CARICOM

CARICOM CENSUS DOCUMENT

GIS and Census Mapping Training Manual

Census GIS & Map Development Guide

CORPORATE GRAPHICS AND COMMUNICATIONS

Census GIS & Map Development Guide

Turkeyen • Georgetown
Guyana, GY
Phone 203.555.0167 • Fax 203.555.0168

Table of Contents

Version

1

Preface

GIS has revolutionized census mapping. Today, virtually every country uses GIS to create census maps, disseminate geostatistical products, maintain databases, and provide wide range of services to its users.

CARICOM held a Census Symposium in 2005 that covered several topics related to geographic information. More specifically the symposium focused on "Conventional Approach to Preparing Maps of Enumeration Districts" (item 6.4.1) with a presentation by Jamaica. This session emphasized that the mapping exercise should be ongoing and used as a dissemination tool. Furthermore, it underscored that countries should not only be able to interpret and read maps but that the map exercise should become an integral part of the regional guidelines. Another significant session directly addressing the area of geographic information punctuated "Experiences with the Use of Geographic Information System Capabilities in the Preparation of Census Maps and Implications for Data Dissemination" (item 6.4.2).

In the early 90's it was realized that hand drawn maps were inadequate in guiding enumerators. During the mid 90's census Geographic Information System (GIS) developments were just beginning to be conceptualized and implemented by some countries. In mid-2000, through collaboration with other agencies with a GIS in place, data sharing agreements were formed to allow accurately georeferenced raster and vector spatial datasets to be shared. The UN meetings on geospatial technologies for census also reflect the importance of geographic information particularly for census. In 2007, The UN Statistics Division, taking into account recent geospatial technological advances, organized several workshops on the use of geospatial technologies in census-mapping operations. One of these workshops was held in the Caribbean. Another similar workshop was held in the Pacific Island region where similar geographic issues and conditions exist.

At present, the popularity of the use of GIS for mapping and for data dissemination is increasing within the Caribbean Region. As a consequence, it was further noted that efforts should be made at the regional level to investigate the use of these systems for the 2010 Round of Population and Housing Censuses - adhering to the Mission Statement of the Secretariat that states "To provide dynamic leadership and service, in partnership with Community institutions and Groups, toward the attainment of a viable, internationally competitive and sustainable Community, with improved quality of life for all."

Today, virtually every country uses GIS to create maps, disseminate geostatistical products, maintain databases, and provide a wide range of services. The GIS software most dominantly used is ArcGIS by Environmental Systems Research Institute or ESRI. The global trend shows that this software becomes one of key tools for statistical offices as more and more geospatial products and services will be demanded by users to view analyze and share statistical data far beyond just the census. The census data often simply represents the common denominator of a GIS dataset and other data is integrated into the relational database management system and GIS to develop more specific analyses and products to a wide variety of users.

CARICOM has supported capacity building and best practices in the use of GIS in statistical activities. CARICOM has put forth to the Caribbean countries opportunities to attend the conferences and meeting and receive in-country guided assistance. The objective is to enhance and exchange lessons learned in creating the datasets for census and statistics exercises.

These exercises have addressed a wide spectrum of mapping and GIS activities currently in place in the 2010 Census worldwide. They have provided a unique, interactive discussion platform with a focus on sharing best practices and lessons learned among governmental agencies, regional organizations, private sector, civil society organizations, and academia. GIS experts in the Caribbean now have a platform for the continued discussion on the newest developments in the GIS field and to jointly address the present challenges as well as to identify and provide methodologies to develop key geospatial datasets through the application of commonly used geospatial software and tools.

These opportunities have allowed Caribbean countries to further develop their capacities to implement an efficient census and statistics GIS. The 2009-2010 conferences, meetings and in-country guided assistance have helped identify the ways to address the current specific needs for the 2010 Census through training and exercises and practical applications, enhancing the national, as well as regional geospatial infrastructure.

CARICOM hired a consultant, Charles Reese Brigham, to assist in the preparation of the CARICOM Census Mapping Manual through experiences visiting all CARICOM member countries and several associate member states and assisting in the map development process. In October CARICOM country GIS personnel participated in the review of the draft CARICOM Census Mapping Manual before submission. The CARICOM Census Mapping Manual reflects the mapping practices and recommendations of all the countries GIS personnel, regional GIS workshops, and CARICOM census meetings discussions on GIS and census mapping. Those bodies emphasized the need for countries to approach the use of census geography programs as a continuous process rather than merely a sequence of mapping and dissemination operations.

It was emphasized that the CARICOM Census Mapping Manual should demonstrate how the use and apply contemporary geospatial technologies and geographical databases that are not only beneficial at all the stages of the population and housing census process but more importantly reflect how they are used among CARICOM countries. For instance, the CARICOM Census Mapping Manual should show how those technologies improve efficiency in the preparatory, enumeration, processing and dissemination phases of the census for CARICOM countries through practical examples of their use.

In this regard, it is important that the CARICOM Census Mapping Manual put into the hands of GIS and mapping officers and census mapping managers and related personnel a technical guide on the contemporary methods, tools and best practices in current use by CARICOM countries that would enable them to better articulate their needs and deal with census-mapping operations more efficiently. In short, the CARICOM Census Mapping Manual covers both map and census GIS development in considerable detail and briefly discusses managerial and operational needs. It addresses GIS tasks that concern mapping personnel and addresses technical and practical issues that concern census cartographers and takers; it also briefly touches on organizational and institutional issues that concern statistical agency heads and other managers; and it explicitly addresses technical and practical issues that concern census cartographers and census takers.

ACKNOWLEDGEMENTS
I would like to express my very great appreciation to the Statistical Offices who facilitated the visits of the CARICOM Consultant. I would like to offer my special thanks to the Statistics Programme personnel of the Caribbean Community (CARICOM); Advice given by the GIS programmers of the Secretariat of the Pacific Community which has been a great help in my work and understanding of the challenges of small island states. I am particularly grateful for the assistance given by the ESRI Global Affairs Account Manager, DigitalGlobe, MDA and GeoEye. Assistance provided by Scott Pontifex of the Secretariat of the Pacific Community was greatly appreciated. I wish to acknowledge the help provided by mapping personnel in all CARICOM countries. Dr. Philomen Harrison provided me with very valuable Information regarding the challenges with this work. I would like to thank the following agencies for their assistance with the collection of my data and information needed: UNDP, UNSD, ECLAC, SPC, and many others. My special thanks are extended to the staff of CARICOM for their continued support and my wife and mother who read all of my work.

It is important to point out that although these countries are small in comparison to other countries they suffer from the same census mapping issues of larger countries such as large distances and scale, demanding physical geography and scattered populations. Staffing, lack of training and expertise and funding are the dominant issue plaguing most census mapping activities worldwide.

Introduction

All National Statistics Office (NSOs) in the Caribbean use Geographic Information System (GIS) in their census operations to a large extent. International and regional workshops have concluded that the use of GIS and geospatial technologies plays a critical role in the success of a census. A UN workshop in 2007 concluded that the three main reasons GIS was introduced (and still is) were to improve information processing, improve the quality of data collected, and to facilitate better decision-making. A workshop in 2010 was arranged by CARICOM and held in Belize on Census Mapping and GIS and focused on the technical aspects of Enumeration District (ED) mapping. It also emphasized these benefits. This demand for census mapping and related data in GIS format came from an increased request by data users as well. The reasons for the use of GIS for census reflect the main reasons GIS was originally introduced - to better understand geography and the human/environment relationship.

With the evolution of GIS have come multitudes of commercial and non-commercial software providers. As expressed in many of the CARICOM meetings and as reflected by CARICOM itself, as well as the international community in general, the goal has been to provide access to GIS technology and methods, which strengthen national capacity to all interested parties to the extent that the user has training to execute, and within the available budget. This is a very delicate process that should not be considered lightly. Specific comprehensive GIS software is largely within reach today and easily competes or outweighs costs of traditional approaches if implemented in the correct fashion with proper support. National Census and Statistics offices of CARICOM use software provided by ESRI, although versions and suites in use vary (ArcView3.2 to 10 and ArcInfo/ArcMap and ArcIMS/ArcSDE and ArcServer). Other GIS and associated software such as MapInfo or AutoCAD is rarely used by countries and if so, is usually in conjunction with ESRI products. Free and open source software is increasingly generating interest but has not been mainstreamed into the census GIS process. Aside from marketing, this may be largely due to the inconsistent support community when compared to the ESRI suite of products and related significant support and user communities.

The primary aim of the census Mapping Manual is to facilitate the preparation and production of Census Enumeration District (ED) maps that would be used for the location and enumeration of households at a National level during the data collection phase of Population and Housing Censuses in the Caribbean Region. It focuses on the development of a census GIS for map production and the incorporation and use of available data (e.g. national base maps, boundary, and road and river datasets) so that ED boundaries are clearly delineated and correspond to physical ground features to ensure that there are no omissions or duplications of land coverage. The mapping manual also focuses on standardizing ED maps in the Caribbean region through a common template and tools for map creation.

The goal of the census mapping manual is to provide technical and some managerial assistance to CARICOM member states and other interested parties that are producing national, supervisory, community and, most importantly, ED maps. And to facilitate the preparation and production of Census Enumeration District (ED) maps that would be used for the location and enumeration of households at a National level during the data collection phase of the Population and Housing Censuses. Field exercises related to census GIS development and map production are also discussed.

The census GIS is the core geodatabase that is used to create map products. The design of the geodatabase caters to the present and future needs of the census and the stakeholders involved. For example, in some countries various agencies participate in the census GIS development in order to build upon this database further after the census. The census GIS contains and takes care of the various attributes as defined in the technology of the Geographic Information System. In a population and housing census for example, apart from the spatial location of the building, dwelling and/or household, demographic statistics captured before during and after the census are also stored in the database for the benefit of the GIS end user. This database in the ESRI suite of products is commonly referred to as the geodatabase. The geodatabase is a collection of geographic datasets for use by ArcGIS. There are various types of geographic datasets, including shapefiles, attribute tables, topologies and many others.

GEODATABASE

It is evident that all the required information for the census GIS need not be built as one project, particularly since other agencies collect different types of geographic data that compliment the census GIS and vice versa. The census GIS can be designed on an evolving basis. The priority information needed by the Central Statistics Office, commonly in charge of the census exercises, can be collected first followed by obtaining the remaining information and attributes in a phased manner. This line of practice has led the field of census GIS development to create the initial essential information for census through proper satellite or aerial imagery and photography, field work surveys for ground control points and geographic feature coordinates as well as photogrammetry. These all act as a foundation from which to link any other spatial information collected in subsequent phases. Later collaborationinvolves the building of relationships with National and Regional mapping - or spatial data gathering agencies.

The census GIS in many way serves as the baseline GIS database from which to build the national GIS infrastructure from its goal is to develop national and sub-national boundary datasets utilize demographic data and disseminate/house statistical outputs and demographic variables. Often demographic variables alone serve as the common denominator in the development of new datasets and of understanding population and housing trends and their dynamic nature. The incorporation of geographic weights now adds a more easily understandable spatial spectrum to understanding and interpreting statistical trends (e.g. geographic weighted regression and spatial autocorrelation).

A significant issue with implementation of a fully functional census GIS in the region was the lack of adequate, continuous and comprehensive training. While some national statistical authorities were able to ensure the proper training on some occasions, in general, this was the biggest hurdle to overcome in taking full advantage of the GIS technology and data. The workshop underscored that there is no lack of labor force with more than adequate technical background available for training in GIS – it is simply the inability to ensure continuous training that prevents national authorities in faster implementation of GIS. This manual addresses this need by addressing the Census mapping aspect. However, the technological expertise needed to develop and maintain a GIS spans several books and are software and application dependent. The manual does not extensively cover the broader management aspects of census mapping, as this has been widely available for some time. Rather it focuses on the creation of the census Maps themselves and the specific needs for the creation of ED maps for the census. The ED maps represent the backbone of the census and without a comprehensive census GIS database ED maps cannot be accurately produced, hence poor representation on the ground. It is hoped that this manual will help people in charge of census mapping to adapt and develop their own workflows, ones that can best suit their needs based on what is presented, particularly since many of these methods have been used in the region and are meant to be further improved over time.

General Technical & Institutional Issues

The following section describes the base for technical and institutional issues. These include the critical design issues that determine the nature of the census GIS and the range of applications that it will support. Success of data inventory, data importing and data conversion processes depends on a well-designed institutional environment and a well planned operational strategy. The planning steps are divided here into institutional issues such as the organizational structure for geographic support, the explicit delineation of census geography, and the design of the geospatial database. These stages can be carried out more or less simultaneously using organizationally approved methodologies, and many of the choices depend also on the chosen data input strategy.

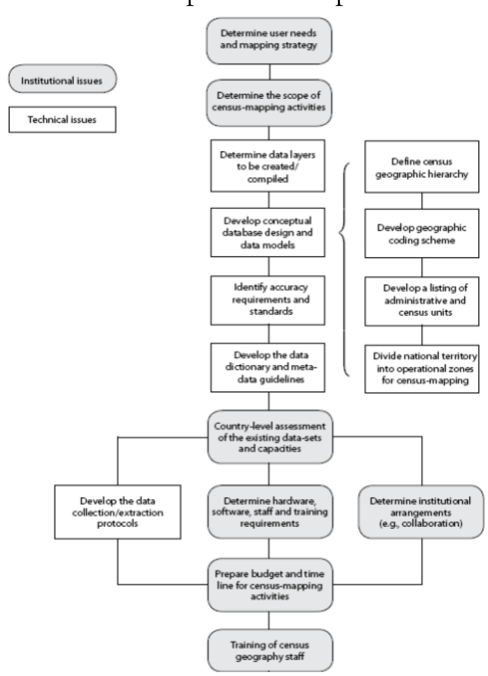


Figure 1 Stages in planning geographic work for the census

Internationally recognized architecture for evaluating enumeration and ED map preparation and production is shown in Figure 1. It employs a logical arrangement (in stages) with a focus on a "geocentric" census. A "geocentric census" means organizing the census process around the geography. For many NSOs on the forefront of fully embracing digital capabilities, often the investment is in analogue-to-digital conversion of paper enumeration area mapsheets and other basemaps, which involves careful scanning and correction procedures so that maps can be used as a basis for new digital geographic databases. These databases can then be compared against remotely sensed imagery, and then field-corrected using GPS. Incorporating such technology as scanning, imagery, and GPS allows the mapping unit to focus efforts on areas most in need of updating since the previous census.

Further to the institutional process of creating ED maps, it is important to have the correct policies in place so that the NSOs are up to speed on currently employed technologies, identifying common problems and solutions, and tailoring mapping tasks according to the performance potential of the personnel available. This will help evaluate new methods and save time on normal census mapping processes (e.g. heads-up digitizing). It also requires improving the larger geospatial infrastructure while creating the necessary mapping products needed for the 2010 round of censuses.

Thoroughly implementing all of these stages is difficult since many NSOs often have multiple pressures that hinder the complete exercising of all the stages in their entirety. Compromises are often made so that the most geographic work can be done within the time allowed. Each country has to weigh its ability to execute each stage in order for the geographic work for census can be completed. Aside from the biggest and most common administrative hurdles such as budgetary constraints, staffing and training, the most common geographic obstacles are implementing a geodatabase/data model and maintaining metadata guidelines/data dictionary. These geographic areas of planning receive the least attention due to the rapid nature of the census GIS exercise. The geodatabase allows for greater technical performance in the GIS and is vital in keeping data organized and transparent in the GIS. The metadata and data dictionary are critical for the sustainable development of the geodatabase and its use beyond the census since they provide the ability to log and capture information about the dataset such as use, creation, structure etc.

Common Needs & the Mapping Strategy

Current investigation by each NSO of the geographic data capture activities will help identify common technical areas in need of assistance to aid the mapping strategy and census GIS development. From an evaluation of current activities, mapping/GIS managers can determine the expansion and/or contraction of geographic practices. This will assist in reconciling user expectations with what is feasible given available resources, working backward from final products and services to requirements.

What is the purpose of the census in relation to spatial analysis?	What is the spatial extent (total area) and grain size (ground resolution) currently and what is preferred?
What type of spatial data do you need to achieve your goals?	What are the sources of these data, and what are the appropriate types of data to enhance census mapping activities?
What constitutes the ED layer currently?	What road, roads, river and complimentary layers such as landmarks etc. exist?
What geographic hierarchy exists/has to be created/modeled?	Do initial census mapping methodologies employ a flexible approach?
What is the strategy and best method for evaluating the accuracy and validity of the census results?	What is the ED and reporting unit structure for the census? ED/supervisory
What existing GIS staff can be used for the census GIS and map development	How will the documentation of the data be cataloged?
What are the costs for Upgrades of hardware or/and existing software?	What are the staff training needs?
What is the ease and difficulty as far as installation of hardware/software as well as maintenance?	What infrastructure documentation, manuals, Help-line, and vendor support (patches etc.)?
Is the existing workforce adequate?	What methods are/will be used in generating GIS datasets?

Table 1. Critical questions within the geographic work for the census that each NSO should address.

In creating census geographic products it is important that the common needs of the main users of the information produced are identified as early as possible. This helps to identify workflows within the mapping strategy that will enable more rapid dissemination before the conclusion of the enumeration phase. This includes consultation on geographic content, i.e. geographic structures such as administrative hierarchies or summary levels, and also geographic base products that support analysis of census data. This normally consists of 3 user groups: Persons and institutions participating in the census operations; Census geographic data product users; the general public. User-privacy concerns are also an important consideration in the successful deployment of products and applications. The administration of the geostatistical datasets and applications infrastructure used for this must be kept minimal because of the large number of services that need this information and the resolution restrictions that maintain privacy.

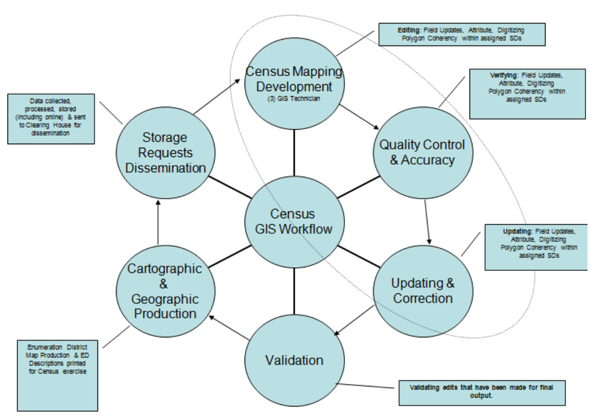


Figure 2 Census GIS workflow within the GIS unit of the Bahamas Statistics Office

Bahamas GIS Unit identified common technical areas in need of assistance, to aid the mapping strategy and census GIS development. A streamlined workflow was outlined that greatly assisted the office in keeping the mapping activities in simplified order. Workflows can vary and it is best to keep the level of complexity transparent for all personnel involved. Complex workflow architectures are often not the optimal solution. The workflow displayed in Figure 1 should not be used on its own and instead should be adapted to best fit the office needs. Redesign and re-evaluation of these workflows are critical.

Essential Data Layers

Raster Data: continuous, rectangular cells laid out on a grid

Layers are referred to here as visual representations of a geographic dataset in any digital map environment. Each country has various layers of information that are required for the final map product. Weighing the user requirements against available resources in the budgetary cycle (i.e. human and financial resources, existing data, size and terrain, timeframe etc.) becomes critical particularly as the day of enumeration arises. Assessments must be conducted on the compiled enumeration area maps and geographic boundary files in a digital format for all statistical reporting units for feasibility in the use and generation of ED maps for the 2010 Round of Population and Housing Censuses. This will also include listings of all statistical and administrative reporting units; raster data layers such as mapsheets and satellite imagery; vector data layers containing feature data such as geographic boundaries, streets and landmarks; other geographic features.

Vector Data: coordinate-based data that represent geographic features and polygons, lines & points

It is necessary that GIS personnel understand the critical data layers for the enumeration exercise. These are the primary layers that are to be included in the final ED map. Imagery is a paramount resource in censuses of today. To harness this trend most Caribbean countries have invested in personnel with knowledge on the use of aerial photography and satellite remote sensing data for the development of essential data layers in a GIS. As imagery is often a costly purchase, it is essential that these staff have the necessary required understanding of procuring imagery; identifying focus areas for ordering/Areas of Interest (AOIs); principles of remote sensing; resolution of remote-sensing data; high resolution sensors; satellite imagery versus aerial photography; scene orientation and imagery information content. With knowledge of these components the mapping team can be current on significant improvements in efficiency brought by these imagery datasets to aid GIS development.

Often the NSO does not have the expertise to determine its imagery needs exactly and in this situation the country must seek assistance from expert in the geographic community, preferably from regional organizations with expertise in contemporary census mapping methods. The use of imagery for census is commonplace and also a new method for the creation of census maps for many countries in the Caribbean. It can be seen as a threat to those that are unfamiliar with its use and benefits and require clear communication of its use to long term census personnel who might not be familiar with the use of imagery. The use of imagery in the census aims to make everyone's work more accurate and accountable. It should not be used as a tool to remove officers from the field tasked with map development, yet it should encourage field officers to capture finer resolution data when deployed to conduct fieldwork. This means activities such as verifying new developments, identifying front door locations, street direction, landmarks, etc.; this does not take away work from census personnel. Most importantly, NSO management should not interpret the use of imagery as a failsafe but instead as a way to achieve a smaller margin of error in the final census count.

Different countries need different amounts and types of imagery for their purposes. Targeting populated areas for imagery purchase is a common strategy to reduce costs but ideally imagery of the entire country is preferred. Purchasing imagery for populated areas required the creation of AOIs. These AOIs are created using a GIS and creating polygons around areas. An alternative is the use of proprietary tools that allow the user to check for imagery from a given provider.

These methods essentially allow the interested buyer to save the areas of interest as separate AOI shapefiles and send them to the imagery provider. This would normally involve the creation of an area of interest (e.g. shapefile) with as few vertices as possible to minimize the pricing per area of interest. Having fewer vertices make it easier and faster for the imagery provider to clip and specify (an) area in the imagery product.

Imagery allows for the mapping office to generate various data layers with the ultimate goal for census being a building footprint layer with all associated census attributes. Many countries are attempting to capture this level of detail and there is increased interest by NSOs to develop geographic database strategies and modeling for these large spatial datasets. Structuring these geographic databases require creating standard datasets for census and related workflows; geodatabase structuring and loading attributes and managing and creating vector datasets effectively; automating geographic database processes such as census map production through programming methods; creating topology rulesets.

Use of Mapsheets and Georeferencing

Many government map series such as National Government Survey Offices quadrangle Mapsheets are available in a scanned format. Offices often scan the maps, trim them to neatlines, georeference the Mapsheets corners, and mosaic them across an administrative area such as a Country or Parish. The Mapsheets are often the starting point data layer in the development of a GIS. Mapsheets are essentially Digital Raster Graphics scanned from National Government Survey Offices, which show terrain, names, and major features in an image format. The integration of Mapsheets in a GIS allows the ability to create GIS datasets such as geographic boundaries, streets and landmarks through digitizing from the image (often in tiff format) and also assist in editing and updating existing geographic data. Mapsheets can also be processed using scanning abilities in software that allow for the extraction of features from the imagery as well. Mapsheets have traditionally been used as the main content for creating census maps, particularly in rural areas when other geographic data may not available or comprehensive.

An example of Mapsheets use in the census is in Jamaica, whereby Mapsheets are used to create ED maps where higher resolution geographic information is not available. All existing Mapsheets are used in the GIS to add value to the creation of the GIS and the development of census maps. These Mapsheets are used as the main content in the census maps where data at higher resolution are not available and/or accessible. The

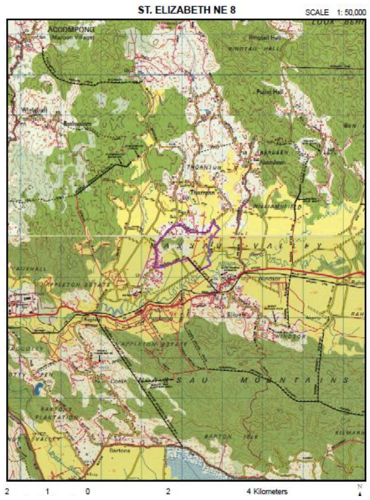


Figure 3 1:50,000 Mapsheets showing a single ED boundary (purple) used in the census GIS for map production

Geographic Services Unit (GSU) of the Statistical Institute of Jamaica, converted GPS waypoints into X, Y coordinates relative to their local coordinate system (WGS_1984 GCS_JAD_2001). The team identified several (commonly greater than 5) waypoints on the Mapsheets and moved the maps accordingly, using the known point as a base point to the corresponding coordinate. After this alignment, the scale was checked and then rotated where necessary so that all of the waypoints fell in the right place. This process is often referred to as georeferencing. Georeferencing is commonly used to bring existing Mapsheets and other geographic datasets into a GIS.

Georeferencing is the process of assigning geographic information to an image. Knowing where an image is located in the world allows information about features contained in that image to be determined. This information includes location, size and distance. There are three common georeferencing tasks that are conducted when using raster imagery in a GIS which are (1) Un-georeferenced Imagery (2) Georeferenced Imagery (3) Orthorectified imagery.

An image can be thought of as a grid of colored dots called pixels. If the information on the size of a pixel is known, then objects in the image can be measured. The location of a feature can only be determined relative to another feature. The position of an object in the real world is not known with Un-Georeferenced imagery. If this concept is extended to an aerial photograph with a known scale, real objects can be measured.

Georeferenced Imagery is where the Earth has been divided into a grid with lines of Longitude and Latitude. When an aerial photograph is aligned to this grid, it becomes "georeferenced." A georeferenced image allows distances and areas to be measured and the exact position of any pixel can be determined. Imagery from satellite imagery providers is commonly in a georeferenced format. These imagery datasets are frequently being used by CARICOM countries in the census as a critical layer in the GIS for creating datasets and, at times for inclusion in the ED map.To georeference an image, the positions of known points, called "control points," are determined. These control points can be collected in the following ways:

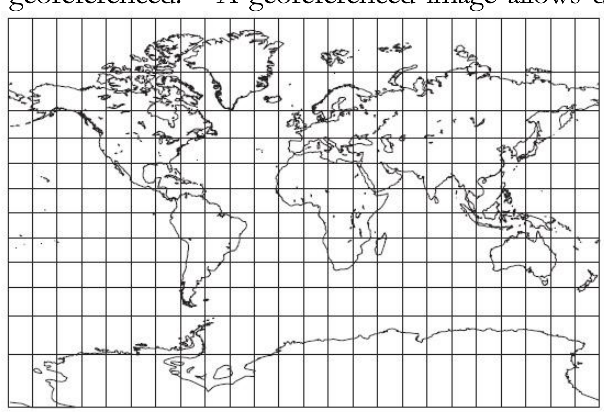


Figure 4 Common Grid with lines of Longitude and Latitude

Method of Collection	Accuracy
Differential GPS Receiver	1/32 - 1 meter
Handheld GPS Receiver	1 - 5 meters
Georeferenced Base Map	5 - 20 meter

Control points should be collected from immovable objects that will be easy to identify in the imagery/aerial photograph. At least three control points are required for georeferencing but 5 or greater are most common. Additional control points add redundancy and help to increase accuracy. Larger areas and higher resolutions can require many more control points to maintain accuracy.

When the control points are plotted on a grid of latitude and longitude, the position and orientation of the aerial photograph can be determined. The aerial photograph is then moved and rotated until it fits the control points. GIS Software creates a statistical "best fit" of the image to the control points and then "re-samples" the image. Once the image has been georeferenced, each pixel will have a coordinate associated to it. The location of any object in the image can now be determined. With knowledge of where the map is in the world, it can be made available to compliment any other geographic dataset within the census GIS.

Orthorectified imagery adds additional accuracy to georeferenced imagery and can be obtained by accounting for distortions caused by camera position and terrain. It is best suited for customers who will be making precise measurements, or when hilly terrain is expected to cause unacceptable distortions. It also tends to be more costly. An Orthophoto is an aerial photograph with a uniform scale and is usually required if maps are to be generated from the aerial photograph. To create an orthophoto, an aerial photograph is projected on to a digital elevation model (DEM). Having a DEM allows for distortions caused by terrain to be eliminated. For mostly flat terrain, orthorectification of images does not serve to increase accuracy significantly.

A DEM is a critical dataset that is often neglected during the development of a census GIS. A DEM is the representation of continuous elevation values over a topographic surface by a regular array of z-values, referenced to a common datum. DEMs are typically used to represent terrain relief. GIS personnel can use the DEM to understand terrain in great detail. The DEM helps in determining ED canvassing times and challenging geographies. Several travel time models can be applied to a DEM to calculate routes for census and survey, which can aid in understanding the costs of such exercises.

There are also other map sources such as Topographic Line Map Sheets (TLMS), which can be scanned and used as backdrop in field maps as well as for census maps in certain cases. More particularly, getting these TLMS sheets into vector formats can be of great assistance, one can say, but also the data itself may exist in vector format. TLM is a lithographic map that portrays the greater detail of topographic and cultural information. Relief is shown by contours and spot elevations measured in meters. The map is a representation of terrain detail. Features are plotted to correct orientation and true location. The map depicts the level of detail required for infantry and reconnaissance units to navigate in various terrain environments including jungle, mountain, arctic, and desert.

Use of Imagery and Aerial Photography: A revolutionary change in the creation of census maps

The widespread availability of commercial and freely available imagery has drastically changed the way census maps are designed in the Caribbean and has proved to be highly effective for census GIS development and map production. All countries are using some form of imagery to guide census GIS development and census map production. Imagery at high resolution of less than 5 meters and obtained within a year of census has fostered development of new, high grade, GIS layers that closely reflect the ground exercise at the time of the census as closely as possible (e.g. building layers, landmarks, rivers, roads). Imagery has also provided the census maps with an error checking methodology whereby imagery can be compared to existing census maps and clear comparisons can be made quickly. Of course, this is entirely dependent upon the date of the census map and the date of the imagery. Also, it is generally more cost effective than the use of any other method of map updating (hand drawn or GPS field deployments) since more rapid identification of new developments and landscape changes are possible through the use of imagery, allowing the field team to identify and prioritize EDs in need of further investigation.

It cannot be emphasized enough that the use of imagery has revolutionized the statistical community in providing greater accuracy for the census. Most countries are using imagery extensively and many have purchased georeferenced imagery that corresponds with the time of the census, normally within a year or less of the enumeration exercise. There are many examples on the use of imagery for census activities in CARICOM. Provided are some examples from CARICOM member states on their experience with the use of imagery in the census and the necessary steps each office took to get the best use out of these data for census and statistics.

A good example that illustrated the use of imagery to improve the census accuracy is taken from St. Lucia. The statistics office mapping team linked the buildings to the questionnaire using a unique id of ED number+Building number (e.g. 14712+63=1471263). The Enumerators write in the building number onto each questionnaire as they visited the households. The office then joins the two fields when in the data processing stage within the database after the questionnaires are scanned. The enumerators were provided with a numbered aerial photo of the corresponding ED. They were also provided with an ED map showing strict location and canvassing information.

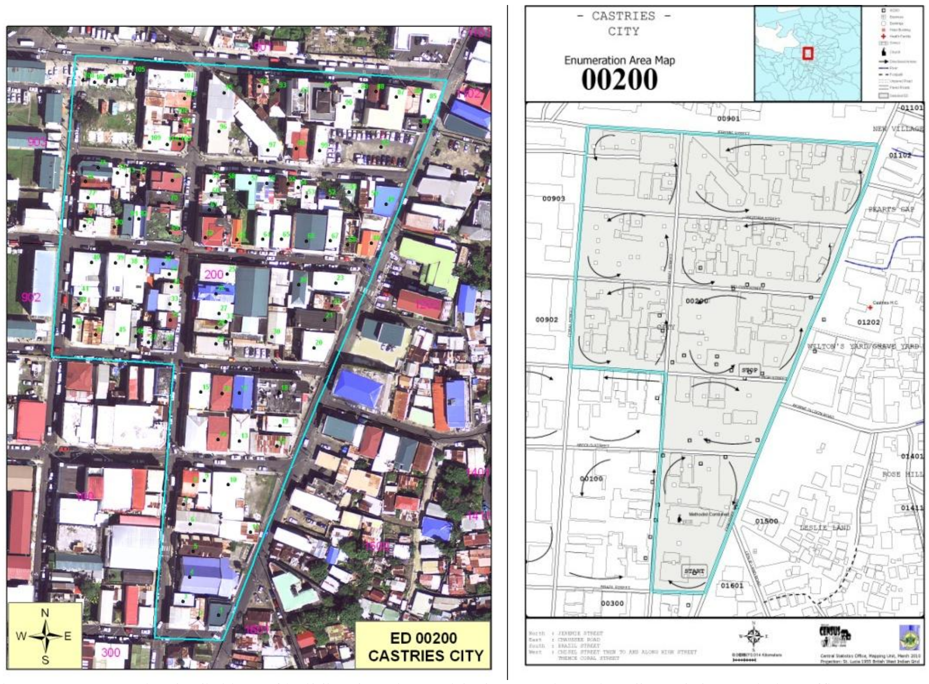


Figure 5 In St. Lucia, the linking of building locations with the questionnaire allowed the Statistics office to aggregate out at any preferred geographic level. This allows the ability to map demographic characteristics with a great amount of detail and fosters more specific targeting of populations in need of services, defining explicit geographic locations where, for instance, a new senior center might be needed.

There was interest by the GIS manager and the Statistics office in the linking of the building GIS layer(s) to the questionnaire to obtain more precise census geostatistical analysis capabilities. This can be done building simple relationship class structure, RDMS queries and geoprocessing tasks into the questionnaire dataset after the census database has been structured. This would allow for mapping at any specified geographic area at a finer scale than the EA. This requires the implementation of a coding system that caters to database development. The use of the Parish code and ED code as numeric attribute field values were recommended for map automation and for RDMS development for future geospatial and geostatistical activities.

Use of Imagery: Belize

Through capacity building efforts, funds were made available to the Statistical Institute of Belize (SIB) for census mapping operations for purchasing of imagery. While the primary goal was to obtain imagery for the entire country, these funds allowed the SIB to obtain imagery for approximately 40% of the country. Therefore they concentrated on urban and densely populated areas. Imagery for specific target areas was acquired through careful decision by the GIS technician and an IT person. Areas of Interest (AOIs) were created that allow different vendors to check for imagery in a simplified format.

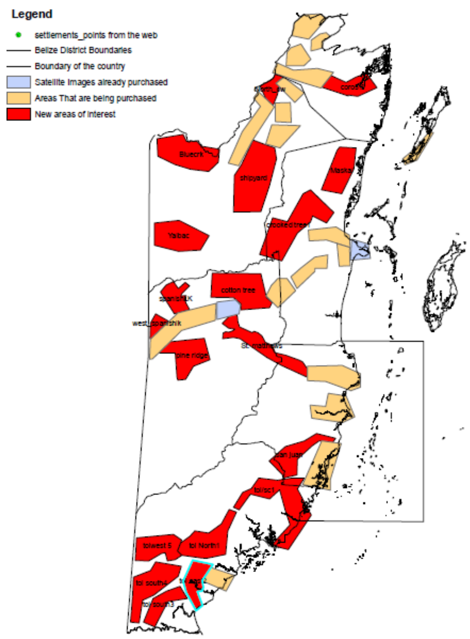


Figure 6Areas of Interest (AOIs) selected by the GIS officer for Belize.

One important tip to remember in creating AOIs is to create the polygons in the shapefile(s)/feature class(es) with as few vertices as possible. This will make it easier for the vendor to clip the imagery to the polygon geometry thus reduce the amount of work needed, lowering the overall price of the imagery itself. Below (Fig.3) is an illustration provided by SIB on their AOI selection. There are also AOI tools provided largely by vendors and proprietary software that allow the ability to save the areas of interest as a separate AOI shapefile for later use. These tools maximize the search efficiency within an imagery provider catalog and to quickly create an area of interest with as few vertices as possible to minimize the pricing per area of interest.

Once the AOIs are selected, pricing and availability and required resolution are determined. There are a multitude of resellers and vendors available that provide different imagery products. It is important that each offer consider these carefully in order to get the highest quality product that will satisfy the need for the exercise. A selection criterion often helps to make the case of which reseller/vendor to choose from. In the case of Belize, the GIS technician derived a selection criteria scheme to aid the imagery purchase and weigh options carefully.

Criteria for selection of Satellite Images

Criteria	Max Score
Resolution	25
Currency	25
Cost	25
Cloud Coverage	25

Satellite images were to be weighted on the following specified criteria: *Resolution*: Any resolution of less than one meter accuracy is acceptable and will be judged an equal weight of 25 points. Where everything else is equal, an image of higher resolution will be chosen over one of lower resolution; even though the two images may have scored equal points for resolution i.e. both have resolutions of 1 m or less.

Resolution	Points
1 m or less	25
<= 1.5 m &>1 m	20
<=2.0 m &>1.5 m	15
<= 2.5 m &>2.0 m	10
>2.5 m	Unacceptable

For every drop of resolution of 50 cm from one meter, the amount of points awarded will be reduced by 5. Images with resolution greater than 2.5 m are unacceptable. This is summarized in the table below:

Currency	Points
2008 or 2009	25 Points
2006 or 2007	20 Points
2004 or 2005	15 Points
2002 or 2003	10 Points
2000 or 2001	5 Points
Earlier than 2000	0 Points

Currency: Current images (2008 or 2009) will be awarded full points i.e. 25. For every 2 year drop in currency from current 5 points will be deducted as shown in the following table. The currency of images that are comprised of a mosaic of different currency will be deemed to be the average of the currency of its parts. Thus an image made up of parts that were taken in 2002, 2005 and 2007 will be equal to (2002 + 2005 + 2007) / 3 = 2004.7 and will be awarded 15 points for currency.

Price ($/km2)	Points
$10.00	25
$10.01 - $11.00	20
$11.01 - $12.00	15
$12.01 - $13.00	10
$13.01 - $14.00	5
>= 14.01	0

Cost: For every polygon, the price quote (per km2) that is lowest will be awarded a full 25 points. For every increase of one dollar from that lowest cost, rounded up to the nearest dollar, the points awarded will be reduced by 5. The example below illustrates the case where the lowest quoted price is $10.00 per km2.

Cloud Coverage	Points
0%	25
5%	20
10%	15
15%	10
20%	5
>= 20%	0

Cloud Coverage: The percentage of clouds, the image has and the location of the clouds will be assessed based on a scale of 0% being the minimum value and 20%, the maximum allowable interference. For every 5% of additional cloud interference, there will be a 5 points deduction.

Criteria for selecting imagery vary by country situation. Creating a selection checklist will speed up decision making and make the process of procuring imagery more understandable to other personnel that may not be familiar with the technical aspects involved.

Use of Aerial Photography: Montserrat

1.0 Summary

Aerial Photography plays an important role in the management of land use on Montserrat. Land use relates to environmental sustainability, disaster management, utility management and a host of other related fields. When aerial photography is combined with other data the analysis possibilities are endless. Examples are the research of geological features, exploration of geothermal energy, resolving conflicts in land usage for development and decision making in areas of conflict between development and areas zoned for conservation, and/or urban and rural growth.

The use of Aerial photograph dataset has proven to have more advantages over satellite imagery in the context of Montserrat. Montserrat has cloud cover most times during the year, therefore it will be reflected in a satellite image. Also satellite imagery does not provide as great a detail, because of Montserrat's size.

2.0 Project Rationale

2.1 Background

The need has arisen to have the aerial photography dataset updated, as the initial data was acquired in 2002. Since 2002, the rate of development on Montserrat has been very fast, in an effort to continue to provide sustainability to an island that has been ravished by a volcanic eruption.

As a result important replacement buildings, such as the Montserrat Volcano Observatory, Bank of Montserrat, St. Patrick's Credit Union, and new housing developments, as in Drummonds and Lookout, cannot be seen in the existing dataset. There have been requests for the department to provide data relating to these areas, and we have been faced with the problem of not being able to provide up-to-date information.

This proposed project is in line with the Strategic Development Plan objectives for 2008-2010 particularly, SDP Objective 3: To conserve Montserrat's natural resources, ensure development is environmentally sustainable, and that appropriate strategies for disaster mitigation are in place.

This project would contribute to this objective through the provision of accurate data, which will assist in the management of land whilst protecting natural resources through prudent management and to ensure effective environmental management, education and participation in decision making, as well as to protect and conserve biodiversity and other natural resources.

This project is also highlighted in the Physical Planning Unit's Business Plan 2008/2009. Pages 6, 8, 14,15,17,19 & 20. It is highlighted that the acquisition of a new aerial dataset will assist the department with regard to the monitoring of development activities to ensure orderly development of the island. Risk assessments can be undertaken to determine areas best suited for human habitation and also track land use changes to ensure compliance with the Physical Development Plan for Montserrat. This will positively impact the poor, vulnerable and marginalized by assisting in providing sustainable building settlements and preventing slum settlements.

Stakeholders and beneficiaries include, but not limited to: The Physical Planning Unit, The Lands and Survey Department, the Department of the Environment, Department of Agriculture, The Inland Revenue Department, the Disaster Management C0-ordination Agency, Public Works Department, Police Department, Montserrat Utilities Limited, Real Estate Agents, Land Owners, Home Buyers etc.

2.2 Policies

No previous actions were taken to address this problem. It is hoped that a policy can be implemented to have aerial photographs flown every 5-10 years depending on the rate of development.

2.3 Project Approach

BLOM Aerofilms previously called Simmons Aerofilms are known as the UKs leading provider for aerial photography and digital mapping solutions. The company has considerable experience of aerial photography, digital mapping and Geographic Information Systems and is involved in a number of national mapping projects throughout the world, with current contracts in Asia, Caribbean and Europe.

BLOM Aerofilms were contracted before by the Ministry of Agriculture, Land, Housing and the Environment, Government of Montserrat to carry out ground control, aerial photography including 1:5,000, 1:7,500 and 1:10,000, orthophotos and digital mapping. The photography was acquired in April 2002 using aircraft mobilized from the UK. They provided services to include: aerial photography, Ground control, Aerial triangulation, DTM capture, Topographic mapping, Orthophoto production and Quality control.

The data produced from this early survey is what has been used in the office to date. We now see the need to update the data we have, and from BLOM's proven track record we have decided to contract them again.

3.0 Financial Information

3.1 Budget

Item	Requirement	Price EC$
1*	a. Mobilization of aircraft, digital camera and GPS/INS. b. GPS and ground control preparation c. Acquisition of 25cm GSD photography with full stereoscopic overlap of the required area. d. Delivery of imagery together with a flight index plan in digital and hard copy format	$
2	a. Ortho-rectified digital imagery to be supplied in the Montserrat National Grid b. Revised DTM data for all of the island with accuracies of: +/- 0.750m in plan and +/- 0.3m in height c. Revised contour data for all of the island	$
3	a. Revised mapping for the areas produced during the 2002 project	$
	Project Total	$
Option	As part of the 2002 mapping project contours at 2m intervals were delivered for the North of the island. Blom Aerofilms can produce contours and DTM for the South of the island from the 2002 archive photography for the following sum:	$

	Project Total + Option	

Note: BLOM Aerofilms are currently liaising with a number of neighboring Caribbean islands that may be interested in their services. Should they be successful this would enable them to reduce the cost of Item 1a in above pricing, as the cost of mobilization to the region would be reduced should we be involved in additional mapping projects.

3.2 Cost recovery steps proposed

The acquisition of the aerial photograph dataset have the potential to improve Government revenues through the actual provision of services, and also through the spin-off effects - increased availability of data can bolster the real estate market, which in turn can result in increased Stamp Duty and Taxation revenue.

3.3 Recurrent Budget Implication

There is no implication on the re-current budget as it is a one-off expenditure.

4.0 Project Implementation

4.1 Management Arrangements

BLOM Aerofilms will be responsible for the actual day to day implementation of the project. The GIS Manager on Montserrat will report on its progress and the use of the project finances to the Development Unit.

4.2 Timing

4.3 Contracting and Procurement

The procurement method proposed is direct contracting of BLOM Aerofilms. The company has considerable experience of aerial photography, digital mapping and Geographic Information Systems and is involved in a number of national mapping projects throughout the world.

4.4 Accounting

The Permanent Secretary and the Chief Physical Planner will arrange the accounting for expenditures.

4.5 Monitoring

Mr. Franklyn Greenaway, the Chief Physical Planner will be responsible for monitoring the progress of the implementation of the project against planned activities, outputs and outcomes.

4.6 Reporting

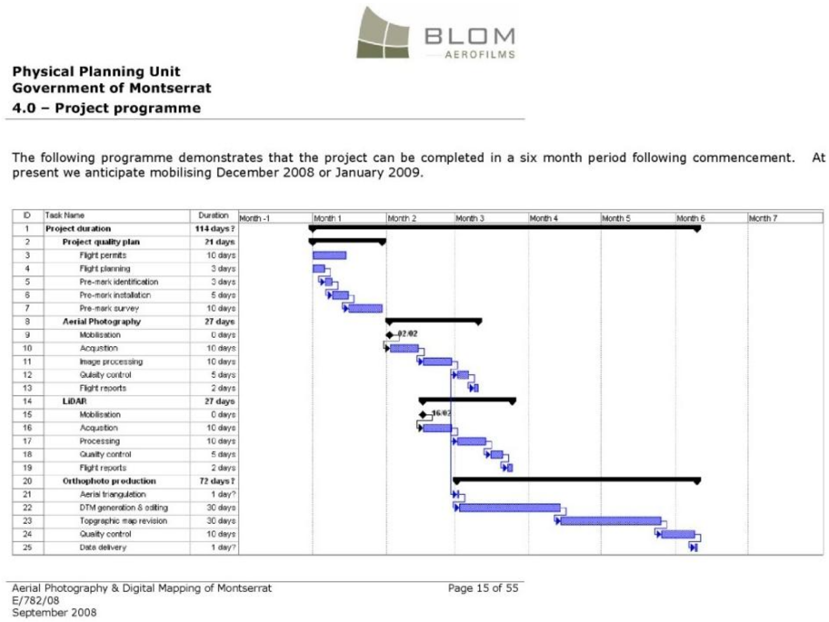
Reports will be submitted quarterly in the required formats.

5.0 Risks and Undertakings

5.1 The risks faced by the project are:

- Acquisition may prove to be particularly difficult if there is a volcanic eruption, as it may cause smoke obscuration and ash dust which may cause problems with the aircraft engine
- Availability of aircraft and equipment
- Ability to mobilize to Montserrat
- Functioning equipment and software
- Implementation

– Implementation


BLOM AEROFILMS

Physical Planning Unit
Government of Montserrat
4.0 – Project programme

The following programme demonstrates that the project can be completed in a six month period following commencement. At present we anticipate mobilising December 2008 or January 2009.

ID	Task Name	Duration
1	**Project duration**	**114 days?**
2	**Project quality plan**	**21 days**
3	Flight permits	10 days
4	Flight planning	3 days
5	Pre-mark identification	3 days
6	Pre-mark installation	5 days
7	Pre-mark survey	10 days
8	**Aerial Photography**	**27 days**
9	Mobilisation	0 days
10	Acqusition	10 days
11	Image processing	10 days
12	Quilaty control	5 days
13	Flight reports	2 days
14	**LiDAR**	**27 days**
15	Mobilisation	0 days
16	Acqusition	10 days
17	Processing	10 days
18	Quality control	5 days
19	Flight reports	2 days
20	**Orthophoto production**	**72 days?**
21	Aerial triangulation	1 day?
22	DTM generation & editing	30 days
23	Topgraphic map revision	30 days
24	Quality control	10 days
25	Data delivery	1 day?

Aerial Photography & Digital Mapping of Montserrat — Page 15 of 55
E/782/08
September 2008

Schedule

Montseratt was successful in implementing the Aerial photography through careful planning and clear outlining of the project. This example is very useful as a sample document for mapping and GIS personnel or otherwise to reference when planning such activities.

Use of GPS

Any census needs a system to link collected population data to location. Imagery can provide this link through digitizing a building layer. GPS can also be used to create this link. Often both are used and greatly compliment each other. In CARICOM, as in many developing countries, no prior location database exists for the Census use. No cadastres are in existence yet, houses do not have addresses or zip codes, etc. The Censuses in the regions needed an independent system to locate households on the map.

The concept of using GPS technology comprehensively to overcome the lack of location information was also an option during the preparation of the Census for CARICOM countries. The consultant and mapping teams agreed in some instances with the idea of using GPS in order to create an independent database containing the location of all households. This largely involved recording GPS locations (so-called 'waypoints') at the doorstep of each house (e.g. Dominica). The census mapping office of Suriname has georeferenced many building locations and maintains a comprehensive building database of these locations that are consistently updated. Using standard grade GPS devices the census mapping personnel visit areas not previously captured using GPS. Only recently they have used imagery to compare their results of the GPS exercise with available georeferenced imagery.

GPSing of households in Timor-Leste

A system of stickers was devised with a unique number printed twice on each one. One part of the sticker with unique number and Census logos was attached to the house, and the second part with only the number was put on the associated questionnaire by the interviewer. The same number was to be used a third time, namely as the Waypoint name. In this way the location of household and the questionnaire could later be correlated with each other. Figure 7 is a slide taken from a Timor-Leste training module on how the household sticker number was to be used three times. Below is a training slide showing how the sticker number is to be used in a three phase process, to establish the link between household, questionnaire and location. While this is not being done in the CARICOM countries, it presents an alternative way of linking census data at finer levels of geography.

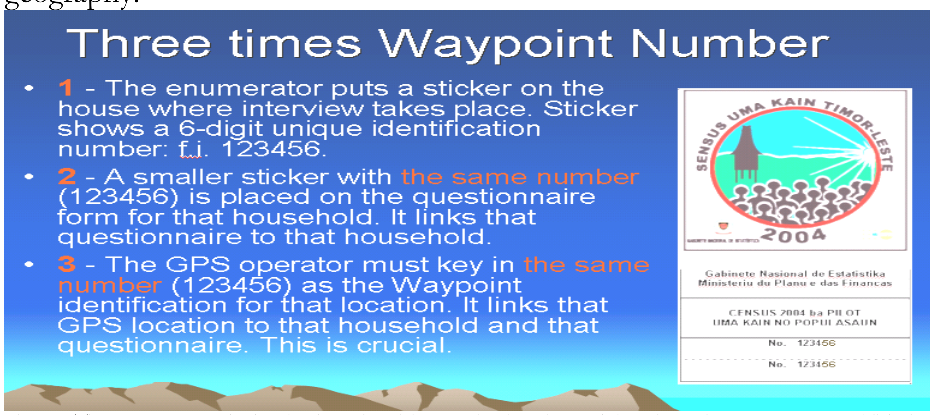


Figure 7 Timor-Leste example showing how the sticker number is to be used in a three phase process, to establish the link between household, questionnaire and location.

The above idea using the GPS to record the position of all households and linking them with the questionnaire was tested in the Pilot Census of 2003 in Timor-Leste and proved successful. Basically, two databases were created, one with main fields being the waypoint number and another latitude and longitude ('location' database), and another with as main fields the sticker number and the contents of the questionnaire ('contents' or 'attributes' database). Upon completion of fieldwork the location and contents databases are combined using the waypoint number and sticker number fields as join fields. This is just one example of feature linking with GPS. Several countries link the buildings in the GIS itself and validate in the field about what type of building/dwelling exists.

Often only a few days training in background and use of GPS need to be provided to Census enumerators, as long as the enumerator groups are accompanied into the field and trained as well. The operation of GPS units should be conducted as a pilot census. When the GPS units come in, and the waypoints are downloaded, and are shown in the GIS, on maps and geo-referenced imagery/arieal photos. It becomes evident that the keying-in of waypoint numbers (e.g. the same number as on the house sticker and on the questionnaire as in the Timor-Leste example) is a straightforward process. The goal is to have the vast majority of houses receive a waypoint, so that is becomes simple to attach census questionnaire data records to the corresponding waypoint record. GPS methods work well also and serve to identify precise locational information. Concerns often lie with purchase of the devices and training.

Obtaining GPS Units

After the decision is made to use GPS for recording of household locations or otherwise hiring an interviewer with GPS knowledge has to be considered. However, it soon becomes apparent that having a GPS for each person is not a satisfactory solution. The cost of GPS devices for 3,000 interviewers, for instance, would be far too expensive for most countries for the Census. Besides being very expensive, training must be considered, and training the interviewers in using the equipment without failure is crucial for the success of the GPS venture.

Rather than provide each interviewer with a GPS, often the solution is to task a Team leader of each Supervisory Area or groups of EDs to follow the interviewing teams round. His/her sole job would be recording GPS locations of all households in his/her Enumeration Region. The number of GPS receivers needed for the job can then be drastically reduced and open up the option of purchasing devices rather than renting people with a GPS for instance.

Purchasing often has more advantages over renting:

1. It allows sufficient time for specialist training of those so tasked;
2. Units could be resold/reused after the census field work has been completed, thus recouping some of the investment.
3. After the Census a number of GPS units could be donated to institutions and agencies, increasing capacity for geographic work and introducing the technology into the country.

Selection of GPS Type

The choice of GPS receiver to be purchased was based on the following characteristics:

1. Easy to use;
2. Able to perform quick and reliable satellite fixes;
3. Rugged and waterproof;
4. Small and light weight;
5. Store at least 500 user waypoints;
6. At least 6-digit waypoint indication (to accommodate up to 1 million unique numbers);
7. User waypoints must be downloadable;
8. The downloaded positional data must be compatible with available software that would allow importing the data into ArcGIS; and
9. The GPS unit has to be low-cost

The use of GPS for Census has an advantage because it allows for the linking of records of the enumeration database to corresponding points on the map. But there are other advantages too, notably census quality control, database improvement and advanced analysis possibilities.

GPS waypoints draped over hi-res geo-referenced images, with individual houses visible, offers a unique quality control opportunity. Census enumerators returning from the field usually claim that they covered all households in their areas. In the past, management had only a limited possibility for checking the completeness of enumeration. The traditional recourse is to perform a post-census survey, to estimate the percentage of under coverage. With GPS, waypoints taken could easily and quickly be plotted over the imagery on which all houses are visible. It will become immediately clear which households have not received a waypoint and most probably have not been interviewed. Not only can the management make an assessment of the rate of coverage, but it also can show enumerators immediately which areas they forgot to visit. GPS, as well as imagery, therefore serves as a powerful quality control mechanism of enumerator work because it provides near-real-time feedback. This has never before been possible.

Essential Datasets: Common File & Storage Types for Geographic Data

There are a multitude of file types that a GIS can consume. Several of the most important file types that are essential to understand when creating datasets for the census GIS are explained. There is a wealth of information available on the many different file types and storage capabilities within a GIS. Therefore only the most widely used are outlined here and only touch on the capabilities of each file and storage type discussed. In any case, consulting the help in the GIS is the best resource for further understanding the environment.

FILE & STORAGE
Shapefile
Feature Class
Annotation
Feature Datasets
Geodatabases

Since the advent of computer-based geographic information systems in the 1960's, many different formats for storing geographic data have been developed. The most common being the shapefile, which was developed by ESRI for previous generations of software. While they have been successful, they now place technical limits on users for modeling geographic systems and handling huge amounts of data. **The term GIS data layers referred to in this manual are represented as individual shapefiles on their own or feature classes in a geodatabase.** The technical definition of a **shapefile** is as follows: An ArcView GIS data set used to represent a set of geographic features such as streets, hospital locations, trade areas, and ZIP Code boundaries. Shapefiles can represent point, line, or area features. Each feature in a shapefile represents a single geographic feature and its attributes (Wade and Sommer 2006).

Somewhat similar, Feature classes are sometimes confused with shapefiles. Feature classes are contained in a geodatabase or feature dataset. The definition of **Feature class** is as follows: Thematic datasets that represent geographic features such as enumeration districts, roads and landmarks. Feature classes are tables with a spatial field. An additional column with a shape field is used to specify geometry of point, line or polygon for the feature class.

Textual information associated with GIS layers in a geodatabase is best contained in annotation feature classes. An annotation is a type of feature that provides a textual description of a place such as a landmark or geographic feature such as a streets or rivers. There are two types of annotations in a geodatabase. **Standard annotation** is not formally associated with features in a geodatabase. **Feature –linked annotations** are associated with a specific feature in a feature class. Annotation layers are a powerful tool used to store textual information and contain symbology including font, size, color, and any other text symbol property. They can also include graphic shapes such as ED canvassing arrows and boxes.

Feature Datasets are contained in a geodatabase and is a container of feature classes and other datasets that enforce a common spatial reference. These datasets are used to spatially integrate related feature classes. The primary purpose of the feature dataset is for organizing related feature classes into a common dataset for building topology, network datasets, or a geometric network. In census work, feature datasets are useful for organizing feature classes for data sharing since many NSOs have multiple GIS personnel.

In the 21st century the geodatabase has become prevalent. The geodatabase was created by ESRI in response to user requirements for increased scalability for large datasets, open access to data, and allows for modeling geographic systems such as census. Geodatabases were borne out of the maturity of relational databases and their increasing ability to understand the specialized requirements of geographic data. The technical definition of a Geodatabase is as follows: The geodatabase is a collection of geographic data stored using a database management system (DBMS) or file system. The three fundamental geographic datasets used in a geodatabase are tables, feature classes, and raster datasets.

There are 3 types of geodatabases - from smallest to largest they are the Personal, File and ArcSDE geodatabase. **Personal geodatabases** are based on Microsoft Access Software and was the original geodatabase available in the beginning after the release of ArcGIS 8. It is designed for single users working with small datasets having a maximum size of 2 gigabytes.

File Geodatabases were introduced in ArcGIS 9.2. It is ideal for GIS projects done by single users and small workgroups. It improves on the personal database in terms of scalability, performance and scale up to handle very large datasets by using an efficient data structure that is optimized for performance and storage. The file geodatabase can support more than one editor at a time, provided the editors work on different tables, stand alone feature classes, or feature datasets. Of all three geodatabases, the file geodatabase is the most efficient for government projects such as census.

ArcSDE geodatabases are designed to be scalable and to support multiple readers and writers concurrently. Each has the ability to manage a shared, multi-user geodatabase. It allows permissions to be set on the data and the geodatabase can be implemented on many relational database platforms. It can be scaled to support any number of users. Implementation of an ArcSDE, or commonly termed enterprise database, will allow the organization to leverage all the geographic data with the most minimal constraints. The SDE geodatabases is the end result of an established, sustainable server GIS created to serve internal and external users simultaneously. These geodatabases take a significant amount of planning and careful consideration before implementing and are normally implemented when a country has a comprehensive GIS dataset that can be leveraged accordingly.

Essential Datasets: Creation of Critical GIS data layers

Hand-drawn maps used in the past are now inadequate for the task of locating housing units for a census of population. Also, recent studies suggest that intra city variations in poverty and health in developing countries, such as select CARICOM countries, may be greater than differences between urban and rural populations (ASPRS) prompting the capture of larger scale geographic datasets and changing the traditional dimension of statistical offices and their once sole focus on EDs. While the ED layer still remains the most critical GIS layer following roads, rivers, landmarks and other complimentary datasets, there has been significant interest in capturing building footprints. Currently in the Caribbean, many countries focus on the capture of ED boundaries, roads, and rivers and their associated annotations. Within the capture of these critical datasets, there is a movement to capture or prepare for the capture and integration of building information and the necessary GIS tasks involved in preparing for the capture of such data. This movement in the Caribbean has proved to solidify GIS as a critical statistical office counterpart once thought to be a somewhat separate exercise.

Contemporary statisticians now realize that Geographic Information Systems do not realize their full potential without the ability to carry out methods of statistical and spatial analysis, and an appreciation of this dependence has helped to bring about a renaissance in the field. Also, straightforward application of standard methods, particularly at the ED level, ignores the special nature of spatial data, and can lead to misleading results (e.g. spatial autocorrelation, modifiable areal unit problem, etc).

ED Layer

The census enumeration, acquired through fieldwork, is the gold-standard methodology for counting people in an area. Delineating spatial units at the Enumeration District scale for which disparate socio-economic, health, development, and environmental data can be optimally summarized to support spatial statistical analysis is of great demand worldwide. Information from censuses serves to affect and influence decisions in these areas and their different issues through the analysis and use of information captured within these spatial units. Failure to accurately delineate these spatial units generally results in a larger margin of error in the enumeration exercise. Thus the Enumeration District boundary layer becomes the most critical guiding layer in a GIS as it forms the extents of the spatial unit used in censuses, to be displayed on the census map itself. Therefore careful consideration must be given in developing this layer as errors in the boundary can cause significant issues during the census.

UN suggested Best practices in the delineation of ED boundaries indicate that EDs should:

- Be mutually exclusive (non-overlapping) and exhaustive • (cover the entire country).
- Have boundaries that are easily identifiable on the ground.
- Be consistent with the administrative hierarchy.
- Be compact and have no pockets or disjoined sections.
- Have populations of approximately equally size.
- Be small and accessible enough to be covered by an enumerator within the census period.
- Be small & flexible enough to allow the widest range of tabulations for different statistical reporting units.
- Address the needs of government departments and other data users.

- Be useful for other types of censuses and data-collection activities as well.
- Be large enough to guarantee data privacy.

Among these criteria are some that facilitate census data collection, while others pertain to the usefulness of EAs in producing output products — i.e., the relationship between data collection and tabulation units. It should be kept in mind that the purpose of a census is to produce useful data for administrators, policy makers and other census-data users. Maximum flexibility and suitability for producing the best possible output products should thus take precedence over convenience of census enumeration. However, EA delineation must also make sense logistically for field operations.

The size of enumeration areas can be defined in two ways: by area or by population. For census-mapping, population size is the more important criterion, but surface area and accessibility also have to be taken into account to ensure that an enumerator can service an EA within the time allotted. The chosen population size varies from country to country and is generally determined on the basis of pre-test results. Average population size may also vary between rural and urban areas since enumeration can proceed more quickly in towns and cities than in the countryside. Under special circumstances, enumeration areas that are larger or smaller than average may have to be defined. For most practical purposes, the population size of an enumeration area will be in the low to mid-hundreds.

Before delineation of EA boundaries, the number of persons living in an area and their geographic distribution need to be estimated. Unless there is information from a recent survey, a registration system or some other information source, these numbers need to be determined by counting the housing units, determining the associated number of households and multiplying by an average household size. The number of housing units can be determined through cartographic fieldwork, cooperation with government officials, extrapolation from previous census results or by means of aerial photography or satellite imagery.

Enumeration area boundaries need to be clearly observable on the ground. Even if they do not have considerable geographical training, all enumerators need to be able to find the boundaries of the area for which they are responsible. Thus, population sizes between enumeration areas may be varied in order to produce an easily identifiable delineation. Natural features that can be used for this purpose are roads, railroads, creeks and rivers, lakes, fences, or any other feature that defines a sharp boundary. Features with more gradual edges, including brush, forests and elevation contours, such as ridges, are less ideal. Administrative boundaries are often not visible. In some instances, it is unavoidable to use EA boundaries that are not clearly visible on the ground. In this case, an exact textual description and appropriate annotation on the EA maps is required. Examples are offset lines and extended lines. For example, an EA boundary may run parallel to a specific road at a clearly defined offset. Or a portion of an EA boundary may be defined as the extension of a clearly visible road to another clearly defined feature, such as a river or railroad.

Specific issues related to EA delineation will be encountered in many countries. For instance, while villages may be assigned to specific administrative units, the actual boundary delineating the village area may not be defined. Also, special populations, such as transient, nomad or military personnel, need to be assigned a geographic reference. Again, naval personnel are often assigned to their home ports. When planning for locating hard-to-enumerate populations, it should be recalled that operational costs are sometimes 10 to 20 times higher than for residential populations in urbanized areas.

Among the criteria for EA delineation is determining the ideal EA size, which is based upon the number of people one enumerator can count in the time period scheduled for data collection. The plan for EA delineation should reflect the overall census plan, stemming from the number of days allotted for enumeration. A census pre-test can determine the number of housing units (HUs) that an enumerator can cover per day.

As an illustrative example, if 16 HUs can be enumerated per day in urban areas but only 10 per day in rural areas, and if the period of enumeration is 10 days, then the ideal urban EA would contain 160 HUs and the rural EA would contain 100 HUs. If the average number of persons in an HU is five, then the ideal population size would be 800 for an urban area and 500 for a rural area. Other factors influencing the size of an EA include administrative area boundaries; the visibility of EA boundary features; the presence of collective living quarters, such as barracks, hotels and school dormitories; and the mode and availability of transportation.

Population estimates are essential for proper EA delineation. Local officials can be called upon to provide small-area estimates, or NSO field personnel can visit the areas in question. In areas that have not experienced dramatic change, estimates can be adjusted from the previous census based on the time elapsed. (UN Handbook P37)

Road Layer

Road infrastructure is in many cases the most permanent geographic feature used to guide enumerators and delineate boundaries in a census. Often, roads layers are polylines representing streets. These are developed by creating a single line for the road with a name. Once all of the lines are created, there are tools available that can create intersections. These GIS tools create a point or vertex at the implied intersection of two segments - implied means that the segments don't have to actually intersect on the map. The tools create a point or vertex at the place where the segments would intersect if extended far enough. For example, suppose you want to create a road intersection. One part of the line (street) must be placed at the implied intersection of two segments of an adjoining line (street). Intersection tools find this implied intersection point and create the intersection vertex of the new streets. Attribute information, when complete contains items such as text map features that are the name of the road at that given geographic location; the name includes the directional, street name, street type, and possibly a post-directional.

In many cases in the Caribbean this type of street layer development methodology and entity type definition described above is not used or available. The street layer is digitized from mapsheets or imagery as a series of polylines. They are given names when available or at a later time, since, in several instances street names may not be determined until field visits and archival census data are compared. The NSO GIS unit digitizes segments between each junction, using snapping to make sure each line segment connects with the next. This method has worked well in creating the necessary road layer for the census but often falls short as an "official" road layer until street name, street type, and direction information are complete. The roads layer is often managed by the national transportation authority in the country and is a difficult endeavor for NSO GIS units where time is rarely available to complete this layer.

River Layer

This layer is often administered by national water, land or environmental authorities and again is often not developed by the NSO GIS unit. This hydrology map layer represents the water features for the country and can be polygon (e.g. lakes and ponds) and/or waterways (e.g. rivers, creeks, channels). This layer would ideally contain line features defining the boundary edges of hydrologic entities such as creeks, lakes, and streams. Most often, in the Caribbean, the river data used in census entail a single polyline layer and is not complete, lacking attributes such as river width.

Building Layer

Data capture through the use of imagery can provide building level information as a single GIS data layer in greater detail for census when compared to the use of a GPS device. This building level layer is key in guiding the enumerator to the necessary dwellings and households. Rather than a GPS point feature, its aim is to provide a unique building outline showing the shape and relative size of the building. This allows for greater detail beyond just the ED. Also, the building layer can be linked to the questionnaire through unique coding. This allows the statistics office to aggregate out from the building level to any specified geography (survey boundaries->ED->Administrative Dist., Parrish etc.). However, linking at this level takes time and the census map creation exercise should always focus first on the EDs and later consider further geographies.

The building layer would consist of a unique ID for each building falling within the ED or associated geographic hierarchy. In the case below the housecode is found to be within the buffer (block groups) District> Locality> Buffer> HOUSECODE since the geographic hierarchy consists of sublevels below the ED.

OBJECTID *	SHAPE *	DISTRICT	LOCALITY	BUFFER	HOUSECOD	SHAPE_Length	SHAPE_Area
1	Polygon	3	029	0198	0029	144.855159179021	1265.02706577436
2	Polygon	3	031	0999	0020	146.895876007486	1275.9272275238
3	Polygon	3	029	0999	0001	53.3379095407213	100.503367906585
4	Polygon	3	029	0175	0010	8.93386930836604	3.47776863547793
5	Polygon	3	029	0175	0009	7.03623550758379	2.95465749849437
6	Polygon	3	029	0175	0008	8.44438649464222	4.41760415214712
7	Polygon	3	029	0175	0007	9.18678687784741	4.91968869875913
8	Polygon	3	029	0175	0006	7.19573825822845	2.57112564366863
9	Polygon	3	029	0175	0005	5.8577714803577	2.09777309001292
10	Polygon	3	029	0175	0004	9.83752071772764	5.89135858068909
11	Polygon	3	029	0175	0003	6.64580167864618	2.75589496306008
12	Polygon	3	029	0175	0002	9.94693800632062	6.04839492162136
13	Polygon	3	029	0175	0001	6.8960845742198	2.9600348169208
14	Polygon	3	001	0128	0006	18.277510240663	12.2590194778029
15	Polygon	3	001	0128	0005	10.2329737847657	6.4925245573814
16	Polygon	3	001	0128	0004	7.44613908842548	3.39455221662578
17	Polygon	3	001	0128	0003	10.8762954172372	6.93771306110147
18	Polygon	3	099	0027	0013	39.116741233256	80.2502404158595
19	Polygon	3	099	0019	0002	13.2895623817784	10.6992168234472
20	Polygon	3	099	0019	0001	16.7671811947315	16.0212869986329
21	Polygon	3	099	0012	0001	15.6963212344907	14.52848667093
22	Polygon	3	001	0125	0003	42.1111038192313	41.1039859735648
23	Polygon	3	001	0125	0002	10.4909375855853	6.56777851649848
24	Polygon	3	001	0125	0001	11.5241424798513	7.71390927978402
25	Polygon	3	001	0124	0003	17.9652183663182	15.6747693911655
26	Polygon	3	001	0124	0002	9.75477380841929	5.92707716152692
27	Polygon	3	001	0124	0001	26.4471822564717	31.2136175431149
28	Polygon	3	001	0123	0002	16.0701729615884	15.9802092229517
29	Polygon	3	001	0123	0001	34.4814783305671	54.3742942815533
30	Polygon	3	001	0122	0004	9.64250375905202	3.98747989624997
31	Polygon	3	001	0122	0004	26.6478952140462	42.3521443835203
32	Polygon	3	001	0122	0002	14.5161059610265	12.7040445100087
33	Polygon	3	001	0122	0001	22.8277537490158	29.9143471482152

0 (of 25047)

Preview: Table

Figure 8 Practical view of an attribute table of a building layer in a GIS showing individual building features and common fields.

These layers are often administered by other National authorities but in many instances these datasets be incomplete or not available and have to developed/created by the NSO GIS unit. It cannot be understated that in many instances the creation of Road and River network and building GIS layers have been developed by NSO GIS units, which puts undue burden on the NSO GIS staff as they are already understaffed, under funded and under-skilled to manage the development of such critical datasets. It also takes away significant time that could be otherwise spent on preparing for geographic and statistical exercises and linking of statistics through RDMS or related tables in GIS. More collaboration and assistance is needed in the development of these datasets by national agencies in the Caribbean. This can be through the establishment of a Mapping sub-committee or data sharing agreements.

It is also important to note the GIS layers described here may not be representative of the data available in the country. These situations are reflective of what NSO GIS personnel ought to have available to them for the census.

Database Design Concepts

The development of a tailor-made sustainable path towards a census GIS for collection and dissemination of geostatistical purposes within the constraints of the country must be carefully assessed. A range of available technical and logistical options are available and is largely is dependent on database integrity, budgets and time. From a set of scalable options, the census office can select the subset of techniques and procedures that best fit the needs of the country. Provided below is a non-exhaustive list of options in increasing order of complexity.

- Production of digital maps created on the basis of existing sketch maps;
- Imagery at a high resolution less than 5 meters;
- Geo-referenced enumeration area layers or spatial data files with proper coding and metadata that can be properly integrated with other digital geographic databases;
- Inclusion of geographic reference layers, showing, for instance, roads, rivers, Landmarks, point features and other features-these can be included as simple images from scanned maps, or designed as a unit's own geographic database
- A digital database of precisely located dwelling units, created with the aid of Geographic positioning systems; and
- A digital postal address registry where addresses are matched automatically or semi automatically to digital road databases (UN Handbook).

The Geographic Database: Census Geodatabase Design

The census GIS and resultant geographic database is not a comprehensive database of all country data. Rather it contains the necessary data to conduct the census and be flexible enough to use in statistical exercises such as sampling and surveys. The optimal situation is where the development of the census GIS compliments the development of all data included. This means that the data contained in the census GIS is further improved for accuracy (i.e. updating attributes and adding features) and then handed back to respective collaborating agencies to be built further, eventually participating in the National Spatial Data Infrastructure (NSDI). In many cases, collaborative agencies such as Physical Planning, Survey Departments, and Natural Resource Management Agencies can significantly build on the geographic data developed/improved upon during the census exercise.

In many cases, based on design of the geographic database (and information) and country specific requirements, each NSO has an established census geographic hierarchy; geographic coding scheme; administrative and census units listing; tiling of national territory into operational zones for census mapping. A practical view of the census geodatabase that countries are developing and implementing is provided in Figure 5.

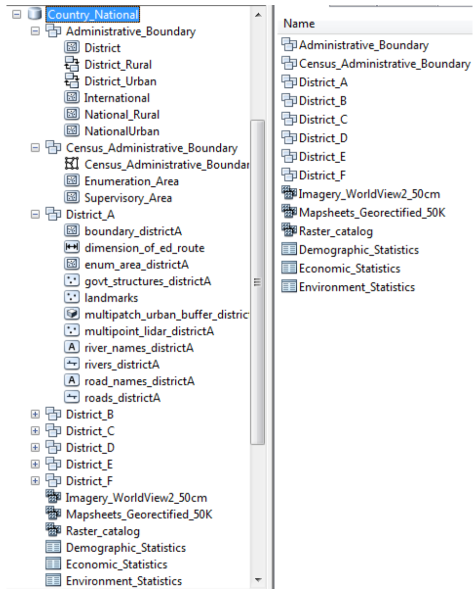


Figure 9 View of a Census Geodatabase displaying the geodatabase and containing feature datasets, feature classes, raster datasets, and tables.

A well structured geodatabase and associated feature datasets (populated with feature classes) ensure sustainable GIS development so that data are more effectively shared and built upon. Depending on the spatial relationships that are most important for the census GIS personnel to maintain, the geodatabase allows more flexibility in terms of data management and editing. It also allows consistent checks for accuracy through the use of topology rulesets and other features. It is faster when used in the GIS, more robust, contains various types of data, and has significantly higher storage capability. An applicable conceptual data model for the geodatabase with various data types included is provided in Figure 10. This model can be referred to as a data model in GIS jargon. The particular model is developed with consideration of the census datasets and hierarchy reflective of CARICOM member states NSOs. The geodatabase should be used to organize and maintain the dataset. Also, an archive of all existing data in the model can easily be exported into another geodatabase or as individual shapefiles.

Data Model: Often refers to the detailed structure of a geodatabase, feature datasets and linkages

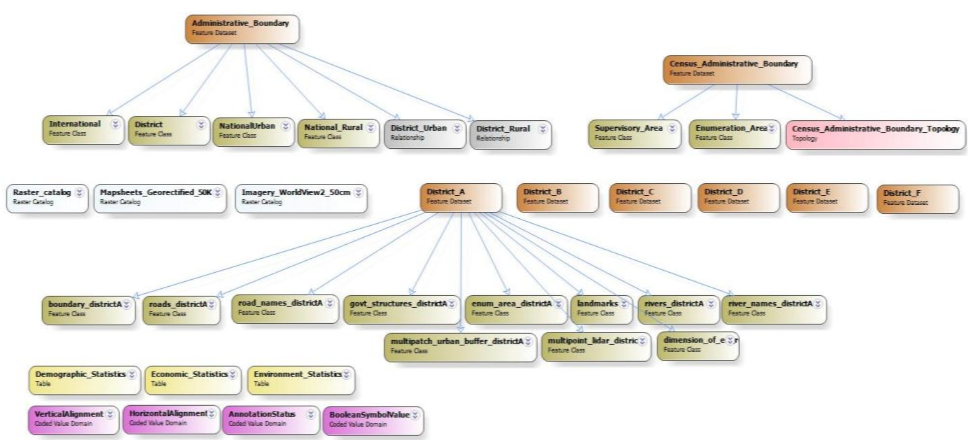


Figure 10 Simplified Data Model diagram displaying the internal linkages and various data types commonly found in a Census GIS.

Special consideration must be given to the spatial relationships in a geodatabase, particularly with regard to the ED layer of which the Statistics office is often responsible for. Transforming geographic data so that ED points become nodes of polygons requires defining them so that they know what they are in relation to other objects. Topology rules govern the polygon relationships of features within a given

GIS layer/ED layer, while others govern the relationships between features in two different GIS layers. Topology rules can also be defined between subtypes of GIS layers in one or another GIS layer. This could be used, for example, to require EDs to be contained within and not overlapping district boundaries in their entirety. Future development and use of GIS for many other purposes beyond topology require the use of the geodatabase structure in order to sustainably development of the geographic infrastructure (such as the creation of annotation classes, relationship classes, geometric networks).

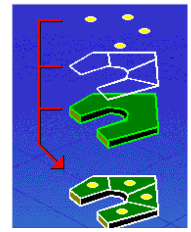


Figure 11. Require layers to be contained within and not overlapping other layers.

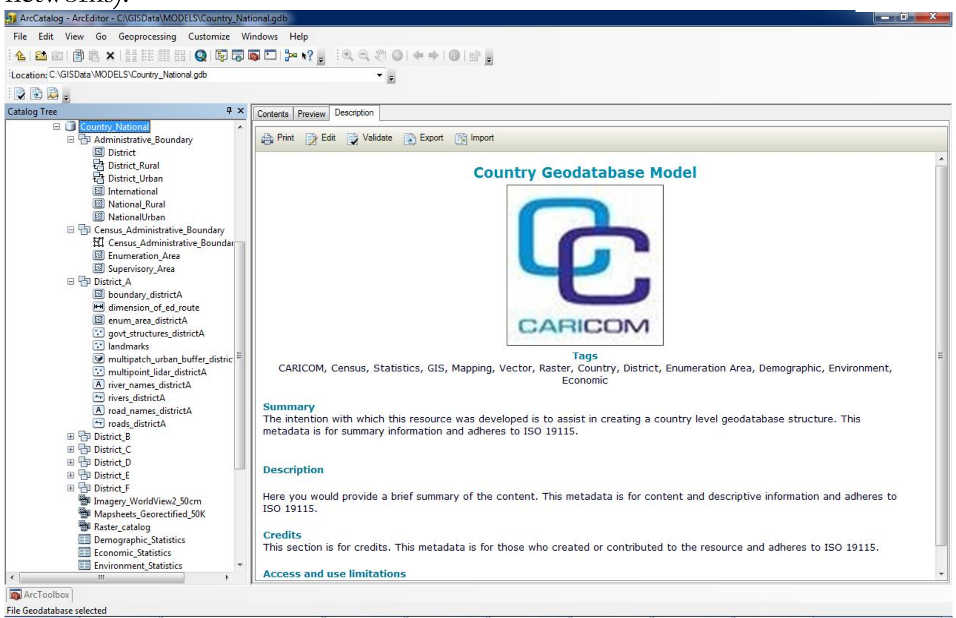


Figure 12 Hypothetical Country Geodatabase census model displayed in GIS software, ArcGIS ArcCatalog.

Harmonized Coding of the Administrative Hierarchy

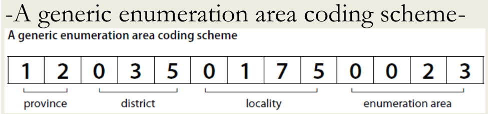
As stated by the UN (2009) Geographic coding or "geocoding" of EAs and the geographic boundary hierarchy is commonly developed using an internal identifier and is not usually accessible by the user and should not be modified externally. A more meaningful identifier is needed that can be used to link the geographic features to the attributes recorded for them. For the enumeration areas and administrative units, this link is the unique EA or administrative identifier, which is listed in the master file of all geographic areas relevant in the census.

How this identifier is entered is software and process specific. It can be added during the digitizing process by entering the identifier before the feature is digitized. Or it can be added at a later stage by selecting the feature interactively and adding the identifier through a menu interface. For polygon features, some systems require the user to add a label point that is contained in each area unit. While conceptually simple, coding may require considerable time and resources.

Indeed, a unique code needs to be assigned to each enumeration area. This code is used in data processing to compile enumerated information for households in each EA and to aggregate this information for administrative or statistical zones for publication. This is the numeric code that provides the link between the aggregated census data and the digital EA boundary database stored in a GIS. Ideally, the coding scheme needs to be determined on a country-by-country basis. However, the rules used to assign codes need to be unambiguous and should be designed in collaboration within the NSO, especially between the mapping/geospatial data and those managing the data core. The most important principles in the design of a coding scheme are flexibility, expandability, intuitiveness and compatibility with other coding schemes in use in the country. The statistical office is often the custodian of coding schemes in the country and should also be the focal point for the design of the census-mapping codes.

A hierarchical coding scheme will usually facilitate the consistency and clarity of the numeric identifiers. In this approach, geographic units are numbered at each level of the administrative hierarchy — usually leaving gaps between the numbers to allow for future insertion of newly created zones at that level. In other words, the gaps that are left will depend on how many units are added. For example, at the province level, units may be numbered 5, 10, 15 and so on. A similar scheme would be used for lower-level administrative units and for enumeration areas. Since there are often, for example, more districts in a province than provinces in a country, more digits may be required at the lower levels. The unique identifier for each smallest level unit — i.e., Handbook on Geospatial Infrastructure in Support of Census Activities the enumeration area — then consists simply of the concatenated identifiers of the administrative units into which it falls.

For example, a country could use the following coding scheme:
-A generic enumeration area coding scheme-



An EA code of 1203501750023 means that enumeration area number 23 is located in province 12, district 12035 and locality 120350175. The unique code is stored in the database as a long integer or as a 13-character string variable. Obviously, the variable type needs to be the same in the census database and in the geographic database. Storage as an integer variable has the advantage that subsets of records can be selected easily, using standard database query commands in any database management system or GIS package.

Storage of the code as a character variable, on the other hand, can improve consistency, for example through the use of leading zeros. In this case, the code is considered a name made up of American Standard Code for Information Interchange (ASCII) characters rather than a sequential number.

In cases where administrative and reporting units are not hierarchical, special coding conventions need to be developed. In any case, it is very important to be completely consistent in defining and using the administrative unit identifiers since they are the link between the GIS boundaries and the tabular census data. The census office should therefore maintain a master list of EA and administrative units and their respective codes, and should commit any changes made to the master list to the GIS and census databases. The NSO might also consider publishing a list of EAs, including relevant coding and the latitudes and longitudes of centroids (central points), possibly with populations enumerated as well.

Fortunately, the numbering of EDs in the Caribbean region is relatively straightforward and manageable in the countries do not have to have trailing zeros to accommodate significantly high (e.g. 1000) more EDs in the near future. In many cases there are normally 3-4 geographic hierarchies that are commonly nested. The common hierarchy is the National Parish/District Community/□ Enumeration District□ Neighborhood/Block. The most common geographic boundaries in the GIS that make up the boundary coding in CARICOM countries are the District□ Parish□ Enumeration District and is comprised of 5 digits: 2 for the District/Parish level and 3 for the ED number. At the building footprint layer level this coding may be an additional 3 digits for the building number (8 digit coding to link building to the geographic boundary hierarchy).

Any other geographic features such as landmarks or structures are not enumerated during the census.Census data itself are commonly related to these geographic boundaries by way of a relationship class within the GIS. A table relate is created so that one can query and select features in one GIS layer and see all the related features in another layer or table. A relationship class is an object in a geodatabase that stores information about a relationship between two feature classes, between a feature class and a non-spatial table, or between two non-spatial tables. Both participants in a relationship class must be stored in the same geodatabase. Cardinality controls how relates and relationship classes are set up. Cardinality is a description of how records in two different tables are related to one another. Cardinality may be one-to-one, one-to-many, many-to-one, or many-to-many. Table relates are created between tables that have one-to-many or many-to-one cardinality. Relationship classes support all cardinalities. When you create a table relate or a relationship class, you specify the field in each table that the relationship will be based on. The fields must be the same data type (i.e., text, short integer, long integer, object ID, etc.).

With the exception of a few CARICOM member states, many countries in the region do not have an established addressing system. In many instances street names do not exist for many parts of a given country let alone address numbers. To work around these inconsistencies and issues countries have moved toward using GIS to create a building footprint that can be indexed and queried. This has empowered countries to better understand the needs of the country through having the ability to aggregate at any geographic level and better understand statistical trends. It should be noted that countries are currently moving fast to obtain this level of accuracy before the census but in many cases this work is still not well understood and proper priority has not been given to this task. Therefore, this building footprint currently exists only in a few countries.

Accuracy Requirements and Standards

A determined level of accuracy is needed for the ED maps at an early stage, preferably early in the inter-censal period. The level of accuracy is country specific and varies due to geographic and demographic elements. Early work in catering the design of the cartographic elements and their levels of accuracy to the needs of the enumerator; field statistician; district manager; regional manager etc. is valuable for determining map standards.

Standards vary widely and standards such as nested geographic hierarchies or a bar code system for map identification, in which the application uses this information to select and retrieve relevant maps from a database, can automate processes that used to take more time and resources to. Field work also takes significant time and the processes can often be standardized to maximize efficiency and speed up the validating of the field work in the geographic database. This can be though minimizing the field GPS data collection activities through improved accuracy or adding features to the field worker maps for quick orientation.

It is important that NSOs work closely with other offices to maintain accuracy requirements set out by national GIS community whenever possible. Government Agencies may provide these standards in the datasets themselves in the form of standard fields in an attribute tables for GIS data layers or simply through a symbology, sets included in feature classes or the metadata. It may also be through verbal/paper/electronic communication of law. The most commonly held standard with concern to census is the requirement for nested geographies.

This requires the precise delimitation of ED boundaries and their related and often overlapping datasets. More rigorous standards are often set on the ED layer with respect to the nested geographic boundaries that, for instance, can be at the Parish level where the census geographies must adhere to the other geographic hierarchies of the country as defined by government agencies.

One example is from the census mapping activity of Belize whereby boundaries for the districts in the country were required to adhere to statutory instruments. These statutory instruments are in legislation set by the government that defines town boundaries. These boundaries have been approved through consultation with the Town Boards but did not previously exist with high accuracy in the GIS. As EDs are nested within these boundary standards it was critical that the geographic accuracy be captured in the census GIS. The broader process is of delineating these boundaries are outlined below:

Correct Urban Boundaries using the Statutory Instruments:

- Legislation by Govt. of Belize approved with consultation with the Town Boards on establishing the town boundaries.
- Plot the Statutory Instrument Coordinates in ArcGIS
    - Distance and direction information is captured from the Statutory Instrument and input into the GIS
    - Maintain as much node/vertex information as possible according to the SI directions
    - Adjust Urban/Rural Town Boundaries (NAD 27)
- Ensure that the boundaries suit the information from the Statutory Instrument
    - Taking into consideration the routes of streets so that the EDs are easier to follow and more defined (e.g. if the line passes though a house then include/exclude the house rather than go through it)
    - Use the roads, rivers and other geographic features as much as possible
    - Add the attribute information on the towns/cities boundaries
    - Save the layer

- Conduct the Splitting, where possible, of Enumeration Districts and Create the draft ED Layer according to the delineated Statutory Instrument
    - Export the Urban/Rural layer and use the newly created Urban/Rural .shp to delineate the EDs according to the new Urban/Rural boundaries as determined by the Statutory Instruments.
    - These EDs are to be guided by 2000 census boundaries

While administrative hierarchies used in census are commonly nested, there are boundaries that are arguably controversial to the census as a statistical exercise itself. This arises when ED boundaries/census hierarchies are determined through the use of electoral or constituency boundaries. Most CARICOM countries are not required to nest the census hierarchies/EDs within the electoral or constituency boundaries. These constituency boundaries are entirely political, largely subjective in nature and arguably take away from the statistical exercise of the census. Such political boundaries often do not adhere to international standards for census boundary delineations; do not follow geographic features; tend to be more related to socially accepted relationships rather than statistical exercises; and tend to change often. Furthermore, while the office may be capable of accommodating the development of these new boundaries and their effect on ED boundaries, it takes away time for other critical GIS activities and development during the census and should therefore be an entirely separate activity.

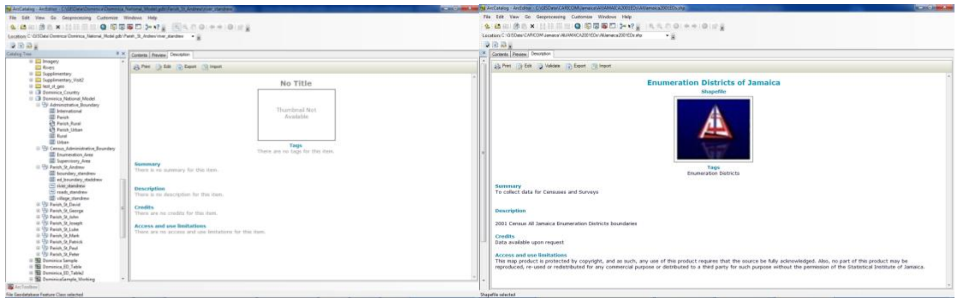
These accuracy requirements fall in line with the careful evaluation of the specific goal and anticipated result of the geographic exercise. Improving workflows in the field in order to reach the required targets in an effective, thorough and timely manner and keeping in mind continuous development of geographic activities for future application accurately is critical.

Metadata Guidelines

Metadata creation is an ongoing process.Integrating information (e.g. data set names, format, sources, narrative content etc.) into the census geography workflow is an important factor that is often taken lightly even though most GIS packages allow for metadata to be easily entered using a standardized, streamlined interface to do so. The documentation that will accompany the data set distribution or disseminated data (i.e. meta-data) does not have to be as comprehensive as the in-house information that is compiled for all databases. Simple ASCII ("read-me") text files can be used. Data users will usually not need detailed information on data lineage or processing steps, and ease of interpretation is more important for external users. Thus, the documentation should contain a clear, concise and complete description of those aspects of the database that are relevant to a user. Provided that the census office maintains a comprehensive meta-database, the user-targeted data documentation can be compiled very quickly.

All too often empty…

NO Metadata   Metadata Completed



Documentation is descriptive information supplied by a person when she/he edits information in the description tab such as the units of measure for data stored in a field and the information that data represents. Good documentation protects your investment in the resources you have created. You can have more confidence in your decisions when you know your data is accurate, current, and derived from a reliable source.

Properties describe inherent characteristics of the item such as the extent of features in a feature class or the location of a text file. By default, ArcGIS derives an item's properties and adds them to its metadata automatically when metadata is viewed, validated, exported, and imported. Properties can also be updated in an item's metadata using the Synchronize Metadata tool whenever it is appropriate to do so. Adding properties to metadata automatically supplements good documentation and helps reduce metadata maintenance costs. Because ArcGIS automatically handles the properties, all you need to worry about is completing the documentation.

Each GIS data set should be accompanied by a data dictionary that provides information for each individual GIS data layer or data table. The term "data dictionary" is an older term that refers to the specific formatting and field names in the data set. The information provided includes file names, formats, feature types (points, lines or polygons) etc. These fields are also commonly found in GIS today and can be easily entered. The data documentation and data dictionaries can also be incorporated into a comprehensive users guide. A user guide might contain a more detailed explanation of database content, data lineage and quality. Step-by-step explanations of example applications or copies of census maps created with the database can also be included.

A typical example of a data dictionary is provided below. This is a sample data dictionary for the distribution of a census geographic database of localities for the hypothetical country and further comprehensive information can be found on pp 195 of the 2010 UN Handbook.

Field name	Description	Field definition	Range	Codes	Missing values
AREA	Area of the locality in km²	Real, 6.1	Positive value	None	-999
FLAG	Indicates whether the polygon is the major one for the locality. For localities that consist of two or more polygons, only the biggest or most important will have a value of 1	Int, 1	0-1	0-minor 1-major	
URBAN	Indicator whether the locality is classified as urban or rural.	Int, 1	0-1	0-rural 1-urban	-1
LOC_NAME	Name of locality	Char, 25	None	None	"n. a."
DIST_NAME	Name of district	Char, 25	None	None	"n. a."
PROV_NAME	Name of province	Char, 25	None	None	"n. a."
AREA_TOTAL	Total area of locality in km²	Real, 10.3	Positive value	None	-999
AREA_LAND	Area of locality covered by land in km²	Real, 10.3	Positive value	None	-999
AREA_WATER	Area of locality covered by waterbodies in km²	Real, 10.3	Positive value	None	-999

POP.DBF—census population indicators

Field name	Description	Field definition	Range	Codes	Missing values
LOC_CODE	Official locality code. Provides the link to the GIS attribute data tables loc.dbf and hh.dbf.	Int, 8	Positive value	None	-999
POP_TOT	Total enumerated population	Int, 7	Positive value	None	-999
POP_DENS	Population density in persons per km² (POP_TOTAL / AREA)	Real, 5.1	Positive value	None	-999
...	...	...	...	...	...

HH.DBF—census household indicators

Field name	Description	Field definition	Range	Codes	Missing values
LOC_CODE	Official locality code.	Int, 8	Positive value	None	-999
HH_NUM	Number of households.	Int, 7	Positive value	None	-999
HH_HEAD	Sex of head of household.	Int, 1	0-1	0-male 1-female	-1
...	...	...	...	...	...

Table 2 Sample data dictionary for the distribution of a census geographic database of localities for the hypothetical country.

Existing Datasets and Institutional Arrangements

Governmental administrative units are often most relevant for the census office. Clearly, enumeration areas need to be consistent with the boundaries that form the administrative hierarchy in the country and population distribution. But data layers, such as transportation and hydrography, are also very important for census-mapping since roads and rivers form a natural delineation for enumeration areas.

This requires NSOs to work together with and the larger geographic community to develop the data collection/extraction protocols for CARICOM Member States. This includes identifying external datasets that add value to census map creation and the mapping workflow such as boundary delineations and feature identification data. In past census activity most Caribbean countries relied heavily on assistance from their resident Lands and Surveys or Urban Planning Departments; it is critical that government agencies be identified that produce digital geographic databases. National mapping agencies increasingly use fully digital techniques in the entire map-making process. Other Governmental Departments, including transportation, health, and environment, also use geospatial technology to manage the information they collect and use for analysis and planning. Additionally, private sector in the field of utilities, telecommunications and mining sectors, have realized the advantages of managing their information needs in digital geographic form.

The establishment of a Mapping Subcommittee consisting of a wide range of expertise from different government agencies increases accuracy and speed of the ED boundary delineation process. This committee works closely together to come to agreement on proposed boundary changes not strictly determined by the NSO and needed for the census. This activity saves significant time, reduces duplication and creates transparency throughout all parties in use of the geographic boundaries. Each member brings a range of specific knowledge and skills to the table. Normally consisting of 1 or 2 people from each sector, there are govt. architects, GIS specialists, statisticians, policymakers, development strategists and associated managers all working together for the national common interest in creating a coherent geographic hierarchy for use in the census and for the development of the most efficient geographies for future use. This committee meets an average of 1 or more times a week depending on workload until the boundary changes are finished according to the census timeline. Such a committee allows a sharing of knowledge and education on the intricacies of the census and, more specifically, the geography of the country.

The Mapping Subcommittee proposed boundary changes would be approved by the necessary bodies - during the meeting - through the initial sketches, photography of the proposed changes, and through a GIS map outputs paper format and/or digitally. With these mediums the process is relatively straightforward, transparent and the changes are adopted in a more efficient manner. The Mapping Committee efficiency is due to the tremendous knowledge of what is on the ground as noted by the committee members in their regular work.

The mapping Subcommittee meeting often focuses on identifying EAs to be split that fall within other jurisdictional boundaries that are not directly census related but have to be nested (e.g. constituency boundaries) and proposing splits through explicitly delineated geographic areas. This process can also involve aligning boundaries nested within the appropriate geographic hierarchies and making the splits align with parcel boundaries. A GIS can be used to view/project on a wall or screen the different geographic layers involved in guiding the split (e.g. imagery and boundaries) and to identify boundaries and other geographic features. At the same time a larger sized paper map can be used to trace the proposed split for later approval. Real-time digitizing can also be a solution to speed up the splitting of the EAs within the geodatabase to allow a quicker view of the proposed ED layer.

Broad scopes of the Terms of Reference of a Mapping Sub-committee are provided below. These terms must be catered to the country's geographic hierarchy and the responsible parties that create their own statistical reporting units. Not all of these TORs may pertain to a subcommittee of a specific country. For example, the focus currently might be on the subjective electoral boundaries and their inconsistencies with the ED layer currently in use. It is important to keep the exercise simple, with minimal steps, involving other parties only when needed.

Common Terms of Reference of a Mapping Sub-committee

1. Review existing Enumeration Area (EA) maps and respective household listings to identify those that have more than 100/150 households and would require splitting; new boundaries;
2. Review maps of EAs that fall within two districts and assist in determining new boundaries;
3. Review EAs that have boundary lines that dissect block and parcel and assist in determining

new boundaries;
4. Review EAs that have a small number of households and are adjacent to determine if these should be merged;
5. Review district maps to determine clusters of EAs for field supervision maps;
6. Assist in determining the canvassing route for EA maps;

Of most importance are the institutional arrangements and collaboration that exists within National Statistics Offices of Dominica, a small country. These relationships are critical for the success of the Caribbean census. Often understaffed and under funded, CARICOM countries greatly benefit from national and regional relationships during a census. All CARICOM countries collaborate at a national or regional level on ways to improve the census and best practices. CARICOM frequently conducts regional meetings on census activities and supports countries financially and physically through funding and professional support for census. Several other agencies such as ECLAC, UNFPA, UNDP and similar, also support census activities in the region. Since smaller countries are often hindered by lack of staff and training, it is critical that the government agencies collaborate on data and resources needed for the census. The government collaboration that exists in the CARICOM member states is paramount to success. It cases where this support is minimal or nonexistent, the given country is often overwhelmed and takes a task by task approach to its preparation for the census.

This is of utmost importance to the success of the census GIS and map development in Dominica. The CSO worked closely with other government agencies involved in geographic work and also tapped into academic resources by using people from areas of Information Technology and Computer Science. The agencies the CSO directly approached and harnessed the strengths through staff mobilization from was varied and included: Forestry, Wildlife & Parks, Physical Planning, Ministry of Agriculture, Land and Surveys, Ministry of Public Works, Energy and Ports, Ministry of Agriculture.

Hardware, Software, Staff & Training Requirements

Training has traditionally been one of the greatest challenges in census offices. In the past training in census mapping has been either non-existent or minimal. The training is often decentralized and off-site, making it difficult to maintain training over the span of the following census. It requires a geocentric approach to overcome the idea that the census is not a dynamic activity. It requires a long term planning strategy for best results to be achieved. It is important that training modules in preparation and production of maps and in GIS be tailored according to the geospatial infrastructure development as a whole and that it require specific strategies. The following checklist consists of major requirements for catering training. It is important to help identify quickly or build the capacity of the NSO to learn new methods, absorb information and build on existing knowledge.

- Assess available geographic information and data conversion needs
- Address methods such as digitizing and scanning of paper maps orincorporating/using CAD based technology or data lacking topology
- Address Critical Geographic data sources for EA delineation
- Types of maps required
- Inventory of existing sources
- Importing existing digital data
- Geographic data conversion: analog to digital
- Address cartographic design elements in census map products
- Address field work techniques and practices and needed training and/or incorporation of new methods
- Evaluate information technology and provide assistance in maintaining/developing system administration strategies and infrastructure.
- Identify country-specific requirements embedded in census mapping strategies (e.g. remote sensing, geographic tiering, and web deployment)
- Determine levels of training needed and ensure training is accessible afterwards (i.e. redo exercises, provide electronic versions of presentations and reports)

Budget & Timeline for Census Mapping Activities

One of the key goals of the census project is to improve quality and increase output simultaneously while keeping costs within reason. This is associated with identifying areas of timely work and high costs and minimizing the time/cost it takes to complete a given task. It is important to note that critical events can occur and significant changes can take place in a given country's status with regard to environmental and social conditions, making it important that a flexible plan will provide cost-effective, acceptable outputs and prevent any undesirable compromises and outcomes.

Training of Census Geography Staff

It is widely recognized that the ESRI suite of software is commonly employed for census mapping purposes globally. More specifically, many CARICOM countries using this software suite of products have expressed an overwhelming interest and great need to learn more and take further advantage if its capabilities for managing geographic databases in support of census activities. Expertise needs to be provided in order to take advantage of and employ the ESRI products and software. This also applies to other geospatial/non-geospatial software tools and technologies to provide technical assistance and training will where determined. Offices require GIS Managers who are trained in the use and management of the software that includes the technical and organizational components.

Training needs to be tailored according to the census mapping methods employed/decided upon by the NSO and conducted through hands-on exercises, presentations and direct assistance as decided by the mapping team. It needs to be country specific and focus, where possible, on the training of trainers. Training materials should be obtained that work within the capacity of the country to execute them based on currently employed methods, hardware, and software. It is recognized that many countries have maps in non-spatial formats either hand drawn or in a CAD based system and lack topology or projection information. In these instances it is critical that a careful and thorough evaluation be completed in order that an efficient and sustainable method for the production of ED maps is created for use in the census. In many instances, the geospatial data that exists in countries is a patchwork of regions and tends to be at a larger city level. These data are a critical starting point if available and the remaining non-digital geographic data can be incorporated into the workflow over time. Often this translates into creating ED maps for populated placed in a GIS when available and using available non-georeferenced or georegistered maps for surrounding areas.

Training priorities should target activities related to the completion of enumeration area maps and boundaries according to the determined timeline. In the main office, this may involve manual and heads up digitizing of boundaries using satellite imagery and geographic data or other rudimentary methods that satisfy the needs of the NSO. Where unclear boundaries exists and the delineation of the boundary has not been georeferenced, specific housing units in the ED will be identified to guide the enumerator to households in their respective EA. Boundaries can be updated later through georeferencing and digitizing with minimal impact on the enumeration process. As ED divisions are approved, they can then be sent to the district office for checking and then to the main census and statistics office.

It is critical to identify needed training in field methods is as early as possible since it requires a thorough investigation of data collection and validation techniques in order to fully understand its timeliness and the workload of the field personnel. Fieldwork exercises often take significant time so it is necessary to maximize efficiency in order to free up time for further activities (e.g. fixing edits and further canvassing of country). Coupled with fieldwork, the process of digitizing and updating the maps with incoming field data and verifying them may also take significant time. With concern to hand drawn maps, low cost commercial scanning products may help speed up the importing of ED division maps into the GIS database. If these techniques are tested and managed carefully they will minimize the amount of digitizing needed and help to free up time for other geospatial infrastructure development activities. Overall, early identification of the methods used for field verification and the number of available non-spatially referenced maps will help identify and define strategy for training needs prior to the actual field work.

Figure 13 GIS, GPS and Census Mapping Training at the National Training Center, Dominica

DOMINICA (example): Training of census geography staff

Training provided in Dominica was not exhaustive nor was it long enough and has required the census mapping team to compromise in areas where certain topics were not covered or understood clearly. However it reflects the situation in the Caribbean and shows the true situation many countries face with regard to training. This training was the only targeted census mapping training that was provided to Dominica in-country. 18 staff was engaged together in a specific training in GIS for the census. Within 2 week the following topics were covered and you can see that the training covered a significant amount in a very condensed period. Normally this amount of material is covered over a period of months or more. The consultant, Charles Brigham, conducted training.

CENSUS GIS DEVELOPMENT TRAINING IN DOMINICA AGENDA
GIS Training Center, Roseau, July 30-August 13, 2010
Organization of Work

CENSUS GIS DEVELOPMENT TRAINING IN DOMINICA AGENDA
GIS Training Center, Roseau, July 30-August 13, 2010
Organization of Work: Consultant/Facilitator: Charles Brigham

TIME	AGENDA
Friday JULY 30, 2010	
8:30-1:00	- General Introduction - Briefing meeting with Census GIS Trainees - Identification of Sessions - Presentation on GIS Concepts
1:00-1:30	Lunch (in-house)
1:30-5:00	- Presentation on GIS Concepts (cont.) - Presentation/Introduction to GIS fundamentals and census GIS - Software Installation Procedures
Monday AUGUST 02, 2010 (National Holiday)	
9:00-4:00	- Preparation of Exercises (facilitator only) - Preparation of Available Datasets (facilitator only) - Review of Preparatory Activities and Timeline for Census Map Production and Delivery (facilitator only)
Tuesday AUGUST 03TH, 2010	
9:00-1:00	- GIS Evaluation Software Installation Assistance/Confirmation - GIS Evaluation Software Use Overview - Explanation on the importance of using geodatabases - File Geodatabase and populating the File Geodatabase with geographic data exercises
1:00-1:30	Lunch (in-house)
1:30-4:00	- Exercises and discussion on Creating and Enumeration District Area Using GIS

TIME	AGENDA
	- Working with the GIS Data of Dominica/Data Organization/ED-Census-NDSI & General Workflows/Goals - Editing/Updating the Enumeration District Area Layer Using GIS - Understanding/Editing Attributes - Imagery Use & Vector Data Creation
Wednesday AUGUST 04th , 2010	
9:00-1:00	- Working with the GIS Data of Dominica (cont.) - Heads up digitizing of the ED layer with available imagery
1:00-1:30	Lunch (in-house)
1:30-4:00	- Editing/Updating the Enumeration District Area Layer Using GIS - General GIS Exercises - Imagery Use & Vector Data Creation - Heads up digitizing of the ED layer for West Side of Dominica
Thursday AUGUST 05th , 2010	
9:00-1:00	- Heads up digitizing of the ED layer for West Side of Dominica (cont.)
1:00-1:30	Lunch (in-house)
1:30-4:00	- Heads up digitizing of the ED layer for West Side of Dominica (cont.) - General GIS Exercises - Mapping Dominica Household Data
Friday AUGUST 06th , 2010	
9:00-1:00	- Editing/Updating (Merging GIS data) the Enumeration District Area Layer Using GIS - Exercise on Splitting Enumeration District Boundaries - Understanding/Editing Attributes - Understanding Topology Rules - Demonstration on creating Topology rules to manage GIS data
1:00-1:30	Lunch (in-house)
1:30-4:00	- Discussion on automated mapping and census map creation - Presentation and Demonstration of the CARICOM Map Template
	- WEEK 2 - WEEK 2 - WEEK 2 - WEEK 2 - WEEK 2 - WEEK 2 - WEEK 2 - WEEK 2-
Monday AUGUST 09th , 2010	
TIME	
8:30-1:00	- Exercise on automated census mapping and the use of the Map Template - Applications of the Map Template in Census work of Dominica

TIME	AGENDA
	- Making a draft ED MAP for Census of Dominica - Digitizing from available imagery sources for Dominica - Splitting EDs of Dominica using freely imagery and GIS resources
1:00-1:30	LUNCH
2:00-5:00	- Discussion on Fieldwork/Field Deployment and GIS development - Organization of Field Activities - Reporting Structure
Tuesday AUGUST 10th, 2010	
9:00-1:00	- Fieldwork Activities - Updating ED Procedures - Capturing other information (Household Data, etc) - GPS Use/interface programming & development
1:00-1:30	Lunch
1:30-4:00	- Fieldwork Activities (cont.)
WEDNESDAY AUGUST 11th, 2010	
9:00-1:00	- Mapping Subcommittee Discussion/Development and tasks - Field Teams - GPS setup/interface development - Identification of specific issues and concerns (e.g. map editing and updating)
1:00-1:30	Lunch
1:30:4:00	- Discussion on office/field geographic workflow at Stats office - Training / Discussion on geographic data used in census - Software assistance (migrating GIS data to Stats Office)
Thursday AUGUST 12th, 2010	
9:00-1:00	- Presentation on Automated ED Map Production - Practical field structuring with 4 teams deployed to a 4 Enumeration Districts (1 ED each) - Automated Mapping and GIS - Template Development for Map Production
1:00-1:30	Lunch
1:30:4:00	- Discussion on GIS Data needs of the CSO for Future Activities for Census - Provision of Exercises for further training
Friday AUGUST 13th, 2010	
9:00-1:00	- Production of wall maps for assisting census activities - Wrap-up of specific issues and readdressing of pending issues - Wrap-up Meeting with Director - Brief review of activities - Award Ceremony/ Closing

More training was desired in Dominica and most Caribbean countries greatly encourage further training of their staff in GIS for census and statistics. This will continue to remain a demand in the region and in the future as GIS becomes more incorporated in the workflows of statistical activities.

Workshops also contribute significantly to enhancing the capacity for countries to developthe geospatial infrastructure for census. One example is the CARICOM Workshop on Census Mapping and GIS for the 2010 Round of Population and Housing Census: Identifying Imagery Needs, Use and Data Development Strategies for the Census, held in Belize City, Belize, 23-26 February, 2010. As outlined in the agenda (see below for agenda and Annex for full report). The purpose of the workshop was to harness the best practices in the use of imagery to create the necessary datasets for census maps and develop the census and statistics Geographic Information System (GIS). The workshop outlined map production practices currently in place for the 2010 census round through the sharing of country experiences. It also identified key geospatial datasets for census mapping and GIS and provided methods for the creation of such datasets through the use of commonly used geospatial software and tools.

Figure 14 Belize workshop participants. From left to right: Anguilla, Belize, Guyana, Belize x3, Virgin Islands, Montserratt, St. Kitts & Nevis, Jamaica, Antigua & Barbuda, Anguilla, St. Vincent & the Grenadines, CARICOM x2, Barbados, CARICOM x2, Grenada, Suriname, Trinidad & Tobago, St. Lucia (not shown).

There was focus on providing countries with catered imagery needs and a platform for sharing methods and collaborating with different country offices on most efficient practices for creating enumeration area maps for the census. Exercises on common methods for creating a structured census geodatabase were also provided. More specifically, the workshop: (1) presented country experiences and current situations on the best way to create maps and other outputs for the 2010 Round of Census maps in the region; (2) provided an overview of the different satellite products and options, with an overview of requirements with regard to census and statistics; (3) presented an overview of critical datasets for the enumeration exercises and illustrated the development of a census and statistics GIS; (4) facilitated a dialogue among participants on collaborative agreements and best practices led by country examples in creating census maps and developing the census and statistics GIS; (5) provided focused and specialized training and assistance in specific areas of census GIS and map production per country situation with exercises.

CARICOM Workshop on Census Mapping and GIS for the 2010 Round of Population and Housing Census: Identifying Imagery Needs, Use and Data Development Strategies for the Census
Belize City, Belize, 23-26 February, 2010
Organization of Work

TIME	TOPIC
	TUESDAY FEBRUARY 23, 2010
	Opening Ceremony
8:30-9:00	**Registration of Participants**
9:00-10:00	**Session 1 – Opening Remarks** - Mr. Glen Avilez – Director-General, Statistical Institute of Belize - Ms. Barbara Adams – Deputy Programme Manager, CARICOM Secretariat - Mr. John Mensah – Regional Census Coordinator, CARICOM Secretariat Administrative Matters **Introduction of Workshop Consultant/Facilitator** Mr. Charles Brigham
10:00-10:30	**Coffee Break and Group Photo**
	Review of Census GIS and Map Production Practices (where countries are) Objective: To present country experiences and current situations - a discussion on the preparation of 2010 Round of Census maps in the region. Countries will address their specific challenges and share ideas on the best way to create maps and outputs for the census.
10:30-12:00	**Session 2** – **Country Presentations** Presentations on the preparation of census map and the development of census and statistics GIS - Opening by CARICOM - Presentation by each participant of the situation in his/her country - General Discussion
12:00-13:30	**Lunch**
	Imagery Products Objective: To present an overview of the different satellite products and options and an overview of requirements.
13:30-15:00	**Session 3 – Common Geospatial datasets for Census** Objective: To provide an overview of the dataset normally used for census maps - The Geocentric Census - Census and Statistics GIS - Presentation on the fundamental datasets for census GIS - Application and Presentation of Geospatial Datasets for Census- A Practical Example
15:00-15:15	**Coffee Break**
	Best Practices in Creating Geospatial Datasets for Census Objective: To present an overview of critical datasets for the enumeration exercises and present and discuss the development of a census and statistics GIS.

TIME	TOPIC
15:15-16:45	**Session 4–Overview of Different satellite products** Presentations on the available imagery products, requirements and critical considerations for the purchasing of Imagery - Presentation on necessary requirements for procuring imagery - Identifying focus areas, Ordering, AOIs - Presentation by Country - General Discussion
	<u>WEDNESDAY FEBRUARY 25, 2010</u>
8:30-10:00	**Session 5 – Developing Critical Datasets for Census** Presentation on the most important datasets required for creating census maps/ED maps and beyond to the building level. - Discussion on workflows for standard census GIS datasets - Exploring Topology and Creating Rulesets - Structuring Geodatabases/Data Models & Loading Feature Classes - Methodologies for streamlining development
10:00-10:15	**Coffee Break**
	Country Presentation and Country Specific Focus Objective: To provide countries the opportunity to collaborate on best practices led by country examples in creating census maps and developing the census and statistics GIS
10:15-12:00	**Session 6 – Home Country Presentation** The Statistics Office of St. Lucia presentation on census mapping and GIS. - Presentation by St. Lucia - General Discussion
12:00-13:30	**Lunch**
13:30-15:00	**Session 7 – Country Specific Training and Exercises** Objective: Provide a collaborative focus and specialized training and assistance in specific Areas of Census GIS and map production per country situation/exercises - Exercises Related to Specific GIS Tasks - Targeted Training and Assistance based on Country Situation - Field Work and the use of Imagery for identifying features and validating existing maps - Exercises and Specialized Training - General Discussion
15:00-15:15	**Coffee Break**
15:15-16:45	**Session 8 – Country Specific Training and Exercises** - Exercises and Specialized Training - General Discussion
	<u>THURSDAY FEBRUARY 26TH, 2010</u>
	Regional and Sub-regional Capacity Building/Commercial Supplier Presentation Objective: To provide countries the opportunity to discuss collaboration and staffing and CARICOM/Member States Capacity Building practices and GIS community building networks

TIME	TOPIC
8:30-10:00	**Session 10 – Collaboration and Staffing** - Roundtable on Staff Retainment - Roundtable on Hiring Requirements General Discussion
10:00-10:15	**Coffee Break**
10:15-13:00	**Session 9 – Commercial Supplier Presentation** - Presentation by Supplier - General Discussion
13:00-14:00	**Lunch**
14:30-16:45	**Session 11 – Structuring a GIS for Census/Automated Mapping/CARICOM Capacity Building** - Presentation on Automated Mapping of Census Enumeration Areas - Presentation on How to Structure, Design and Evaluate Capacity for the use of GIS for Census Activities - Preparation for Future Meetings and ESRI User Conference, Caribbean GIS, CARICOM GIS Meeting - General Discussion
	FRIDAY FEBRUARY 26TH, 2010
	HOST COUNTRY VISIT & Con
9:00-10:30	**Adoption of the final report, recommendations and conclusions**
11:00-1:00	**Statistics Office Visit** - Visit to Statistical Institute of Belize, Belmopan - General Discussion **Conclusion and Remarks (from Conference)** Closing remarks and community group announcement (Google group on census and statistics GIS signup) - Remarks by CARICOM - Remarks Director-General of the Statistical Institute of Belize - Closing by Trinidad & Tobago

Outputs of the Workshop

The workshop provided practical applications for the conduction of mapping and GIS activities in respective countries through looking at available GIS tools and technologies and imagery options both free and commercially available using specific software tools and direct line communication with vendors. Further assistance and outputs were in the creation of an online community of census and statistics GIS personnel that will actively share exercises, data sources, software, programming and other related GIS experiences and best practices and develop capacity for a more accurate census alongside development of a statistics GIS through the use of current methods. Hands on training was also provided on a group and per-country basis throughout the workshop.

The Enumeration Layer Development Process

Establish Fields and attributes to be included in the ED layer

After creating a feature class or shapefile it is required that certain fields be used to identify attributes of any geographic data. Establishing common fields in the ED layer such as EAID (Enumeration Area/District ID), PNAME (Parish Name), PCODE (Parish Number) EALABEL (Label Field for each EA). As long as you have sequential coding for your EDs it is simple to add and calculate new fields. Instead of populating fields manually, use the field calculator tools in the GIS software to accelerate the process. Many of the field calculations are executed through Visual Basic and python programming languages. Below is an example of the attribute table in a working ED layer for Grenada.

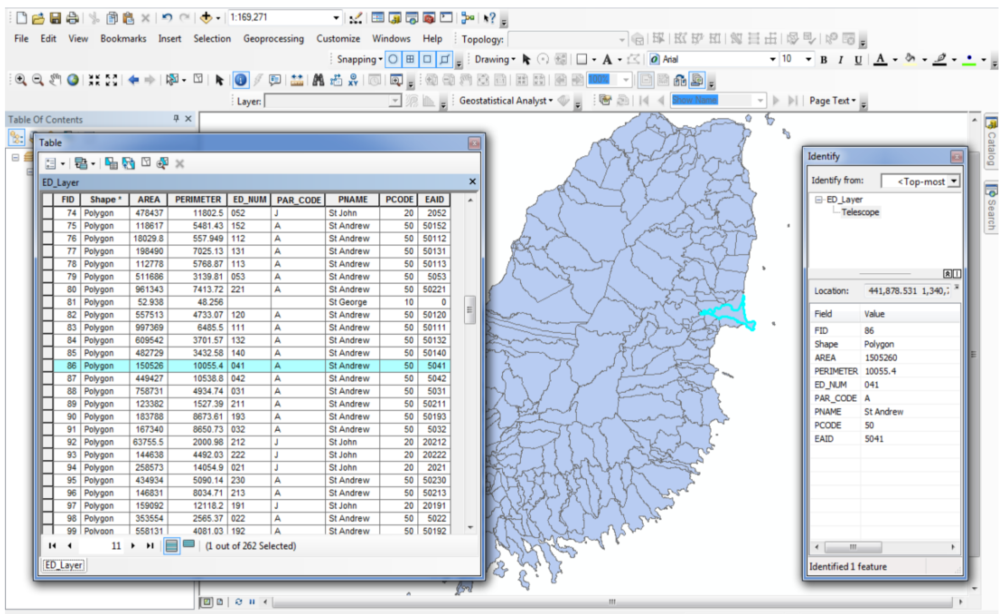


Figure 15 Illustration of ED field names and associated attributes for an ED layer of Grenada.

Common fields must be established to maintain a harmonized geographic dataset. In creating/editing the ED layer it is critical to compile GIS layers with like attribute field for each dataset. Establishing a common field name dataset will allow for easy migration into new geodatabases and across geodatabases.

Digitizing and Editing of the ED layer

Before you begin creating or editing the ED layer, overlay the ED layer or boundary datasets onto existing sources of information such as Georectified mapsheets, imagery, and other geographic datasets that exist at the largest resolution possible in the GIS. All of this information will guide in delineating boundaries more accurately and quickly. A wealth of freely available information exists online through different GIS data resources such as through web map services/web coverage services or by directly download.

In the last few years the available geographically referenced data has rapidly increased. Freely available imagery at high resolution, particularly in urban areas and in WGS84, is becoming increasingly available for immediate consumption and use in a GIS through basemap wms services. These services allow the GIS user to bring up various sources of imagery and geographic data at a global level and this allows for an excellent starting point if no current sources exist which have good resolution. The countries of 2010/2011 censuses for CARICOM have harnessed the freely available imagery for use directly in the GIS to aid in the ED editing and delineation process through direct updating and identifying fieldwork priorities. Previously(within the last 5 years), offices were moving between the use of Google Earth and GIS. Such basemap services provided in the GIS packages such as Bing Maps/Virtual Earth, World Imagery Layers by ESRI, Delorme Basemaps) have revolutionized ED maintenance.

WGS84: The most widely used geocentric datum & geographic coordinate system. GPS measurements are based on WGS84.

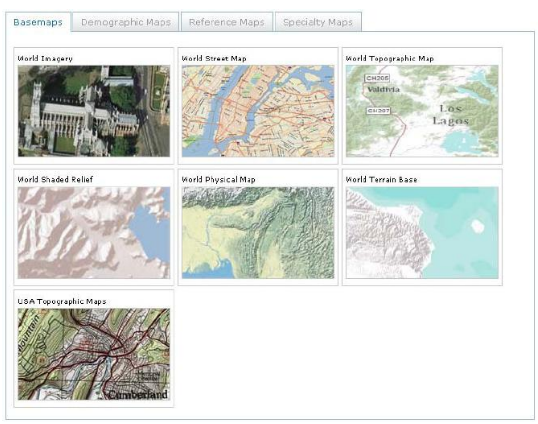


Figure 16 Common basemap layers directly accessible by GIS software

It is now widely agreed that imagery is the most cost effective method for ED layer creation and census GIS data development than the use of any other method (hand drawn or GPS field deployments). The reasoning for this is that more rapid identification of populated developments are made possible through the use of imagery, allowing mapping personnel to easily prioritize EDs in need of further boundary investigation. This identification of the number/increase in development(s) is noted through the use of the ED layer and its associated prioritized list of EDs with likely or planned development overlaid on the imagery. Furthermore, imagery allows the enumeration exercises to be more easily verifiable since current imagery reflects the count at the time of the census as closely as possible (E.G. building layer, landmarks, rivers, roads). Extracting information through digitizing from imagery is a straightforward, regimented and streamlined exercise and a common practice for GIS officers and increasingly for field officers.

Digitizing and Editing of the ED layer

Topology (math): The branch of geography that deals with the properties of a figure that remain unchanged even when the figure is bent, stretched otherwise distorted.

One of the primary reasons topology was developed was to provide a rigorous, automated method to clean up data entry errors and verify data. The typical digitizing procedure is to digitize all lines, build topology, and label polygons and then clean up slivers, dangles, and under- and overshoots and build topology again, repeating the clean and build phases as many times as necessary. What if the process did not start with the tangled mess of "cartographic spaghetti"? By approaching digitizing from a feature-centric perspective and enforcing topology when each ED boundary or other feature is digitized and labeled, sliver polygons, dangling nodes, missing labels, and multi-labeled features would be eliminated. GIS software platforms support a feature-centric digitizing approach that requires on-the-fly calculation of geometric intersections and today's computers are powerful enough to support feature-centric digitizing for most GIS users.

Topology (GIS): The arrangement that constrains how point, line, and polygon features share geometry - defined and enforced by integrity rules (e.g. no gaps)

ArcGIS supports such feature-centric digitizing through the Append Polygon, Split Polygon, and Split Line tools through topology creation and rule-sets. With these tools users can add a polygon (or line) adjacent to an existing polygon and have boundaries match perfectly. Today's GIS also supports topological editing of shared boundaries or nodes through the manipulation of vertices. The importance of map topology in the ED layer development and editing processes cannot be understated. In many instances, people tasked with geographic work on boundary delineation and adjustment understand that topology is important but often do not know how to properly apply it to their work.

Figure 17 Topology Editing: Moving shared vertices at once. (courtesy, NSO Dominica)

Topology and associated rules implemented on the geodatabase feature classes depend on what spatial relationships are most important for the mapping personnel to maintain. Some topology rules govern the relationships of features within a given feature class, while others govern the relationships between features in two different feature classes. Topology rules can also be defined between subtypes of features in one or another feature class. This could be used, for example, to require street features to be connected to other street features at both ends.

The first step in enforcing topology rules in a feature class is to remove twisted or self-intersecting polygon rings and to ensure that the "inside" of the polygon is on the correct side of the polygon boundary. Next, gaps are identified by creating a rectangle that encompasses all the polygons of interest and serves as a backdrop. The polygons are subtracted from the rectangle containing all polygons. The remaining areas are gaps. A gap polygon is removed by merging it into an adjacent polygon or by making it a legitimate polygon. Overlaps are found by intersecting each polygon with all other polygons. If an intersection is found, then the polygon representing the overlap is created. Overlaps can be removed by deleting the overlapping area from one of the involved polygons. Once boundary changes have been made, the area and perimeter of each polygon should be recalculated.

Topology rules in ArcEditor or ArcInfo are created in a geodatabase feature dataset through geoprocessing of data

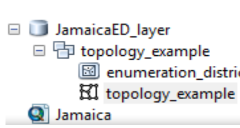


Figure 18 Topology ruleset example in a feature dataset

through a wizard-based interface. You can define the layers you want to participate in the topology and specify the various rules that will be included to model the integrated behavior of different data types. With concern to ED layer maintenance, topology greatly assists the GIS officer in identifying gaps and overlaps, providing a count, listing and location tools for quick editing. Without topology, these gaps and overlaps go undetected during map creation and can lead to errors during enumeration whereby the enumerator may canvass an area larger that identified (overlap) or smaller than identified originally (gap).

Now that the ED features class is coincident (occupying the same space), the GIS officer will want to ensure future changes to these features do not break this connection. ArcEditor and ArcInfo licenses (not ArcView) can take full advantage of the 25 topological rules that geodatabases provide. With an ArcView license, you can only edit map topologies by imposing topology (e.g. sharing ED vertices) upon simple features on a map during an edit session, allowing you to simultaneously edit simple features that overlap or touch each other. You can use the Topology Edit tool on the Topology toolbar and the Modify Edge and Reshape Edge edit tasks to edit the features in a map topology. The features can be in one or more feature classes and may have different geometries. Line features and the outlines of polygon features become topological edges when you create a map topology. Point features, the endpoints of lines, and the places where edges intersect become nodes. A map topology can be applied to simple features in a shapefile or to simple feature classes in a geodatabase. The feature classes that participate in the map topology must be in the same folder or geodatabase.

Once the geographic integrity of the ED layer is ensured, the future editing process can be streamlined and the splitting of EDs based on household, adjusting of ED boundaries, merging of EDs into supervisory districts and other boundary related editis can be conducted in rapid fashion. Any future editing of the layer become easily manageable and if errors occur (for instance when another person modifies the ED layer and does not use topology) they can be quickly remedied.

Updating ED boundaries through Fieldwork (boundary and attributes)

Fieldwork often involves making sure the ED contains the correct number of households and that this number conforms to the established guidelines (normally ~150) and adjusting boundaries accordingly. Other common fieldwork activities involve visiting areas where imagery was not available and checking the maps against the available or original data used to create them or for identifying whether features used as boundaries still exist. This is often the most time consuming as it normally consists of traveling to rural areas where small scale Mapsheets or low resolution imagery were used to aiddevelopment of the ED layer.

Use of GPS for Field Mapping and other parallel GIS activities in Dominica

The office conducted a household listing. To make best use of the available hardware, personnel and time available GPS was used to execute this household listing. An interface was created to facilitate the data capture and training in the development of this interface was provided. The mapping team was trained in the use of GPS and is now able to capture coordinates as well as attributes of the coordinates (e.g. building type, number of dwellings, number of households etc.). This information will assist the office with both GIS development as well as census preparatory exercises.

Through the training of census mapping personnel, the Central Statistics office of Dominica developed a census GIS for 2010 census map creation. This includes cost effective methods to minimize duplication and errors and to maximize efficiency and ensure sustainable growth of a GIS. The Capturing of information was done through the use of GPS and field teams deployed to particular EDs. The GPS locations and attribute information were imported in a GIS. Recent imagery for the west side of the island was procured and is used to aid GIS development and fieldwork. The imagery provides the GPS team a snapshot of the situation on the ground near the time of the census and will be a historical reference for any future discrepancies in post enumeration exercises and after the census it can be used for other critical work.

Imagery is now providing building level information as a single GIS data layer in greater detail and is being augmented by the use of GPS to capture more precise locations such as front door locations and other location information. The short-term goal is to best use this geographic information in the maps to guide the enumerator to the necessary dwellings and households. The long-term goal during and beyond the census is to develop a comprehensive building level layer that will allow for mapping of statistical information at any geographic level and strict identification of demographic characteristics at the building level.

The office is now trained to capture this level of information in 2 ways which complement each other- (1) through the creation of a GIS layer (polygon feature class which provides a unique building outline showing the shape and relative size of the building) through the use of imagery and (2) as through capturing GPS points of building locations and information during the household listing exercise. The latter GPS exercise is currently being conducted in great detail and training on the workflow was provided during the assistance visit. The culmination of these exercises will allow the building layer to be linked to the questionnaire through unique coding. This allows the statistics office to aggregate out from the building level to any specified geography (survey boundaries->ED->Administrative Dist., Parrish etc.). However, since linking at this level takes time and the census map creation exercise should always focus first on the EDs and later consider further geographies.

Use of GPS for Field Mapping and other parallel GIS activities in St. Lucia

The production of digital ED maps has evolved since the introduction of GIS at the Statistics Departmentin St. Lucia. From hand drawn maps to updated computer generated maps. ED maps were continuously being used by the Survey Section of the Department for their Quarterly Labour Force Survey. Enumerators use the maps in the field regularly to collect data. The regular updating of roads and building using GPS remains priority for the Mapping Unit. Staff would go out with the handheld GPS units to collect data. An average of three EDs would be targeted each week for updating. Upon returning to the office the data would be transferred to the computer using the Pathfinder office software which came with the Trimble GPS units. Differential corrections were made using base files obtained from the Cable and Wireless base station. The corrected files were then exported as shapefiles ready to be imported into the GIS project. The Department has Trimble GeoXM handheld GPS units. These units combine the Trimble GPS with Microsoft Windows CE operating system, - a mobile GIS unit.

Upgrading of the unit meant upgrading of the software. The office also purchased licenses for TerraSync and ArcPad 7.0.(Arcpad 8.0 was recently released), which made updating of roads and buildings just got more accurate. The Arcpad software whose primary use is for the updating of buildings, allows the one to import actual ArcMap projects to the unit. The navigation facility allows someone to navigate easily. The TerraSync software is used mainly for the updating of roads. Corrections can be made to EDs in the field where new buildings can be added and non existing buildings removed. Verification of ED boundaries can also be done in the field to ensure that mapped roads and tracks coincide with the actual roads and tracks on ground. Most EDs were updated and the buildings renumbered sequentially to accommodate the re-listing of the EDs prior to the 2010 census.

Following the last census a number of EDs had to be subdivided to accommodate the increase in development in certain areas and there by maintain an average of 100 households per ED. Preparatory work has concluded and the maps have been produced to a high level of accuracy resulting in an efficient and highly accurate census in St. Lucia.

Census Field Mapping and Data Capture

The Population and Housing Census is an extensive exercise covering all of the population and geographic areas of a given country. Census Field Mapping personnel must see themselves as quality control agents, who, through painstaking efforts, will positively influence the outcome of the census. This requires census mapping personnel to be well trained in the use of Global Positioning Systems needed to retrieve co-ordinate data for geographic features and more regularly, building locations since many countries strive to achieve highly detailed maps and a level of data capture beyond the ED in order to link building locations in the GIS and with other statistical data. The captured information is used to produce updated Enumeration District (E.D) maps and/or provide preliminary population and housings counts. It is therefore, extremely important for Mapping personnel demonstrate their commitment to the exercise and that these personnel be carefully chosen.

Field Mapping personnel must adopt a uniform and standardized approach to this exercise, by which the entire field staff will be in a position to carry out their duties as expected by the Central Statistical Office. Guidelines set out in this manual as representative of guidelines used in CARICOM countries, and are intended to minimize conflicts in the interpretation of key concepts employed in a visitation record (or similar) and to facilitate the information acquisition and transfer in an orderly and effective way. Training sessions executed by statistics offices and parties involved in the census mapping activity must center on:

1. Geographic concepts and how they are employed in the collection of GPS data;
2. Techniques in using the Global Positioning System GPS;
3. Procedures to be followed during field work that include:
    a. The use of maps to locate areas, buildings and households;
    b. The questions that will be asked ofrespondents in gathering of information;
    c. System of communication & document/sheet control;
4. Techniques in conducting field work and dealing with common issues

In order to efficiently carry out Field Mapping assignments an area map, ED maps (existing draft and/or old) and descriptions of boundaries, visitation record/listing sheet, and a GPS Handheld Unit are commonly provided

Field Mapping personnel use maps during the exercise for locating the Enumeration District (E.D) and identifying the Enumeration District boundaries. The maps orient field personnel/enumerators with respect to his/her E.D. They also identify a landmark or well-known street that is selected to begin orientation. Optimally, the route of travel to be followed using the directional arrows or building numbers in consecutive fashion may be included. The Field Mapping personnel also use the maps to identify housing units with reference unique markers such as electric pole identification.

Enumeration district maps (E.Ds) are distinctly different from field maps and focus on showing boundaries of census enumeration districts, which are established to help administer and control data collection - geographic areas assigned to each Mapping Assistant/Enumerator represent a specific portion of the country.

As developmental changes occur on a district and country level - - new roads are constructed, new settlements appear and disappear, and urban areas expand, there is a constant need to update these maps. For censuses and surveys, maps that are as up-to-date as possible, are needed to assure that E.Ds are properly identified and features used as boundaries still exist.

Field Mapping personnel maps are used together with an assignment sheet and the visitation/listing record, first locating households listing the entries in the sheet. The accompanied maps help to completely canvass all tasked areas/E.Ds and households in his/her assignment and must not leave out any areas. It is therefore critical to follow any indicated routes on the maps to ensure, locate and completely canvass your assigned areas. During these visitation/listing exercises, look out for map symbols and significant landmarks indicated on your maps. These will assist you in establishing your whereabouts and estimating the distance already covered. Also, if significant points of interest or geographic features are not indicated then make a note of them in the map or a comment so that the information can be incorporated into the GIS.

The primary purpose of delineating the ED is to divide the country so that each part is enumerated once. Mapping Field Assistants should be able to locate the ED using peripheral information provided on the map and determine precisely what area is to be covered following the delineated boundary lines and written description. The assignment sheet should be complete with any comments regarding information that could further improve the census GIS and resultant, final ED map. During the visitation/listing the field personnel are required to record or estimate the distances traveled from point to point so that a general understanding of the time it takes to do such work is more clearly understood when the hiring/ED assignment process commences.

Global Positioning System (GPS), is a radio navigation system that allows land, sea, and airborne users to determine their exact location, velocity, and time 24 hours a day, in all weather conditions, anywhere in the world. The capabilities of today's systems render other well-known navigation and positioning "technologies"—namely the magnetic compass, the sextant, the chronometer, and radio-based devices—impractical and obsolete. GPS is used to support a broad range of military, commercial, and consumer applications.

In the case of mapping fieldwork some specialized GPS receivers will store data for use in Geographic Information Systems (GIS) and map making.

Orientation & Scale

In using a "Field ED" or field map, it is important to know how to line up the map with the ground; that is, to locate on the ground certain features found on the map. The act of lining up a map is called "orientation." Orientation is important in census work when combining information from several maps into one map, verifying maps in the field; and conducting the field enumeration.

Orientation is the basic operation in determining direction of a map. To orient a map, the most important pieces of information are the north arrow and the symbols representing the ground features. A north arrow may represent true north (the direction of the geographic pole) or magnetic north (the direction of the magnetic pole). For most census purposes, the difference between true north and magnetic north is not significant.

When maps for several areas are to be combined, the maps should be placed so that the north arrow on each map points in the same direction or so that the parallels and meridians run in approximately the same respective direction on the various maps. Then symbols for any features that the areas have in common, such as roads, streams, parks, bridges and large buildings should be lined up so that their relationship are correct.

In the field, the map should be aligned with the ground. To do this, it is necessary to find objects on the ground, which are represented by specific symbols on the map such as a school or highway. The map should be held so that features on the map line up with their corresponding features on the ground. When the map is properly oriented, it is possible to identify any ground object and its relationship in direction and distance with any other object on the ground.

Orientation has traditionally been a major problem in the creation of census maps in past census. The field mapping personnel are easily disoriented if they are not familiar with an area tasked for their workload. Past census maps illustrate this point throughout the CARICOM countries. One example is provided below showing the large issues with regard to orientation; particularly note the wrongly oriented map (left) and the correctly oriented map (right). As most countries are using GIS to produce their maps orientation is inherited. All information is georeferenced through spatial referencing systems in a GIS based on projections, which are mathematical transformations of an ellipsoidal earth onto a two dimensional surface.

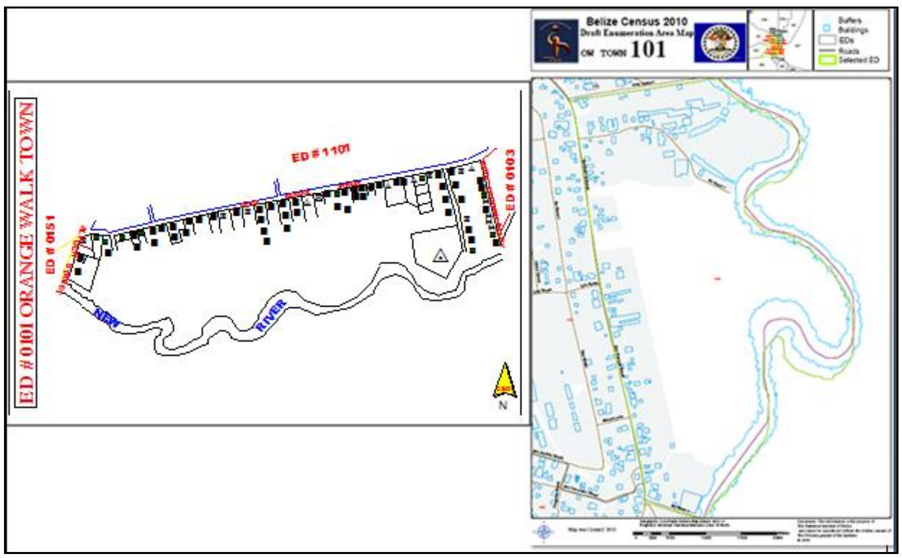


Figure 19 The CAD map (left) is oriented incorrectly with North wrongly pointed East; the map on the right is correctly oriented and scaled geographically with regard to up being north as well as building distances and building, road and river footprint

Scale is the relationship of the distance between any two points on the map to the horizontal distance between those same two points on the actual surface of the earth. Without an expression of scale on a map, it is difficult to determine distance on that map. It is often expressed as a word statement, numerical fraction or ration, or graphic scale.

Often only one type of scale is shown for both field maps and ED maps, it is usually found in a prominent place on the map. Generally, the representation of scale in graphic form is most useful for census purposes. The graphic scale is the only satisfactory form for representing scale on maps. The scale enables the map user to translate distances between points on the map into corresponding distances on the ground. By using the scale, enumerators can determine distances required for field checking/updating the maps, for conducting the enumeration. The scale is always shown on a census map, to realize the extent of travel in a field map or E.D.

There is at least one symbol that appears on most census maps and that is the north arrow; which helps the enumerator locate his position and direction. Census maps have a standardized legend as well. Good symbols are drawn so that they are somewhat similar to the actual feature on the ground. Quick and easy recognition of symbols is an advantage; it decreases the chances of error and provides even the most inexperienced enumerator with a guide he/she can readily understand.

Measuring Distances

Another important task in updating a map is the measurement of distance. Traditional methods of pacing and striding outlined below along with more contemporary methods. Pacing takes time and must be done correctly to be useful. GIS methods are also used and tend to take less time but require the data and attribute information to make the required measurements.

Pacing - If distances are measured on foot, a simple but careful pacing procedure is required. Distance is calculated by counting the number of steps between two points and multiplying by the average length of pace. For example, if a person steps off 150 paces between two points, and the average length of his pace is 0.6 meters, the distance between the points is 150 x 0.6 meters. To determine average length of pace, proceed as follows:

1. Walk a known distance at least three times in each direction (six times). Walk normally; do not take unusually long or unusually short steps. Count the number of steps taken each time.

2. Compute the average number of paces for the six trips. In computing the average, disregard any count that is unusual. For example, if the counts are 154, 157, 153, 141, 159 and 153, disregard 141. The average would be (154 + 157 + 153 + 159 + 153) divide by 5 or 155.2 paces.

More contemporary measurement of distance within an E.D. or field map can be done through the use of GIS, provided the information exists to do so. These estimations of the measurement of distance traveled within an E.D are found by using area and perimeter information coupled with average time spent in a listing/questionnaire exercises and average walking speed. These estimations are comparable to estimations from the field.

More specifically, Travel Time Cost Surface Models (TTCSM) will become more widely available (TTCS calculates travel time from defined point and/or linear locations to other locations within a user defined area of interest (AOI)). The TTCSM are designed to model travel time in geographic areas using readily available geospatial layers such as road, river, footpath, digital elevation models. Output from the TTCSM are point to point specific travel time least cost paths (i.e. the modeled fastest path(s)) and raster maps in which each cell value is the modeled time required to reach the given cell from the specified starting point(s).

The essence of the TTCSM is in deriving a meaningful travel cost surface. Travel cost is a function of the user defined and derived cost and speed surfaces. The cost surface defines the weight or impedance of traveling through a cell, while the speed surface defines the speed at which movement within the cell occurs. Speed is a function of the user-defined average walking speed and slope, except on the road network where speed is equal to the defined road speed limit. Using the derived travel cost surface, travel time calculations are performed using either traditional cost distance or more robust path distance modeling methods. The TTCSM is intended to be used as a tool to facilitate more efficient field and ED sampling and canvassing design and planning. Several TTCSM models are packaged as python script requiring only basic understanding of python programming, and a moderate level of proficiency in GIS.

Administrative and Enumeration District Boundary

E.D boundaries, which are not administrative divisions, observe natural features such as rivers, ravines and man-made features such as roads or paved drains. The feature selected is one that is easily seen and clearly identified both on the ground and on the map. Occasionally, the use of invisible lines for E.D boundaries cannot be avoided. Offset lines and extended lines are two types of invisible or imaginary lines. An offset line is a line that parallels a main road at a specified distance, e.g. at the rear of residential lots along the road. Extended lines are imaginary lines used as boundaries to connect one visible feature with another. An imaginary line is usually not used as an E.D boundary if it passes through an area where households are located in an irregular pattern close to the imaginary line. Imaginary lines also do not cross each other.

If travel conditions exist within an area that will cause an enumerator to spend a disproportionate amount of time traveling between households, then the number of households per E.D must be reduced to compensate for his/her difficulty. Terrain, settlement patterns, weather, poor and impassable roads, households and population size are factors, all of which are used in determining the size of E.Ds.

Visitation Record (Country Example)

The Visitation Record is a very important document in which field workers are required to record basic information about the physical building in which they have located the households.

Visitation Record Use in St Lucia

Importance of the Visitation Record (V.R)

(i)Because we rely on what and how you record information in the V.R to provide information on the characteristics of the population, it is very important that the details should be CORRECTLY recorded.

(ii)The V.R is the FIRST SOURCE of establishing a Population Count for an area and is therefore VERY IMPORTANT.

(iii)The V.R is linked to your E.D map and helps the office to follow your daily progress of enumeration.
(iv)The V.R is used to carry out field checks of your work.

On your V.R there are four important concepts that need to be learned. These are:

(a)**Building**: A building is defined as a physical structure which is separate and independent of any other structure. It must be covered by a roof and enclosed within external walls. A building may be a factory, shop, detached dwelling, apartment building, warehouse, repair shop, poultry pen etc. NOTE that detached rooms relating to main building are treated as part of the main buildings, for example detached kitchens, toilets, servants quarters, garages, etc.

(b)**Dwelling Unit**: A dwelling unit is any building or separate and independent part of a building in which a person or group of persons is living at the time of this activity. It must have direct access from the Street or common landing, staircase, passage or gallery where occupants can enter or leave without passing through anybody else's living quarters.

(c) **Private Household**: A private household consists of one or more persons living together (i.e. sleeping most nights of a week, 4 out of 7) and sharing at least one daily meal. It is important to note that a member of a household need not be a relative of the main family. For example, a boarder or a domestic servant who sleeps in most nights of the week is a member of the household. It is possible for a household to consist of just one person, or of more than one family, as long as they share living arrangements. A group of unrelated persons living together can also comprise a household.

(d) **Business Unit**: A business unit consists of a building or part of a building used for producing goods and services to be sold either for profit or voluntary/non-profit making e.g. Repair shop, Bakery, Shop, Red Cross or Lions Building.

How to Fill Out the Visitation Record

1. Look first at the Cover page. The top right of the V.R bears the Identification Codes of the District and Enumeration District to which you are assigned. You will then be required to enter the date you started and completed your V.R and include your name and signature in the appropriate box/line at the bottom of the page.

2. Completing Page 3 of the VR

Col. (1) – Date of First Visit: Enter here the date you first visited each building, dwelling unit or business place during field visits.

Col. (2) – Building Number: The purpose of this column is to provide a precise count of the number of buildings each enumeration district. Each building that you visit must be given a number in a serial order as you visit it, starting from 001, 002, 003 etc.

Col. (3) – Household Number Following an identical procedure to that used for Col.2, the numbering of households must be in Serial order, starting from 001 and continuing 002 etc. Always check back to verify the last number used in order to avoid repeating numbers. The last number appearing in this column should represent the total number of households in the enumerationdistrict.

Col. (4) – Dwelling Unit: You are required to number the dwelling units; number just as you did for the building and household number, in serial order starting from 001. Remember that it is possible to have more than one household in a dwelling unit in which case the same dwelling unit number is to be repeated for each such household. Thus it will be possible to identify each household with the dwelling unit, which accommodated it. The last number appearing in this column should represent the total number of dwelling units in the E.D.

Col. (5) – Full Name of Head of Household: You are required, in the case of a household, to Household/Establishment write in BLOCK LETTERS the name of the head of the Establishment Or Institution in these columns, putting the surname first and the given (Christian) name or names after.For example, MILLER, THOMAS. This is because surnames quickly identify members of a family, which is represented by the head of the household. With respect to a business place, you must write both the name of the proprietor and the name of the establishment, e.g. "John Black, Green Café". In the case of an Institution write the full name of the institution e.g. St.Rose's Nursing Home.

Col. (6) – Full Address:You must record as much detail as possible. In urban and semi-urban areas where houses are numbered, you must record the number of the house, the name of the road or street, and the town, village or locality. In rural and remote areas, where addresses may be vague, the name of road, track (if named), mile post, electricity pole (lamp post) or other permanent landmarks.

Cols. (7), (8) and (9): Number of Persons, both sexes, male and female respectively: The total number of males in col. (8) and the total number of females col. 9). Include visitors who are expected to stay for at least one month and persons temporarily away or in institutions.

Col. (10) - Business Unit Number: As in cols. (2), (3) and (4), you are required to number the business units in serial order starting from 001. Remember more than one business may be operated in the same building in which case use a separate line for each, giving each such business in col. (6) a number. The last number should be the total number of business units in your assigned area.

Col. (11) – Type of Business Activity: You must enter in detail the type of activity carried on in the Business. Examples are dry goods store, grocery, boutique or motorcar repairs, etc. This information is important in allowing the department to properly classify businesses by the type of activity they engage in.

Col. (12) - Number of Paid Employees: The total number of paid employees i.e. persons entered on the pay-sheet of the business place during the last pay period (week, fortnight, and month) must be recorded in this column

The assignment is to be completed within a prescribed period. After the GPS points have been collected, the field personnel must update all their assignment control lists and complete their Visitation Records and return all materials to the Central Statistics Office.

If you look at the GIS development process as a process in which data and information are analyzed and processed it can be easier to picture the workflow involved. Often it is best to simplify the process and the scale the activities accordingly. This will help keep the project manageable and clearly understood. Below is an example that provides a starting point for a census GIS project. It can be applied to any geographic project that aims the develop geographic information that can be used in any given project.

Simplified Census GIS Analysis Process (Analysis=ED map creation)

Geographic officers perform analysis on the data within the census GIS. GIS analysis within the census GIS is a process that follows a basic set of steps to create ED maps. The actual methods you use can be simple or complex. They can involve selecting features having a given value to building a model to combine many layers of geographic data for the census. For simple methods, such as a selection to extract features from a dataset (such as ED centroids), you might intuitively implement the process as a single operation rather than as distinct steps. This is often the case but not always.

Frame the question:

You start an ED Map analysis by figuring out what information you need. This is often in the form of a question. Where are the most households located in the region? How many households have been added in each ED since the last census? Which new households cross census boundaries or are within a critical range of a boundary and planned for further development. Which EDs have the highest number of households? Being as specific as possible about each question you are trying to answer will help you decide on how to approach the analysis, which method to use and how to present the results in your new ED layer or any related geographic layer.

Other factors that influence the analysis are how it will be used and who will use it. The obvious use is to develop the ED map using the highest quality of geographic and statistical information possible. An NSO might simply be exploring the data on its own to get a better understanding of how a place developed or how things behave; or you may need to present results- to census staff or other decision-makers or the public for discussion, for scientific review or other setting. It either case your methods need to be rigorous, and results focused.

Prepare your data

Preparing GIS data is often an iterative process requiring redesign and reevaluation of methods. For example boundaries may need adjustments upon the purchase of need imagery or acquisition of new datasets.

The type of data and features you are working with helps determine the specific method you use (for example, your forest and river data might be stored as polygons or rasters). Conversely, if you need to use a specific method to get the level of information you require, which is often the case for census data, you might need to obtain additional data. You have to know what you have (the type of features and attributes), and what you need to get and/or create. Creating new data may mean calculating new values in the data table or obtaining new layers or in more common a complex cases obtain and create layers. Understanding the data that goes into the analysis will help you interpret the results.

Choose a method

There are almost always two to three ways of getting the information you need. Often, one method is quicker than the other and gives you more approximate information. Others may require more detailed data and more processing time and effort, but provide more exact results. For example, you can find households within 500 feet of a road as the crow flies by creating a buffer -or - within 500 feet walking along the streets and paths (a more involved network analysis and often more costly too). You decide which method to use based on you original question and how the results of your ED Maps (analysis) will be used.

Process the Data

Once you've selected a method, perform the necessary steps in the GIS. This often involves running several functions in a sequence. For example, to find the amount of forest in an ED (to inform field personnel), you might extract the area of interest from the forest area, convert it from raster to polygons, overlay it with the ED boundaries, then do tabular analysis to calculate the total forest in each ED (to get a better idea of ED geography). In the case of statistical analysis, you will also want to calculate the statistical significance of your initial results.

Investigate and Analyze the Results

Looking at and questioning the results help you decide whether the information is valid or useful, or whether you should rerun the analysis using different parameters or even use a different method. Using GIS makes it relatively easy to make these changes and create a new output. Older census field methods did not allow this type of flexibility. You can compare the results from different ED map analyses used by others and see which method provides the most accurate information for your purpose.

Critical Considerations: Census Data Resolution in Disaster Management

In the Caribbean region the frequency and occurrence of natural disasters and more particularly, flooding is of great concern. In addition to the wide scale devastation caused, these disasters have a direct negative impact on people and population. To mitigate the devastation caused by these catastrophes, it is imperative to have accurate and complete information on population distribution at the desired demographic scale.

Risk assessment is the central pillar of the hazard risk management framework. Risk is often defined as the probability of a loss occurring. It depends on the frequency and intensity of the hazard, the people and structures exposed to those hazards, and their vulnerability. The information provided by the assessment contributes to an informed decision making process that reduces the chances of surprises, and enables consequences to be managed and planned for in advance.

The census information in many countries is not frequently updated. Moreover, the available information does not have information on the micro level spatial distribution of population. This becomes a potential bottleneck for effective disaster management analysis and mitigation strategies.

There is a need to explicate the close and critical relationship between census data and disaster management and preparedness must be further understood. Many countries are embarking on collecting geographic information beyond just the ED, at the building level. Collecting geographic information at the building level will inform management with some of the best information possible when a disaster strikes.

Data Required for Vulnerability Mapping

The exposure and vulnerability modules of disaster management calculate the effect of a hazard with respect to the assets and the population of the affected area.

A critical component of any disaster risk model is the exposure component. In this component the elements at risk are categorized in a way that lends itself to the estimation of their vulnerability to that particular hazard. This requires the categorization of the components at risk by their structural types, their height and their periods of construction. The structural typology greatly influences the vulnerability of elements at risk, e.g. buildings outside floodplains, building type, construction (timber or steel), while buildings in a floodplain may be weaker to the flood hazard. The period of construction may also influence the vulnerability of the components at risk because design regulations are gradually improved taking into account local and global hazard loss experiences. A significant amount of this information is census based such as building location, occupancy, etc.

Exposure components analysis uses the following data for its analysis:

- Building Construction types – Steel, Concrete, Masonry
- Size, No. of stories and age of the Buildings
- Occupancy and Type of Occupancy
- Type of Building – Residential, Commercial, Industrial
- Type of Structure
- Built area and non built area of the building/locational information

Vulnerability of the area is calculated in terms of both the physical and social functions. Social vulnerability is the susceptibility of populations to death and injuries, the assessment of which involves casualty modeling to compute mortality and injury rates associated with various catastrophic events. The physical vulnerability refers to the degree to which an asset would get damaged or destroyed in a hazardous environment caused by catastrophic events. The vulnerability module quantifies the damage susceptibility of each asset class with respect to varying levels of ground motion and collateral hazards. Imagery greatly helps to understand these functions, particularly as countries have access to freely available archived imagery through virtual earth or ESRI world imagery in a GIS that can be compared to imagery acquired for census and beyond.

Census Data in Disaster Management

Disaster models are preceded by data collection activities where by relevant and accurate data with respect to the hazard model requirement was generated, collated or sourced. i.e. calculation of loss and damages pertaining to flooding caused in the Caribbean. While the vulnerability of the areas is often calculated in modeling initiatives, the availability of data for countries varies widely. Some countries may have accurate block or community level data that is/was available within the census department and which can imply that there are no assumptions that can be made with respect to distribution of buildings and the population that can be affected by earthquakes. On the other hand, the ED level data or the lack of block level data often poses a challenge for accurately predicting or distributing the risk or losses across the buildings within the specified location. Hence, it is assumed that all the buildings were or could be equally impacted and would suffer similar losses and damage by earthquake. Therefore, it is critical that census attempt to take the activity as an opportunity to define and capture adequate building level information.

Advantages of Having Reliable and Complete Data:

1. Accurate distribution of buildings in the area insures the accurate distribution of losses. This is of particular importance in case of flood hazards where the banks of the river, more prone to disaster are farther from central congested areas of the city. If the data was not accurate at the specified scale, the losses presented by the models would be skewed.
2. The accuracy of population data of the buildings is important in case of casualty modeling. Casualty modeling uses the population data to stimulate the extent and number of casualties a particular hazard can cost.
3. As per our observations while modeling disasters, most of the census data which is available is from residential areas while the more disaster prone commercial and industrial areas are not covered well. Lack of data for commercial and industrial zones result in inaccurate distribution of losses and therefore reduce the applicability of the models.

The problem of not having reliable data for modeling purposes can be circumvented by using remote sensing techniques to generate the desired data.Using imagery to understand land use is a very important input for multiple applications. It is widely used in the telecom sector, urban and utility planning, environmental/hazard planning, etc. The land use map only depicts the relative density and sometimes relative height of buildings, but does not cover the number of people that are residing in the area. On the other hand, the census information is compiled at an administrative unit level; therefore, it lacks the spatial distribution or concentration of population within the administrative unit. This leaves a gap in needed information for a given disaster. Countries can now push beyond the collection of administrative level information to capture building level information in a GIS for the census. This level of information capture was often left to outside businesses and agencies due to lack of capacity and resources for countries to carry out such information capture. Today, countries are developing the capacity to capture this level of information through purchasing their own imagery near the time of the census.

Geographical Information System (GIS) helped in integrating the two datasets to derive the number of people living in each land use unit within an aggregated administrative unit. The information thus derived is very helpful in urban utility planning and disaster management.

Traditionally, for disaster management, information about the demographic details and its spatial distribution is imperative for assessment of human population at risk before the disaster and even for quantification of affected population after any disaster event. The maps generally available are either in the form of population density map or population maps depicting absolute population with respect to administrative boundary. In such maps, within the administrative boundary, where and how many people are residing is not known. This is now beginning to change since many census efforts have started to focus on capturing geographic information at the building level and linking that information to the census questionnaire so that demographic characteristics can be mapped at any aggregated level beyond the building. This is a revolutionary change in the practices for the census in the CARICOM region.

Often the opportune time to collect geographic and statistical data is during the census and, in many countries, it is the only comprehensive exercise with the opportunity to do so. Special considerations should be given to activities not directly related to just the census exercise and should span across many themes from disasters to economy to environment. The example of how census data at varied geographic levels is critical to flooding and other disasters in only one example of the great importance of having an established and accurate GIS of which to use as a resource.

ED Map Design, Pattern and Style Elements

The mapping process has traditionally been an extensive manual process prone to large human error. Creating each map manually or digitally on an individual basis takes time and a template speeds up this process enormously. This Enumeration District mapping template is one of the latest versions of the numerous existing mapping template styles that have been used throughout history. It shares common components, yet it is fundamentally different from the versions used normally. Many different world and country map templates exist, particularly within GIS software that do not have a strict focus on the ED map. Included in this document is accompanied a unique best practices approach which focuses on an applied GIS programming approach used by some CARICOM countries. This best practices approach assists them to standardize and streamline the mapping process which has traditionally been very time consuming. The required accuracy of census statistics today prompts a more rigorous geographic approach that centers itself on geographic accuracy and geographic information systems and their integration.

Templates are now found most commonly in electronic format and used widely in GIS programs or Graphic design software. The use of software driven templates has many possibilities and allows the user to save a template and reuse the exact template to display any preferred data layers immediately without recreating the map template components. Further to this, digital templates can go far beyond the purpose of streamlining the creation process to allow automating the map creation process. A brief introduction to an applied GIS programming approach used by some CARICOM countries that assists in standardizing and streamlining the Enumeration District mapping process follows.

Map template a pre-developed page layout in electronic or paper format.

A map template is a pre-developed page layout in electronic or paper format that can be used to make new map outputs with a similar design, pattern, or style. The page layout is the part of graphic design that deals in the arrangement and style treatment of elements or content on a page. In this context, the creation of Enumeration District census maps components will be presented and outlined that are necessary for creating standardized ED maps. Discussion will also focus on other map related tasks with regard to ED map development and production.

TheED mapping template is geared for the CARICOM countries and was developed with the consideration and collaboration of its members and associate members. There is no one template that all countries employ. The ED map template is a dynamic document that can be flexible according to users needed but also have common elements that are universal in nature.It aims to specify the critical elements needed for census enumerators and these elements are often case specific since countries geographies and land use patterns differ widely, particularly with regard to urban/rural or developed/undeveloped characteristics. Some countries may be heavily urban with concentrated, gridded land use pattern while others are widely dispersed, rural and have no uniformland use pattern. Therefore, the template must be a flexible document that the user can cater to suit his/her needs. If you do not need a primer on Enumeration District Design, Pattern and Style Elements you can skip directly to the templates located online through the CARICOM website in electronic format (Scalable Vector Format.SVG) or for import into GIS programs, .MXD (For Import into a GIS).

International Recommendations and Common Standards

Many census offices attempt to adhere to internationally recognized standards as much as possible. Unfortunately, enumeration area map design standards are not explicitly defined within any one document with specific examples of country use. Perhaps the most relevant resource for such recommendations and standards is found in the UN Handbook on Digital Mapping and GIS (2000) and the proceeding UN Handbook on Geospatial Infrastructure for Census and Statistics (2009), both published by the United Nations Press. These two handbooks explicitly reference ED map recommendations and standards and are a good resource to start from. Unfortunately, neither includes a working template or a source for retrieving an ED template to use in a real world adaptation. These needs were the primary drivers for the creating this document.

The primary purpose of a census ED Map is to indicate the area and its boundaries which must be covered by the enumerator during the canvass of his/her individual ED (see BUCEN, 1978, p.20). ED maps should be simple because they will be used by enumerators who have limited experience with maps. On the other hand, they must contain enough information to allow easy orientation. The following are some sample components of what an ED Map should contain according to the UN handbooks described above.

Sample components of a digital EA map

• The entire area to be enumerated, defined by a clearly indicated boundary line;

• Some parts of the neighboring areas (i.e., the peripheral area) to facilitate orientation;

• Any geographic and text information contained in the census cartographic database that will facilitate orientation within the EA: streets and roads, buildings, landmarks, hydrological features, and so on;

• A consistent map legend, including the exact names and codes of the administrative and enumeration zones, a north arrow, a scale bar and a legend explaining the symbols used for geographic features

The illustrations shown below in Figure 1 are the main individual components of a hypothetical urban EA map. All features are stored in separate map layers in the same spatial reference system or as graphics templates. The main components are the street network, buildings and EA boundaries layer. In addition, annotation, symbols, labels and building numbers are stored in separate data layers, although these could also be added dynamically. The last component is a template consisting of neatlines and a legend that is consistently used for all EAs.

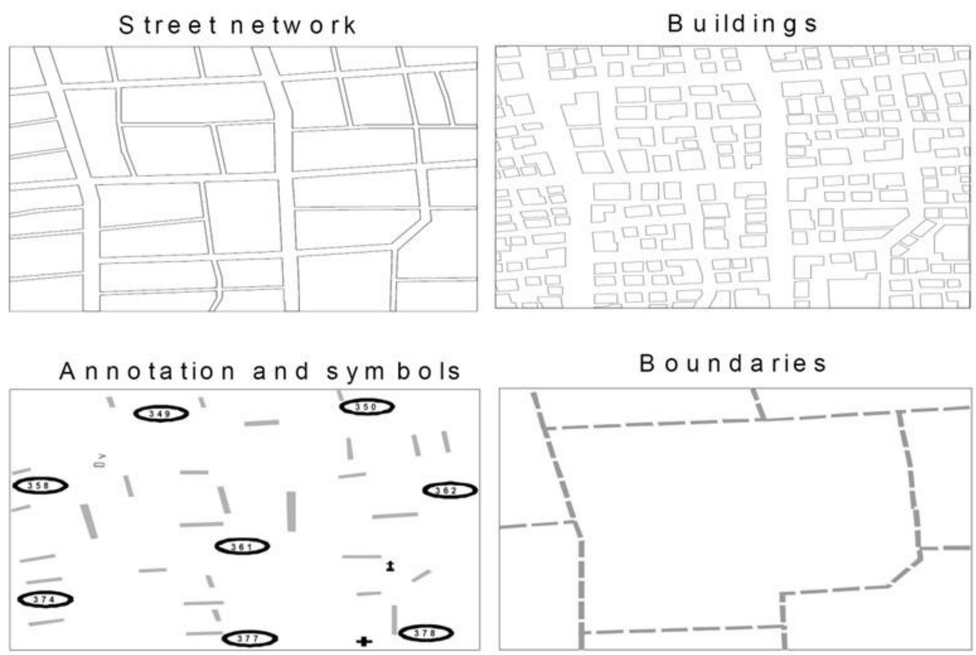


Figure 20 Main sample components of a digital EA map include street networks, buildings, annotations and symbols, boundaries, building numbers

UN Handbook (2000 p 77-78)
Sample components of a digital EA map.

To hold all of these layers and to indicate and reference these layers correctly, information must be contained in the ED map that explicitly calls out the components and their elements in each layer of information. For instance, legend items indicating a color and or size/type of line would be needed in order to distinguish which lines on the ED map were streets buildings or boundaries. Another example would be the need for orientation and scale to clearly understand where the enumerator is and what is there. This is the juncture of the ED map development process whereby the template needs to be considered. If you were to include other layers such as rivers and historical landmarks you would consider the information needed in your template. As maps get increasingly dense the importance of the ED template layout becomes more important. Often in urban areas roads (street centerlines) often overlap the boundary. To mitigate this issue a dotted line may represent a road and a hash line the ED boundary so that you would be able to distinguish the boundary from the road and clearly view the ED.

Figure 21 is an ED map template with neat lines and a legend. The legend is a key element in and ED map. Map Legends are used to help in the identification of boundaries and features from the symbols and names printed on the ED map. It is important to note in the template that all of the elements are in the legend area. This is one way of laying out your information that will reference the components in the template but there are many ways to arrange an ED template to make it work for your purpose. To restate, the primary purpose of the ED map is that it must contain enough information to allow easy orientation. Too much or too little information will hinder the enumerator's ability to perform.

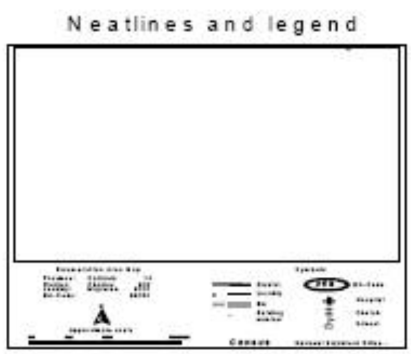


Figure 21 ED Map template illustrating neatlines and legend with no data shown in the data frame.

When the ED map template is populated with data and information it becomes a finished product. Often these maps, whether created by hand or digitally, take significant time to produce. Most countries are using GIS methods to create their census maps and on average the time it takes to produce the maps themselves after the geographic data or census GIS is complete or ready is approximately 2 months or greater. This is often far too long and frequently overlaps with other activities. Nevertheless, this is the case and having a template ready to use can help reduce time spent on design elements and allow more focus on GIS tasks, customization or other census preparatory work.

After the completion of verification and quality assurance procedures for all base maps and EA delineations, census cartography staff will prepare to print the final maps. This preparation process involves declaring a template for use that houses all necessary country information for each ED map. Figure 3 shows the complete EA map with all components overlaid in one map display. Depending on the scope of census mapping activities and the complexity of the enumerated area, EA maps may contain less or more information than this sample map. As GIS and digital mapping methods become widespread, more and more geographic information is contained in the ED maps due to the sheer amount of data that can fit on a paper using electronic means versus hand drawing.

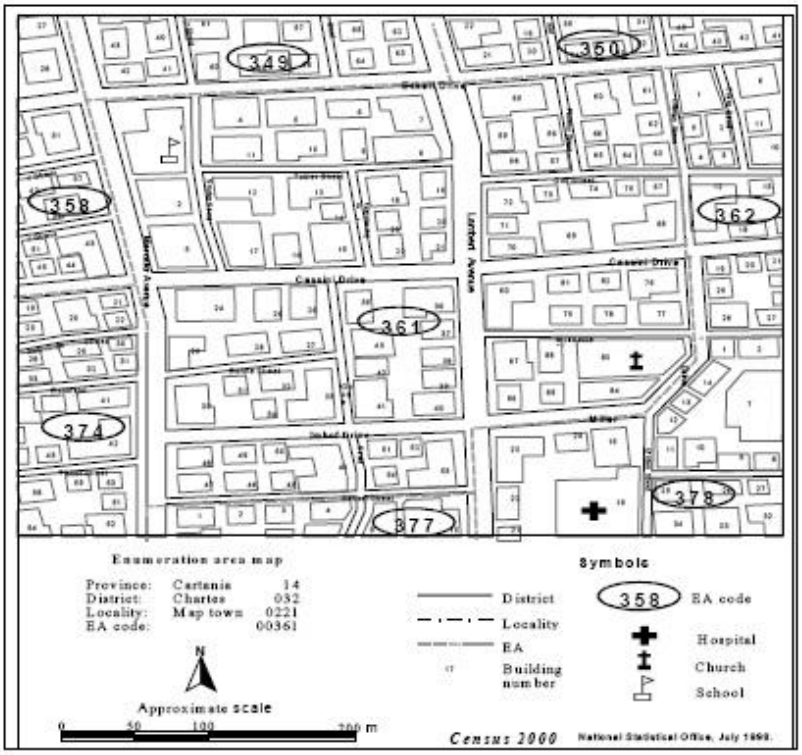


Figure 22 Example of an urban enumeration district area map (UN Handbook 2000)

ED Map Template Elements

Elements: referred to here as components of a model or in this case a template.

The elements you choose to include in the ED map template are essentially what will guide all enumerators in your work. It is important that detail oriented approach is applied in preparation of the map template. With design based ED templates it is critical to understand direction and scale when creating the ED maps. These two components are the most difficult to consistently maintain in design based ED map template creation environment. GIS based template environments often prevent this from happening since the geographic data is georeferenced on the globe within the view so any scale bar or north arrow added automatically adjusts to the respective view.

It is highly recommended that GIS based templates be used to avoid these issues and cannot be understated enough. Unless your template can be automatically georeferenced you will have a significant amount of extra work to do in adjusting the template for each map. This adjustment process is software specific and is not covered within the scope of this document. Most countries are using GIS programs to create their census maps and within these programs a seamless template can be created that links the data layer features with the legend, the north arrow with the template view, and so on. Below is a list of elements and other common considerations that require addressing in the development of an ED map template. Many of these elements apply to all census maps and maps in general.

Essential Elements

- ☐ North Arrow
- ☐ Legend
- ☐ Title Block
- ☐ Scale
- ☐ Boundary Lines
- ☐ Adjoining EDs
- ☐ Legend
- ☐ Mapping Symbols
- ☐ Disclaimer

Recommended Elements

- ☐ Start and Stop
- ☐ Canvassing Arrows
- ☐ Boundary Description (on the map itself)
- ☐ Names of Geographic features (e.g. Names of schools and Public Buildings)
- ☐ Peripheral Information
- ☐ Copyright Year

Additional/Optional Elements

- ☐ Map Inset
- ☐ Imagery or Aerial Photography
- ☐ Disclaimer
- ☐ Author
- ☐ Copyright Date

Scale Considerations

- ☐ 1:2500 is often optimal for best visual display on 11x17 ED map for densely populated areas
- ☐ 1:10,000 is often optimal for best visual display on 11x17 ED map for densely populated areas
- ☐ Most fluctuate between these ranges

Including Textual Information

Descriptions

Descriptions can be included in the map itself or on another sheet. The option to have the description on another sheet allows for more information to be provided to the enumerator rather than having a condensed description on the map itself. However, if there is room on the map to allow a description with the map becoming too cluttered with small textual descriptive information then descriptions should be included. This is also dependent on the level of training that the enumerators have of their selected areas for canvassing during the census. In some instances a description may not be needed

Disclaimers

Disclaimers are an important way to specify warnings or expectations to the general public (or some other class of persons) on the use right of maps produced.

Authoring

Authoring provides a source that links the process of preparing, assembling to the original creator.

Map Layout/Layout Styles

Data Frame

The data frame defines the context for the data with which you work; these include the coordinate system, measurement units, scale, the drawing order of layers, and so on. Often in a GIS, this data view isolates the contents of a data frame for you to edit or work with. The data frame is normally bound by a border that contains all the information in the map.

Insets

Inset maps are sometimes used to show related themes of data at smaller scales. Inset maps will either be open or closed forms. A closed form inset uses an immediately recognizable outline (a province, a country, the world, etc.) An open form is a portion of any of these that is a little harder for the reader to locate without additional information, unless your audience is already familiar with how you have cropped the area shown in the inset. Be sure to show a clear distinction between the inset and the main map (a border, drop shadow, etc.). You may also want to include orientation features as well such as a north arrow.

Map dimensions

11X17 is the most common map dimension used for census purposes. They tend to be portable and fold well with other documents that are carried by census workers. Depending on the scale of the map there may be instance where more than one 11x17 map is needed for a given census district. Multiple 11x17 maps are sufficient provided the maps clearly identify the areas of interest. In some cases larger maps might be the suggested choice. If larger maps are used it is critical that these map be portable enough for the enumerator and not cumbersome to fold and unfold during their work. Also, Census managers, alongside 11x17 maps might choose to have developed ANSE sized maps at 34 × 44(inches) or 864 × 1118 (mm). These maps are standard wall, general plotter sized maps that can contain a significant amount of information.

Most census mapping projects take advantage of as much canvas space as possible. The more canvas space available equals better (or larger scaled/more detail) geographic information in the ED map to guide the enumerator. Careful consideration must be given as to how many elements are needed for the ED map to be functional for the enumerator.

Regional Enumeration District Map Styles: Country Examples

Included in Annex I are examples of ED maps from different countries from the Caribbean region. There are numerous examples globally and regionally. These specific examples are dated circa 2009 and represent some of the trends in ED maps of the time. What is becoming more widespread is the use of two maps: one including imagery and one with vector data only. This has proven to be useful in providing the greatest amount of information to the enumeration so that they can quickly identify the building- dwelling unit- household. However, in many instances countries may only have the ability to produce one type of map (the maps may only have boundary data, streets, and building locations or vice versa/coupled with imagery).

As you can see in the examples provided many ED templates exists and tend to vary widely. What is common is that many appear to be produced electronically and most of them in a GIS. The fact that most are produced in a GIS reveals that since many countries are moving this direction some common flexible template framework is needed. With these maps in mind and the technologies used to create them CARICOM urged that a flexible template would be a strong way forward in setting the standard before a large deviation in template styles occurs. Already, as you can see, many of the ED maps vary with size, style and content within the different elements and components. However some commonalities exist. Before jumping to the template, without going into great detail, it is pertinent to quickly discuss briefly the evolution of the ED map creation process at this stage in order to clearly understand the template needs of countries and the content discussed further.

Where we stand/Evolution of the census mapping process

With information technology guiding the field of geography, increasingly adding deductive theoretical reasoning to the discipline, have come changes in the way geographic information is collected, stored and displayed. Geographic information systems have changed the way in which statistics offices data are stored and are produced for census purposes. With the exception of very few worldwide, countries have adopted GIS into their census mapping process in some extent. In the past there were often issues of timeliness that arose from being reliant on other departments, agencies or commercial suppliers for the production of census maps. This issue is being largely resolved through countries acquiring their own GIS program within the government or statistics office. This move has allowed all the census maps to be created in house and if preferred, out-sourced to a private firm for printing. With GIS, The statistics offices are given the right to be the data owners and cater the style of data to its working partners and clients.

Most countries use GIS to produce census maps or produce map outputs. This is often done on a per map basis whereby each map is created in the GIS the saved electronically one at a time. In many cases countries have large geographic datasets and use these datasets exclusively to produce census maps. The maps can include but are not limited to enumeration maps, census manager maps, field supervisor maps, regional maps. Some countries have set up full-operable systems whereby all statistics and geographic information are housed.

Census Mapping Template

This template was created in a GIS. As you see below, it is a straightforward template with many of the elements reflected in maps common to the region and the world. It is not meant to be complex or difficult to design. While templates are relatively simple to develop in a GIS for a technician, they may be confusing to develop for the outside user. The benefit of having the template, aside from common standards for Caribbean countries to follow and adapt, is the ability to import various software suites and start adding data immediately to the template without having to think about all of the map elements and design. For the novice ED map producer perhaps tasked with paper maps, the ED map template can be scanned or imported directly as an .mxd file (ArcGIS). For the GIS user, the template can be imported into GIS software and adapted programmatically to suit their needs. This template has been widely used by the CARICOM community.

Enumeration District Map Template

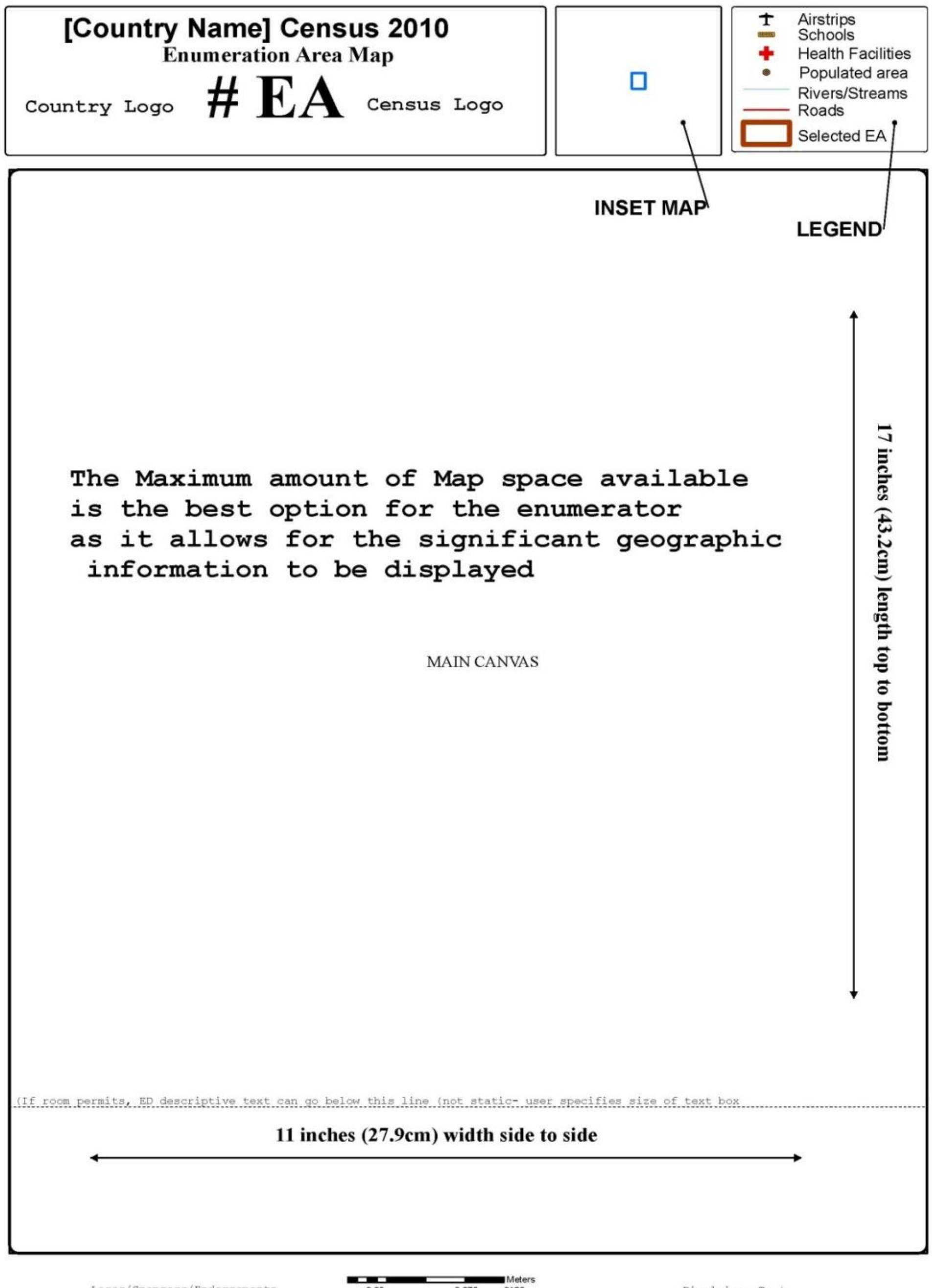


Figure 23Enumeration District Map Template Used as a starting point reference for census.

Increasingly, GIS users are more and more overwhelmed with daily office requests for census and field maps. This prompts further investigation into the needs of GIS users in charge of census. With this in mind it was necessary to move beyond just the template and briefly discuss some automation routines that can speed up the production of the hundreds and thousands of maps that must be produced by offices that normally have heavy workloads and differing resource constraints.

Automated Mapping through Programming Methods

The aim of this section is to show some capabilities of GIS for ED mapping beyond the use of the template. It is not a lesson on programming, object oriented languages, or GIS. The automated mapping through Programming Methods described here is shown using country specific examples. It is not an endorsement of software products, companies or country personnel. You are welcome to download the code and use it but beyond that you will be moving outside the scope of this document. If you choose to do so there are many sources available in geography forums, GIS websites, programming forums and specific user groups that can guide your excitement and interest in developing similar census products (that's the underscored aim here).

To better understand the needs of countries in terms of producing ED maps it is necessary to have a firm grasp of the dominant technologies used to create such outputs. The most dominant suite of mapping products worldwide for census mapping and GIS for that matter is the ESRI suite of products. With this in mind it is necessary to produce products that would easily integrate into this suite of these products. This document does not cover the ESRI suite of products in detail or the programming languages used to create the map **form**. Rather, the form overview is meant for countries that may not have access and/or expertise to develop a catered census map form and want something to work from immediately.

Form: foundation of a user interface and are guided through the use of controls with which users interact.

The form described here is a user interface that has controls that interact with the mapping software or more particularly in this case, the map template. Controls are the various widgets with which users interact. The form is aimed to quickly provide a given country with a working product that can develop capacity, initiate the learning process and be used quickly. It is hoped that this template will be a building block for countries to move more rapidly to technology driven methods to achieve a result. It is also hoped that the templates of this nature created by countries can be shared among users for others to develop and build upon. It is not meant as an endorsement for any products but for an extension of tools that are often not put to use in common mapping products due to various reasons.

Lastly, programming is complicated. Everything is so interrelated that it's difficult, if not impossible, to isolate each programming concept and then present the material in a linear fashion. That is precisely why it is used here to show the potential and be used as a way to teach the capabilities within a software, capabilities that are normally neglected but available to others to use to prepare for future census map projects.

The Need/Rationale

Most countries now use GIS and/or digital mapping to create series of maps for enumeration in a census. However, the production of these maps is still a manual process. Average time to complete could take up to 2-3 months or longer. This is often an unacceptable time frame for most National Statistical Offices charged with production of maps.

Constraints include:

- Digital map data for census work is of varying quality and formats differ from Census requirements
- Time: short timeframe of census lifecycle
- High staff turnover
- Lack of skilled staff and IT literacy
- Financial Issues

Addressing the Need

Integration options in a GIS allow for the creation of a geodatabase to be used within programming languages the create an automated process that is normally thought of as a manual process of adding layers to the view and graphically arranging text and icons to create a single census map. The automation of GIS mapping capabilities within GIS software will not limit other design needs and uses of the maps and overall display of data. Design needs such as adding an ED number or label for each map is streamlined.

Enumeration Map Form Layout Application Characteristics

Geodatabase: a collection of geographic datasetsfor use by GIS software (mainly ESRI)

User controls in the form define level of geography and map extent – from census EA, Village, Island, Province. These controls are fully integrated with the GIS software and flexible (meaning you can edit them). This integration is flexible to any workload such as fieldwork and preliminary mapping tasks as it automates map production against any geodatabase. Below is an illustration of the Form accompanied by the GIS data it is referencing in the map. Once the user selects the EA to preview, the GIS zooms to the ED and it is displayed within the template.

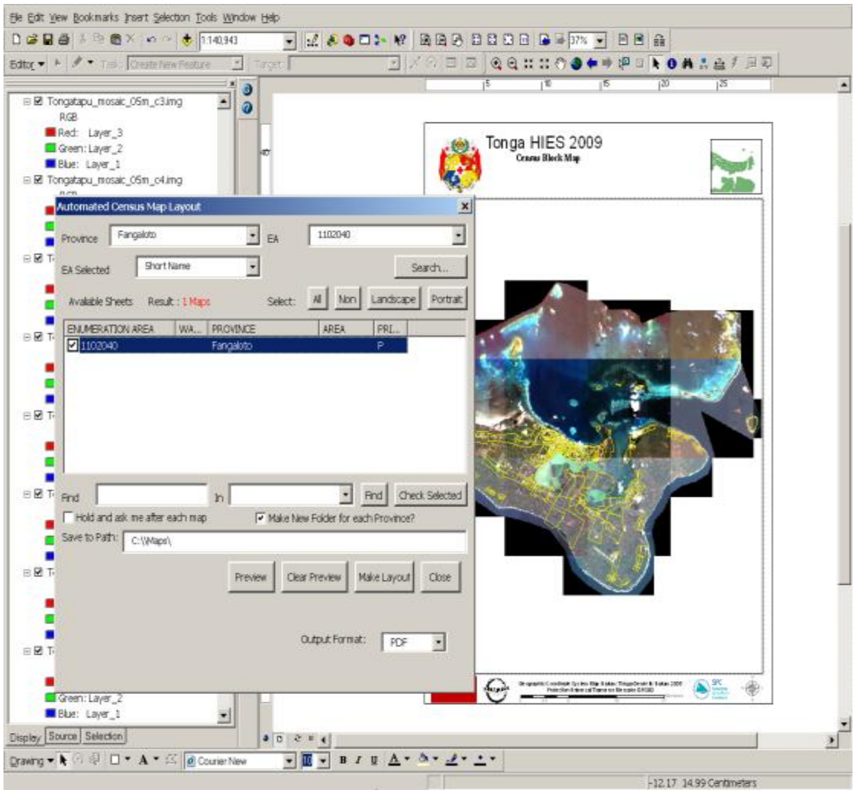


Figure 24 Illustration of the Form accompanied by the GIS data it is referencing in the map layout and template.

Once the ED is selected and preview is clicked the map layout zooms to the selected ED. From this point on the map can be viewed, checked, edited, and/or exported for print. You can select or unselect layers in the table of contents as you go. If the ED map appears to cluttered you can unselect a layer and vice versa. The illustration below shows the layout view with the selected ED displayed and layers selected to be included in the map (the selected ED is outlined in a lighter green).

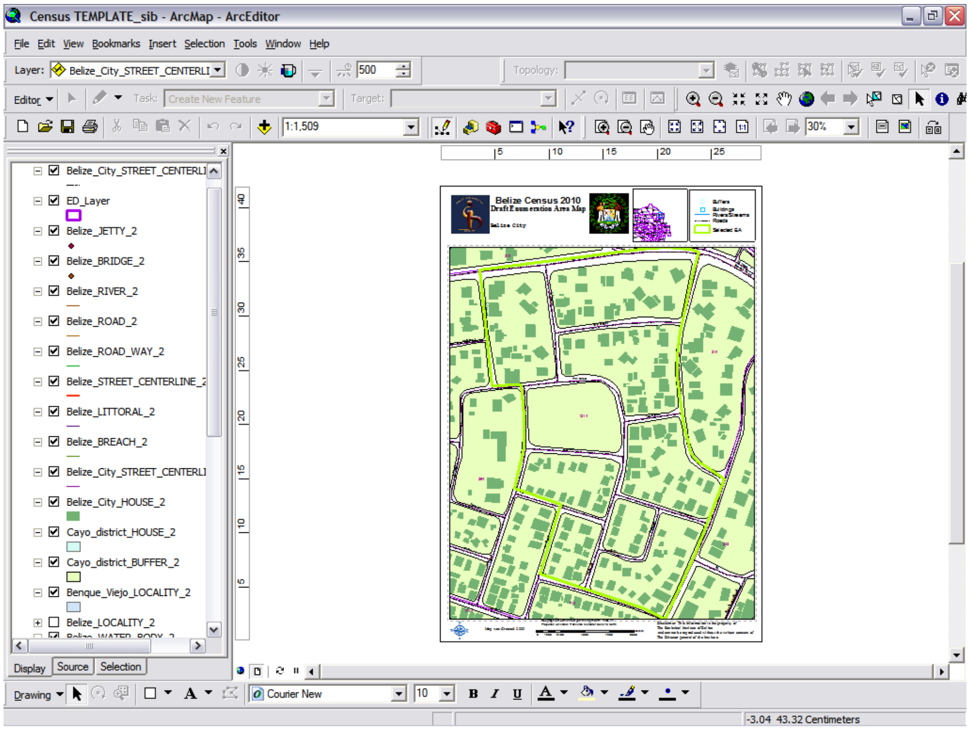
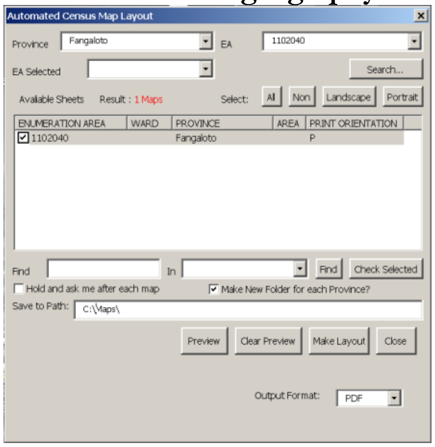


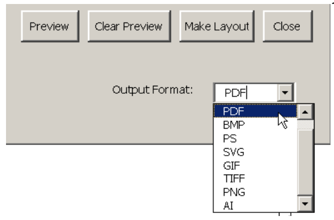
Figure 25 Illustration showing the layout view with the selected ED displayed and layers selected to be included in the map

The controls are easy for non-GIS specialists to use. However, to edit and create such controls requires a jump into understanding programming! The controls are designed to operate within ESRI's ArcGIS to enable automated mapping of EA areas at different levels of census geography.

Selection: Choose geography to map



Selection: Choose the map output type



The interface also allows users to select groups of EAs for census mapping by selecting the relevant geography of the country. The created maps can then be exported to a range of different formats such as PDF BMP EMF JPEG AI SVG GIF. The output image quality of the export can also be changed by assigning a value for example in the range 1 to 5. 1 corresponds to "Best" (larger file), 5 corresponds to Fast (smaller file).

Automated Mapping through Programming Methods developed through the use of forms save time and improves census office responsiveness and ability to create EA maps on demand for fieldwork and census execution. It is less labor intensive and increases focus on data quality of EA maps while not requiring staff to manually spend hours creating maps.

Furthermore, GIS is no longer isolated and users of all levels of GIS experience can produce, print maps ready to go out in the field through such an interface. Such interfaces are currently in use in Barbados, Belize and St. Lucia and others for the census and the applications within NSOs continue to grow. A few examples of the Automated Enumeration District Map Product Prototype are displayed below (It is referred to as a prototype until agreed by the NSO to be used for the enumerator in the census)

Enumeration District Form

The Forms code that have been created by the countries are not included. The illustrations are meant to show the products produced from similar forms. In some cases the form provided here has been used directly to create the products shown or have inspired the production of a new form interface. The Forms code has to be catered to the users' geographic information within the GIS. Therefore no one form can be used for all countries since all different countries geographic data are unique!

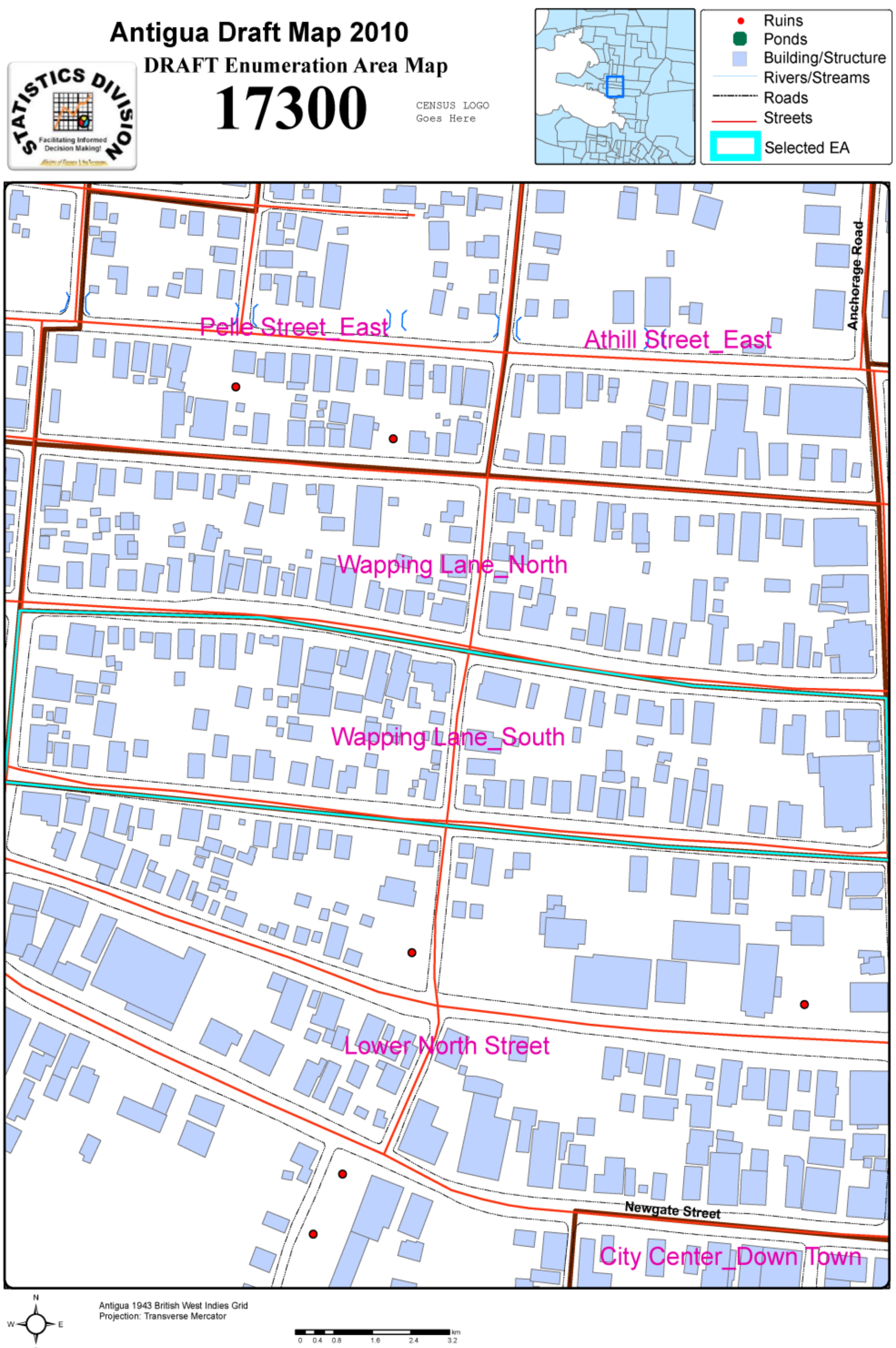
Note: This form template is built through programming ArcObjects with Visual Basic for Applications. The VBA environment has been used widely for building custom applications in ArcGIS products. This template, more specifically, is normally referred to as a form class used in a GIS mapping software. The form class is used to create and show forms at runtime.

Annex I.

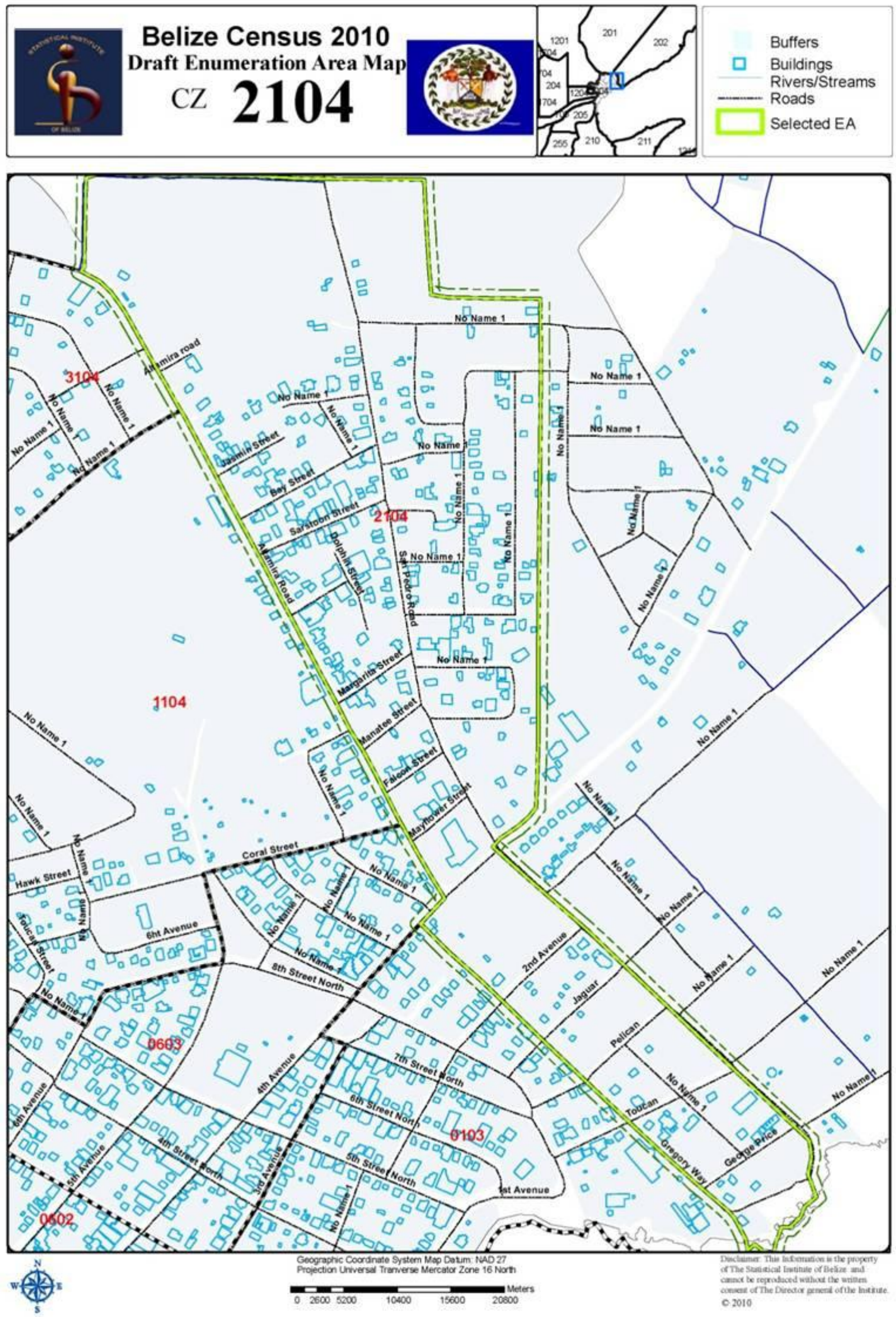
Output Examples from the use of the ED map Template

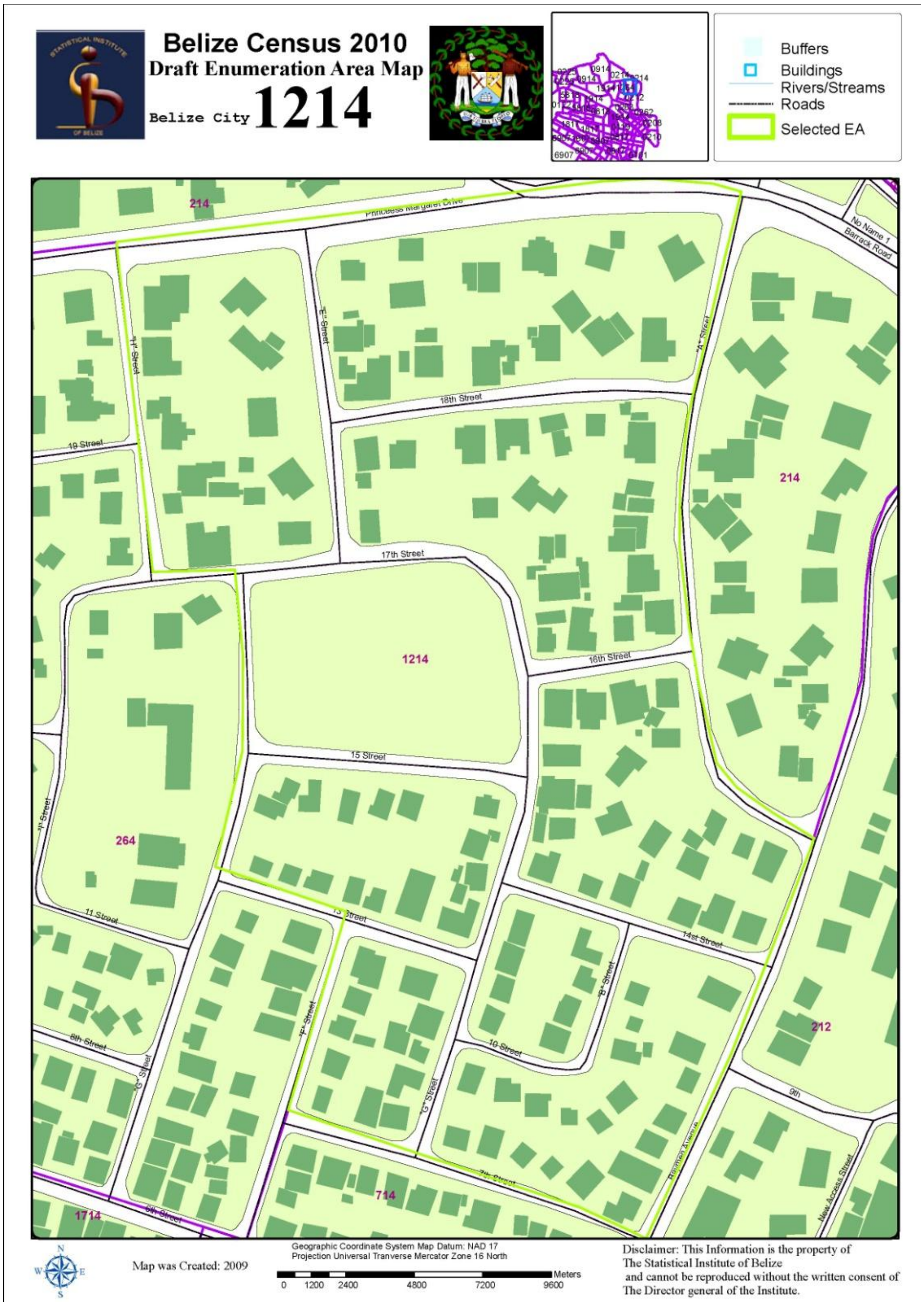
The following examples are some pre-developed page layouts in electronic format used by countries. They are produced largely through automated mapping programming methods. In no way do these examples represent census material that is used during the census. These examples represent the graphical outputs resulting from the template during testing phases of assistance visits or later correspondence. For explicit examples on the ED maps generated from the template contact the countries directly. These images were provided by statistics offices for example purposes only and cannot be used for distribution without consent of the National Statistics Office and/or in conjunction with the permission by CARICOM.


Antigua Draft Map 2010
DRAFT Enumeration Area Map
17300
CENSUS LOGO Goes Here
STATISTICS DIVISION
Facilitating Informed Decision Making!
Ruins
Ponds
Building/Structure
Rivers/Streams
Roads
Streets
Selected EA
Pelle Street_East
Athill Street_East
Wapping Lane_North
Wapping Lane_South
Lower North Street
City Center_Down Town
N
S
E
W
Antigua 1943 British West Indies Grid
Projection: Transverse Mercator
0 0.4 0.8 1.6 2.4 3.2
km



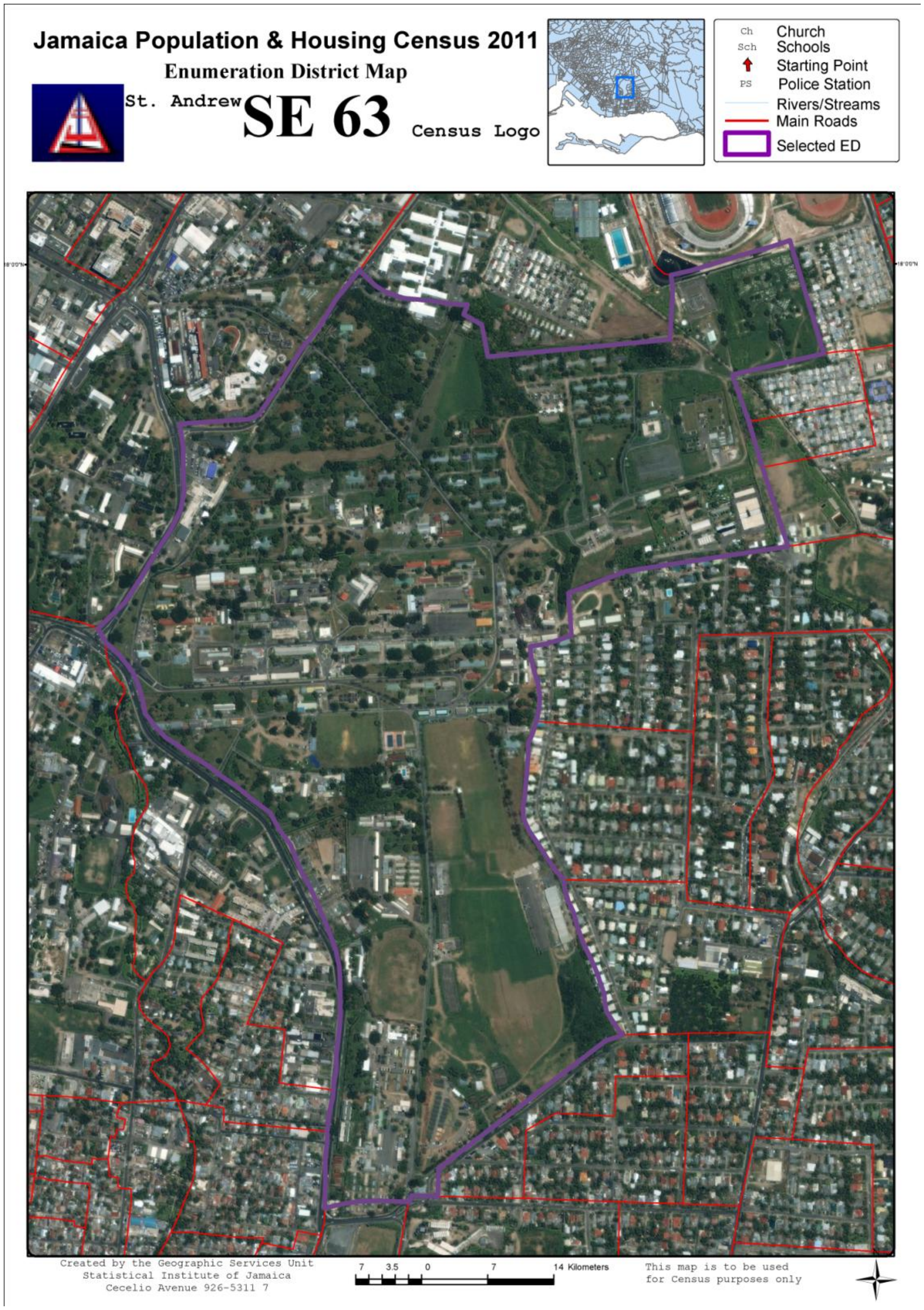
Antigua Draft Map 2010
DRAFT Enumeration Area Map
17300
STATISTICS DIVISION
Facilitating Informed Decision Making!
CENSUS LOGO Goes Here
Ruins
Ponds
Building/Structure
Rivers/Streams
Roads
Streets
Selected EA
Pelle Street_East
Athill Street_East
Anchorage Road
Wapping Lane_North
Wapping Lane_South
Lower North Street
Newgate Street
City Center_Down Town
N
W
E
S
Antigua 1943 British West Indies Grid
Projection: Transverse Mercator
0 0.4 0.8 1.6 2.4 3.2 km

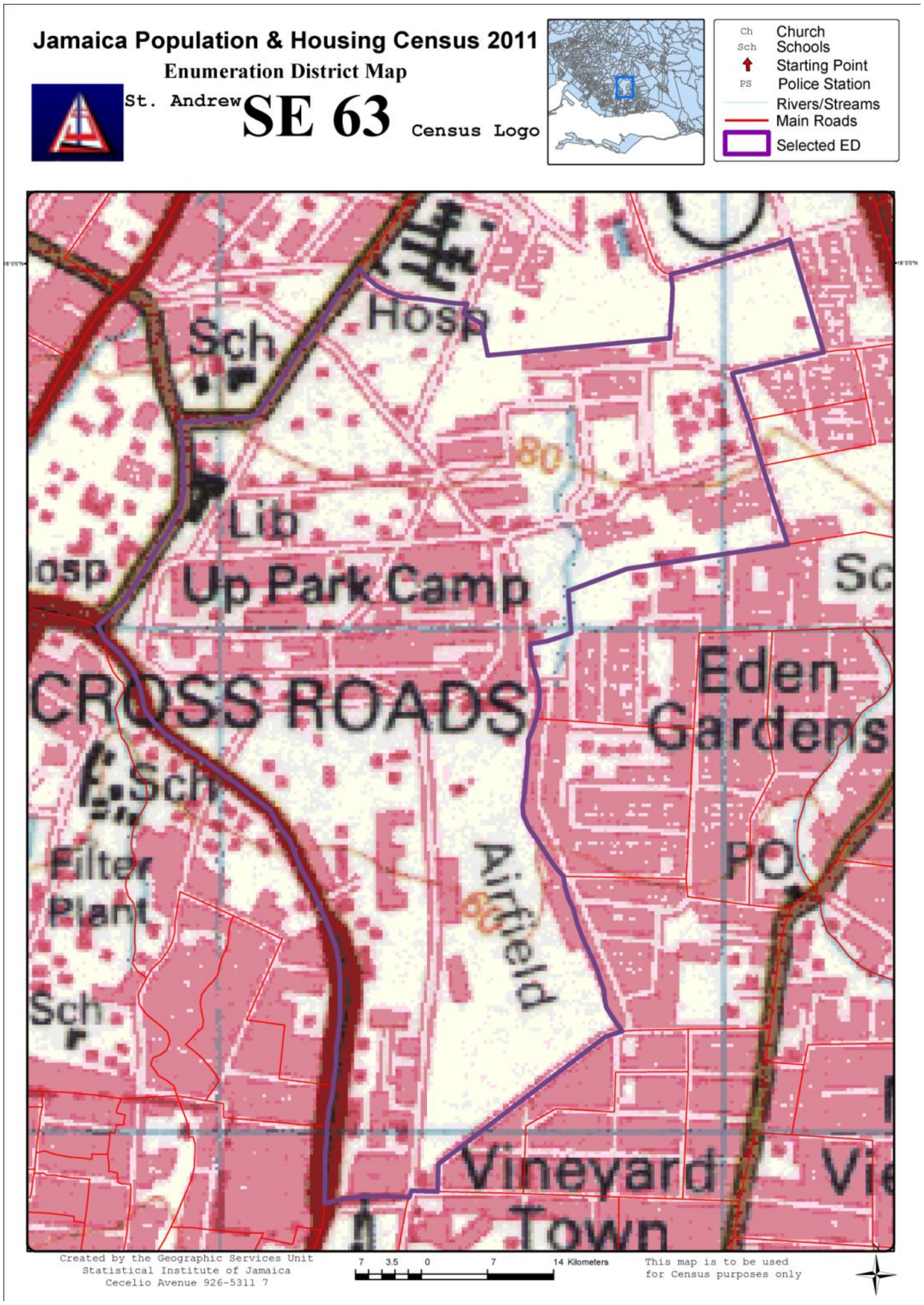

Belize Census 2010
Draft Enumeration Area Map
CZ 2104
Buffers
Buildings
Rivers/Streams
Roads
Selected EA
3104
2104
1104
0603
0103
0602
No Name 1
Altamira road
Altamira Road
Jasmin Street
Bay Street
Saratoon Street
Dolphin Street
San Pedro Road
Margarita Street
Manatee Street
Falcon Street
Mayflower Street
Coral Street
Hawk Street
6ht Avenue
8th Street North
7th Street North
6th Street North
5th Street North
4th Street North
4th Avenue
5th Avenue
1st Avenue
2nd Avenue
Jaguar
Pelican
Toucan
Gregory Way
George Price
Geographic Coordinate System Map Datum: NAD 27
Projection Universal Tranverse Mercator Zone 16 North
Meters
0 2600 5200 10400 15600 20800
Disclaimer: This Information is the property of The Statistical Institute of Belize and cannot be reproduced without the written consent of The Director general of the Institute.
© 2010

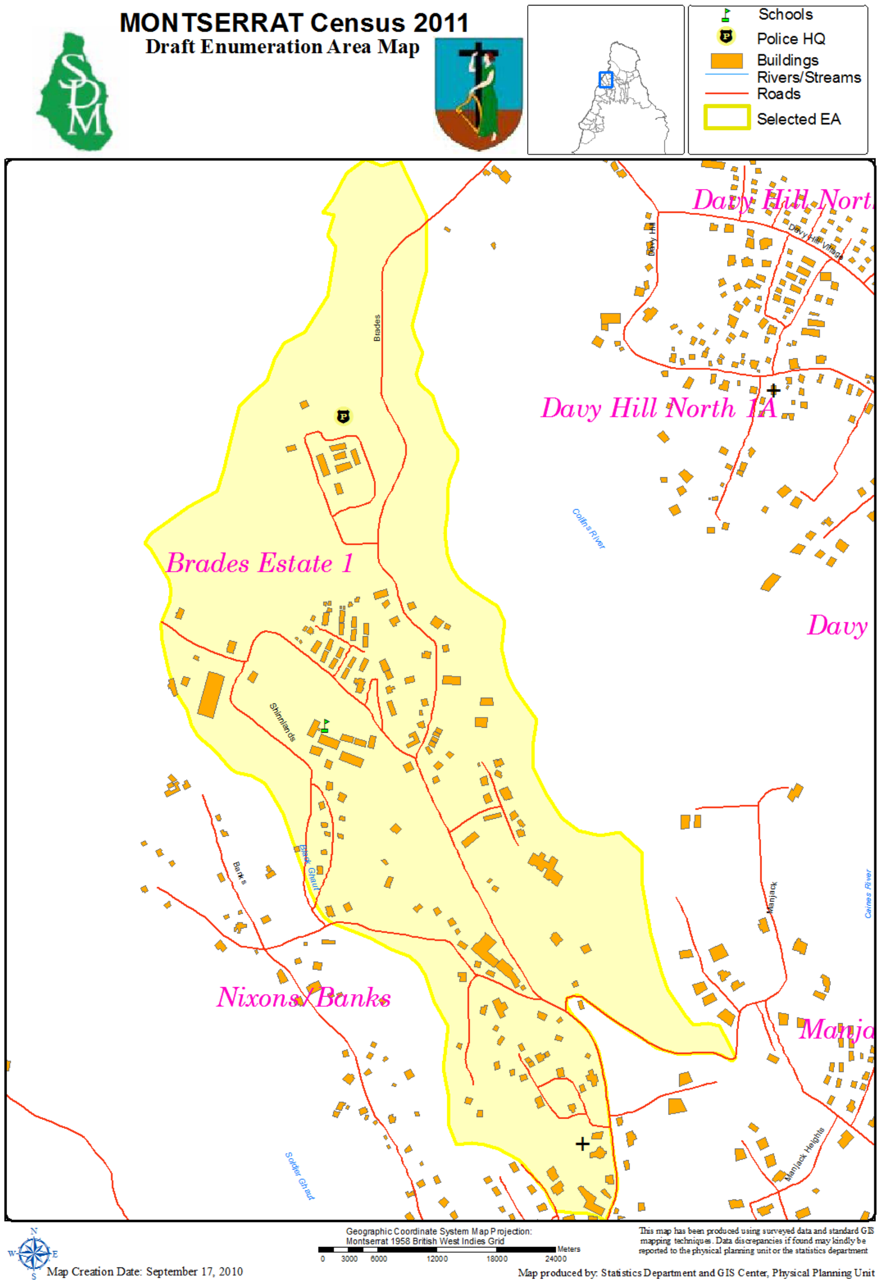

Belize Census 2010
Draft Enumeration Area Map
Belize City 1214
Buffers
Buildings
Rivers/Streams
Roads
Selected EA
Princess Margaret Drive
18th Street
17th Street
16th Street
15 Street
14st Street
10 Street
9th
1214
214
264
212
714
1714
Map was Created: 2009
Geographic Coordinate System Map Datum: NAD 17
Projection Universal Tranverse Mercator Zone 16 North
Meters
0 1200 2400 4800 7200 9600
Disclaimer: This Information is the property of
The Statistical Institute of Belize
and cannot be reproduced without the written consent of
The Director general of the Institute.

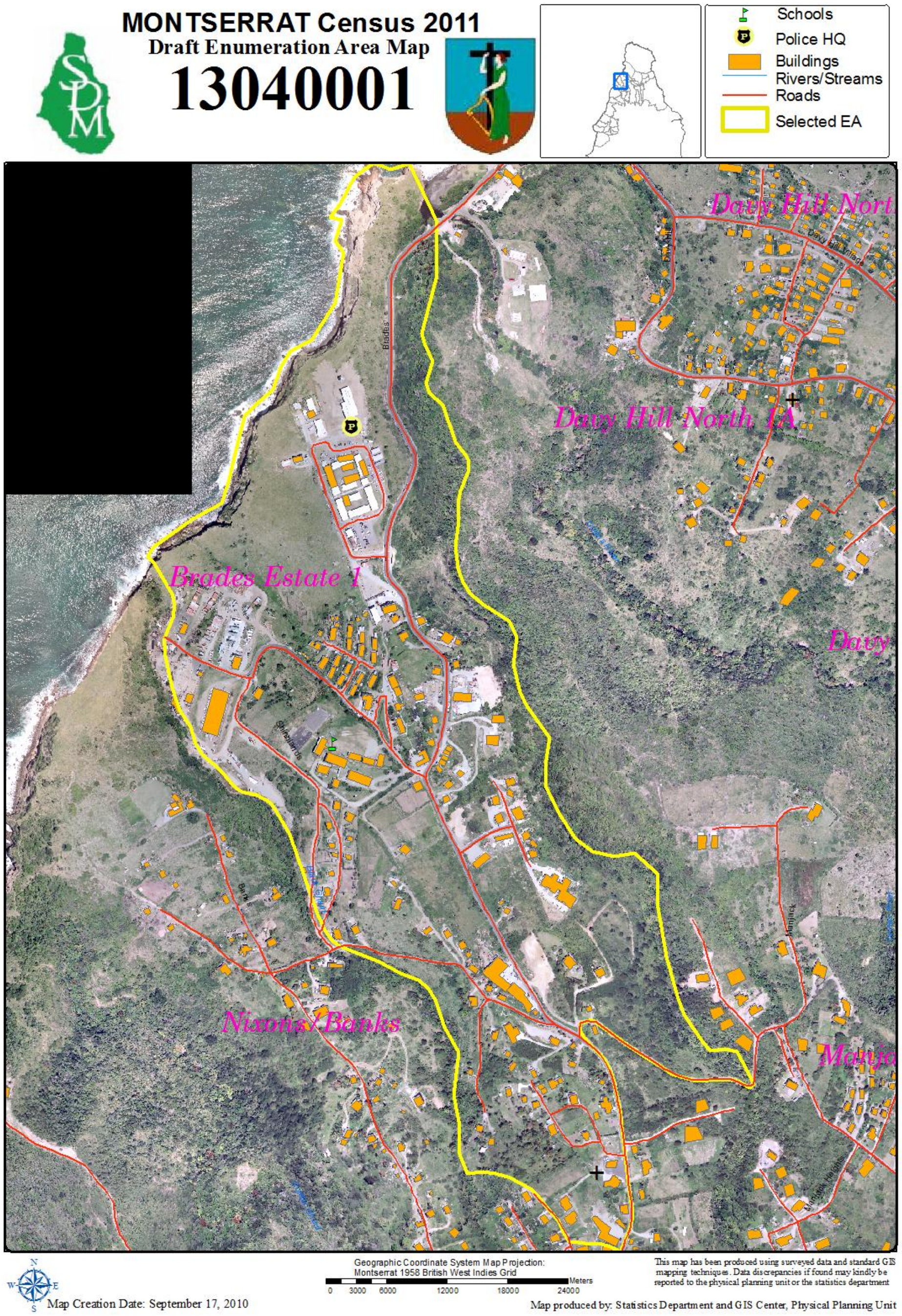

Dominica Census 2010
Enumeration Area Map
10010
Buildings
Dwelling Location
National Park
Rivers/Streams
Roads
Selected EA
300
310
340
30
20
10
350
370
360
60
40
60
St George
Logos/Sponsors/Endorsements
0
3350
6700
13400
Disclaimer Text

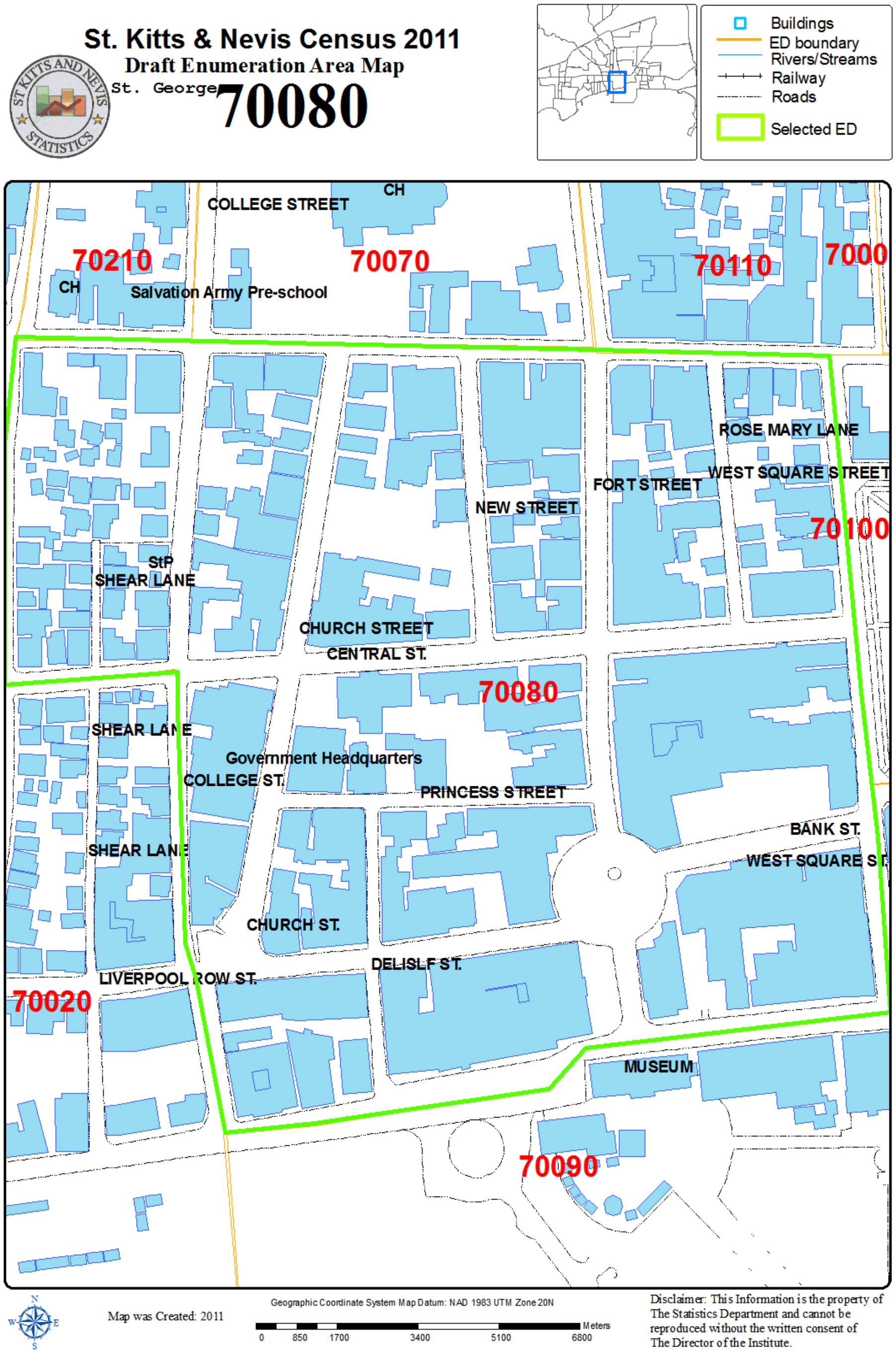

GRENADA POPULATION & HOUSING CENSUS 2011
Enumeration District Map
1200
EXAMPLE LEGEND
Schools
Health Facilities
Populated area
Rivers/Streams
Roads
Selected EA
300
300
1000
1100
1500
1200
1300
1400
1300
Created by the Grenada Statistics Office
0 4500 9000 18000 27000
Census Purposes Only


Jamaica Population & Housing Census 2011
Enumeration District Map
St. Andrew
SE 63
Census Logo
Ch Church
Sch Schools
Starting Point
PS Police Station
Rivers/Streams
Main Roads
Selected ED
Created by the Geographic Services Unit
Statistical Institute of Jamaica
Cecelio Avenue 926-5311 7
7 3.5 0 7 14 Kilometers
This map is to be used
for Census purposes only

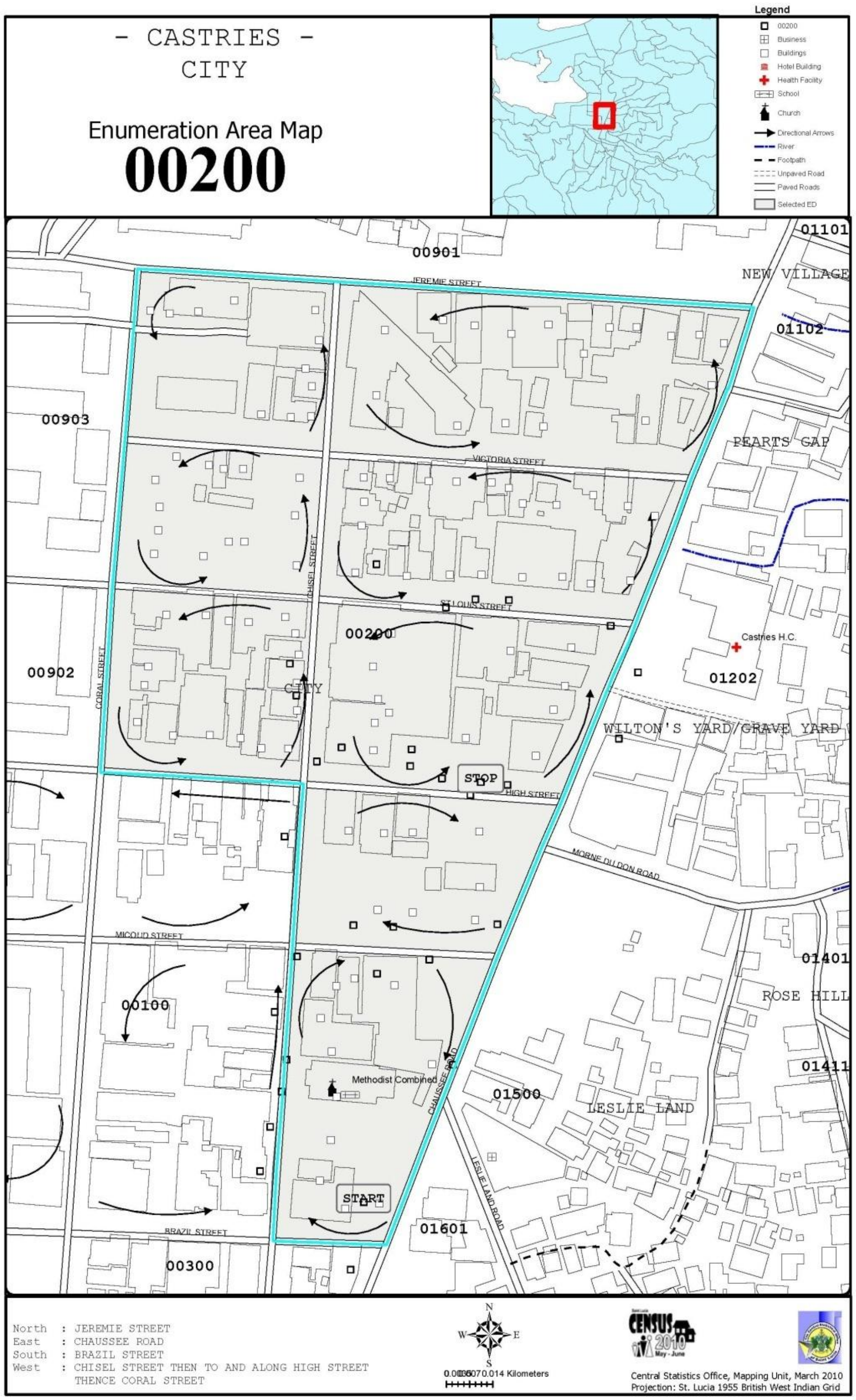

Jamaica Population & Housing Census 2011
Enumeration District Map
St. Andrew
SE 63
Census Logo
Ch Church
Sch Schools
Starting Point
PS Police Station
Rivers/Streams
Main Roads
Selected ED
Hosp
Sch
Lib
Up Park Camp
CROSS ROADS
Eden Gardens
Sch
Filter Plant
Airfield
PO
Sch
Vineyard Town
Created by the Geographic Services Unit
Statistical Institute of Jamaica
Cecelio Avenue 926-5311 7
14 Kilometers
This map is to be used for Census purposes only


MONTSERRAT Census 2011
Draft Enumeration Area Map
Schools
Police HQ
Buildings
Rivers/Streams
Roads
Selected EA
Davy Hill North
Davy Hill North 1A
Brades Estate 1
Davy
Nixons/Banks
Manja
Brades
Shinnlands
Banks
Manjack
Manjack Heights
Soldier Ghaut
Geographic Coordinate System Map Projection:
Montserrat 1958 British West Indies Grid
Meters
0
3000
6000
12000
18000
24000
This map has been produced using surveyed data and standard GIS mapping techniques. Data discrepancies if found may kindly be reported to the physical planning unit or the statistics department
Map Creation Date: September 17, 2010
Map produced by: Statistics Department and GIS Center, Physical Planning Unit


MONTSERRAT Census 2011
Draft Enumeration Area Map
13040001
SDM
Schools
Police HQ
Buildings
Rivers/Streams
Roads
Selected EA
Davy Hill Nort
Davy Hill North 1A
Brades Estate 1
Davy
Nixons/Banks
Manjo
Geographic Coordinate System Map Projection:
Montserrat 1958 British West Indies Grid
0 3000 6000 12000 18000 24000
Meters
This map has been produced using surveyed data and standard GIS mapping techniques. Data discrepancies if found may kindly be reported to the physical planning unit or the statistics department
Map Creation Date: September 17, 2010
Map produced by: Statistics Department and GIS Center, Physical Planning Unit


St. Kitts & Nevis Census 2011
Draft Enumeration Area Map
St. George
70080
ST KITTS AND NEVIS STATISTICS
Buildings
ED boundary
Rivers/Streams
Railway
Roads
Selected ED
COLLEGE STREET
CH
70210
70070
70110
7000
Salvation Army Pre-school
ROSE MARY LANE
WEST SQUARE STREET
FORT STREET
NEW STREET
70100
StP
SHEAR LANE
CHURCH STREET
CENTRAL ST.
70080
Government Headquarters
COLLEGE ST.
PRINCESS STREET
BANK ST.
WEST SQUARE ST.
CHURCH ST.
DELISLF ST.
LIVERPOOL ROW ST.
70020
MUSEUM
70090
Map was Created: 2011
Geographic Coordinate System Map Datum: NAD 1983 UTM Zone 20N
0 850 1700 3400 5100 6800 Meters
Disclaimer: This Information is the property of The Statistics Department and cannot be reproduced without the written consent of The Director of the Institute.


St. Kitts & Nevis Census 2011
Draft Enumeration Area Map
St. George 70080
Buildings
Rivers/Streams
Railway
Roads
Selected EA
COLLEGE STREET
CH
70210
70070
70110
7000
Salvation Army Pre-school
ROSE MARY LANE
WEST SQUARE STREET
FORT STREET
NEW STREET
70100
StP
SHEAR LANE
CHURCH STREET
CENTRAL ST.
70080
Government Headquarters
COLLEGE ST
PRINCESS STREET
BANK ST.
WEST SQUARE ST
CHURCH ST.
DELISLE ST.
LIVERPOOL ROW ST.
70020
MUSEUM
70090
Map was Created: 2011
Geographic Coordinate System Map Datum: NAD 1983 UTM Zone 20N
0 850 1700 3400 5100 6800 Meters
Disclaimer: This Information is the property of The Statistics Department and cannot be reproduced without the written consent of The Director of the Institute.


- CASTRIES -
CITY
Enumeration Area Map
00200
Legend
00200
Business
Buildings
Hotel Building
Health Facility
School
Church
Directional Arrows
River
Footpath
Unpaved Road
Paved Roads
Selected ED
00901
01101
NEW VILLAGE
01102
JEREMIE STREET
00903
PEARTS GAP
VICTORIA STREET
CHISEL STREET
ST LOUIS STREET
00200
Castries H.C.
00902
CORAL STREET
CITY
01202
WILTON'S YARD/GRAVE YARD
STOP
HIGH STREET
MORNE DU DON ROAD
MICOUD STREET
01401
ROSE HILL
00100
01411
Methodist Combined
CHAUSSEE ROAD
01500
LESLIE LAND
LESLIE LAND ROAD
START
01601
BRAZIL STREET
00300
North : JEREMIE STREET
East : CHAUSSEE ROAD
South : BRAZIL STREET
West : CHISEL STREET THEN TO AND ALONG HIGH STREET THENCE CORAL STREET
N
W
E
S
0.014 Kilometers
CENSUS 2010
May - June
Central Statistics Office, Mapping Unit, March 2010
Projection: St. Lucia 1955 British West Indian Grid


ED 00200
CASTRIES CITY
N
W
E
S
200
100
300
901
902
903
1101
1102
1202
1401
1411
1500
1601


ST. VINCENT & THE GRENADINES CENSUS 2010
CD 08 SANDY BAY
Census Logo
08010-SANDY BAY (1)
CD 08 SANDY BAY
Legend
Police Station
School
Road
River/Stream
Track
Building Footprint
8040
SANDY BAY (4)
New Sandy Bay Village
CARIBBEAN SEA
Middle Street
Windward Highway
8010
SANDY BAY (1)
8020
Karo Point
SANDY BAY (2)
N
From Karo Point (where the river meets the sea), along said river to where it is bridged on the Windward Highway; continue along the Winward Highway turning left onto the village road, where it junction with the Middle Street Road at the utility pole #15472; continue along Middle Street to the junction near the police station continuing along the track road to the Cayo River to the sea, back to the starting point.
Logos/Sponsors/Endorsements
0
0.025
0.05
0.1
0.15
0.2
Kilometers
Disclaimer Text

Annex II. Glossary

Accuracy —The degree to which a measurement or representation agrees with the true, real-world values. Determination of an acceptable accuracy requirement and development of an accuracy standard are some of the first steps in a GIS project. Accuracy is not to be confused with precision which refers to the ability to distinguish between small quantities in measurement. E.g., a point location might be measured precisely (e.g., with five significant decimal digits) but inaccurately (e.g., several meters off from it's true real-world position.

Address — a number or similar designation that is assigned to a housing unit, business, or any other structure. Addresses mainly serve postal delivery, but are also important for administrative purposes, for example in civil registration systems and in census taking.

Address matching — the process of matching general attribute information to geographical locations on a street network using a street address. For example, a tabular address register can be matched to a comprehensive digital street map to produce a GIS point layer showing the location of each household. This is sometimes also called geocoding.

Administrative unit — a geographic area that serves administrative and governmental functions. They are usually defined and established by legal action.

Aerial photography — the techniques for taking photographs from an aerial platform, usually a low-flying plane. Also sometimes called vertical photography or orthophotography. Air photos are used for photogrammetric mapping allowing a high degree of accuracy.

Aerial Survey — a cartographic survey by means of aerial photography or other remote sensing technology.

Annotation — text that is used to label features on a map. Annotation can be stored in a GIS and drawn onto maps for display or printing. In contrast to text information in an attribute table, annotation is only used for cartographic display and not for analysis.

Arc — see line.

Area — a bounded, two-dimensional extent of the earth's surface that is represented in a GIS as a polygon.

Areal interpolation — the transfer of an attribute from one set of reporting zones to another, incompatible, set of zones; for instance, the estimation of population totals for ecological regions based on a GIS data set of population by district.

Areal unit — a natural or artificial area that is often used to compile and report aggregate data. Examples: land cover zones or enumeration areas.

Attribute — a characteristic of a geographic feature. For example, a numeric or text field that is stored in a relational data base table which can be linked to the geographic objects in a GIS. Attributes of an enumeration area, for example, could be its unique identifier, the area in square kilometers, total population and number of households. A distinction is sometimes made between geographic and general attributes. The former are stored in a data table that is tightly linked to the geographic coordinate files and contains fields such as the internal identifiers, feature codes, area, etc. General attributes are typically stored in separate data tables that can be linked to the geographic attribute table.

Band — a layer of a multispectral remote sensing image that shows the signals measured in a defined range of the electromagnetic spectrum. See also multispectral image.

Arc second — one second of latitude or longitude, or 1/3600th of a degree.

Bandwidth — the amount or volume of digital data that can be transferred through a communications connection.

Base data — see framework data.

Base map — a map that shows fundamental geographic features that can be used for locational reference. Sample features are roads, administrative boundaries and settlements. Base maps are used to compile new geographic data or for reference in the display of thematic map information.

Base station — a GPS receiver, whose location has been precisely and accurately determined, that broadcasts and/or collects differential correction information for mobile GPS receivers. See also differential GPS.

Binary — made up of or referring to two, as in binary variables (e.g., yes/no). Also, a form of computer encoding that is based on individual pieces of information called bits that can take on two values—i.e., 0 and 1.

Bit — a binary digit that can assume a value of 0 or 1.

Boundary — a line which defines the extent of an areal unit or the locations where two areas meet. A boundary is represented in a GIS as a line feature which may define a side of a polygon. The boundary may or may not be visible on the ground. I.e., it can follow real-world features such as roads and rivers, or it can be defined solely by geographic coordinates.

Bits per second (BPS) — a measure of transfer speed in digital communication networks.

Buffer — a zone or area of a specified distance around a geographical feature (points, lines, or polygons). Buffer operations are one of the fundamental geospatial capabilities.

Byte — a group of eight binary digits or bits that can be processed as a unit by computer programs. A kilobyte consists of approximately one thousand bytes, a megabyte of one million bytes, and a gigabyte of one billion bytes.

Cadastral information — records that describe the past, present and future rights and interests in land ownership for legal and tax purposes. Cadastral maps show the geographic location and extent of land parcels. Cadastral surveys in many countries now use GIS to store this information. Also called land titling information.

Cartesian coordinate system — a system of lines that intersect at perpendicular angles in two-dimensional space. This system provides the framework to precisely reference locations as x/y coordinates.

Cartogram — a map that is constructed by scaling the reporting units according to the value of a variable recorded for them. Also called value-by-area mapping.

Cartographic generalization — the process of abstracting real world features through a reduction of detail for representation on a map. This involves selection, classification, simplification, and symbolization.

Cartography — the art and science of creating a two dimensional representation of some part of the earth's surface. Features represented may be real objects (topographic mapping), or they may represent concepts and more abstract characteristics (thematic mapping).

Census geographic framework — the geographic collection and reporting units used by a census office in census enumeration and data tabulation. This includes the hierarchical structure of census and administrative units, their designations and codes and the relationships between different units.

Central meridian — the longitude that defines the origin of the x-coordinate of a cartographic projection.

Centroid — the mathematical center of a polygon. For irregularly shaped polygons this can be thought of as a "center of gravity".

Channel — the part of a GPS receiver's electronics that captures the satellite's signal. Multi-channel receivers can capture and process signals from several satellites at the same time.

Chart — a map that is primarily designed for sea and air navigation, for example nautical or aeronautical charts.

Choropleth map — a statistical map in which values recorded for reporting units are first assigned to a number of discrete class ranges or categories. The reporting units are then shaded using symbols (colors or patterns) chosen for each category.

Classification — assigning objects into groups that share the same or similar characteristics. In cartography, the process of assigning symbols to map features that are similar or that have similar values. Classification is used to simplify a map in order to improve communication of the cartographer's message.

Client — a computer that uses data or software stored on another, often remote, computer (server).

Code — the alphanumeric characters used to identify geographic objects. Codes are also used to identify attribute categories such as population density ranges, land use classes or industries. See also geographic code.

Column — in GIS, a group of cells or pixels in a grid or raster GIS database that are aligned vertically. In database management systems, a field or item in an attribute table.

Computer Graphics Metafile (CGM) — a standard file format for exchanging image or vector data.

Computer-Aided Design/Computer-Aided Design and Drafting (CAD/CADD) — a software system that provides the tools for drafting and design, specifically in engineering or architectural applications. CAD systems use graphical coordinate system and are therefore similar to geographic information systems.

Conformal projection — a cartographic projection in which all angles are preserved correctly at each point.

Connectivity — in topological GIS, when two or more lines are joined at a single point or node.

Contiguity — if two or more geographical features are neigbors or adjacent.

Continuous geographical phenomena — geographic variables that vary without clearly distinguishable breaks or interruptions, for example temperature or atmospheric pressure—as opposed to discrete geographical phenomena.

Contour — a line on a map that connects points of equal elevation. See also isoline.

Control — see geodetic control.

Control point — a point on a map, an aerial photo or in a digital database for which the x,y coordinates and possibly elevation are known. Used to geographically register map features.

Control segment —a global network of GPS monitoring and control stations that ensure the accuracy of the satellite signals.

Coordinate — two or three numbers that describe the position of a point in two or three dimensions (e.g., x/y or x/y/z, where z indicates height). A two-dimensional coordinate is sometimes called a coordinate pair, a three-dimensional coordinate a coordinate triplet. In GIS databases, coordinates represent corresponding locations on the earth's surface relative to other locations.

Coordinate Geometry (COGO) — term used by land surveyors for dealing with precise measurements of locations.

Coordinate system — the reference system that is used to specify positions on a map or in a GIS database. A cartographic coordinate system is defined by a map projection, a reference ellipsoid, a central meridian, one or more standard parallels and possible shifts of x and y coordinate values.

Coverage — in GIS, coverage sometimes refers to a vector GIS data set that contains geographic features belonging to a single theme such as census units or roads.

Data capture — conversion of geographic coordinate data from hardcopy sources or by means of field measurements into a computer-readable format. Data capture usually involves digitizing or scanning of paper maps or air photos.

Data conversion — the transfer of data from one format into another. Usually data conversion refers to the translation of paper map information into digital form. In a wider sense geographic data conversion also includes the transfer of digital information from one GIS file format into another.

Data dictionary — a data catalog that describes the contents of a database. Information is listed about each field in the attribute tables and about the format, definitions, and structures of the attribute tables. A data dictionary is an essential component of metadata information.

Data format — usually refers to a specific, possibly proprietary, set of data structures within a software system.

Data model — a user's conceptual design of a data set that describes the database entities and their relations to one another.

Data sets — a logical collection of values or database objects relating to a single subject.

Data standardization — the process of reaching agreement on common data definitions, formats, representation, and structures of all data layers and elements.

Data structure — implementation of a data model consisting of file structures used to represent various features.

Data type — the field characteristic of the columns in an attribute table. For example, character, floating point and integer.

Database — a logical collection of information that is interrelated and which is managed and stored as a unit, for example in the same computer file. The terms database and data set are often used interchangeably. A GIS database contains information about the location of real-world features and the characteristics of those features.

Database Management System (DBMS) — a software package designed for managing and manipulating tabular data. A DBMS is used for the input, storage, manipulation, retrieval, and query of data. Most GISs use a relational DBMS to manage attribute data.

Datum — in cartography, a set of parameters that define a coordinate system. More specifically, a datum is a reference or basis for measurements or calculations. For example, a national cartographic datum establishes the reference framework for cartographic activities in a country.

Differential GPS (DGPS) — the set of techniques used to improve the accuracy of coordinates captured with a GPS by calculating the signal error (offset) for a second GPS receiver (the base station) at a location which has been precisely and accurately determined. The correction factor is applied to the coordinates captured by the mobile unit, either in real-time or in post-processing mode (i.e., using a database of time referenced correction information). In some parts of the world, differential correction information is broadcast continuously from a set of permanent base stations.

Digital Elevation Model (DEM) — a digital representation of elevation information for a part of the earth's surface. A DEM is usually a raster data set in which elevation values are stored for cells in a fine grid, but vector formats can also be used to store elevation. A DEM is sometimes also called digital terrain model (DTM).

Digital orthophoto — a digital image or aerial photograph, usually of very high resolution, which has been geometrically corrected. A digital orthophoto, also called orthoimage, combines the detail of an aerial photograph with the geometric accuracy of a topographic map.

Digital Terrain Model (DTM) — see Digital Elevation Model (DEM).

Digitizing table — a computer peripheral used to capture coordinate data from paper maps or similar cartographic materials. Also called a digitizer.

Digitizing — the process of translating geographic feature information on paper maps into digital coordinates. Digitizing usually refers to the manual process of tracing lines on a paper map attached to a digitizing table with a mouse-like cursor that captures coordinates and stores them in a GIS database.

Discrete geographical features — individual entities that can be easily distinguished such as houses or roads—as opposed to continuous geographical phenomena.

Dissolve — a GIS function that deletes boundaries between adjacent polygons that have the same value for a specific attribute. For example, enumeration area polygons can be dissolved based on the code of their supervisory units to create supervisory maps.

Dot map — a map in which quantities or densities are represented by dots. Usually each dot represents a defined number of discrete objects such as people or cattle. The dots can be placed randomly in the reporting units, or they can be placed to reflect the underlying true distribution of the variable.

Drawing Exchange Format (DXF) — an ASCII format for describing a graphic or drawing developed by Autodesk, Inc. (Sausalito, CA). Initially developed for CAD applications, it has also become a standard for GIS data exchange.

Edgematch — a manual or automated editing technique in a GIS that matches shared features that were digitized from adjacent map sheets. Edgematching may be necessary, for instance, to connect roads or administrative unit boundaries after joining maps that were digitized separately.

Ellipsoid — in cartography, the three dimensional shape used to represent the earth. The earth ellipsoid is characterized by a smaller distance from the center to the poles (semi-minor axis) than that from the center to the equator (semi-major axis). Also called a spheroid.

Entity — a real world phenomenon of a given type. In database management systems, the collection of objects (e.g., persons or places) that share the same attributes. Entities are defined during conceptual database design.

Enumeration area — usually the smallest geographic unit for which census information is aggregated, compiled and disseminated. An enumeration area is defined by boundaries described on a sketch map or in a GIS database. These boundaries may or may not be visible on the ground. Also called census block or census tract.

Equal area projection — a cartographic projection in which all regions are shown in correct proportion to their real-world areas.

Equator — in cartography, the reference parallel, i.e., latitude 0° north and south.

Equidistant projection — a cartographic projection which maintains the scale along one or more lines, or from one or two points to all other points on the map.

Feature — a geographic object displayed on a map or stored in a GIS database. Features can be natural or man-made real-world objects (a river or a settlement) or they can be conceptual or defined features (e.g., administrative boundaries).

Field — a column in a database table.

File Transfer Protocol (FTP) — a standard set of conventions for exchanging computer data files in digital communication systems such as the Internet.

Flow map — a map in which movements, for example of goods or people, along a linear path are shown.

Foreign key — in relational database management systems, a field or item in a table that contains a value identifying rows in another table. It is used in joining two tables by defining the relationship between two elements of a relational database. A foreign key is the primary key in the other table.

Framework data — in the context of national GIS activities, a set of general purpose geographic themes or base data, such as administrative boundaries, elevation or transportation infrastructure. Framework or national spatial data infrastructure initiatives aim at coordinating the development and standardization of GIS data sets of framework data in a country.

Gazetteer — a list of place names and their geographic location (usually latitude/longitude).

Generalization — see cartographic generalization.

Geocoding — (a) the process of assigning geographic codes to features in a digital database; (b) a GIS function that determines a point location based on a street address. See also address matching.

Geodetic control — a network of precisely and accurately measured control or reference markers that are used as the basis for obtaining new positional measurements. Also called benchmark points.

Geographic attributes file — a database table that is tightly linked to the spatial objects stored in a GIS coordinate file. The geographic attribute file or table contains specific information on each feature such as its identifier, name and surface area. In some systems this file is also called point, line or polygon attribute table. Data stored in external tables can be linked through a relational database operation.

Geographic code — a unique alphanumeric identifier that is assigned to a legal, administrative, statistical or reporting unit.

Geographic database — a logical collection of data pertaining to features that relate to locations on the earth's surface.

Geographic hierarchy — in the context of census mapping, a system of usually nested area units that are designed for administrative or data collection purposes. For instance, a country is divided into provinces, which are divided into districts, and so on to the lowest level, which may be the enumeration area. See also census geography.

Geographic Information System (GIS) — a collection of computer hardware, software, geographic data, and personnel assembled to capture, store, retrieve, update, manipulate, analyze, and display geographically referenced information.

Geographic object — a user-defined geographic feature or phenomenon that can be represented in a geographic database. Examples include streets, land parcels, wells and lakes.

Geographic reference file — a digital, tabular master file that lists the names, geographic codes, and possibly attributes of all geographic entities that are relevant to census and survey data collection.

Georeferencing — the process of determining the relationship between page coordinates and real-world coordinates. Georeferencing is necessary after digitizing, for example to convert the page coordinates measured in digitizing units (e.g., centimeters or inches) into the real-world coordinate system that was used to draw the source map. See also transformation.

Geospatial — a term that is sometimes used to describe information of a geographic or spatial nature.

Geo-TIFF — see tagged image file format.

Global Positioning System (GPS) — a system of 24 satellites orbiting the earth that broadcast signals that can be used to determine the exact geographic position on the earth's surface. GPS is used extensively in field mapping, surveying and navigation. The

GPS is maintained by the U.S. Department of Defense. See also Differential GPS, Beidou, Galileo, and GLONASS.

Governmental unit — see administrative unit.

Graduated symbols — in thematic cartography, the use of symbols (e.g., circles or squares) to represent the magnitude of a variable at a point or in a reporting unit. The size of the symbol is proportional to the value of the variable.

Graphic Interchange File (GIF) — a graphics image file format developed initially for transmission of images through electronic bulletin boards. The GIF format, which allows efficient compression of file size, is used for most graphics on web pages.

Graticule — in cartography, the grid of longitudes and latitudes drawn on a map.

Greenwich meridian — the longitude of reference, i.e., 0° east or west. It passes through the English town of Greenwich, a suburb of London.

Grid — a geographic data model that represents information as an array of uniform square cells. Each grid cell has a numeric value that refers to the actual value of a geographic phenomenon at that location (e.g., population density or temperature) or it indicates a class or category (e.g., the enumeration area identifier or soil type). See also raster.

Ground truth — information collected in a field survey to verify or calibrate information extracted from remote sensing data.

Heads-up digitizing — a digitizing technique that does not employ a digitizing table. Instead, features are traced with a mouse on-screen either from a scanned image displayed in the background or following features drawn on a clear medium (e.g., mylar) that is attached to the computer screen.

Hydrography — features pertaining to surface water such as lakes, rivers, canals, etc.

Hypsography — features pertaining to relief or elevation of terrain.

Image — a representation of a part of the earth's surface. However, an image is usually produced using an optical or electronic sensing device. For instance, scanned aerial photographs or remote sensing data are usually referred to as images. In terms of data storage and processing, an image is very similar to a raster or grid.

Infrastructure — the system of public works in a country, state, or region including roads, utility lines, and public buildings.

Integration — in GIS, the process of compiling a consistent set of spatial data from heterogeneous sources. Vertical integration refers to the ability of GIS to combine different data layers that are referenced in the same coordinate system.

Internet — a global system of linked computer networks that allows data communication services such as remote login, file transfer, electronic mail, bulletin boards and newsgroups. The internet is also the foundation for the world wide web (www).

Internet Protocol (IP) — the most important set of codes and conventions which enable the transfer of digital data on the Internet.

Interpolation — the process of estimating a variable value at a location based on measured values at neighboring locations. Used to produce a complete grid data set from point sample information, for instance, a precipitation surface from rainfall stations.

Intersecting — a GIS function that is used to topologically integrate or combine two spatial data layers so that only those features that are located within the area common to both are preserved.

Isoline — lines on a so-called isarithmic map that connect points of constant value. The best known example is an isohypse, which shows lines of equal elevation (also called an elevation contour map).

Join — in relational database management systems, the process of attaching values from a database table to another table based on linking a foreign key to its primary instance in the external table.

Joint Photographic Experts Group (JPEG) — a graphics file format used primarily for photographic images that allows significant file size compression.

Land Information System (LIS) — a term sometimes used for a GIS application that contains information about a specific region including cadastral information, land use, land cover, etc.

Latitude — the "y-coordinate" in a polar coordinate system on a sphere. Measured as the angular distance in degrees north or south of the equator. Also called parallel.

Layer — an individual GIS data set that contains features belonging to the same theme, such as roads or houses. The term layer refers to a GIS's ability to overlay and combine different thematic layers that are referenced in the same coordinate system. Also called coverage.

Legend — in cartography, the information on a map that explains which symbols are used for the features and variables that are represented on the map. This includes the symbol key required to interpret the map, for example, the shade colors and corresponding value ranges of a population density map.

Line — a one-dimensional object. A geographic data type consisting of a series of x,y coordinates, where the first and last coordinates are called nodes and the intermediate coordinates are termed vertices. Sometimes also referred to as an arc or a chain. The part of a line between two intersections with other lines is called a line or arc segment.

Line-in-polygon — a GIS operation in which line features are combined with polygon features to determine which lines fall into which polygons. Using this operation polygon attributes can be added to each corresponding record in the line attribute table (e.g., the district into which the road falls), or line attributes can be summarized for each corresponding polygon (e.g., total road length in a district).

Logical accuracy — a term used for the degree by which relationships among geographic features on a map or in a GIS database are represented correctly (e.g., adjacent to, connected to). A GIS database can be logically accurate, even if its positional accuracy is limited.

Longitude — the "x-coordinate" in a polar coordinate system on a sphere. Measured as the angular distance in degrees east or west of the Greenwich meridian.

Map — a representation of some part of the earth's surface drawn on a flat surface (e.g., paper or a computer display).

Map compilation — the process of assembling, evaluating and interpreting cartographic measure-ments and materials in order to produce a new map.

Map composition — the arrangement of map elements to create a cartographic product that is visually appealing and correctly represents the phenomena that are represented.

Map elements — Components of a thematic or topographic map such as title, legend, scale, north arrow, graticule, borders and neatlines.

Map extent — the coordinates in map units that define the rectangle that encloses all features contained in a specific map display or a GIS database; i.e., the minimum and maximum x and y coordinates in a digital database or the part of a database shown in a map display.

Map projection — a mathematical procedure for converting locations on the earth's surface into a planar coordinate system. Depending on the mathematical formulae employed, map projections have different properties. Some preserve the shape of regions on the globe, others preserve relative area, angles or distances.

Map units — the units of measurement in which coordinates in a GIS database are stored; e.g., centimeters, meters or degrees, minutes and seconds.

Meridian — a reference line that is defined by the corresponding longitude. For example the Greenwich meridian.

Metadata — data about data. A collection of information that describes the content, quality, condition, format, lineage, and any other relevant characteristic of a data set.

Multipath — the error introduced to GPS readings as a result of reflection and scattering of GPS signals on neighboring structures such as houses or trees. Multipath error is a problem mostly in high-precision surveying.

Multispectral image — a remotely sensed data set that consists of a number of bands or layers. These are essentially separate images taken at the same time for the same area, each of which shows the signal of a different range of the electromagnetic spectrum.

Nadir — in aerial photography and remote sensing, the point on the earth's surface that is located directly below a camera or sensor.

Network analysis — procedures to analyze relationships between points or addresses on a set of lines in a GIS database that may represent, for example, a street network. Network analysis is used for location decisions and routing such as emergency management.

Node — the start or end point of a line feature, or the point at which two or more lines connect.

Normalization — The conceptual procedure in database design that removes redundancy in a complex database by establishing dependencies and relationships between database entities. Normalization reduces storage requirements and avoids database inconsistencies.

Orthophoto — see digital orthophoto.

Overlay — the combination of two data layers that are in the same geographic reference system. Overlay can be done for cartographic display purposes, or the two layers can be physically combined to create a new GIS data set (e.g., polygon overlay, point-in-polygon, line-in-polygon).

Overshoot — in digitizing, a line that has been extended beyond the point where it should connect with another line. The resulting spurious line segment is sometimes called a dangle.

Panchromatic image — a remotely sensed image that records the signal in a broad range of the electromagnetic spectrum, similar to a black and white photograph.

Parcel — a single cadastral unit or land property.

Photogrammetry — the art and science of extracting measurements and other information from photographs. In the context of mapping, the procedures for gathering information about real-world features from aerial photographs or satellite images.

Pixel — from picture element. Similar to a cell in an image, grid or raster map.

Planar coordinate system — a system for determining location in which two groups of straight lines intersect at right angles and have as a point of origin a selected perpendicular intersection. See Cartesian coordinate system.

Planimetric map — a map that, in contrast to a topographic map, only shows the locations of features, but not their elevation. A planimetric map may show the same features as a topographic map with the exception of terrain or elevation contours, but will usually only show selected features chosen for a specific purpose.

Plotter — a computer peripheral that can draw a graphic file, similar to a printer, but usually for larger format output.

Point — a zero-dimensional object. An x,y coordinate that is used in a digital geographic database to represent features that are too small to be shown as lines or polygons. For example, households, wells or buildings are often shown as points.

Point-in-polygon — a GIS operation in which point features are combined with polygons to determine which points fall into which polygon. Using this operation polygon attributes can be added to each corresponding record in the point attribute table (e.g., health service area information for a survey sample point), or point attributes can be summarized for each corresponding polygon (e.g., number of hospitals in each district).

Polygon — a two-dimensional object. An area feature that is represented in a vector GIS as a sequential series of x/y coordinates. These define the lines that enclose the area; i.e., the first and last coordinate of the polygon are identical.

Polygon overlay — a GIS operation in which two polygon data layers are combined to create a new data layer. The output layer consists of the areas of intersection of both sets of input polygons. The attribute table of the new data layer contains the attributes from both input data sets. Polygon overlay is one of the fundamental GIS operations that is often used to integrate information from heterogeneous sources such as demographic and environmental data.

Positional accuracy — a term used for the degree by which positions on a map or in a GIS database are recorded correctly with respect to their true location on the earth's surface. Logical accuracy, in contrast, only pertains to correct representation of the relationships among geographic features.

Precision — the ability to distinguish between small differences in measurement. In GIS, coordinate precision is determined by the data type used to store the x and y coordinates (usually double precision, or 16 bytes for each number).

Primary key — One or more fields in an attribute table that uniquely identify a specific instance, row or record.

Protocol — a set of conventions that determine the treatment, exchange and formatting of data in an electronic communications system. Similar to a data standard but applied to procedures.

Quadrangle — a rectangular area that is bounded by pairs of meridians and parallels.

Quality control — the steps and procedures in a database development project or cartographic production system that ensure that the resulting data or output comply with specified standards of accuracy and usability.

Quantile — a statistical or cartographic classification method that assigns an equal number of objects into a fixed number of classes. Four class systems are called quartiles, five classes quintiles, ten classes percentiles. For example, the first of the four quartiles of a data distribution would contain the 25% of observations with the lowest values.

Radius — the distance from the center of a circle to its outer edge.

Raster — a geographic data model that represents information as a regular array of rows and columns, similar to a grid or image. Raster cells are usually, but not always, square. Area or line features are represented as groups of adjacent raster cells with the same value.

Rectification — the process by which an image or grid is converted from image coordinates to real-world coordinates. This usually involves rotation and scaling of grid cells, and thus requires resampling or interpolation of grid values. Similar to transformation of vector data.

Registration — the process of matching features in two maps or GIS data layers so that corresponding objects are coincident. Registration is based on a series of ground control points, and is related to transformation and rubber-sheeting.

Reference map — in the context of census mapping, a cartographic product (hardcopy or digital) that displays some portion of the census geographic framework, e.g., an data collection or statistical dissemination unit.

Relational Database Management System (RDBMS) — a database management system that allows the temporary or permanent joining of data tables based on a common field (a primary and foreign key). Each row, record or instance in a database has a fixed set of attributes or fields. Each table has a primary key that uniquely identifies each record. The table may also contain a foreign key which is identical to a primary key in an external table. A relational join is achieved by matching the values of the foreign key to the corresponding values in the primary key of the external table.

Remote sensing — the process of acquiring information about an object from the distance; i.e., without physical contact. Remote sensing usually refers to image acquisition by means of satellite sensors or aerial photography.

Resolution — a measure of the smallest detail that can be distinguished on a map or in a digital database. Resolution determines the accuracy at which the location and shape of a map feature can be accurately represented at the given map scale. In raster GIS and image data, resolution is sometimes used to refer to the cell or pixel size.

Row — in GIS, a group of cells or pixels in a grid or raster GIS database that are aligned horizontally. In database management systems, a record or instance in an attribute table.

Rubber-sheeting — a procedure in which the shape and location of objects in a GIS database are modified in a non-uniform manner. Rubber-sheeting is often used to bring a GIS data set in an unknown coordinate system into a known system. The adjustments are defined by specifying a large number of links from locations in the input data set to their corresponding correct reference or control points in the output coordinate system.

Satellite image — a digital data set that has been recorded from an earth orbiting satellite either photographically or by a scanner on-board the satellite. A satellite image in a GIS is similar to a raster or grid data set.

Scale — in cartography, the relationship between the distance on a map and the corresponding distance on the earth's surface. Scale is reported as a ratio, for example, 1:100,000, which means, for example, that one centimeter on the map equals 100,000 centimeters on the earth's surface. Since scale is a ratio, a *small scale* map shows a relatively large area, while a *large scale* map shows a small area. More generally, scale refers to the level of observation or enquiry, which may range from micro-scale to macro-scale phenomena.

Scanning — a data capture technique in which information on hardcopy documents (e.g., paper or mylar) is captured and converted into a digital image by means of a light-sensitive optical device. For map data, scanning is an alternative to data input by digitizing. After scanning a map, the image data are usually converted to vector format using a raster-to-vector conversion software or on-screen tracing of line and point features.

Schematic map — see sketch map.

Server — a computer that has been set up to provide certain services to other computers (clients), for instance, a web server is a central repository of data, software or content for the world wide web.

Sketch map — a map (often hand-drawn) that shows main features of a given area, but which may not have a high degree of positional accuracy and may thus not correctly represent distances and dimensions of objects. A sketch map may, however, have a high degree of logical accuracy, meaning that relationships between objects are correctly represented. Also called a schematic map or a cartoon map.

Source material — data and information of any type that is used to compile a map or a GIS database. This may include field observations, aerial and terrestrial photographs, satellite images, sketches, thematic, topographic, hydrographic, hypsographic maps, sketch maps and drawings, tabular information and written reports that relate to natural and human-made geographic features.

Space segment — the part of the GPS system that is located in space, i.e., the 24 GPS satellites.

Spatial analysis — the set of techniques for extracting useful information from geographically referenced data. Spatial analysis includes the integration of geographic data sets, qualitative and quantitative methods for evaluating the data, as well as modeling, interpretation and prediction. In GIS, spatial analysis often refers to the methods of GIS data integration such as polygon overlay or neighborhood analysis. In a wider sense it includes, for instance, spatial process models (e.g., migration dynamics) and spatial statistics (e.g., regression models that account for the spatial arrangements and relationships among observations).

Spatial data — information about the location, dimensions and shape of and the relationships among geographic features. In GIS, spatial data are technically classified as points, lines, areas and raster grids.

Spatial data infrastructure — see framework data.

Spatial index — a lookup table or structure within a geographic database that is used by a GIS or database management system to speed up queries, analytical operations, and display of spatial features.

Spatial interaction — interdependence among geographic entities. It often refers to the flow of goods, services, information or people between geographic locations. Spatial interaction analysis is important in the study of human migration.

Sphere — a globular body similar to a ball. The earth in its simplest approximation is a sphere, but in reality is more accurately represented as a spheroid (see ellipsoid).

Standards — in computing, a set of rules or specifications established by some authority that define, for example, accuracy requirements, data exchange formats, hardware or software systems.

Structured Query Language (SQL) — in relational database management systems, a standard syntax used to define, manipulate and extract data.

Surface — A term often used to describe GIS raster or image data that describe a continuous, smoothly varying phenomena such as elevation or temperature. Even population density is sometimes represented as a raster surface.

Symbols — in cartography, the design elements used to represent map features. Symbol types are points, lines and polygons of a certain shape. Symbolization involves the choice of graphic variables such as shape, size, color, pattern and texture.

Table — in database management systems, the set of data elements arranged in rows (records or instances) and columns (fields or items). The number of columns is usually fixed by the definition of the table structure, while the number of rows is flexible.

Tag Image File Format (TIFF) — a standard image or raster file format that can store black-and-white, grayscale or color images in compressed or uncompressed form. Scanners and other devices that create image data often provide output in TIFF format. In GIS, the Geo-TIFF format is defined as a standard TIFF image file that describes a remote sensing image, digital orthophoto or raster GIS data set. It includes an associated file with a .tfw extension that contains information about the image's geographic reference information, cell size in real-world units and other relevant information.

Template — In cartography, a standardized design of peripheral map elements (borders, neatlines, north arrows) that can be used for a standardized map series. In database management systems, an empty table created for multiple purposes for which only the fields or items have been defined.

Thematic layer — see layer.

Thematic map — a map that presents a specific concept, subject or topic. A thematic map can show quantitative or qualitative information.

Theme — in GIS, a set of geographic objects that usually belong to the same subject group (e.g., roads or settlements) and that are stored in the same GIS database.

Tile — in GIS, a term sometimes used to refer to adjacent digital map sheets that are stored in separate files. Tiles can be of regular shape (e.g., square or rectangular) or they can follow irregular boundaries such as district or province borders. Storing all tiles in the same geographic reference system allows temporary or permanent joining of adjacent tiles.

Topographic map — a map of mostly real-world features including elevation contours, rivers, roads, settlements, and landmarks. The standard map sheets created by a national mapping agencies at various scales are typically topographic maps.

Topology — in GIS, a term that refers to the spatial relationships among geographic features (e.g., points, lines, nodes and polygons). A topologically structured database stores not only individual features but also how those features relate to other features of the same or different feature class. For example, in addition to a set of lines representing a road network, the system will store the nodes that define road intersections which allows the system to determine routes along several road segments. Or, instead of storing polygons as closed loops, where the boundaries between neighboring polygons would be stored twice, a topologically structured GIS would store each line only once together with information on which polygon is located to the left and the right of the line. This avoids redundancy and facilitates the implementation of many GIS and spatial analysis functions.

Transformation — the conversion of digital spatial data from one coordinate system to another through translation, rotation, and scaling. Transformation is used to convert digitized digital map data from digitizer units (e.g., centimeters or inches) into the real-world units corresponding to the source map's map projection and coordinate system (e.g., meters or feet). See also georeferencing.

Undershoot — in digitizing, a line that has not been extended all the way to the point where it should connect with another line.

Universal Transverse Mercator (UTM) — a cylindridcal map projection that is often used for large scale (i.e., local) mapping.

User segment — the portion of the GPS system that includes all types of receivers of GPS signals.

Vector data — a GIS data model in which the location and shape of objects is represented by points, lines and areas that are fundamentally made up of x,y coordinates.

Vector product format (VPF) — a vector GIS format developed by the U.S. National Map and Imagery Agency (formerly Defense Mapping Agency) intended to become a universally accepted vector data exchange format.

Vertex — one of a series of x,y coordinate that defines a line. The first and last vertices of a line are usually called nodes.

Additional glossaries and dictionaries can be found in Padmanabhan et al. (1992), ASCE (1994), McDonnel and Kemp (1995), and Dent (1998). Online resources include the following:

Canada Centre for Remote Sensing http://www.ccrs.nrcan.gc.ca/

Geographer's Craft Project (University of Texas) http://www.colorado.edu/geography/gcraft/contents.html

GPS World magazine http://www.gpsworld.com/gpsworld/static/staticHtml.jsp?id=8000&searchString=glossary

Perry-Castañeda Library, University of Texaswww.lib.utexas.edu/Libs/PCL/Map_collection/glossary.html

U.S. Census Bureau www.census.gov/dmd/www/glossary.html

U.S. Geological Survey http://interactive2.usgs.gov/learningweb/explorer/geoglossary.htm

Useful Addresses and URLs

GIS packages			
Autodesk Inc.	San Rafael, Calif., CA	AutoCAD	www.autodesk.com
Bentley Systems Inc.	Huntsville, AL	MicroStation	www.bentley.com
ESRI, Inc.	Redlands, CA	ArcGIS, ArcInfo, ArcView, ArcExplorer, Atlas GIS	www.esri.com
Intergraph	Huntsville, AL	GeoMedia	www.intergraph.com
MapInfo Corp.	Troy, NY	MapInfo GIS	
Microsoft Corp.	Redmond, WA	MapPoint	www.microsoft.com
Oracle Corp.	Redwood Shores, CA	Oracle Spatial	www.oracle.com
UNSD Software Project	New York, NY	PopMap	www.un.org/Depts/unsd/softproj/index.htm
Siemens	Munich, Germany	SICAD Spatial Desktop	www.siemens.com
Smallworld Systems Inc.	Englewood, CO		
PCI Geomatics Group	Richmond Hill, Ontario, Canada	SPANS and PAMAP	www.pci.on.ca
ThinkSpace Inc.	London, Ontario, Canada	MFWorks	www.thinkspace.com
Vision* Solutions	Ottawa, Ontario, Canada	Vision*	
Specialty software			
Blue Marble Geographics	Gardiner, ME	Coordinate management and GIS development tools	www.bluemarblegeo.com
Caliper Corp.	Newton MA	Maptitude, GIS+, TransCAD	www.caliper.com
Core Software Technology)	Pasadena, CA	TerraSoar (distributed geospatial databases), ImageNet (online geo-spatial data distribution)	www.coresw.com
Quantum GIS		Open source software	(http://qgis.org)
Thuban		Open source software	http://thuban.intevation.org
Open EV		Open source software	http://openev.sourceforge.net
Remote sensing image processing systems			
Leica GeoSystems/Erdas	Atlanta, GA	ERDAS Imagine	www.erdas.com
Earth Resource Mapping	San Diego, CA	ER Mapper	www.ermapper.com
Clark Labs	Worcester, MA		
MicroImages Inc.	Lincoln, NE	TNTmips	www.microimages.com
PCI Geomatics Group	Richmond Hill, Ontario,	EASI/PACE,	www.pci.on.ca

	Canada	OrthoEngine		
Research Systems Inc	Boulder, CO	ENVI visualization software	www.rsinc.com	
High resolution satellite imagery and digital orthophotography				
GeoEye	Thornton, CO	Carterra and Ikonos satellites	www.spaceimaging.com	
Digital Globe	Longmont, CO	QuickBird and EarlyBird satellites	www.digitalglobe.com	
Orbital Imaging Corp.				Dulles, VA
EROS Data Center	Sioux Falls, SD			
Spot Image		Spot satellites	www.spot.com	
Maps Geosystems	Munich, Germany	Aerial surveys (Africa, Middle East)	www.maps-geosystems.com	
EarthSat	Rockville, MD	Satellite and mapping services	www.earthsat.com	
Global Positioning Systems				
Magellan Corp.	Santa Clara, CA		www.magellangps.com	
Ashtech	Santa Clara, CA		www.ashtech.com	
NovAtel Inc.				Calgary, Alberta, Ca
Sokkia Corp.	Overland Park, KA		www.sokkia.com	
Trimble Navigation Ltd.	Sunnyvale, CA		www.trimble.com	
Garmin				
Journals				
GeoWorld, GeoAsia, GeoEurope, GeoInformation Africa, Mapping Awareness, Business Geographics	GeoWorld, Fort Collins, CO		www.geoplace.com	
GPS World			www.gpsworld.com	
International Journal of Geographical Information Science	Taylor & Francis, London, UK			
GeoInfosystems	Advanstar Pub., Eugene, OR			
Journal of the Urban and Regional Information Systems Association	URISA, Park Ridge, IL		http://www.urisa.org/	
Miscellaneous				
National Center for Geographic Information	Santa Barbara, CA	GIS research center	www.ncgia.ucsb.edu	

and Analysis			
International Institute for Aerospace Survey and Earth Sciences (ITC)	Enschede, Netherlands	GIS Training Courses	http://www.itc.nl/
European Umbrella Organization for Geographic Information (EUROGI)	Netherlands		www.eurogi.org
U.S. Federal Geographic Data Committee	Reston, VA		www.fgdc.gov
Permanent Committee on GIS Infrastructure for Asia & the Pacific			www.permcom.apgis.gov.au/

Note: Many of these glossary terms and URLS overlap across different manuals and guides. The glossary of terms provided above are largely a selected condensed version of census geography terms provided in the 2010 UN Handbook on Geospatial Infrastructure for Census Activities.

References

United Nations Statistics Division. 2000. Studies in Methods. Series F No. 79. Handbook on geographic information systems and digital mapping. United Nations. New York.

United Nations. 2010. United Nations Handbook on Geospatial Infrastructure in Support of census Activities. F, No. 103.

United Nations. 2000. United Nations Handbook on Digital Mapping and GIS. United Nations Statistics Division. F, No. 79.

U.S. Census Bureau. 1997. Mapping for Censuses and Surveys, United States Bureau of the Census, International Statistical Programs Center. Agency for International Development.

UN Economic Commission for Europe. 2009. Making Data Meaningful. Part 1: A guide to writing stories about numbers ECE/CES/STAT/NONE/2009/4; Part 2: A guide to presenting statistics. ECE/CES/STAT/NONE/2009/3.

Peters, A. and MacDonald, H., 2005. Unlocking the Census with GIS. ESRI Press.

Sommer S. and Wade T., 2006. A to Z GIS: An Illustrated Dictionary of Geographic Information Systems. ESRI Press.

Rogerson, P., 2010. Statistical Methods for Geography. Sage Publications Ltd; Third Edition.

Mitchell, A., 2005. The ESRI Guide to GIS Analysis: Volume 2: Spatial Measurements and Statistics.