Data Layer: The data layer (or coverage) theme should be clearly described. Be sure to include any exceptional entities which should or should not be included.

Feature Type(s): The type of geographic feature should be specified as point, line or area. Areas are also referred to as polygons. If a coverage has two types of features (points and lines, or lines and polygons), the attributes for each feature type should be in separate lists.

Data Source(s): The source of the attribute data should be clear, <u>including the date of the information</u>. Different attributes may have different sources, so the attribute source may need to be included with individual attribute definitions. Include a bibliographic reference to a document and the scale of the source data if appropriate.

V. ATTRIBUTE

The definition for each attribute should include the following information. An

DEFINITIONS

example is given in Section VI.

Item or Field Name: The actual name used by the computer, with no spaces. Underscores can be used, but dashes are not compatible with dBase names. The display of attribute values in tables may be improved if short names are used for narrow attributes.

Generic names like "ID" and "CODE" should not be used because of potential conflicts between names when data layers are overlayed. In particular, item names "AREA", "PERIMETER" and "LENGTH" should be avoided as these names are generated by ARC/INFO.

Attribute: Several words describing the attribute.

Attribute Source: Used when different attributes come from different sources.

Description: Describes in detail (1) the contents of the attribute, (2) the possible values for the attribute, (3) all codes and their meanings (listed or footnoted), and (4) any notable exceptions. Upper, lower or mixed case should be specified for character fields. Often the description can be quite simple, but it should be complete.

Type: The type of data to be stored. The type can be used for both INFO and xBase, unless otherwise noted. Several special purpose INFO types are not listed here.

C	=	Character (includes digits and punctuation)
I	=	Integer (positive, negative or zero), INFO only; may be represented as type N with no decimal places, as in xBase
N	=	Real Number (decimal points are part of the width), xBase only.
B	=	Binary format integer, INFO only. The width is the number of bytes (generally 2 or 4).
F	=	Binary floating point real number (INFO only). The width is the number of bytes (generally 4 or 8).
L	=	Logical (True or False), xBase only. Logical items in INFO should be 1 character items, set to 'Y' or 'N'.
M	=	Memo field, for variable lengths of text, xBase only. Useful for notes and descriptions. In INFO, memo-type fields can be handled by having a separate documentation textfile.
D	=	Date

Width: The number of characters or bytes used by the attribute. See "Type". Note that the decimal point for type N must be included in the width.

Output Width: The number of spaces used for output. This is only required for INFO types B and F.

Decimal Places: The number of decimal places for types N and F, if used.

Example: An example of a typical value, if needed.

VI. EXAMPLE

SAMPLE ATTRIBUTE DEFINITIONS FOR WATERFALLS, CASCADES AND GORGES

Data Layer: Waterfalls, Cascades and Gorges from ANR publication (see Data Source)
Coverage Name: WATCASGO
Feature Type: Point locations
Data Source: All attribute information is contained in "Waterfalls, Cascades and Gorges of Vermont", 1988, VT Agency of Natural Resources (unless otherwise noted).

Attribute	Item Name	Description	Type	Width	Decimal Places	Example
Identifier	WCGID	basin#report# (basin = 1 to 17)	I	5	0	17999
Type	WCGTYPE	W, C or G (upper case)	C	1		W
Name	SITENAME	Site name (upper/lower)	C	25		Lana Falls
Length (ft)	WCGLENGTH	Horizontal	I	4	0	9999
Width (ft)	WIDTH	Horizontal	I	3	0	999
Total drop (ft)	TOTDROP	Vertical	I	3	0	999
Slope (degrees)	SLOPE	Avg, degrees	I	2	0	99
Wall height (ft)	WALLHT	Vertical	I	3	0	999
Rare plants	RAREPLANT	Y or N	C	1		Y
Flow regulated	FLOWREG	Y or N	C	1		N
Privacy	PRIVACY	see Codes	I	1	0	5
Tidiness	TIDY	see Codes	I	1	0	3
Swimmability	SWIM	see Codes	I	1	0	4

Codes

PRIVACY:
1 = not secluded or wild
2 = somewhat secluded
3 = wild
4 = very secluded
5 = private

TIDY:
1 = clean
2 = fairly clean
3 = a mess

SWIM:
1 = good bathing
2 = fair swimming
3 = good swimming
4 = great swimming

GEOGRAPHIC AREA CODES

I. PURPOSE

This document describes and provides standard codes for the various geographic areas located in the Lake Champlain Basin, including towns, townships, villages, cities, counties, census blocks and others. Any tables of data compiled by these geographic areas should use the codes listed here for GIS compatibility.

A primary use of these standard codes is to store and retrieve socioeconomic data. One function of GIS is to produce thematic maps (maps in which towns or other areas are shaded in based on the values for the different areas). An example would be a thematic map of population density for Lake Champlain Basin towns. Thematic maps will be easier to produce from various data sources if users adhere to these standard codes.

Standard codes and computer field names and formats are described for the following geographic areas:

- counties
- cities towns and villages
- census tracts, block groups and blocks
- regional planning commissions

The townships in Quebec in the LCB are similarly coded based on Statistics Canada census codes.

II. FIPS AND CENSUS CODES

Whenever possible, Federal Information Processing Standard (FIPS) codes and census codes set by the federal government will be used. A primary use for these codes will be to link geographic areas with the census data. The FIPS and census codes described below are detailed in *Census of Population and Housing, 1990; Public Law (P.L.) 94-171 Data Technical Documentation* (prepared by the Bureau of the Census, 301-763-2074).

The U.S. Bureau of the Census divides up the country with this hierarchy:

United States
  Region
    Division
      State[/Province]
        County
          County Subdivision
            Place (or part)
              Census Tract/Block Numbering Area (or part)
                Block Group (or part)
                  Block

The divisions listed above in **bold** are described in this document.

Some FIPS codes are used by the census for certain geographic areas. However, in some cases a *Census code* is used rather than a FIPS code to meet the needs of the census. The following table lists standard fields (i.e., attributes) developed by the census bureau to be used in New York, Quebec, and Vermont for the census divisions listed above.

Approximate geographic division	Field description	Computer field name	Field width/ type[1]	Examples
State	FIPS State	STUSAAB	2 N	50 = Vermont
County	FIPS County	CNTY	3 N	001 = Addison county
Town/city	FIPS MCD	COUSUBFP	5 N	09025 = Bristol town
Town/city	Census MCD[2]	COUSUBCE	3 N	015 = Bristol town
Village/city	FIPS Place	PLACEFP	5 N	08950 = Bristol village
Village/city	Census Place	PLACECE	4 N	0400 = Bristol village
Census tract/ block numbering area		TRACTBNA	6 N	6489
Census block group		BLCKGR	1 N	3
Census block		BLCK	4 C[3]	301

[1]Field types N = Numeric, C = character; numeric types are used (instead of the census "alpha/numeric") where possible.
[2]MCD = Minor Civil Division; generally a city or town.
[3]It may be that BLK can be a numeric (N) field; if there are no alphabetic suffixes to the 3-digit codes, users may prefer a numeric field type.

For a given town, city or village, census tract, block numbering area, block group, or block the above fields must be concatenated (joined together) with the state/province code to form a unique code within the basin. Vermont has a FIPS6 code. This is simply a concatenation of the CNTY & COUSUBCE fields.

The "approximate geographic divisions" of town, village and city are those defined by the Bureau of the Census, and are not necessarily incorporated areas. Census places (PLACECE) include both incorporated villages, cities and Census Designated Places (CDPs). The CDPs are identifiable by name, but are not legally incorporated places. The non-incorporated village of Middlebury VT is an example of a CDP.

Note that incorporated cities have both a unique 6-digit FIPS6 code, a PLACECE, and 5-digit COUSUBFP code. For example, Burlington VT has a unique FIPS6 code of 007 015, and a unique FIPS10 code of 007 015 0460 and a unique COUSUBFP code of 10675.

III. HISTORIC FIELD NAMES

Some census codes are called FIPS codes in common usage, although they are not actually FIPS codes. For example, the 6-digit field commonly used for Vermont towns, FIPS6, is made up of a 3-digit *FIPS* code for the county and a 3-digit *Census MCD* code for the town within the county. A federal FIPS MCD field, FIPSMCD, exists with a unique 5-digit code for each minor civil division in each state, but it has not been in common usage for GIS data in Vermont.

Census data have been historically distributed in Vermont with field names other than the standard names listed above. The user must exercise caution in determining which coding schemes imbedded in an historical database. These historic names are non-standard and are not recommended for use, but they are listed in the table below as an aid in understanding the available census data.

Standard field name	Historic name
CNTY	FIPSCNTY
COUSUBCE	FIPSMCD (note: this is *not* a FIPS code)
PLACECE	FIPSPLAC (note: this is *not* a FIPS code)

IV. CENSUS TRACTS, BLOCK NUMBERING AREAS, BLOCK GROUPS AND BLOCKS

Census tracts are statistical subdivisions of a county delineated by local census committees. Block numbering areas (BNAs) are subdivisions of a county for grouping and numbering blocks where local census committees have not established tracts. Tracts and BNAs are identified by a 4-digit basic number and may have a two-digit suffix separated by a period (the period is implied in digital data, and the suffix may be blank). Tract numbers range from 0001 through 9499.99, while BNAs range from 9501 through 9989.99. Both tracts and BNAs are coded in the field TRACTBNA (6 digits). The TRACTBNA code is unique within each county.

A census block group (field BLCKGR) is a cluster of blocks having the same first digit of the blocks' identifying numbers within their tract or BNA. For example, block group 3 includes all blocks numbered between 301 and 397. Block groups are only unique within each census place (PLACECE). Block groups generally contain between 250 and 550 housing units.

Census blocks (field BLCK) are small divisions for census data collection, numbered uniquely within each census tract or BNA. A block is identified by a 3-digit number, sometimes appended by a single alphabetical suffix. Census block codes are only unique within each census place (PLACECE).

V. LISTING OF STATE/PROVINCE & COUNTY CODES

State/Province FIPS codes (2-digit field STUSAAB):

36 New York
24 Quebec
50 Vermont

County FIPS codes (3-digit field CNTY):

New York:
019 Clinton
031 Essex
033 Franklin
113 Warren
115 Washington

Quebec:
None

Vermont:

001	Addison	015	Lamoille
003	Bennington	017	Orange
005	Caledonia	019	Orleans
007	Chittenden	021	Rutland
009	Essex	023	Washington
011	Franklin	025	Windham
013	Grand Isle	027	Windsor

VI. LISTING OF GEOGRAPHIC AREA CODES

GEOGRAPHIC AREA CODES

NAME	STUSA (PROV)	CNTY (CEN	COUSUBFP (CENSUBD)	PLACEFP	COUSUBC	PLACECE	FIPS6	FIPS10	RPC	PATHCODE
QUEBEC CANADA										
Bedford	24	46	40							
Bedford Village	24	46	35							
Bolton	24	46	65							
Dunham	24	46	50							
Frelighsburg	24	46	10							
Hemingford	24	68	10							
N.D.du Montreal-Carmel	24	56	25							
N.D.-de-Stanbridge	24	46	100							
Noyan (St. Thomas)	24	56	15							
Potton	24	46	30							
Rainville	24	46	110							
St. Armand West	24	46	20							
Phillipsburg Village	24	46	15							
St. Georges-de-Clarencevill	24	56	10							
St. gnace-de-Stanbridge	24	46	95							
St. Pierre-de-Veronne	24	46	25							
St. Sabine	24	46	105							
St. Sebastien	24	56	50							
Stanbridge	24	46	45							
Stanbridge Station	24	46	30							
Stuckley-Sud	24	45	105							
Sutton	24	46	778							
Sutton Village	24	46	55							
Abercom Village	24	46	5							
Venise-en-Quebec	24	56	5							

GEOGRAPHIC AREA CODES

NAME	STUSA	CNTY	COUSUBFP	PLACEFP	COUSUBC	PLACECE	FIPS6	FIPS10	RPC	PATHCODE
NEW YORK										
Clinton County	36	19								
Altona town	36	19	01583		5					
Altona CDP	36	19	01583	01572						
Au Sable town	36	19	03221		10					
Keeseville village (pt.)	36	19	03221	39089						
Beekmantown town	36	19	05485		15					
Black Brook town	36	19	06761		20					
Champlain town	36	19	13750		25					
Champlain village	36	19	13750	13739						
Rouses Point village	36	19	13750	63979						
Chazy town	36	19	14113		30					
Clinton town	36	19	16397		35					
Dannemora town	36	19	19653		40					
Dannemora village (pt.)	36	19	19653	19642						
Ellenburg town	36	19	23921		45					
Mooers town	36	19	48252		50					
Mooers village	36	19	48252	48241						
Peru town	36	19	57375		55					
Peru CDP	36	19	57375	57364						
Plattsburgh city	36	19	58574		60					
Plattsburgh city	36	19	58574	58574						
Plattsburgh town	36	19	58585		65					
Cumberland Head CDP	36	19	58585	19408						
Morrisonville CDP (pt.)	36	19	58585	48538						
Plattsburgh AFB CDP	36	19	58585	58596						
Plattsburgh West CDP	36	19	58585	58601						
Saranac town	36	19	65211		70					
Dannemora village (pt.)	36	19	65211	19642						
Schuyler Falls town	36	19	65717		75					
Morrisonville CDP (pt.)	36	19	65717	48538						
Essex County	36	31								
Chesterfield town	36	31	15330		5					
Keeseville village (pt.)	36	31	15330	39089						
Crown Point town	36	31	19246		10					
Elizabethtown town	36	31	23833		15					
Essex town	36	31	24768		20					
Jay town	36	31	38396		25					
Keene town	36	31	39067		30					

GEOGRAPHIC AREA CODES

NAME	STUSA	CNTY	COUSUBFP	PLACEFP	COUSUBC	PLACECE	FIPS6	FIPS10	RPC	PATHCODE
Lewis town	36	31	42114		35					
Moriah town	36	31	48428		45					
Mineville-Witherbee CDP	36	31	48428	47707						
Port Henry village	36	31	48428	59333						
North Elba town	36	31	51935		55					
Lake Placid village	36	31	51935	40761						
Saranac Lake village (pt.)	36	31	51935	65233						
North Hudson town	36	31	53077		60					
Ticonderoga town	36	31	73891		75					
Ticonderoga village	36	31	73891	73880						
Westport town	36	31	80775		80					
Westport village	36	31	80775	80764						
Willsboro town	36	31	82271		85					
Wilmington town	36	31	82315		90					
Franklin County	36	33								
Altamont town	36	33	01528		5					
Tupper Lake village	36	33	01528	75671						
Bellmont town	36	33	05716		15					
Brighton town	36	33	08213		30					
Franklin town	36	33	27243		65					
Harrietstown town	36	33	32314		70					
Saranac Lake village (pt.)	36	33	32314	65233						
Santa Clara town	36	33	65178		90					
Warren County	36	113								
Bolton town	36	113	07234		5					
Hague town	36	113	31335		20					
Horicon town	36	113	35639		25					
Lake George town	36	113	40519		35					
Lake George village	36	113	40519	40508						
Lake Luzerne town	36	113	40662		40					
Lake Luzerne-Hadley CDP	36	113	40662	40667						
Queensbury town	36	113	60356		45					
Glens Falls North CDP	36	113	60356	29338						
West Glens Falls CDP	36	113	60356	80082						
Warrensburg town	36	113	78300		60					
Warrensburg CDP	36	113	78300	78289						
Washington County	36	115								
Argyle town	36	115	02561		5					
Argyle village	36	115	02561	02550						

GEOGRAPHIC AREA CODES

NAME	STUSA	CNTY	COUSUBFP	PLACEFP	COUSUBC	PLACECE	FIPS6	FIPS10	RPC	PATHCODE
Dresden town	36	115	20885		15					
Fort Ann town	36	115	26715		25					
Fort Ann village	36	115	26715	26704						
Granville town	36	115	30037		35					
Granville village	36	115	30037	30026						
Hampton town	36	115	31885		45					
Hartford town	36	115	32457		50					
Hebron town	36	115	33040		55					
Kingsbury town	36	115	39650		65					
Hudson Falls village	36	115	39650	35980						
Putnam town	36	115	60092		70					
Whitehall town	36	115	81633		85					
Whitehall village	36	115	81633	81622						

GEOGRAPHIC AREA CODES

NAME	STUSA	CNTY	COUSUBFP	PLACEFP	COUSUBC	PLACECE	FIPS6	FIPS10	RPC	PATHCODE
VERMONT										
Addison Coun y	50	1								
ADDISON	50	1	325	0	5	0	1005	10050000	AC	ADDISON
BRIDPORT	50	1	8575	0	10	0	1010	10100000	AC	BRIDPORT
BRISTOL	50	1	9025	0	15	0	1015	10150000	AC	BRISTOL
BRISTOL VILLAGE	50	1	9025	8950	15	400	1015	10150400	AC	BRISTLVL
CORNWALL	50	1	16000	0	20	0	1020	10200000	AC	CORNWALL
FERRISBURG	50	1	26275	0	25	0	1025	10250000	AC	FERRSBRG
GOSHEN	50	1	28600	0	30	0	1030	10300000	AC	GOSHEN
GRANVILLE	50	1	29575	0	35	0	1035	10350000	TR	GRANVILL
HANCOCK	50	1	31525	0	40	0	1040	10400000	TR	HANCOCK
LEICESTER	50	1	39325	0	45	0	1045	10450000	AC	LEICSTER
LINCOLN	50	1	40075	0	50	0	1050	10500000	AC	LINCOLN
MIDDLEBURY (CDP)	50	1	44350	44275	55	1528	1055	10551528	AC	MDDLBRYP
MIDDLEBURY	50	1	44350	0	55	0	1055	10550000	AC	MIDDLBRY
MONKTON	50	1	45550	0	60	0	1060	10600000	AC	MONKTON
NEW HAVEN	50	1	48700	0	65	0	1065	10650000	AC	NEWHAVEN
ORWELL	50	1	53725	0	70	0	1070	10700000	AC	ORWELL
PANTON	50	1	53950	0	75	0	1075	10750000	AC	PANTON
RIPTON	50	1	59650	0	80	0	1080	10800000	AC	RIPTON
SALISBURY	50	1	62575	0	85	0	1085	10850000	AC	SALISBRY
SHOREHAM	50	1	65050	0	90	0	1090	10900000	AC	SHOREHAM
STARKSBORO	50	1	70075	0	95	0	1095	10950000	AC	STRKSBRO
VERGENNES CITY	50	1	74650	74650	100	2670	1100	11002670	AC	VERGENNS
WALTHAM	50	1	76075	0	105	0	1105	11050000	AC	WALTHAM
WEYBRIDGE	50	1	83275	0	110	0	1110	11100000	AC	WEYBRDGE
WHITING	50	1	83800	0	115	0	1115	11150000	AC	WHITING
Benning on coun y	50									
DORSET	50	3	17725	0	15	0	3015	30150000	BC	DORSET
PERU	50	3	55000	0	30	0	3030	30300000	BC	PERU
RUPERT	50	3	61000	0	45	0	3045	30450000	BC	RUPERT
Caledonia Coun y	50									
GROTON	50	5	30550	0	20	0	5020	50200000	NV	GROTON
HARDWICK VILLAGE	50	5	0	0	25	1110	5025	50251110	NV	HARDWCKV
HARDWICK	50	5	31825	0	25	0	5025	50250000	NV	HARDWICK
PEACHAM	50	5	54400	0	45	0	5045	50450000	NV	PEACHAM
SHEFFIELD	50	5	64075	0	60	0	5060	50600000	NV	SHEFFELD
STANNARD	50	5	69925	0	65	0	5065	50650000	NV	STANNARD
WALDEN	50	5	75700	0	75	0	5075	50750000	NV	WALDEN

GEOGRAPHIC AREA CODES

NAME	STUSA	CNTY	COUSUBFP	PLACEFP	COUSUBC	PLACECE	FIPS6	FIPS10	RPC	PATHCODE
WHEELOCK	50	5	83500	0	85	0	5085	50850000	NV	WHEELOCK
Chi enden Coun y	50									
BOLTON	50	7	6550	0	5	0	7005	70050000	CC	BOLTON
BUELS GORE	50	7	10300	0	10	0	7010	70100000	CC	BUELSGOR
BURLINGTON CITY	50	7	10675	10675	15	460	7015	70150460	CC	BURLNGTN
CHARLOTTE	50	7	13300	0	20	0	7020	70200000	CC	CHARLOTT
COLCHESTER	50	7	14875	0	25	0	7025	70250000	CC	COLCHSTR
ESSEX JUNCTION VILLAGE	50	7	24175	24400	30	840	7030	70300840	CC	ESSEXJCT
ESSEX	50	7	24175	0	30	0	7030	70300000	CC	ESSEX
HINESBURG	50	7	33475	0	35	0	7035	70350000	CC	HINESBRG
HUNTINGTON	50	7	34600	0	40	0	7040	70400000	CC	HUNTNGTN
JERICHO VILLAGE	50	7	36700	36625	45	1295	7045	70451295	CC	JERICHOV
JERICHO	50	7	36700	0	45	0	7045	70450000	CC	JERICHO
MILTON	50	7	45250	0	50	0	7050	70500000	CC	MILTON
MILTON VILLAGE	50	7	45250	45175	50	1580	7050	70501580	CC	MILTONVL
RICHMOND VILLAGE	50	7	0	0	55	2140	7055	70552140	CC	RICHMNDV
RICHMOND	50	7	59275	0	55	0	7055	70550000	CC	RICHMOND
ST. GEORGE	50	7	62050	0	60	0	7060	70600000	CC	STGEORGE
SHELBURNE	50	7	64300	0	65	0	7065	70650000	CC	SHELBURN
SOUTH BURLINGTON CITY	50	7	66175	66175	70	2420	7070	70702420	CC	SBURLGTN
UNDERHILL	50	7	73975	0	75	0	7075	70750000	CC	UNDERHIL
WESTFORD	50	7	80350	0	80	0	7080	70800000	CC	WESTFORD
WILLISTON	50	7	84475	0	85	0	7085	70850000	CC	WILLISTN
WINOOSKI CITY	50	7	85150	85150	90	3160	7090	70903160	CC	WINOOSKI
Franklin Coun y	50									
BAKERSFIELD	50	11	2500	0	10	0	11010	110100000	FG	BAKRSFLD
BERKSHIRE	50	11	5425	0	15	0	11015	110150000	FG	BERKSHRE
ENOSBURG	50	11	23875	0	20	0	11020	110200000	FG	ENOSBURG
ENOSBURG FALLS VILLAGE	50	11	23875	24025	20	820	11020	110200820	FG	ENSBRGFL
FAIRFAX	50	11	24925	0	25	0	11025	110250000	FG	FAIRFAX
FAIRFIELD	50	11	25225	0	30	0	11030	110300000	FG	FAIRFLD
FLETCHER	50	11	26500	0	35	0	11035	110350000	FG	FLETCHER
FRANKLIN	50	11	27100	0	40	0	11040	110400000	FG	FRANKLIN
GEORGIA	50	11	27700	0	45	0	11045	110450000	FG	GEORGIA
HIGHGATE	50	11	33025	0	50	0	11050	110500000	FG	HIGHGATE
MONTGOMERY	50	11	45850	0	55	0	11055	110550000	FG	MONTGMRY
RICHFORD VILLAGE	50	11	59125	59050	60	2120	11060	110602120	FG	RICHFRDV
RICHFORD	50	11	59125	0	60	0	11060	110600000	FG	RICHFORD
ST. ALBANS CITY	50	11	61675	61675	65	2240	11065	110652240	FG	STALBNCY

GEOGRAPHIC AREA CODES

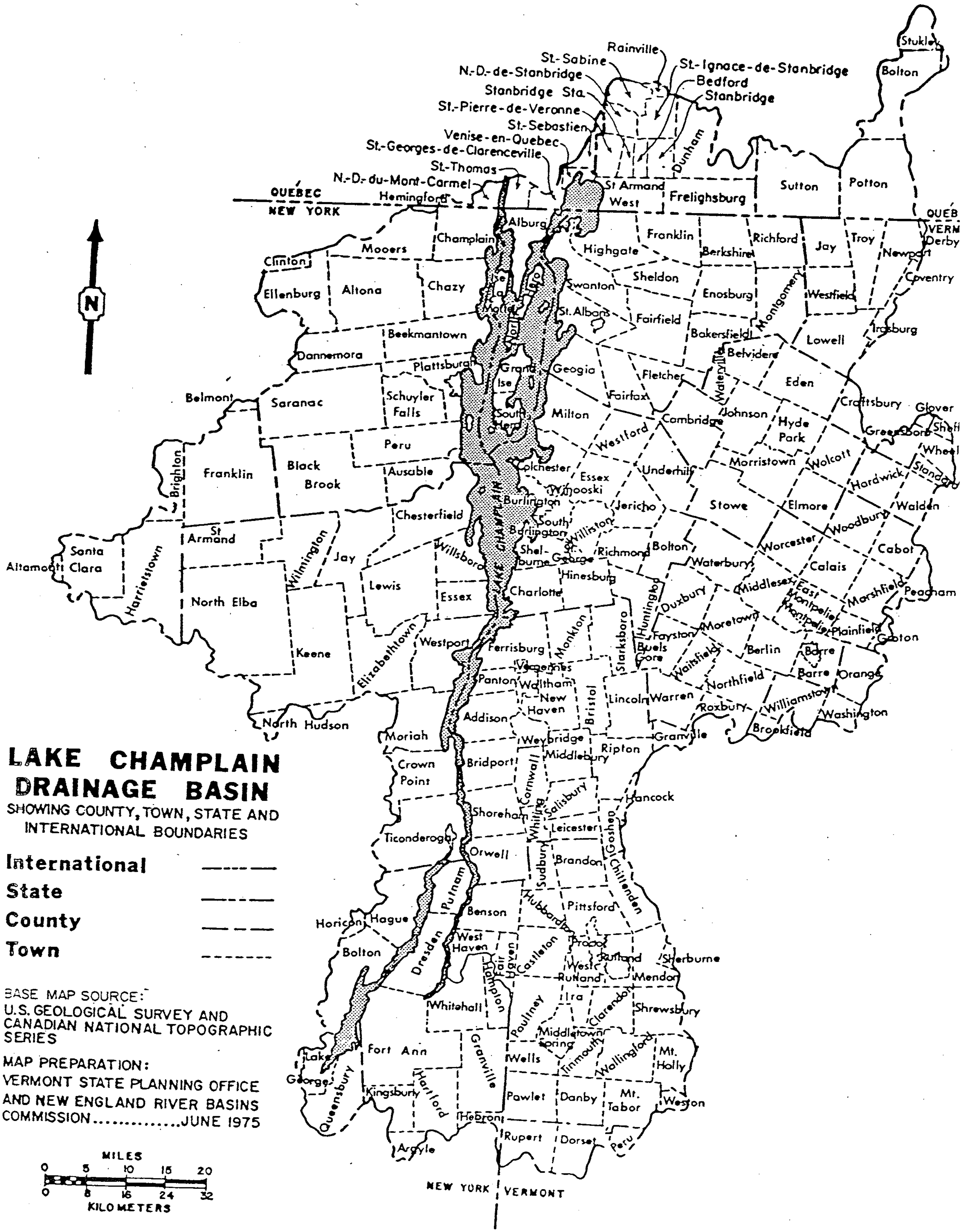
NAME	STUSA	CNTY	COUSUBFP	PLACEFP	COUSUBC	PLACECE	FIPS6	FIPS10	RPC	PATHCODE
ST. ALBANS TOWN	50	11	61750	0	70	0	11070	110700000	FG	STALBNTN
SHELDON	50	11	64600	0	75	0	11075	110750000	FG	SHELDON
SWANTON	50	11	71725	0	80	0	11080	110800000	FG	SWANTON
SWANTON VILLAGE	50	11	71725	71650	80	2580	11080	110802580	FG	SWANTONV
Grand Isle Coun y										
ALBURG	50	13	700	0	5	0	13005	130050000	FG	ALBURG
ALBURG VILLAGE	50	13	700	625	5	50	13005	130050050	FG	ALBURGVL
GRAND ISLE	50	13	29275	0	10	0	13010	130100000	FG	GRANDISL
ISLE LA MOTTE	50	13	35875	0	15	0	13015	130150000	FG	ISLAMOTT
NORTH HERO	50	13	50650	0	20	0	13020	130200000	FG	NHERO
SOUTH HERO	50	13	67000	0	25	0	13025	130250000	FG	SHERO
Lamoille Coun y										
BELVIDERE	50	15	4375	0	5	0	15005	150050000	LC	BELVIDER
CAMBRIDGE VILLAGE	50	15	11500	11425	10	510	15010	150100510	LC	CAMBRDGV
CAMBRIDGE	50	15	11500	0	10	0	15010	150100000	LC	CAMBRIDG
JEFFERSONVILLE VILLAGE	50	15	11500	36475	10	1280	15010	150101280	LC	JFFRSNVL
EDEN	50	15	23500	0	15	0	15015	150150000	LC	EDEN
ELMORE	50	15	23725	0	20	0	15020	150200000	LC	ELMORE
HYDE PARK VILLAGE	50	15	35050	34975	25	1200	15025	150251200	LC	HYDEPRKV
HYDE PARK	50	15	35050	0	25	0	15025	150250000	LC	HYDEPARK
JOHNSON VILLAGE	50	15	37075	37000	30	1310	15030	150301310	LC	JOHNSONV
JOHNSON	50	15	37075	0	30	0	15030	150300000	LC	JOHNSON
MORRISVILLE VILLAGE	50	15	46675	46825	35	1650	15035	150351650	LC	MORRISVL
MORRISTOWN	50	15	46675	0	35	0	15035	150350000	LC	MORRISTN
STOWE VILLAGE	50	15	70525	70450	40	2510	15040	150402510	LC	STOWEVL
STOWE	50	15	70525	0	40	0	15040	150400000	LC	STOWE
WATERVILLE	50	15	77425	0	45	0	15045	150450000	LC	WATERVIL
WOLCOTT	50	15	85375	0	50	0	15050	150500000	LC	WOLCOTT
Orange Coun y	50									
BROOKFIELD	50	17	9325	0	15	0	17015	170150000	TR	BROOKFLD
ORANGE	50	17	53425	0	40	0	17040	170400000	CV	ORANGE
WASHINGTON	50	17	76750	0	75	0	17075	170750000	CV	WASHNGTN
WILLIAMSTOWN	50	17	84175	0	85	0	17085	170850000	CV	WILLMSTN
Orleans Coun y	50									
CRAFTSBURY	50	19	16300	0	30	0	19030	190300000	NV	CRFTSBRY
GLOVER	50	19	28075	0	40	0	19040	190400000	NV	GLOVER
GREENSBORO	50	19	30175	0	45	0	19045	190450000	NV	GRNSBORO
IRASBURG	50	19	35575	0	55	0	19055	190550000	NV	IRASBURG
JAY	50	19	36325	0	60	0	19060	190600000	NV	JAY

GEOGRAPHIC AREA CODES

NAME	STUSA	CNTY	COUSUBFP	PLACEFP	COUSUBC	PLACECE	FIPS6	FIPS10	RPC	PATHCODE
LOWELL	50	19	40525	0	65	0	19065	190650000	NV	LOWELL
NEWPORT TOWN	50	19	48925	0	80	0	19080	190800000	NV	NEWPRTWN
TROY	50	19	73525	0	85	0	19085	190850000	NV	TROY
NORTH TROY VILLAGE	50	19	73525	52075	85	1790	19085	190851790	NV	NTROYVIL
WESTFIELD	50	19	80200	0	90	0	19090	190900000	NV	WESTFLD
Ru land Coun y	50									
BENSON	50	21	5200	0	5	0	21005	210050000	RR	BENSON
BRANDON	50	21	7750	0	10	0	21010	210100000	RR	BRANDON
BRANDON (CDP)	50	21	7750	7675	10	319	21010	210100319	RR	BRANDONP
CASTLETON	50	21	11950	0	15	0	21015	210150000	RR	CASTLETN
CHITTENDEN	50	21	14350	0	20	0	21020	210200000	RR	CHITTNDN
CLARENDON	50	21	14500	0	25	0	21025	210250000	RR	CLARNDON
DANBY	50	21	16825	0	30	0	21030	210300000	RR	DANBY
FAIR HAVEN (CDP)	50	21	25375	25450	35	877	21035	210350877	RR	FAIRHAVP
FAIR HAVEN	50	21	25375	0	35	0	21035	210350000	RR	FAIRHAVN
HUBBARDTON	50	21	34450	0	40	0	21040	210400000	RR	HUBBRDTN
IRA	50	21	35425	0	45	0	21045	210450000	RR	IRA
MENDON	50	21	44125	0	50	0	21050	210500000	RR	MENDON
MIDDLETOWN SPRINGS	50	21	44800	0	55	0	21055	210550000	RR	MIDLTNSP
MOUNT HOLLY	50	21	47200	0	60	0	21060	210600000	RR	MTHOLLY
MOUNT TABOR	50	21	47425	0	65	0	21065	210650000	RR	MTTABOR
PAWLET	50	21	54250	0	70	0	21070	210700000	RR	PAWLET
PITTSFORD	50	21	55600	0	80	0	21080	210800000	RR	PITTSFRD
PITTSFORD VILLAGE	50	21	0	0	80	1940	21080	210801940	RR	PTTSFRDV
POULTNEY VILLAGE	50	21	56875	56800	85	2000	21085	210852000	RR	PLTNEYVL
POULTNEY	50	21	56875	0	85	0	21085	210850000	RR	POULTNEY
PROCTOR	50	21	57250	0	90	0	21090	210900000	RR	PROCTOR
RUTLAND CITY	50	21	61225	61225	95	2210	21095	210952210	RR	RUTLNDCY
RUTLAND TOWN	50	21	61300	0	100	0	21100	211000000	RR	RUTLNDTN
SHERBURNE	50	21	64825	0	105	0	21105	211050000	RR	SHERBURN
SHREWSBURY	50	21	65275	0	110	0	21110	211100000	RR	SHRWSBRY
SUDBURY	50	21	71050	0	115	0	21115	211150000	RR	SUDBURY
TINMOUTH	50	21	72925	0	120	0	21120	211200000	RR	TINMOUTH
WALLINGFORD (CDP)	50	21	75925	75850	125	2738	21125	211252738	RR	WLNGFRDP
WALLINGFORD	50	21	75925	0	125	0	21125	211250000	RR	WALNGFRD
WELLS	50	21	77950	0	130	0	21130	211300000	RR	WELLS
WEST HAVEN	50	21	80875	0	135	0	21135	211350000	RR	WESTHAVN
WEST RUTLAND	50	21	82300	0	140	0	21140	211400000	RR	WRUTLAND
WEST RUTLAND (CDP)	50	21	82300	82375	140	2978	21140	211402978	RR	WRUTLNDP

GEOGRAPHIC AREA CODES

NAME	STUSA	CNTY	COUSUBFP	PLACEFP	COUSUBC	PLACECE	FIPS6	FIPS10	RPC	PATHCODE
Washing on Coun y	50									
BARRE CITY	50	23	3175	3175	5	140	23005	230050140	CV	BARRECTY
SOUTH BARRE (CDP)	50	23	3250	66025	10	2418	23010	230102418	CV	SBARREPL
GRANITEVILLE EAST BARRE (C	50	23	3250	29425	10	1010	23010	230101010	CV	EBARREPL
BARRE TOWN	50	23	3250	0	10	0	23010	230100000	CV	BARRETWN
BERLIN	50	23	5650	0	15	0	23015	230150000	CV	BERLIN
CABOT	50	23	11125	0	20	0	23020	230200000	CV	CABOT
CABOT VILLAGE	50	23	11125	11050	20	480	23020	230200480	CV	CABOTVIL
CALAIS	50	23	11350	0	25	0	23025	230250000	CV	CALAIS
DUXBURY	50	23	18550	0	30	0	23030	230300000	CV	DUXBURY
EAST MONTPELIER	50	23	21925	0	35	0	23035	230350000	CV	EMONTPLR
FAYSTON	50	23	25825	0	40	0	23040	230400000	CV	FAYSTON
MARSHFIELD	50	23	43600	0	45	0	23045	230450000	CV	MARSHFLD
PLAINFIELD VILLAGE (p .)	50	23	0	0	45	1960	23045	230451960	CV	PLNFLDVL
MARSHFIELD VILLAGE	50	23	43600	43525	45	1510	23045	230451510	CV	MRSHFLDV
MIDDLESEX	50	23	44500	0	50	0	23050	230500000	CV	MIDDLSEX
MONTPELIER CITY	50	23	46000	46000	55	1610	23055	230551610	CV	MONTPLR
MORETOWN	50	23	46225	0	60	0	23060	230600000	CV	MORETOWN
NORTHFIELD	50	23	50275	0	65	0	23065	230650000	CV	NORTHFLD
NORTHFIELD VILLAGE	50	23	50275	50200	65	1770	23065	230651770	CV	NRTHFLDV
PLAINFIELD	50	23	55825	0	70	0	23070	230700000	CV	PLAINFLD
PLAINFIELD VILLAGE (p .)	50	23	0	0	70	1960	23070	230701960	CV	PLNFLDVL
ROXBURY	50	23	60625	0	75	0	23075	230750000	CV	ROXBURY
WAITSFIELD	50	23	75325	0	80	0	23080	230800000	CV	WAITSFLD
WARREN	50	23	76525	0	85	0	23085	230850000	CV	WARREN
WATERBURY	50	23	76975	0	90	0	23090	230900000	CV	WATERBRY
WATERBURY VILLAGE	50	23	76975	76900	90	2800	23090	230902800	CV	WTRBRYVL
WOODBURY	50	23	85525	0	95	0	23095	230950000	CV	WOODBURY
WORCESTER	50	23	86125	0	100	0	23100	231000000	CV	WORCSTER


LAKE CHAMPLAIN DRAINAGE BASIN
SHOWING COUNTY, TOWN, STATE AND INTERNATIONAL BOUNDARIES
International
State
County
Town
BASE MAP SOURCE:
U.S. GEOLOGICAL SURVEY AND CANADIAN NATIONAL TOPOGRAPHIC SERIES
MAP PREPARATION:
VERMONT STATE PLANNING OFFICE AND NEW ENGLAND RIVER BASINS COMMISSION.............JUNE 1975
MILES
0 5 10 15 20
0 8 16 24 32
KILOMETERS
N
QUEBEC
NEW YORK
VERMONT
LAKE CHAMPLAIN
Rainville
St-Sabine
St-Ignace-de-Stanbridge
N.-D.-de-Stanbridge
Bedford
Stanbridge Sta.
Stanbridge
St.-Pierre-de-Veronne
St.-Sebastien
Venise-en-Quebec
St.-Georges-de-Clarenceville
St.-Thomas
N.-D.-du-Mont-Carmel
Hemingford
Dunham
St Armand West
Frelighsburg
Sutton
Potton
Stukley
Bolton
Alburg
Champlain
Mooers
Clinton
Ellenburg
Altona
Chazy
Highgate
Franklin
Berkshire
Richford
Jay
Troy
Newport
Derby
Coventry
Swanton
Sheldon
Enosburg
Montgomery
Westfield
St. Albans
Fairfield
Bakersfield
Lowell
Irasburg
Beekmantown
Dannemora
Plattsburgh
Grand Isle
Georgia
Fletcher
Belvidere
Waterville
Eden
Belmont
Saranac
Schuyler Falls
South Hero
Milton
Fairfax
Johnson
Hyde Park
Craftsbury
Glover
Greensboro
Cambridge
Westford
Peru
Morristown
Wolcott
Hardwick
Brighton
Franklin
Black Brook
Ausable
Colchester
Essex
Winooski
Underhill
Burlington
Jericho
Stowe
Elmore
Walden
Chesterfield
South Burlington
Williston
Woodbury
St Armand
Santa Clara
Altamont
Harrietstown
Wilmington
Jay
Willsboro
Shelburne
St. George
Richmond
Bolton
Waterbury
Worcester
Calais
Cabot
Lewis
Essex
Charlotte
Hinesburg
Huntington
Duxbury
Middlesex
East Montpelier
Montpelier
Marshfield
Peacham
North Elba
Keene
Elizabethtown
Westport
Ferrisburg
Monkton
Starksboro
Fayston
Moretown
Plainfield
Groton
Buels Gore
Waitsfield
Berlin
Barre
Vergennes
Panton
Waltham
New Haven
Bristol
Lincoln
Warren
Northfield
Orange
Addison
Roxbury
Williamstown
Washington
North Hudson
Moriah
Weybridge
Granville
Brookfield
Crown Point
Bridport
Cornwall
Middlebury
Ripton
Salisbury
Hancock
Shoreham
Whiting
Leicester
Goshen
Ticonderoga
Orwell
Sudbury
Brandon
Chittenden
Putnam
Benson
Hubbardton
Pittsford
Horicon
Hague
West Haven
Fair Haven
Castleton
Proctor
Rutland
Sherburne
Bolton
Dresden
Hampton
West Rutland
Mendon
Whitehall
Poultney
Ira
Clarendon
Shrewsbury
Middletown Springs
Fort Ann
Granville
Wells
Tinmouth
Wallingford
Mt. Holly
Lake George
Queensbury
Kingsbury
Hartford
Pawlet
Danby
Mt. Tabor
Weston
Hebron
Rupert
Dorset
Peru
Argyle
NEW YORK
VERMONT

PART 2: GUIDELINES - SECTION B
DIRECTORY STRUCTURE AND FILE NAMES

I. PURPOSE

These procedures are primarily for ARC/INFO system managers and technicians. The directory and file naming conventions recommended here are used by the LCBP GIS Service Center. Communications with the Service Center and user support will be simplified if these conventions are followed. Any applications, software or SMLs/AMLs developed by the Service Center will use these conventions. Most importantly, if these conventions are followed, this document can serve as part of your system documentation for future users at your site.

Although this document uses pc-DOS pathnames, the discussion applies to any operating system. The term "<COVER>" is used throughout to denote an ARC/Info coverage name.

II. DIRECTORY STRUCTURE

1. Several considerations dictate the choice of a directory structure:
   a. Subdirectories should not be too large. A good rule of thumb is to have no more than 25 to 50 files in a subdirectory.
   b. ARC coverages are often divided up geographically into tiles. For example, soils data for Vermont are divided up by county.
   c. Commonly used ARC coverages, such as logos and scale bars, can be grouped together for simple access.

2. Community directories:
   The LCBP GIS Service Center uses a separate top-level directory for each community. These pathnames are listed in "Geographic Area Codes", *LCBP GIS Handbook*, Part 1, Standards, Section E. For example, coverage RDS (roads) for Plainfield might be stored with pathname PLAINFLD\RDS.

3. Directory D:\LEGEND:
   Directory D:\LEGEND is used to store north arrows, scale bars, and other legend coverages. Variations are serially numbered (NORTH1, NORTH2, etc.) or named by the scale (SCL5000, SCL1250, etc.).

4. Directory D:\NY or D:\VT:
   The directory D:\VT or D:\NY is used for statewide coverages of general utility. Many SMLs reference this directory. Some of the coverages in D:\VT include:

TB250	Town boundaries at 1:250,000
CNTY250	County boundaries at 1:250,000
STATE250	State boundary at 1:250,000
RPC250	Regional planning commission boundaries, 1:250,000
ORGRID	1:5000 orthophoto edges
GRID	U.S.G.S. 7.5 minute quadrangle edges
DLGTILES	Digital Line Graph tile boundaries

5. Regional directories:
   LCPB GIS Service Center foresees the use of regional directories LCBP for basinwide coverage; ACRPC, BCRC, CCRPC, CVRPC, FGIRPDC, LCPC, NVDA, RRPC, SWCRPDC, TRORC, UVLSC and WRC for each of the Vermont Regional Planning Commissions and New York's County Planning Offices. The

corresponding Regional Service Centers are named in "Geographic Area Codes" *LCBP GIS Handbook*, Part 1, Standards, Section E.

6. Directory D:\PROJECT
Files used with the ARC/Info PROJECT command are kept here. Their names reflect the source and destination coordinate systems. For example, file LLDMS-SP.SML is for projecting from Latitude-Longitude (in Degrees-Minutes-Seconds) to State Plane coordinates. These files are available from the Service Center, as described in the *LCPB Data Catalog*.

III. FILE AND COVERAGE NAMES

1. Several considerations dictate the choice of ARC file names:

a. All files directly associated with a coverage should have the same name as the coverage with a different extension. When coverages are copied or EXPORTed, all files with the same name will automatically be copied with the coverage. Also, you will know which coverage the file is associated with when you (or someone else) finds it much later.

b. pc-DOS 5.0 limitations should be followed since most LCBP GIS sites use pcARC/Info.

c. The limitations of dBase should taken into consideration, since most sites will be using pcARC/Info version 3.4D.

d. INFO file names should have no extension, since dBase files will be given the .DBF extension. The exception is for INFO files related to ARC/Info coverages; these INFO files can have extensions up to 8 characters long since the extension becomes the dBase file name (for example, INFO file <COVER>.DATA becomes dBase file DATA.DBF in the coverage's subdirectory).

e. It is generally unnecessary to pluralize a coverage name with 'S', since the plural can be assumed. For example, use coverage name LAKE instead of LAKES.

f. File naming conventions (n is a digit from 1 to 9)

g. Coverage names: <COVER>
Coverage names follow pc-DOS 5.0 conventions: 8 characters or less, not including . " / \ [ ] : | < > + = ; ,. If a data layer exists at two different scales (rarely the case), the smaller (less accurate) scale may be indicated in its coverage name. For example, TB250 is used for town boundaries at 1:250,000.

i. Lookup tables: <COVER>.LUn
Note that one coverage may have several lookup tables. A single lookup table for a coverage may be named <COVER>.LUT.

j. Documentation files:
Documentation files exist for the coverage, attribute tables, annotations and coverage updates. See "Data Layer Documentation", *LCBP GIS Handbook*, Part 1, Standards, Section B for detailed information.

i. FREQUENCY files: <COVER>.Fn

j. STATISTICS files: <COVER>.Sn
Files generated by the FREQUENCY and STATISTICS commands are often left behind without being erased. These extensions make it easy to identify them.

k. Temporary files: <FILE>.TMP
These are for temporary use and can be deleted when you've forgotten what they're for.

l. Files JUNK.* or coverage JUNK:
If you want a temporary file or coverage with no future use, name it JUNK. Other users will know they can delete it.

LOCATIONAL DATA STANDARDS

I. PURPOSE

The purpose of this document is to establish standards for developing locational data and metadata for inclusion in the Lake Champlain Basin Program Geographic Information System (LCBP-GIS). This document addresses a very narrow set of circumstances for which the more general standards detailed in Sections A-E of the handbook may not be appropriate. This document does not replace or supersede those sections, and should only be used under the circumstances described in the section called "Applicability of This Document," which appears below.

II. APPLICABILITY OF THIS DOCUMENT

The Technical Advisory Committee (TAC) of the Lake Champlain Management Council (LCMC) has adopted standards for the development and documentation of spatial data in the Lake Champlain Basin. Those standards are detailed in Sections A-E of this handbook, and they should be adhered to whenever a project funded by the LCMC will result in the development of spatial data. However, the TAC recognizes that certain projects will result in data that, while it is not primarily spatial in nature, can and should be described in spatial terms so that it can be included in the LCBP-GIS. As an example, water quality monitoring stations in the Lake are producing a stream of data that is highly valuable to water chemists and biologists. If reasonably accurate locational data can be collected for the stations in a coordinate system that is compatible with the LCBP-GIS, then the data being generated by the stations could be tagged to a mappable location and played easily against other geographically coincident data. The water quality data would thus be enhanced in a way that would make it more easily available for environmental modelling applications.

The locational data standards described herein are intended to be used, then, in those situations where location is an enhancement of the data, rather than a primary feature of the data. The standards aim to ensure that any spatial attributes that are collected will be of the highest quality that can be attained within reasonable limits, while not introducing so many additional procedures and rules as to become impractical. The intention is to provide a set of standards that will facilitate the inclusion of as many data layers as possible into the LCBP-GIS, without introducing into it unacceptable levels of error and uncertainty.

Because linear and areal features (lines and polygons) present more complex data management problems than does simple point data, the term "locational data," as used in this document, refers to x-y coordinates (and attendant attributes) that describe point locations on the ground. These locational data standards should not be applied to line or polygon data.

III. OVERVIEW

The standards described in this document cover four broad topics: **data collection**, which includes a discussion of common field collection techniques as well as common data conversion techniques; **cartographic representation**, which includes map coordinate systems and map projections; **data documentation**, a slimmed-down version of Section B of this Handbook; and **digital file formats**, which discusses acceptable formats for importing locational data into the LCBP-GIS.

IV. DATA COLLECTION

For purposes of this discussion, we will consider the two most common ways that locational data can be collected for inclusion in a GIS. The first, which is probably most relevant here, is direct field collection. Stated simply, this method involves going out to the field and using some kind of a survey instrument to measure where a thing is. The thing might be a building, a stream gaging station, a manure pile, or what have you. The survey instrument used may range in complexity from a tape measure to a global positioning system (GPS) receiver. It doesn't really matter which instrument is used, the concept is the same: the position of the thing being measured is fixed on the surface of the earth relative to some number (preferably more than one!) of known positions. The accuracy of the resultant locational data then, is dependant on two factors: how well-defined are the locations of the known or reference positions, and how accurately is the measuring done.

Of course, there are practical differences among the various types of survey instruments that can be used, but these differences can be thought of quite simply in terms of how much the machine does versus how much the operator does. If you use a tape measure to determine location, you will be obliged to find well-defined reference points and to do all the necessary trigonometry involved in deriving an x-y coordinate from a series of lengths and angles. If you use a GPS receiver, the receiver "knows" the location of the reference points (satellites in well-defined orbits) and does all the math for you. The tape measure and the GPS receiver represent opposite ends of the spectrum; there are many hues in between: simple optical theodolites vs. "total stations," LORAN vs. GPS, and on and on.

But our concern at the moment is to define the minimum requirements for obtaining locational data that is acceptable for inclusion in the LCBP-GIS. For our purposes, locational data derived with a tape measure is just as acceptable as locational data derived with a GPS receiver, or any other instrument that gives the desired result. The important criteria number only two:

1. Locational data must be expressed in a coordinate system that is compatible with the LCBP-GIS. Either the State Plane Coordinate System (SPCS) or the Universal Transverse Mercator (UTM) system is acceptable. In either case, the standard unit is meters. Latitude and longitude values are also acceptable. (Note: these coordinate systems will be discussed in more detail in Section V of this document.)

2. The accuracy of the locational data must be well-defined and clearly documented. THIS IS IMPORTANT! Your data will be less than useless and not acceptable for inclusion in the LCBP-GIS unless its accuracy can be well-understood by all who will use it. The only way to make it so is to document it well.

Besides direct field measurement, locational data can also be obtained in some cases by a process called data conversion. This process requires that the location of the object in question has already been fixed on a map. The problem then becomes converting the paper map to a digital form that is compatible with the LCBP-GIS. The best way to accomplish the conversion is by digitizing the map, a process that is described in detail in Section A of this Handbook. In some cases, it may be acceptable to simply record positions on a map by interpolating from the map's coordinate grid, type the resulting locational data into an appropriately structured computer file, and submit the file to the LCBP-GIS Service Center for inclusion in the LCBP-GIS. In either case, meticulous documentation of the process you use is paramount. Without it, your data simply cannot be included in the LCBP-GIS.

V. CARTOGRAPHIC REPRESENTATION

The earth is a sphere; maps are flat. There you have the fundamental complication involved in map-making. This document is not the place for a detailed discussion of how cartographers go about dealing with that fundamental complication, but knowing nothing at all about the topic will almost guarantee gross errors and useless data so grab a cup of your favorite beverage and read on: you need to know this.

In order to make a map, a cartographer projects some portion of the earth's spherical surface onto a flat plane. There are various algorithms for doing so. Mercifully, we need not be concerned with the details of any of them here. Suffice it to say that different projections serve different purposes; some are useful for navigation, some for showing clearly the relative areas of land masses or other objects, and others are useful for showing clearly the relative distances between objects. The important thing to know about map projections, for our purposes, is that they make it possible to put coordinate systems on maps that reference locations on the earth's surface. What we are mainly concerned with here is a basic understanding of coordinate systems and how they are used.

But first a brief word about latitude and longitude (lat-long). Although lat-long obviously can be, and frequently is, displayed on maps, it is not really a map coordinate system. Lat-long is a spherical coordinate system that describes the angle of an imaginary line drawn from the center of the earth to a point on the surface of the earth. As a spherical coordinate system, lat-long lines must always be somewhat distorted when they are drawn on a flat map. GPS and LORAN systems frequently give coordinates in lat-long. The LCBP-GIS can accept lat-long coordinates, and project them into one of the standard coordinate systems described below.

There are two main coordinate systems used in the Lake Champlain Basin: the State Plane Coordinate System (SPCS) and the Universal Transverse Mercator (UTM) coordinate system. Both have their mathematical roots in the transverse mercator projection. There are several important things to know about them. First, they are traditional cartesian systems. That is to say, they are based on an imaginary grid of north-south running lines and east-west running lines. Coordinates are given as cartesian pairs representing the intersection of one imaginary north-south line with one imaginary east-west line. Familiar enough to anyone who ever took high school geometry.

The second important point is that they are both zonal systems. This takes a bit of explaining. Consider the example of a coordinate given as east 800 meters by north 1300 meters. Another way of stating that position would be "at the intersection of a north-south running line that is 800 meters east of the origin and an east-west running line that is 1300 meters north of the origin." But what is the origin? As it happens, the origin is just an arbitrarily determined point on the ground. Well, almost arbitrary. In the interests of keeping life simple, surveyors long ago decided that, a. they only wanted to deal with "eastings" and "northings," never "westings" and "southings" and, b. no negative numbers. That implies that the origin always has to be somewhere southwest of the area being mapped. It also implies that, as soon as you start mapping far away from the origin, you start having to deal with very large numbers. The solution, at least for the coordinate systems of interest to us, was to create several zones, each with its own origin. That way it is always possible to select an appropriate zone to ensure that every location within the area being mapped can be expressed as a reasonably small positive easting and a reasonably small positive northing.

It is readily apparent, then, that knowing which zone a map is using is critical to interpreting coordinates derived from that map. This becomes especially important in areas that are covered by multiple overlapping zones. As it happens, that's us. Three zones overlap and are commonly used within the basin: the New York SPCS East, The Vermont SPCS and the UTM Zone 18. Here is an example that hopefully will clarify the importance of understanding map coordinate systems. There is a small airstrip in Shelburne, VT. Table 1 describes its (very) approximate location expressed in meters using three separate coordinate systems, all of which are entirely valid for doing work in the basin.

	Coordinates (In meters)	
Zone	North	East
VTSPCS	207,000	94,000
NYSPCS-East	484,000	237,000
UTM-Zone 18	4,135,000	641,000

Table 1: Approximate locations of Shelburne Airstrip given in 3 different coordinate systems.

The table makes clear the fact that large errors and uninterpretable data can easily result from a misunderstanding of coordinate systems. There is a single reason for all this droning on about cartographic theory and coordinate systems: to make the reader understand the importance of knowing, and documenting clearly, which coordinate system is being used to express a given set of locational data.

VI. DATA DOCUMENTATION

Section 1:B of this Handbook describes in detail the data layer documentation standard for the LCBP-GIS. Most of that standard is Arc/Info specific, and does not necessarily apply to the sort of locational data we are concerned with here. Nevertheless, parts of the standard are applicable and can be used as is. Doing so will ensure that locational data documentation is automatically in a format that is compatible with all the other data documentation in the system. What follows, then, is a slightly modified extract of the full standard. The assumption throughout is that locational data consists of a set of x-y coordinates and identifiers (possibly with attendant attribute information) that describe a corresponding set of point locations on the ground.

Locational data cannot be accepted for inclusion in the LCBP-GIS unless it is documented. At a minimum, that documentation should fully describe the data, its attributes, its development and its updates over time. Data layer documentation should be maintained in a database management system. The documentation file formats described here assume the use of xBase software on a PC but, with slight modification, they can be fit into virtually any database management software on any computing platform.

The main documentation file is the DOC file. Using the xBase convention, it would be named DOC.DBF. A definition of the file structure follows:

FILE STRUCTURE FOR DOC.DBF:

Field	Field Name	Type	Width	Dec	Description
DATA LAYER SUMMARY FIELDS:					
1	ENTRY_DATE	Date	8		Date of this entry; updated when any changes are made to the documentation or to the data. (The data layer *version* is given by the LASTUPDATE field.)
2	COVERNAME	Char	8		Maximum of 8 characters; the name by which the GIS data layer containing this data will be known.
3	DATALAYER	Char	60		A one line description of the data layer.
4	FEATURE	Char	50		The GIS feature type; for locational data, this will always be "point".
5	STATUS	Num	2	0	1: Available for distribution at LCBP Data Clearinghouse; meets all applicable LCBP-GIS standards. 2: Not available for distribution due to some restrictions (check with data manager listed in the documentation for more information). 4: In development. Data conversion to GIS format is in process. 5: These data are planned but data conversion has not yet begun. 6: Available for distribution in a pre-release version from LCBP Data Clearinghouse. More thorough data checking or other revisions are expected with the final release.
DATA MANAGER FIELDS:					
6	DATAMGR	Char	40		Name and Title of the person responsible for this data layer; this person should be contacted if there are questions about the source or use of the data.
7	ORGANIZATN	Char	80		Data Manager's Organization (Part of Data Manager's mailing address)
8	STREET_ADD	Char	50		Street address of D.M.'s organization
9	TOWN_ST_ZP	Char	40		Town/state/zip of D.M.'s organization
10	PHONE	Char	16		Phone number of D.M.'s organization
GEOGRAPHIC AREA FIELDS:					
11	GEOG_AREA	Char	60		Area covered by the data, e.g., New York, Franklin County; use DESCRIPTN if more than 60 characters are needed.

DATA SOURCE FIELDS:				
12	SRC_MAP	Char	60	Source of coordinate information, e.g. "GPS," or "Digitized from USGS Quad Sheets"
13	SRC_DATE	Char	30	Published date(s) of the source maps, if applicable, including original compilation and latest revision.
14	SRC_SCALE	Char	30	Scale(s) of all source material, if applicable; more than one scale may be given if the areas covered by each scale are clear. If need be, give details in SRC_DESCRP.
15	SRC_MEDIA	Char	40	Media of the source maps, if applicable - paper, mylar, etc.
16	SRC_DESCRP	Memo	10	A detailed "notes" section for a full description of the data collection techniques and source materials. Anything about how the data were collected that doesn't fit neatly into other fields in this section can be put here.
DATA FORMAT FIELDS:				
17	DATAFORMAT	Char	20	Format used to store these data (e.g. ASCII generate file).
18	COORDSYS	Char	30	Coordinate system these data are stored in (UTM Zone 18, Vermont SPCS, New York SPCS East, or Lat-Long.)
19	PROJ_UNITS	Char	30	Should be meters (for SPCS or UTM) or degrees (for Lat-Long)
20	PJ_PRMTR	Memo	10	If these data are in a projection other than UTM or SPCS, name the projection and give details of the parameters needed to transform from lat/long to this projection, e.g. zones, shifts, horizontal control datum, reference spheroid, etc.
21	H_DATUM	Char	10	Horizontal control datum (NAD 1927, NAD 1983, or other).
ACCURACY FIELDS:				
22	FEATACC	Memo	10	Statement of the likely maximum differences between the feature coordinates and true ground location. (Specify meters or feet). See Appendix A (p. 10) for help in assessing feature accuracy. Attribute accuracy is given in the PAD file.
23	QC_POSACC	Memo	10	Description of method(s) used (or not used) to assess the positional accuracy of data layer features; for example, "All point locations checked by mylar proof plots", or "point locations not yet verified". This field can be added to over time, and used as a record of the QC checks used.
24	QC_ATTACC	Memo	10	Description of method(s) used (or not used) to assess the encoding of attributes associated with data layer features; for example, "All attributes had two independent checks" or "attribute code values not verified". This field can be added to over time, and used as a record of the QC checks used.
25	UPDATE_SCH	Char	50	Planned updates for distribution.
26	LASTUPDATE	Date	8	Date of last update of the data for distribution. This must agree with the .UPD file (described below).
DETAILED DESCRIPTION FIELDS:				
27	DESCRIPTN	Memo	10	Feature definitions, methodology, data collection methods, etc. This should enable the proper use of the data, and include any pertinent information not found elsewhere in the documentation.
28	ASSOC_FILE	Memo	10	Describes files associated with the coverage, such as lookup tables, symbol sets or related attribute tables.
29	LIMITATION	Memo	10	Limits on data use, interpretation and use with other layers. Red flags should go here.

30	DISCLAIMER	Memo	10		This is a disclaimer to appear on all maps that use this data layer, if so desired. It is prepared by the data originator or manager.

The next file of importance for locational data is the PAD, or point attribute documentation, file. Again, using the xBase convention, it would be called PAD.DBF. The purpose of this file is to document any attributes associated with the point locations.

FILE STRUCTURE FOR PAD.DBF:

Field	Field Name	Type	Width	Dec	Description
1	ENTRY_DATE	Date	8		Date of this documentation entry; updated when any changes are made to this item.
2	COVERNAME	Char	16		Identical to the .DOC file's COVERNAME.
3	ITEMNAME	Char	16		The name of the attribute being described.
4	ATTRIBUTE	Char	60		Short description of the attribute.
5	ATT_SOURCE	Char	80		The source of the attribute values.
6	TYPE	Char	1		C, for character, N for number, D for date
7	WIDTH	Num	3	0	How wide a database field is needed to hold the attribute.
8	DEC	Num	2	0	If the attribute is numeric, how many places there are after the decimal point.
9	CONTENT	Memo	10		Detailed description of the attribute, including any codes used and the accuracy of the item values.

The final file that will be discussed here is the UPD, or update, file. The purpose of this file is to document, over time, any updates to the data or its associated attributes.

FILE STRUCTURE UPD.DBF:

Field	Field Name	Type	Width	Dec	Description
1	DATE	Date	8		Date of this entry to the .UPD file, when the update occurred.
2	ACTION	Memo	10		A description of the changes or additions made.

VI. FILE FORMATS

The file formats for data documentation have been addressed above. All that remains is to define the acceptable formats for the data itself. The LCBP-GIS data standard is Arc/Info. The most desirable situation would be to have locational data delivered as full-blown Arc/Info point coverages. In most cases involving locational data, however, that will be impractical. There are many file formats that can be converted easily into Arc/Info coverages. The simplest of these is the "generate" format, which is a simple delimited ASCII file. Its format is very simple, each record consisting of three numbers: a unique integer identifier, followed by an x coordinate, followed by a y coordinate. Here is an example of an acceptable generate file:

```
1,93017,200045
2,91750,198950
3,89975.5,212270
END
```

Note that the last line of the file must contain the key word "end". Note also that the file itself contains no information regarding coordinate systems. The data file would be useless unless accompanied by complete documentation files.

A crucial feature of the generate file is that the identifier in the first field must be unique. That is to say, each point being described must have one and only one identifier, and each identifier must be associated with one and only one point. This is so because, given the simplicity of the format, the identifier is the only thing that *can* be unique, and without some kind of unique identifier to act as a key, there would be no way to associate attributes with their corresponding points. As an example, suppose that the sample file above represents the locations of water quality sampling stations in Lake Champlain expressed in the Vermont SPCS coordinate system. In order to make any sense to a potential user of the data in a GIS environment, the attributes associated with the stations would have to reference the same identifiers. For example, let us say there is an attribute file maintained by the researchers in dBase format that contained information on mercury concentrations over time at each of the three stations. The stations would have to be identified, somewhere in that attribute file, as stations 1,2 and 3, or else there would be no link between the locations and the mercury concentrations.

But what if a particular researcher has an extensive database that has been developed over many years that does not use integers as identifiers, but uses some sort of an alpha-numeric code, like "E9013.CHA, "E9013.DCW, "F1074.BCD," etc. Suppose that the researcher, being a cooperative and interdisciplinary sort, goes back with a GPS receiver and captures locations for all 5,000 of her monitoring stations. Must she then recode her entire database in order to share her data with GIS users? No, but what would be needed in such a situation is yet another file, called a lookup table. This could also be a simple delimited ASCII file, or it could be a database file. In the example given here, the file, if it were an ASCII file, would look like this:

```
"E9013.CHA",1
"E9013.DCW",2
"F1074.BCD",3
(etc.)
```

Now, armed with a generate file, plus a lookup table, plus the researchers' database, plus complete documentation files (!), the LCBP-GIS Service Center could easily produce a full-blown GIS data layer that would be accessible to any GIS user who wished to use the data for spatial modelling purposes (or anything else, for that matter).

VII. SUMMARY

To sum up, then, we have described here a set of standards for capturing, recording, documenting, and transferring locational data for inclusion in the LCBP-GIS. The important points are these:

- these standards apply in situations where location is an enhancement of the data, rather than a primary feature of the data

- these standards apply to point data only, not lines or polygons
- locational data must be expressed in a coordinate system that is compatible with the LCBP-GIS: UTM-Zone 18, VTSPCS, NYSPCS-East, or Lat-Long
- the accuracy of the locational data must be well-defined and clearly documented
- the minimum acceptable documentation includes the DOC and PAD files described in this document
- the preferred file format for locational data is the "generate" file described in this document
- locational data must be referenced by unique, integer identifiers; look-up tables can be used to match numeric identifiers to non-numeric keys

VIII. THE SERVICE CENTER

The LCBP-GIS Service Center is operated by the Vermont Center for Geographic Information, on the University of Vermont campus. Our mailing address is:

VCGI
206 Morrill Hall
University of Vermont
Burlington, Vermont 05405-0106

We can also be reached by phone or fax:

Voice: (802) 656-4277
Fax: (802) 656-0776

The author of this document is Robert Dana, who is the Service Center Project Manager. He can be reached at the address and numbers above, or by email to: robertd@vcgi.uvm.edu.