

**TECHNICAL GUIDES FOR
OPERATIONALIZING HOUSEHOLD FOOD
SECURITY IN DEVELOPMENT PROJECTS**

Food Consumption and Nutrition Division

March 1999



**International Food Policy Research Institute
2033 K Street, N.W.
Washington, D.C. 20006 U.S.A.**

FOREWORD

Development practitioners often face an "information constraint" when it comes to designing, implementing, monitoring, and evaluating food security and nutrition projects. That is, information is lacking on the nature of food security and nutrition problems facing a country, region, or locality, the location of food insecure areas, and the causal links between potential interventions and outcomes of interest.

These 11 guides aim to assist in overcoming this constraint. The principal audience is an operational one: multilateral or bilateral aid agencies, nongovernmental organizations (NGOs), developing country governments, and other development practitioners who are actively engaged in food security and nutrition issues. The guides provide a framework for thinking about what projects would be most appropriate in a given situation and indicate what types of information are needed in order to maximize impact on food security. They can also assist by making development practitioners fully conversant with food security and nutrition concepts.

These guides are neither meant to be exhaustive surveys of all issues or methods with respect to food security. Rather, they are designed to provide a suite of useful methods at different points in the project cycle. Although each document has been prepared as a "stand-alone" product, it is recommended that interested readers begin with Guide #1, which provides an overview of the issues discussed in these guides.

By showing how information constraints can be eased, it is hoped that these guides will assist in improving the targeting of interventions, in designing interventions that will have a maximum impact on food insecurity, and in developing methods for monitoring and evaluation.

CONTENTS

Technical Guide #1: *Operationalizing Household Food Security in Development Projects: An Introduction*, John Hoddinott

- 1.1 Introduction 1-1
- 1.2 The Links between Development Interventions, Household Food Security, and Nutrition 1-2
- 1.3 Summary of Technical Guides 1-11

Technical Guide #2: *Food and Nutrition Security Data on the World Wide Web*, Alison T. Slack

- 2.1 Introduction 2-1
- 2.2 Indicators of Development and Food Security: Process Versus Outcome 2-3
- 2.3 Vulnerability Analysis Mapping (VAM) 2-4
- 2.4 Africa Data Dissemination Service (ADDS) 2-8
- 2.5 Demographic and Health Surveys (DHS): Data and Case Study Example 2-14

Technical Guide #3: *Classification and Regression Trees: An Introduction*, Yisehac Yohannes, John Hoddinott

- 3.1 Introduction 3-1
- 3.2 A CART Example: Indicators of Household Food Insecurity in Northern Mali 3-2
- 3.3 Technical Details: Computing Requirements and Programming 3-6
- 3.4 Further Applications, Strengths, and Weaknesses of CART 3-9

**FOOD AND NUTRITION SECURITY DATA ON
THE WORLD WIDE WEB**

Alison T. Slack



International Food Policy Research Institute
2033 K Street, N.W.
IFPRI Washington, D.C. 20006 U.S.A.

March, 1999

1. INTRODUCTION¹

This guide provides the reader with a discussion of the types of data concerning food security indicators that are available via the Internet for individual developing countries on a disaggregated level, and where on the Internet they can be obtained. It provides this information in the context of discussion on "process" versus "outcome indicators" for use in the identification of food insecure areas and vulnerable populations within countries.

Based on our experience in developing this guide, once having registered and obtained the data from the Internet sources, an analyst with only a very basic knowledge of the Internet, familiarity with a standard spreadsheet or statistical software package, and access to sufficient disk space, would be able to indicate, in less than three days of desk-based research, which areas of a particular country could be targeted for various projects.

The guide begins, in this section, with an overview of the Internet and World Wide Web, how to search the Internet for data available, and how existing subnational level data that is available can be useful for various development projects in terms of targeting and evaluating various food security needs. The second section provides a discussion of process versus outcome indicators as they relate to development and food security. This is followed by sections describing three World Wide Web sites that provide various types of data for use in food security studies: the Vulnerability Analysis Mapping (VAM) project, the Africa Data Dissemination Service (ADDS), and the Demographic and Health Surveys (DHS). For each of these sites, a list is provided of both process and outcome variables available within the sites, the countries for which data are available, and where the data can be located. The last of these three case study sections, DHS, provides, as an example, a detailed account of the steps needed to obtain, download, and transform the data available, and create files ready for analysis at the subnational level. These steps are illustrated by showing how data from a particular country, in this case Mali, can be used to identify which areas of the country should be targeted for various types of

¹ Funding for data collection and analysis of these data has been supported by the International Fund for Agricultural Development (TA Grant No. 301-IFPRI). We gratefully acknowledge this funding, but stress that ideas and opinions presented here are our responsibility and should, in no way, be attributed to IFAD.

development projects. Following the conclusions section, an appendix provides the names and Internet addresses of other sources of subnational-level data sets for developing countries.

The Internet and the World Wide Web

The Internet is essentially an assembly or network of computers all connected to each other via telephone lines. Through this connection, and by using a common computer language to communicate, a wealth of information can be shared by these computers, and virtually any others that join the network.

The Internet is made up of several parts. The most commonly known (aside from e-mail), and probably the fastest growing part, is the World Wide Web (WWW), or “the Web.” The Web uses a computer language called “hypertext” that enables different areas or “sites” of information to become linked together by a simple click of a mouse. It is this linking of sites that is so critical in helping the Internet user search for information—allowing one to jump from one part of a Web site to another, or from one Web site to a different one altogether.

Searching for Information Using the Web

The most straightforward way to get onto the Web is through a browser. A browser is a software program that, once installed onto a computer, allows the Internet user to search or “surf” the Web by selecting from various categories of information presented by the browser. Two browsers were used in developing this guide, Netscape Navigator and Internet Explorer, although many different browsers are available on the market. The browsers also provide a selection of “search engines” (such as Alta Vista, Infoseek, Lycos, etc.) where the user can enter key words, and receive a list of Web sites that have information related to the key words.²

Key words or terms such as “food security” or “regional data” will bring up a list of sites that contain these words (starting with sites that use the words most often and that have the “best fit” for the search terms used). Selections of sites are usually made by clicking, with the mouse,

² Each site, or Web page, on the Internet has an “address” that identifies it from others. These addresses can be typed in and pulled up directly from the browser (and then “bookmarked” for quick access in the future). For example, IFPRI’s Web site, or home page is <http://www.cgiar.org/ifpri/>.

on the links provided on the list, usually in the form of icons or text that has been highlighted and/or underlined. These links will pull up the designated Web site. Once the Web sites are pulled up, searches can often be done within those sites as well. The names of institutions can also be used as key words in searches.

Data on the Internet for Use with Food Security Projects

At an early stage in a development project cycle, there is a need to obtain secondary data at the subnational level on household food security for both targeting and for needs assessment. This is true, for example, in drawing up a country background study, in developing project pre-planning papers, and in the formulation of projects. The Maxwell and Frankenberger (1992) publication provides guidance on the many food security indicators that could be used in this context. They suggest that obtaining many of these data is difficult and time consuming. In recent years, however, the Internet, and specifically the World Wide Web, have become important sources of information on a wide range of issues relating to food security and development, including data for individual countries on a national and subnational level.

The attractions of using this type of data found on the Internet are two fold: (1) these sources appear to provide accurate, detailed, and up-to-date data on a wide number of countries and their regions, states, or other subnational-level entities; and (2) the financial and time costs associated with obtaining this information are minimal. Those interested in obtaining these data need only log onto the Web sites providing the data, and select and download data sets and maps of interest, although some sources require that the user register with the provider in order to obtain the data.

2. INDICATORS OF DEVELOPMENT AND FOOD SECURITY: PROCESS VERSUS OUTCOME

In Simon Maxwell and Timothy Frankenberger's report, *Household Food Security: Concepts, Indicators, Measurements* (1992), indicators that are used for assessing household food security can, and should, be divided into two general categories: process indicators and outcome indicators. Process indicators are those that reflect an area's food situation with respect

to both supply and access, whereas outcome indicators can be used as proxies for food consumption. Food supply-related indicators include those related to agricultural production, agroclimatic data, access to natural resources, institutional development and market infrastructure, and exposure to regional conflict and social unrest. Access indicators include the individual household strategies to overcome food deficits and to meet food requirement needs, such as coping ability indicators. Agricultural strategies (risk management), rationing, and even migration would be included in this category.

Outcome indicators can also be subdivided into two types: direct and indirect. Direct indicators include household budget and consumption surveys, food frequency assessments, and other direct measures of intake and consumption. Indirect indicators would include "second best" estimates (based on cost and time limitations) such as storage estimates, subsistence potential ratios, and nutritional status assessments.

The next three sections describe important Internet sites that include data that can be used for examining food security issues on a subnational level. The discussion includes a list of countries and the types of indicators that are available in each site, dividing them into process and outcome indicators.

3. VULNERABILITY ANALYSIS MAPPING (VAM)

In 1995, The World Food Programme (WFP) created the Vulnerability Analysis and Mapping (VAM) project to provide assistance to its country offices on two fronts: (1) "analyzing the vulnerability of target populations to food insecurity and their capacities for coping with disaster," and (2) "presenting the analysis with digital maps" (from Web site). The inputs that VAM uses are geared towards aiding in long-term strategic and contingency planning, as well as towards improving both the identification and design of various projects. In addition, VAM assists in guiding decisions relating to programming and operational concerns. The WFP/VAM unit works in close cooperation with both the Food and Agriculture Organization of the United Nation's (FAO) Global Information Early Warning System (GIEWS) and the U.S. Agency for International Development's (USAID) Famine Early Warning System (FEWS). The data that are

available include indicators that fall into both process and outcome indicator categories. The list of indicators available for each country vary somewhat.

Process Indicators

Process indicators offered in the VAM Web site include the following.

- General: areas afflicted by poverty, food insecurity, and floods; emergency and development figures; population by district and population flows; refugees
- Agricultural: average production of certain crops; diversity index (crops); land use; crop use intensity; hectares to drought resistant crops
- Meteorological: drought resistance; rainfall (recent and monthly); flood risk
- Food: food systems; stock positions (for WFP food deliveries)
- Infrastructure: market access; distance to major roads and bad roads; density and maximum hectares per person; rails, ports, and airports; overland, sea, air corridors

Outcome Indicators

The outcome indicators available in this site include the following.

- Health care: persons per clinic (unit); vaccination rates; persons per bed; at risk pregnancy rates
- Illness: diarrhea cases (under fives); malnutrition cases; malaria cases; stunting rates
- Sanitation: no water access
- Education/Literacy: female and overall literacy rates; no education rates/attendance at secondary schools

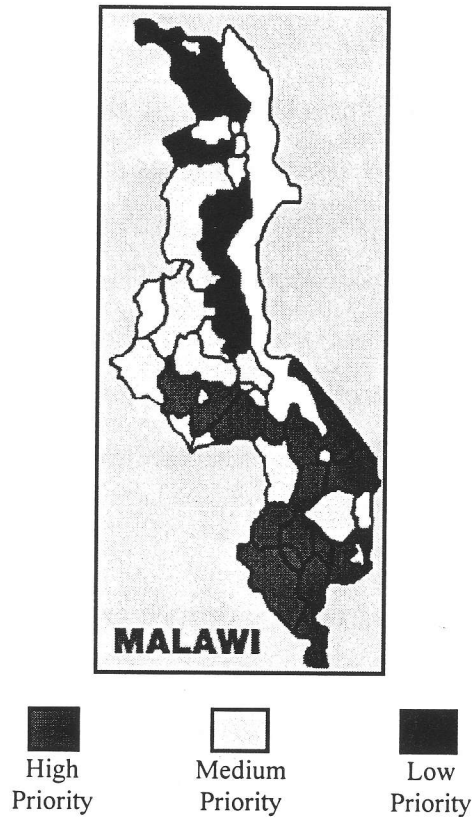
The regions and countries for which the above data are available include the following.

- Africa : African continent; Greater Horn; Great Lakes Region (Rwanda, Burundi, Tanzania, Uganda, and Zaire); Southern Africa (Lesotho, Malawi, Mozambique, Swaziland, Zambia, and Zimbabwe); Chad; Ethiopia; Liberia; Malawi; Senegal; Tanzania; Zambia; Zaire
- Eastern Europe : The Caucasus (Armenia, Azerbaijan, and Georgia); Former Yugoslavia Region
- Asia: Cambodia; Democratic People's Republic of Korea; Pakistan

This data provided by VAM can be found from the WFP World Wide Web home page, <http://www.wfp.org> (available by clicking on the "**Disaster Mitigation**" icon for VAM), or can be accessed directly from the following address: <http://www.wfp.it/vam/vamhome.htm>. Once on the home page of VAM, the user will find a map (whose countries can be selected by clicking on them with a mouse), along with a list of options to review. The user can get information on individual countries either by selecting a country from the map image (with the mouse) or by going into the selection called "**Search the VAM data base**" under the "**Content**" heading, where countries and geographical regions can be selected by typing the names into boxes provided. Information is available in the form of maps, data sets, and reports, and once a country is selected, the user may choose to view a map or download the data used to produce the maps. For example, if a project controller were interested in finding out the overall vulnerability levels within Malawi, he or she could select the country with the mouse from the outline on the map image on the VAM home page, which would bring up the web site for "**Malawi**." The controller is then presented with several options: "**Reports**," "**Maps and Data Holdings**," and year of information within these categories (1995 or 1996).

When the option "**1995 Malawi Vulnerability Analysis**" under "**Reports**" is selected, a regional map of Malawi is presented, illustrating, with a range of colors, the "Areas Identified for Targeting" ranging from high to medium to low priority, along with a summary description of the data used (see Figure 1.) Under "**Maps and Data Holdings**," the "**1995 Maps and Data**" link provides access to maps and data sets for various indicators, including average maize produced,

Figure 1—Areas identified for targeting (Malawi Summary Vulnerability Analysis Working Progress)



female literacy, malaria cases, persons per clinic, etc. The data sets themselves can be downloaded by selecting "**Download**" and clicking on the "**Save File**" option to transfer the data to a local directory. These files are downloaded as "zipped," meaning that they have been compressed to save space, and that they must be "unzipped" to be pulled up in the appropriate software. *PKzip* is a software package that can be used to do this transformation.

Once the files have been unzipped, the data files can be pulled up into the appropriate software, or by using software to transform the files into the desired format. For example, the stunting data set for Malawi includes several individual files. One file is called **HIS9192.dbf**. By using a transferring software package such as DBMS Copy, one can copy the data set from

the .dbf form to another form, such as a spreadsheet package like Lotus or Excel, thereby gaining access to a working data set with information on weights and heights on a subnational level.³

The link to the 1996 information for Malawi provides information under the heading "**A Quest for Causality**" under "**Reports.**" From this link, a controller can review discussions on the overview of the project in Malawi, the data sources, and the methodology for both cross sectional/snapshot and time series information (located on the left-hand portion of the screen). In addition, under the "**Time Series VAM**" selection, the user can click on "**Tables/Maps**" to obtain information on more aggregate indicators. The selections here include maps for poverty index, food deficit, and composite vulnerability index. Figure 2 shows a print out of the Composite Vulnerability Index Map for Malawi, illustrated with shading for each region, ranging from *least* to *most* on the vulnerability scale.

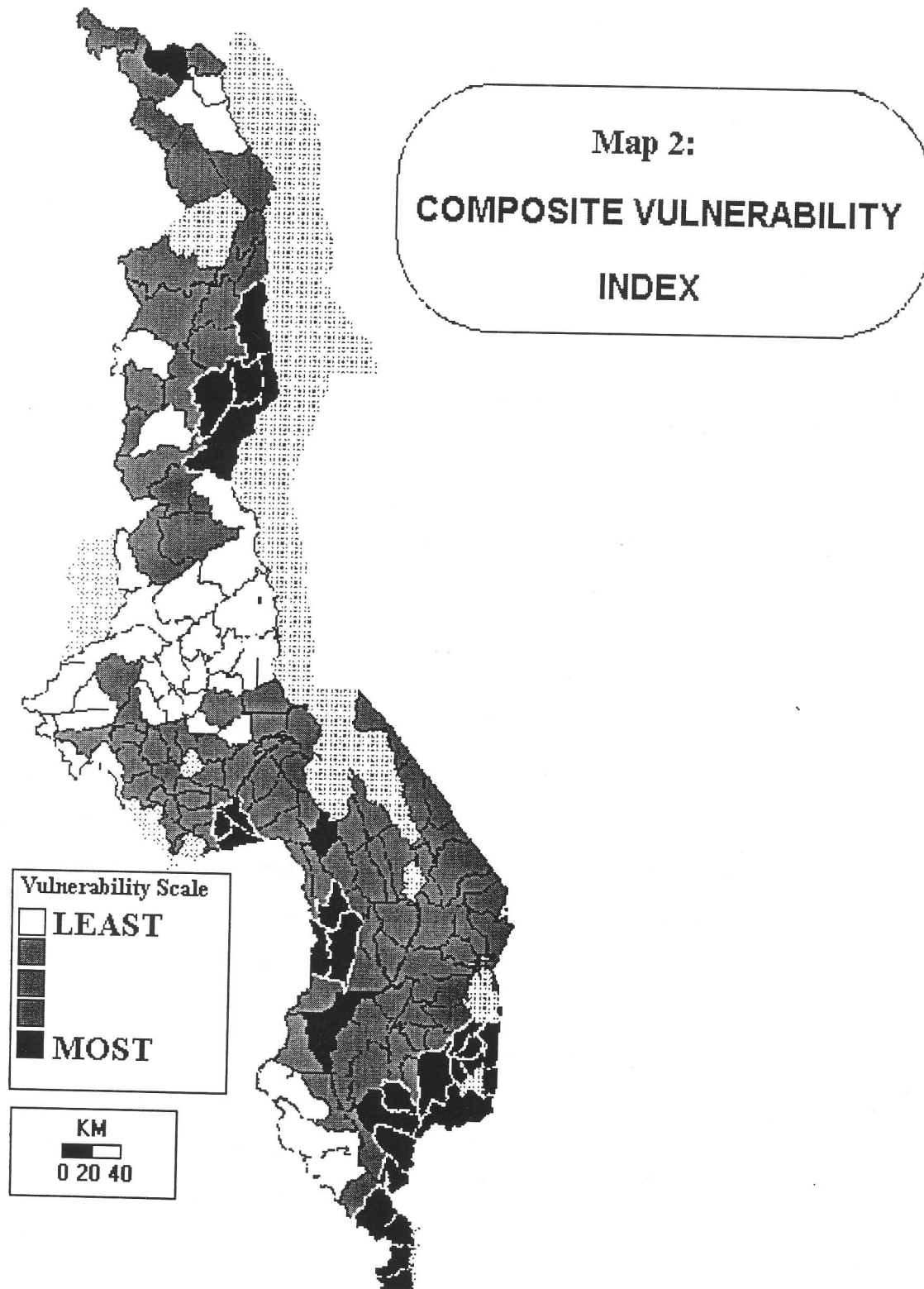
4. AFRICA DATA DISSEMINATION SERVICE (ADDS)

The United States Agency for International Development (USAID) operates the Famine Early Warning System (FEWS), which is a "multidisciplinary project that provides subnational information to decision makers about potential famine situations, allowing them to authorize timely initiatives to prevent famine outbreaks" (from the FEWS Web site). FEWS offers links to other sites such as the "Climate Prediction Center," with data and maps on meteorologic analysis, cloud top graphics, and precipitation estimates on a global level.

ADDS (along with FEWS: Greater Horn of Africa), which can be accessed from the FEWS home page, is a project under the umbrella of FEWS where the focus is on regions and individual countries within Africa. Maps and data files are available, providing information on various indicators relating to weather, agriculture, infrastructure, and vulnerability to food insecurity. This information is available by geographic area, by region or country, or by data theme. Information is also available for each country on the administrative boundaries, other

³ In order to obtain a detailed description of the variables, etc., it appears that the user must contact VAM personnel.

Figure 2—Composite vulnerability index map for Malawi



Map produced by FEWS Malawi for 1996 VAM

related Web sites, and data providers. Finally, there is up-to-date information on the effects and predictions related to the El Niño events of 1982-83 and 1997-98.

Process Indicators

Most of the indicators available in ADDS and FEWS fall into the process indicators category, and the list of indicators available for each country vary somewhat. Below is a list of these indicators.

- Agricultural: general agricultural statistics; cropland use intensity;
- Meteorological: agroclimatic zones; digital elevation model; hydrology; rain stations; vegetation; rainfall
- Infrastructure: railroads and roads
- Markets: prices
- Demography: population clusters (WASAP)

Outcome Indicators

In the FEWS link to the Greater Horn of Africa, maps and data are provided for indicators of "chronic vulnerability to food insecurity." These indicators represent "a long-term condition that predisposes a particular population group or a region to food insecurity" (from FEWS Web site). A vulnerability analysis was done for Kenya in 1995, and a map displaying the range of vulnerability levels on a subnational level is available for downloading. The regions and countries within Africa for which the above data are available include the following.

- Regions: African continent; Sahel West Africa; Central Africa; Greater Horn of Africa; Southern Africa
- Countries: Algeria; Angola; Benin; Botswana; Burkina Faso; Burundi; Cameroon; Cape Verde; Central African Republic; Chad; Comoros; Congo; Côte d'Ivoire; Djibouti; Egypt; Equatorial Guinea; Eritrea; Ethiopia; Gabon; The Gambia; Ghana; Guinea;

Guinea-Bissau; Kenya; Lesotho; Liberia; Libya; Madagascar; Malawi; Mali; Mauritania; Mauritius; Morocco; Mozambique; Namibia; Niger; Nigeria; Rwanda; Sao Tome and Principe; Senegal; Seychelles; Sierra Leone; Somalia; South Africa; Sudan; Swaziland; Tanzania; Togo; Tunisia; Uganda; Western Sahara; Zaire; Zambia; Zimbabwe

The indicators available for these countries and regions, presented both in data files and maps, can be located and downloaded by going through the FEWS home page, <http://www.info.usaid.gov/fews/fews.html> (and selecting either “USGS Africa Data Dissemination Service” in the center of the page under “Other FEWS Information Sources,” or “Africa Data Dissemination Service” toward the bottom of the page). The user may also go to the ADDS Web site directly at <http://edcintl.cr.usgs.gov/adds/adds.html>. The information available from this home page Web site for ADDS is presented in three basic categories: (1) “Overview”: overview of the project; (2) “Data Query by...”: data presented by geographic areas or data themes; and (3) “General”: general information. Under the last category, information is available that compares the extent of natural vegetation coverage for regions in Africa between the periods 1982-83 and 1997-98—reflecting two major El Niño events.⁴ Vegetation coverage, according to Maxwell and Frankenberger (1992), can be used as a process indicator for food supply, and hence be used for potential targeting of projects designed to improve food security. For an image of vegetation coverage of Southern Africa both in 1983 and 1998, the user can go to the following ADDS site and pull up this image (both a month and *dekad*—a 10-day period—will need to be selected)⁵:

<http://edcintl.cr.usgs.gov/adds/elnino/elnino.htm>.

⁴ This page also gives a summary description of the images and how to maneuver around the various web pages in the ADDM site.

⁵ The month/*dekad* menu bar can be pulled up higher on the screen for easier viewing, if necessary, by placing the mouse cursor on the top horizontal line of the menu bar and lifting the image upward.

These maps can be printed or downloaded by selecting the PC or UNIX buttons for either of the two time periods. The files can then be saved on the users local directory and pulled up, using various imaging software packages.

Data on individual countries and regions with respect to specific indicators can be obtained from the second category, "**Data Query by...**". From here, the user can select "**Region/Country**," and then click on a particular country for maps and data. For example, if a researcher were interested in how intensively land is being used for agricultural production from one region of Zambia to the next, he or she would select "**Zambia**," and then under "**Digital Map Data**," select "**Cropland Use Intensity**." This selection brings up a choice of viewing the map or reading about the data and the files that produce the maps. Figure 3 shows the map that is generated from the "**View GIF File**" selection for Zambia. From this map, one can see the degree to which land in each of the country's regions is being used for agriculture.

Also under the "**Data Query by...**" heading is "**Data Theme**." From this link, the user can select from among several choices of indicators, including hydrology, vegetation, agroclimatic zones, and West African Spacial Analysis Project (WASAP). WASAP is "a cooperative effort by Macro International, the World Resources Institute (WRI), and the U.S. Bureau of the Census (BUCEN) to develop a prototype for using geographic information system (GIS) technology to integrate diverse socioeconomic data sets and to facilitate spatial analysis of those data" (from the WASAP site). For example, if Mali is chosen from the WASAP link, the user can opt to view a GIF file that produces a map of population clusters for the country superimposed on the regional boundaries (see Figure 4).

Also under the "**Data themes**" heading are several selections for indicators available in data sets that are not available for viewing on the Web site, but can be downloaded. These include agricultural statistics, administrative boundary files related to agricultural statistics, precipitation, and price. Data on agricultural statistics are available for Mali on both a regional and *Cercle* level (the next administrative level down from region in Mali). This information must be downloaded to a PC and unzipped before it can be viewed.

Figure 3—Cropland use intensity in Zambia

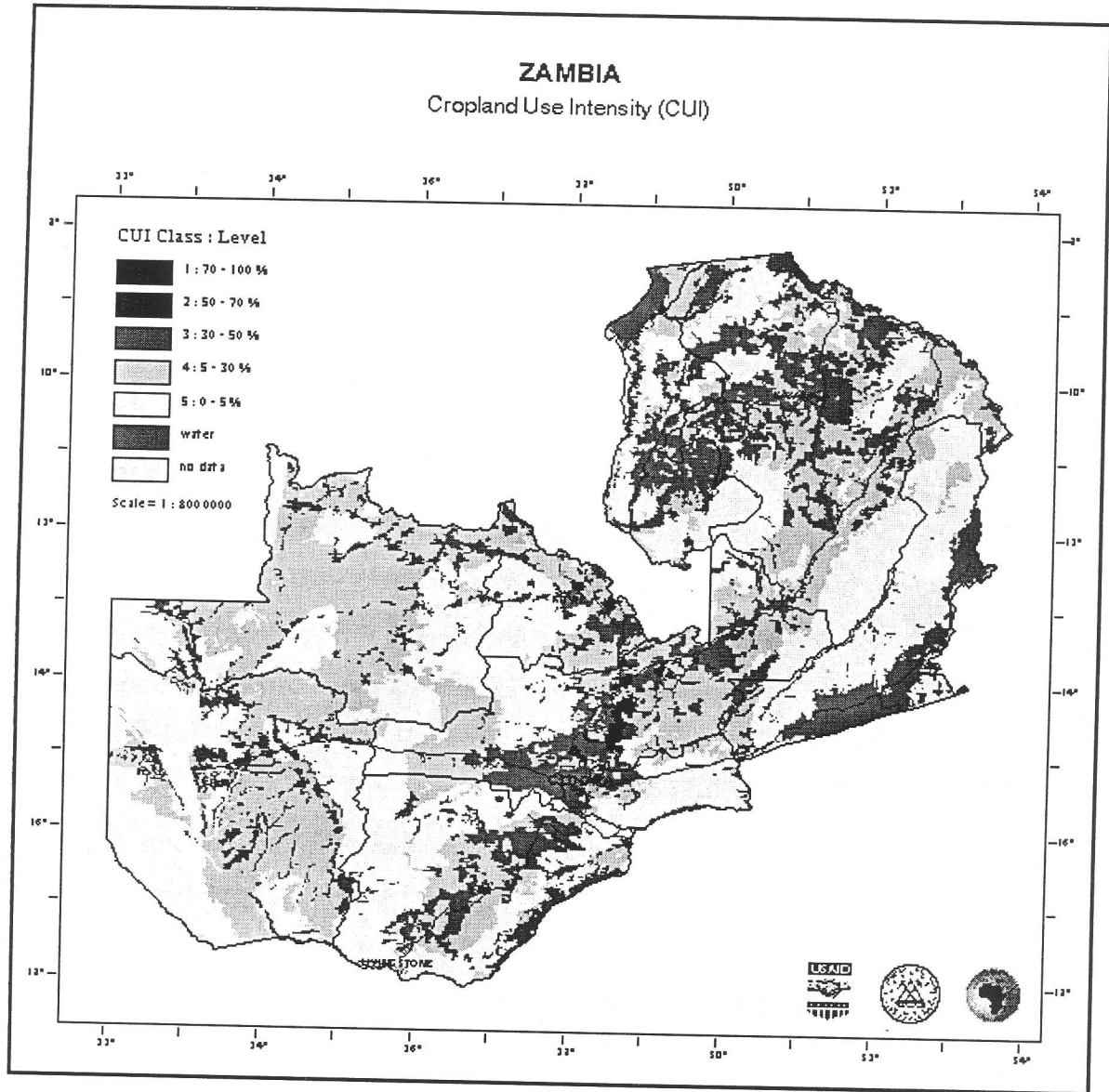
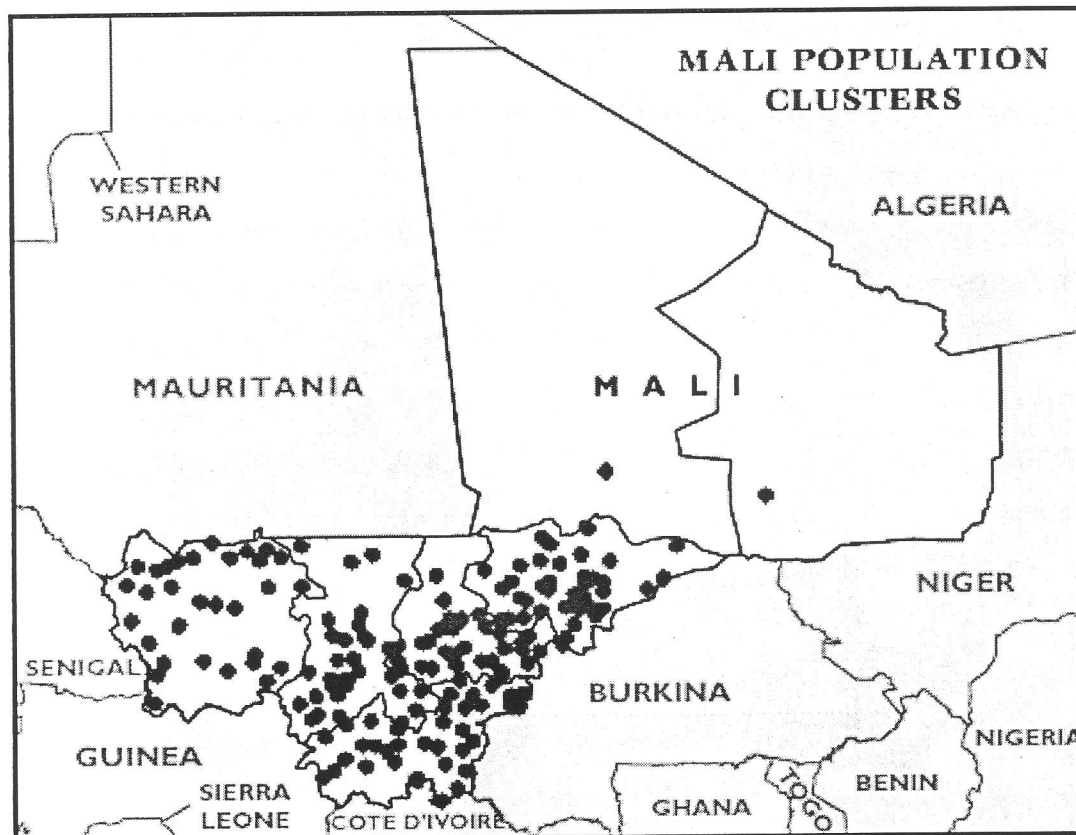


Figure 4—Mali population clusters



5. DEMOGRAPHIC AND HEALTH SURVEYS (DHS): DATA AND CASE STUDY EXAMPLE

The Demographic and Health Survey (DHS) program is a 14-year projected funded by the United States Agency for International Development (USAID). Macro International, Inc., a private U.S.-based company (but with offices worldwide) is responsible for overseeing the collection, processing, and dissemination of DHS data.

Each survey is designed to be nationally representative. The core data collection instrument is a questionnaire typically administered to adult women aged 15 to 49. Sample sizes range from 4,000 to 28,000, the exact size varying substantially from country to country. A

second questionnaire is administered to the husbands of these women. In some countries, additional household- and community-level data are collected. Data are collected principally on health and demographic topics. Data collected that are relevant to food security and poverty targeting fall primarily into the category of outcome indicators and include the following.

- child anthropometric status (height and weight)
- maternal anthropometric status (height and weight)
- male and female education levels
- quality of housing
- holdings of consumer durables
- work status
- access to safe water and sanitation
- community characteristics such as presence of schools, access to health facilities, markets, telephones, and post offices.

Some of the community characteristics here would be used as process indicators, as well, since schools and access to markets would be categorized under institutional development and market infrastructure. Technical Guide #5 provides an introduction to the measures of anthropometric status available in the DHS data set.

Tables 1, 2, and 3 indicate which DHS surveys are available on the World Wide Web. Note that for some countries, more than one survey has been conducted.

Accessing DHS Survey Data Available on the Internet

The primary data obtained from the surveys listed in Tables 1 to 3 are available by accessing the DHS home page, registering with DHS by e-mail, and selecting and downloading the data sets.

Table 2—DHS surveys, SubSaharan Africa

Country	Survey Date												
	1988	1987	1986	1985	1984	1983	1982	1981	1980	1979	1978	1977	1976
Benin													x
Botswana													
Burkina Faso	x						x						
Burundi													
Cameroon								x					
Central African Rep.										x			
Chad													x
Comoros													x
Congo													
Côte d'Ivoire	x												
Eriteria													x
Ghana													
Kenya	x												
Liberia													
Madagascar													
Malawi													x

(continued)

Table 2 (continued)

	1988	1987	1986	1985
Mali				
Morocco				
Mozambique				
Namibia				
Niger				
Nigeria				
Rwanda				
Senegal				
South Africa				
Sudan				
Tanzania				
Togo				
Uganda				
Zambia				
Zimbabwe				

- The first step is to go to the DHS Web site or home page. The WWW address for the DHS home page is <http://www.macoint.com/DHS/>.
- Once "logged on" to the DHS home page, the user is presented with a menu with options that include links to "**about DHS**," "**data collection**," "**data sets/FAQs**," "**statistics**," "**publications**," and "**news**." For obtaining data sets containing subnational-level data, the two data options are the most useful. The user can also get results from some countries in tabular form from the "**statistics**." From this selection, national-level indicators can be compared across countries.
 - **Data collection.** This link provides the user with a brief statement about the surveys and data collection. There is also a link called "**Survey Search**" to identify the countries that have been surveyed by DHS, and to pull up detailed information about any of these surveys. Information is available here for sample size, starting and ending dates for data collection, number of respondents, and special survey modules, as well as survey content (anthropometry, mortality, education, rate of AIDS, etc.). The user may select a region, which brings up a list of countries where surveys have been conducted in that region, or select individual countries.
 - **Data sets/FAQs.** After deciding which of the data sets are of interest, the user may go into this link to browse further the data sets available, register with DHS (which allows eventual authorization for obtaining data sets), and select individual data sets to download.
 - **DHS Data Archive.** This option allows the user to become familiar with (1) the contents of the data archive, (2) the contents of the questionnaires that are used in the country surveys, (3) the different forms of data sets, (4) the different media types, and (5) how to gain access to the data.
 - **Direct Data File FTP.** This option allows the user to (1) browse the data from a map of world regions and then countries within each region, (2) register to become a user, and (3) continue on to download data as a registered user.

- **Browsing the data sets.** Here one can select, from a map image, a region of the world, and then an individual country, which will generate a list of data files—all in zipped form—that are available for that particular country.
- **Data file help page.** This selection provides information on the file naming conventions (country codes, data file types, file version numbers, and the file formats). It also provides information on the three types of DHS data file formats (flat, rectangular, and hierarchical files). And finally, there is a list of the contents of each of the zipped files (file extensions such as .dat, .sps, doc, etc.) along with an explanation of the file (data file, SPSS data description file, etc.).
- **Registering.** To register, which must be done in order to download any data sets, the option "**register**" must be selected, and the program will then guide the user through this process. The user will need to provide certain information, such as name, institution, address, project title, other researchers involved in the project, duration of the project, countries of interest, and a brief description of how the data will be used. Finally, the user will be assigned a user name and then will be asked to select a password (both are case-sensitive). Both the user name and the password will be needed each time the user wishes to download data. Final authorization will be granted (by e-mail) within a few business days, at which point the user can obtain the data sets for the countries of interest.⁶
- Once authorization is granted, the user can go directly to the "**registered users**" options from the "**Direct Data File FTP**" page. The following steps are provided as a guide to selecting and downloading data sets. Data files available for Mali will be used for demonstration purposes.

⁶ It was this authors' experience that authorization was granted before an e-mail arrived. Hence, it is worth trying to download data after a day or two even if the authorizing e-mail has not yet arrived.

- After clicking on “**registered users**,” the user will be prompted to select a region and then a country from world and regional maps.
- Select “Africa” and then “Mali” by clicking on these outlined areas on the respective maps with the mouse. These selections will generate a table with a list of data files from surveys conducted in Mali that can be downloaded. The table provides the data file name, the size of each file, the survey year, the file format, and the file type. As of July 1997, the data from surveys conducted in 1987 and 1995 are available for Mali.
- The list of data sets for Mali include "Couple's Recode," "Household Raw," "Household Recode," "Individual Recode," "Children's Recode," "Male Raw," "Male Recode," and "Service Availability Raw." The household recode files contain data from surveys used to identify members of the household and to select eligible respondents for the individual interview (women in the household). The individual recode files are from in-depth interviews with individual women within a given household. The male raw files are from surveys of individual male members of a household.
- The data files come in three types of file formats: rectangular, flat, and hierarchical. For most PC users, the rectangular files are the most appropriate. These contain a fixed number of records for each case in the data file and are designed to be used with SPSS and other PC-based statistical software packages. They typically require 8-50 megabytes of disk space.
- The results of surveys with individual women are designed to be comparable with other country studies and other survey years for the same country. DHS recommends that users who wish to do comparative surveys download a hard copy of the “**Recode Manual**, which gives a list and description of each of the variables in the data sets.” This option is available in the upper right section of this web page. For some survey years, such as the 1996 Mali data set, there is a recode manual for the documentation specific to that country and survey that is located within the

zipped file and has an extension ".doc". These manuals should be downloaded and even printed out.

- In order to download individual data sets, the user will need to have software that allows for unzipping files, such as *PKzip*.
- At this point the user may select the file types desired. For example, to get the 1996 Household Recode file in rectangular format for Mali, select **MLHR32RT.ZIP**. The user will then be prompted for both the user name and the password chosen during registration. Then a text box will appear labeled "Unknown File Type" noting that a zipped file has been selected for saving. At this stage, all of these files exist together in the one zipped file, and they must be saved and downloaded together as one file. To proceed with downloading the file, select "Save File," which will pull up a "Save As" box for a local directory. Once the file has been saved, the screen will return to the DHS list of files.
- The user may now exit the Internet and the files can be accessed by unzipping them once software such as PKZip has been installed. There are two ways to unzip the files. The user may go to the "DOS" prompt at the directory where the files are located and type **pkunzip MLHR32RT.ZIP**. Or, if the user has access to PKUNZIP in icon form, the user can simply click on the icon, select "open," then "extract," the "selected files" (highlighting those to download), and then click on "to disk" as an option.
- The unzipped files for Mali will include those with extensions .MAP, .DAT, .FRQ, FRW, .SPS, and .SAS. For the 1996 Mali data, the files are available in both SPSS/PC and SAS formats. The files MLHR31RT.SPS and MLHR31RT.DAT are SPSS/PC command and SPSS/PC data files, respectively. For the 1987 data, however, only SPSS files are available.
- The user should then bring up SPSS/PC software, and "**open**" the command file, MLHR32RT.SPS. This file will be used to turn the raw data in MLHR32RT.DAT into a working SPSS/PC data set with identified variables and figures. In order to do this, a few modifications in the command text must be made.
 - The following is the first line of text as it appears in this file.

DATA LIST FILE='MLHR31RT.DAT' RECORDS=73

This command needs to be edited to include the drive and file path such that the local file can be located with the program. An example might be the following:

DATA LIST FILE='d:\mali\data\MLHR31RT.DAT' RECORDS=73

where the local drive (d:) and path (mali/data) of the file is identified.

- A "record" is the section of command text between slashes ("/"). In earlier data files (such as the 1987 Mali data), the "RECORDS=" command does not appear and will need to be added to the first line of command text. Appendix 2 gives an example of an 1987 file, and how the changes should be made. The individual records in this file total 25 (note that only the records before the section with "variable labels" are included). As a result, the first line of the program for the 1987 data set would look like this.

DATA LIST FILE='d:\mali\data\MLHH01RT.DAT' records=25/

- Since this program only manipulates the data temporarily, a final line of command must be added at the end of this text file in order to save the newly created data file. In addition, the new data file will be easier to access with SPSS/PC if the extension is ".SAV" instead of ".DAT". So, for example, with the 1996 data set,

.save outfile='d:\mali\data\mlhrnew.sav' .

placed at the very end of the command file text will save the new data set under a different name (to distinguish it from the original raw data set) and make it usable for SPSS/PC statistical manipulation. (See Appendix 1 for additional data manipulation needed to generate certain information.) Once these changes have been made, the program can be implemented by selecting "run," and then "all" when the cursor is placed at the very beginning of the program.

Output from the Mali 1996 DHS Survey

This section provides the reader with some examples of how the DHS data, once obtained and modified, can be used for regional targeting of various development projects. Table 4 presents data on wasting (weight-for-height) derived from the Mali DHS Individual Recode data set. The figures are broken down to the *Cercle* level—one level down from the Regional level in Mali—and are ranked from worst to best in terms of percentages of wasted children.⁷ Figure 5 also shows the percentage of wasting on a *Cercle* level, but in a geographical display of the data where the highest levels of wasting appear as the darkest areas on the map.⁸

From both Table 4 and Figure 5, it is immediately clear where the highest percentages of wasting are across the country. In Youvarou, which is in the Region of Mopti, located at the top right of the map in Figure 6, the rates of wasting are 70 percent—much higher than the average levels throughout the country of about 20 to 25 percent. If project controllers were interested in identifying specific sites for health care centers for women and children based alone on the percentage of wasted children, then Youvarou might be an ideal choice. In making decisions of site selection, however, consideration should also be given to targeting on the absolute number of cases of a given indicator. For example, Youvarou clearly has the highest percentage of wasting. When the population density of the Regions and *Cercles* are taken into account, however, as can be seen from Figure 6, Youvarou falls into the second category out of four in terms of population density (between 7 and 18 people per square kilometer). A better choice for site selection might be Djenne (also in the Mopti Region), with the second highest percentage of wasted children, but where the population density falls into the highest group with between 24 and 46 people per square kilometer.

⁷ The Northern regions of Toumbouctou, Gao, and Kidal are not included in the data analysis or discussion here, either because the data were not collected by DHS in those regions or because only one of the *Cercles* within a region was surveyed.

⁸ This map was created using *Epi Map*, a software program produced by the Centers for Disease Control and Prevention (CDC). See Technical Guide #4 for using this program with case study data, also from Mali.

Table 4—Ranked percent of wasting (weight-for-height) in Southern Mali (percent)

Region name	Cercle name	Wasting (weight-for-height)	Region name	Cercle name	Wasting (weight-for-height)
MOPTI	Youvarou	70.0	SEGOU	Bla	23.4
MOPTI	Djenne	37.8	MOPTI	Koro	23.1
KOULIKORO	Banamba	37.8	SIKASSO	Sikasso	22.8
SIKASSO	Kolondie	37.2	SEGOU	Macina	21.7
MOPTI	Douentza	34.6	SEGOU	Tominian	21.4
MOPTI	Mopti	32.0	KOULIKORO	Dioila	20.2
KOULIKORO	Nara	29.4	KOULIKORO	Kangaba	19.0
KAYES	Diema	27.9	KOULIKORO	Kati	18.9
SIKASSO	Koutiala	27.3	MOPTI	Bankass	17.5
KOULIKORO	Koulikoro	26.6	SEGOU	Niono	16.5
SIKASSO	Kadiolo	26.3	KAYES	Yelimane	15.0
KOULIKORO	Kolokani	25.5	KAYES	Kenieba	14.6
SEGOU	Segou	25.3	KAYES	Kita	14.2
MOPTI	Bandiaga	25.2	SEGOU	Baroueli	13.6
SIKASSO	Bougouni	24.7	KAYES	Kayes	13.4
SIKASSO	Yanfolil	24.6	MOPTI	Tenenkou	13.1
SEGOU	San	24.5	SIKASSO	Yorosso	12.1
KAYES	Nioro	24.0	KAYES	Bafoulab	11.6

Figure 5—Southern Mali: Weight-for-height (wasting) at Cercle level

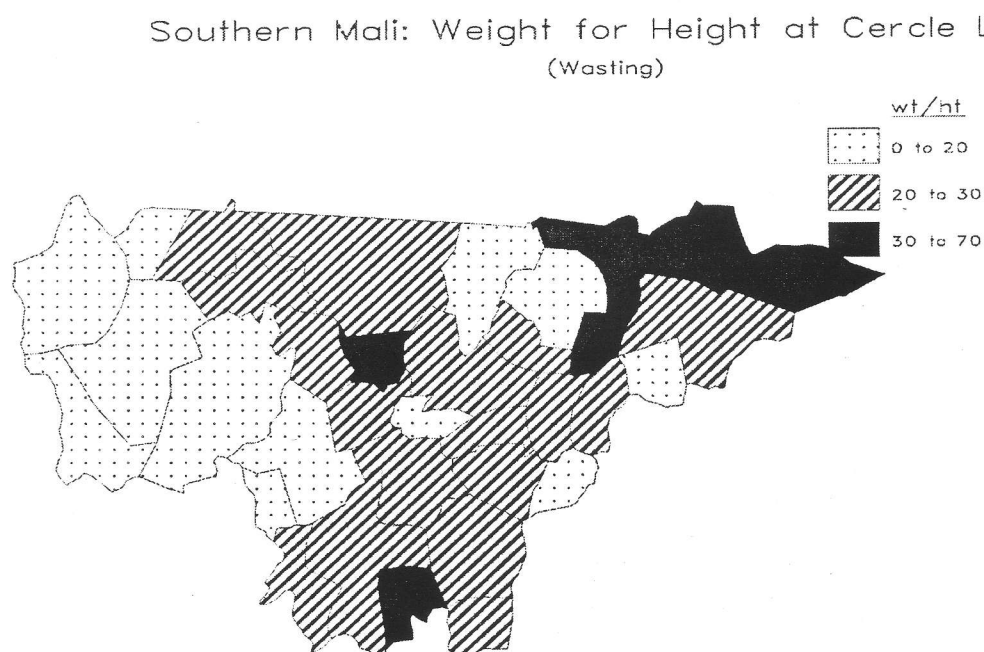
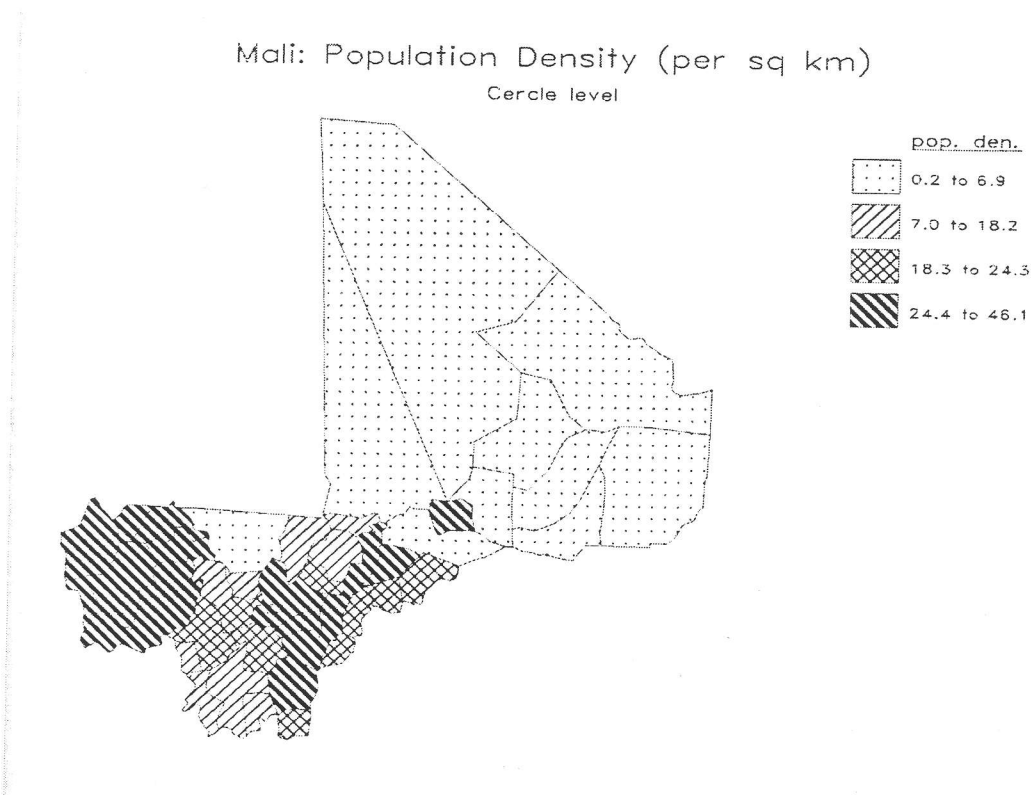


Figure 6—Mali: Population density at *Cercle* level (per square kilometer)



APPENDIX 1

MODIFICATION OF SPSS/PC SYNTAX FILES

Creating New Child Files from the Individual Recode DHS Data Sets

To Recode DHS Data Sets

Data on the children within the households (child health- and nutrition-related information) are available from both the Individual Recode files and, for some countries, from separate child files, called Child Recode files (for example, in Mali, there are Child Recode rectangular data files for both 1987 and 1996: MLKR01RT.ZIP and MLKR31RT.ZIP, respectively). If the Child recode files are not available for a country of interest, the individual recode data file can be used to obtain information on children, but it must be modified to do so. There is a program available on the DHS Web site to do the modifications, but the following method, using SPSS, is a fairly straightforward one.

The reason for creating a separate child data file is that the individual recode data sets for women are presented with individual women as *cases*, and data on each child (up to 6 children for each woman) are included as *variables* in different columns for each women (case). These therefore need to be converted to a new data file to ensure that each child is a *case* and not a *variable*. This process involves creating a new data file by creating individual data files for each child, and then stacking them one on top of the other.

While constructing this file, transformations of variables using appropriate divisors should be done. The DHS recode manual needs to be checked to find the divisors for the variables that might need to be transformed. A general "sample weighting variable" should also be transformed at this point. This variable is v005 in all data files and it has to be divided by 1,000,000 to create the weight. *In most cases, the data must be weighted using this variable before any analysis is conducted.*

As with creating any new data file, variable labels and value labels need to be specified for the newly created variables. Variable labels can be created by the data analyst but the value labels need to be checked carefully from the value labels on the variables in the original data file.

Cleaning the New Child Data File

1. Some of the newly created variables will need to be recoded to ensure that codes for "flagged" cases are not analyzed along with existing data (e.g., 9998 is a flagged value for HAZ in the original data set, which enters the analysis if not recoded in the newly created child file). SPSS will retain the "missing value" codes for variables when it creates the new variables, but it will not record codes for certain "flagged" values. These codes for "flagged variable" are available from the variable definitions in the original data file.
2. Before analyzing this data, the cases must be weighted by the weighting variable. It is also advisable to filter the cases by the variable for the index to the birth history of the child. The advantage of filtering using this child variable is that only data on existing children will be read, and figures for missing values will reflect actual missing values from the data-gathering stage and not for missing values that get repeated as files are stacked one on top of the other. Some variables that should definitely be included when creating a child file are *caseid*, *hlthindx*, *hwindx*, *childage*.
3. Before beginning any analysis, it is advisable to run some descriptives and check the sample sizes on those variables with those reported in the hard copy DHS reports to ensure that the stacking, weighting, and recoding process has not resulted in any errors. For the child files, it is advisable to check the sample sizes for the HAZ, WHZ, and WAZ variables, as these are the variables reported in the DHS reports. It should be noted that since, in some cases, the data sets were released only after the DHS reports for that country were published, it will be normal to have sample sizes that are not exactly the same as in the published report. The sample sizes can be off by a margin of up to 5 cases for some countries: this is because of the recoding process involved in creating the individual recode files.

APPENDIX 2
OTHER SOURCES OF SUBNATIONAL LEVEL DATA AVAILABLE
ON THE INTERNET

The World Wide Web, with its mind-boggling quantity of information, has become an important source of information and data on developing countries. It is now possible to do searches for information on almost any subject and download data on almost any country. This section provides the names and Internet addresses of many different sources of demographic, agricultural, and health data available on a subnational level for developing countries. The data listed below in the Table 5 can be viewed from the web sites, or can be downloaded onto personal computers for further analysis. The actual indicators available may vary from one country to the next within each web site.