

# GIS NEWS AND INFORMATION

April 1996 (No. 9)

*Dedicated to CDC/ATSDR scientific excellence and advancement in disease control and prevention using GIS*

**Selected Contents:** GIS Lab at Ft. Collins (p.3); GIS Users assist Coastal Health District (p.7); GIS dengue project (p.10); CDC/ATSDR GIS survey results (p.12); USGS news, and Census 2000 and TIGER (p.15); Excerpts, National Committee on Vital and Health Statistics (p.16).

**Current Subscribers:** 450

## I. Public Health GIS Training Opportunities

\* NCHS Cartography and GIS Guest Lecture Series: The lecture "An Epidemiological and Geographical Approach to the Study of Lyme Disease" was presented on February 29 by Jay Morgan (Towson State University), and Greg Glass and Dwight Wilson (Johns Hopkins University). It was envisioned to several CDC/ATSDR locations. A case-control study led to a risk-based map of lyme disease in Baltimore county. Logistic regression analysis of environmental covariates and remotely sensed information was included. For copies of transparencies used in the lecture, you may contact Dr. Morgan at <e7g4mor@toe.towson.edu>.

\* The advanced GIS short course "Spatial Data Analysis Using GIS", presented this past September at NCHS by Luc Anselin will be ready for viewing at CDC Anchorage. All interested staff should contact coordinator **Lisa Bulkow**, NCID, in Anchorage (907) 271-4011, for scheduling information. [Editor: Appreciation is extended to **Jay Kim** (NIOSH), **Chet Moore** (NCID), **Shelba Whaley** (PHPPO), **Peter Kilmarx** (NCPS), **Louise House** (ATSDR) and **Lisa Bulkow** (NCID) for making this tape viewable, in their respective work locations, to interested staff.

## II. News from GIS USERS

(Please communicate directly with colleagues on any issues)

## A. General News

1. From **Barbara Tempalski**, Hunter College: I was unaware of CDC's involvement in Guinea worm eradication. Thank you for the information. Sara McLafferty and I have collaborated on projects such as low birthweight in New York City. We are presently writing a grant for a lead study in NYC using a GIS. I also worked with Sara on a CDC rat bite project where we used a GIS to predict where rat bites were most likely to occur in NYC. I have also worked as a consultant for USAID in Cairo using GPS and GIS to track schistosomiasis. Most of my research involves GIS and health. If I can be of any further assistance please let me know. Sincerely, Barbara Tempalski, GIS/Research Analyst, Spatial Analysis & Remote Sensing Lab, City University of New York, Hunter College, 695 Park Ave, New York, NY, 10021.

2. From **Steve Macdonald**, NCEH: CDC National Center for Environmental Health (NCEH), in collaboration with the Council of State and Territorial Epidemiologists (CSTE) and the National Environmental Health Association (NEHA), is hosting a series of workshops, whose goal is to develop a set of guidelines for environmental public health surveillance (EPHS) system development. The first of these workshops focused on the structures necessary for state-based EPHS systems, the second will focus on the EPHS process, and the third on the output of EPHS

systems. By "process" we mean mechanisms for data collection and analysis, which include the use of GIS software. I am interested in learning about experiences which others have had in using GIS in routine analyses of surveillance data (not for special studies). Please reply to: Steven Macdonald, Surveillance Branch, 770-488-7086, Email <[swm3@cehdeh1.em.cdc.gov](mailto:swm3@cehdeh1.em.cdc.gov)>.

3. From **Jay Devasundaram**, State Department of Health, Baltimore, Maryland: I wanted to share with the others on the List, my experience with trying to customize ArcView 2.1 to use Epi-Info, as a statistical engine in the background, to return frequency tables to AV2.1 for mapping. The AV-Epi-Info interface seems to be working really smoothly- Epi-Info creates two dbf files one of frequencies of occurrence of zipcodes and the other frequencies of occurrence of county fips. These two tables are then used in ArcView to create freq distribution maps. I will be glad to share my experiences with others if they wish to contact me in this regard. My work number is (410) 225-6664.

4. From **Steve Game**, NIOSH: At the last Public Health Conference on Records and Statistics, you were doing a study regarding color and patterning for mapping. Could you point me in the direction of either printed results, or some general recommendations that you are using based on your studies? The kind of information I'm interested in are color schemes (ie categorize light to dark, cold colors to hot colors, etc.), pattern recommendations (relating to color blindness), background colors (ie use white for background null or another color)..Wasn't there a past seminar that addressed these issues and if so, do you know if it is available in any form ( video, slides, handouts, etc.). I appreciate your assistance.

[**Editor:** A copy of the ORM/ NCHS Cognitive Methods Staff Working Paper Series, No. 18, June 1995, *Cognitive Aspects of Statistical Mapping* (EDs. **Linda Pickle** and **Doug Herrmann**), was sent to Steve]

5. From **Linda Pickle**, NCHS: Meeting notice: Modelling longitudinal and spatially correlated data: Methods, applications, and future directions. Oct. 15-18, 1996, in Nantucket, MA. Attendance limited. Invited speakers: David Brillinger, Ray Carroll, Noel Cressie, Peter Diggle, Nan Laird, Peter McCullagh. For more info, see <http://tgg.fw.vt.edu/Nantucket/or> send Email to Tim Gregoire, [tgg@vt.edu](mailto:tgg@vt.edu)

6. From **Joseph Scarpaci**, Virginia Tech: A group of Cuban medical geographers is interested in receiving any information (reprints, reports) or software that addresses GIS and health care. No materials are considered too dated or insignificant. If you have any materials (books, reprints, software for PC platform) that you can spare, please forward them to me and I will take them to Havana in June. Thank you. Joseph L. Scarpaci, Urban Affairs and Planning, Virginia Tech, Blacksburg, VA 24061-0113; Office: 540-231-7504 FAX: 540-231-3367 <http://www.arch.vt.edu/SERVICE/UAP/cuba.html>

7. From **Tom Usselman**, National Academy of Sciences: The Mapping Science Committee's activities are now represented on a home page, including some of its thoughts on a spatial data and health activity. The URL is <<http://www.nas.edu.cger/besr/msc.html>>

## B. Technical News

8. From **Jay Kim**, NIOSH: We are actively looking for a PC based GIS program that will work on Windows 3.1 now and on Windows 95

in the near future. There is one very important function that the software must have --- the program must display all 50 states on one page with county boundaries and with Alaska and Hawaii brought in below U.S. - Mexico border line. The size of AK should be reduced and that of HI increased. It does not matter whether the program gives the above map layout by default at the beginning or users are able to move the two states after a thematic map has been completed. If any of your programs meets this criteria, please fax me a copy of a simple thematic map, population by U.S. counties, for example, describing the step by step method how AK and HI were brought under the south of border. Thank you. Jay H. Kim, Ph.D., National Institute for Occupational Safety and Health, 1095 Willowdale Road, Morgantown, WV 26505, 304-285-6111(fax).

9. From **Steve Campbell**, Baltimore City Health Department: Would any GIS Users have a clue as to where one can find census tract boundary maps for Epi-Map? I'd also like (as you know) to get my hands on zip code boundary maps for Epi-Map. Any assistance would be greatly appreciated. (The background to that is I'm working with the Women's Cancer Screening Project at Baltimore City Health Department and we need to present information on our 7 community health centers in terms of their catchment areas and the locations of patients already served. We'd also like to match known morbidity and mortality data to the same areas to see if we're serving areas with high rates of cancer morbidity/mortality.) Steve Campbell, <TQPZ46A @ prodigy.com>

10. From **Chet Moore**, NCID: After seemingly endless delays, DVVID's GIS facility in Fort Collins, Colorado, is up and running. All of the equipment has been consolidated in a single room which was remodeled specifically for this

use. Projects are underway on arboviruses (La Crosse encephalitis), plague, and *Aedes albopictus* (the Asian tiger mosquito). At this point, our equipment and software resources consist of the following:

Dell 566 GIS station with 1GB hard drive, 32MB RAM, CD ROM

1.2GB read-write optical drive

Altek AC41 digitizing board, 36"x48"

Uninterruptable power supplies (2)

Hewlett-Packard HP-750C color printer, D-size (24")

Agfa & Montage slide makers and associated computers

Canon CJ-10 color scanner/printer (8.5" x 11" format)

Dell DGX 590 computer with Atlas GIS software

CMT Global Positioning System (1)

Magellan ProMARK-X GPS unit (3)

TNTmips (MicroImages, Inc.) GIS software (raster/vector/CAD)

MapInfo, v.3.0 GIS software (vector-based desktop mapping)

IDRISI, GIS software (raster-based GIS software)

Map case, 16-drawer

We are particularly interested in exploring possible collaborative projects with groups in the following states: Alabama, Colorado, Midwestern US (Mississippi/Ohio River Valleys), New Mexico, North Carolina, and West Virginia.

11. From **Linda Pickle**, NCHS: The new Spatial Statistics module for S+ is in the hands of beta testers and should be shipping soon. The module is an add-on to S+ and can analyze data from ArcInfo and Geo-Eas. Functions are available to analyze continuous geostatistical data, lattice (neighboring point) data, and spatial point data, where the location is the variable of interest. Analytic tools include the traditional

kriging and variograms, as well as spatial autoregression models and exploratory graphics.

12. **Editor:** (a) SAS/SPECTRAVIEW™ software (Version 6, First Edition) offers interactive surface and volume visualization for viewing, exploring and analyzing two and three-dimensional data. SAS data requirements include at least four numeric values- an X variable, Y variable, Z variable and response variable. An optional fifth BY variable shows how response variables change over time. For more information, call SAS at 919-677-8000 or Net connect at <http://www.sas.com> (b) Of related interest is the use of Spearman's Correlation Coefficient for the nonparametric measurement of log-transformed three-dimensional grids (see Rodriguez, O., A quantitative approach for comparing three-dimensional environmental data grids. *Geo Info Systems* 1996 (March); 6,(3):40-43).

### C. Internet News

(New section on GIS items picked up from Internet by GIS Users)

13. Picked up from Multiple recipients of list STAT-L by **Jay Smith**, NCEH: I'm analyzing spatial variation in disease incidence. I've started using variograms (omnidirectional and directional) as tools for exploratory data analysis. I think I understand what's going on. I'm now using cross-variograms to look at joint spatial relationships between 2 diseases. Can anyone suggest a review article or an introductory text which specifically addresses cross-variograms? I'm trying to get Isaacs and Srivastava "An introduction to geostatistics" - does this book address this topic? Thanks in advance. Anthony Staines, E-Mail - [A.Staines@leeds.ac.uk](mailto:A.Staines@leeds.ac.uk) Dial +44 (0)113 Tel - 2443517 Fax - 2426065 S-Mail - LRF Epidemiology Unit, 17

Springfield Mt, Leeds, LS2 9NG, UK. [Editor: I found the following reference in Cressie: Clark, I., Basinger, K.L., and Harper, W.V. (1989). MUCK: A novel approach to co-kriging. In *Proceedings of the Conference on Geostatistical, Sensitivity, and Uncertainty Methods for Ground-Water Flow and Radionuclide Transport Modeling*, B.E. Buxton, ed. Batelle Press, Columbus, OH, 473-493; Also, an illustration is provided in Isaacs and Srivastava, *An Introduction to Applied Geostatistics*, Oxford University Press, 1989, New York, pp. 175-177]

14. Picked up from [gis-l@urisa.org](mailto:gis-l@urisa.org) by **Mike Mungiole**, NCHS: WWW Access to GIS Conference Proceedings: The University of Maine Library and the National Center for Geographic Information and Analysis, Orono, Maine are pleased to announce Spatial Odyssey, WWW access to the full text of GIS conference proceeding articles <<http://www.odyssey.maine.edu/gisweb/>>. Leading professional organizations involved in publishing GIS conference proceedings are contributing files for the years 1994 and forward: American Congress on Surveying and Mapping / American Society for Photogrammetry and Remote Sensing (ACSM/ASPRS); Automated Mapping and Facilities Management (AM/FM International); European GIS Foundation (EGIS); Association of American Geographers (GIS/LIS); Urban and Regional Information Systems Association (URISA). Spatial Odyssey also provides WWW access to the NCGIA GIS Annual Bibliography Series which lists the contents of numerous major GIS conferences and published collections of articles annually, from 1991 to present. Hypertext links from the bibliographic citations to the full text of articles are in place for the 1994 proceedings. Project Contacts: Marilyn Lutz, Assistant Director Information Systems, University of Maine System, Raymond H.

Fogler Library 5729, University of Maine, Orono, Maine 04469-5729, Phone (207)581-1658, FAX (207) 581-1653, E-mail: lutz@maine.maine.edu and Harlan J. Onsrud, Chair Dept. of Spatial Information Science and Engineering Scientist, National Center for Geographic Information and Analysis, 5711 Boardman Hall, Rm. 348, University of Maine, Orono, Maine 04469-5711, Phone (207) 581-2175, FAX (207) 581-2206 E-mail: onsrud@spatial.maine.edu, <http://www.spatial.maine.edu>.

15. Picked up from Multiple recipients of list SASPAC-L by **Arlene B. Siller**, NCHS: CIESIN (Consortium for International Earth Science Information Network) of Saginaw, MI invites users to experiment with a new WWW-based tool for the processing of geo-based data. This application is being developed under contract with the author (John Blodgett - U of Missouri St. Louis) under the supervision of Hendrik Meij. The project is funded by NASA as a SEDAC contract at CIESIN. The URL for the application at CIESIN is: <http://testserver.ciesin.org:8801/sedac-2.2/geocorr/index.html>; The name of the application is "MABLE/Geocorr". MABLE is an acronym for Master Area Block Level Equivalency file. The application is driven by a massive 7-million observation database (in SAS format - the application is written in SAS with a Perl cgi-bin interface) containing geographic information about every 1990 census block in the 51 states.

"Geocorr" is the geographic correspondence engine that can access the MABLE data and produce output in the form of on-screen reports and comma-delimited format files that define the relationship between user-specified geographic coverages for user-specified geographic areas (universes). Specific examples of what the engine can create

are: 1. For any state or group of states: A file showing each ZIP code in the state(s) and what counties the ZIP intersects with, the total 1990 population of that intersection and what proportion of the ZIP's total population is in each of the counties it intersects with. Such a file is referred to as a "ZIP to county correlation list". 2. For any metropolitan area (1990 definition): A file showing how census tract / BNA codes relate to ZIP codes. Similar to example 1: the correlation list would show the total 1990 population of the intersection and would have an "allocation factor" variable indicating the decimal portion of the tract that is in each intersecting ZIP. 3. The block records on the database contain the lat-long coordinates of a point within the block (a geographic center point). The geocorr application has an option to calculate a weighted average of these coordinate values to be kept on the output. If the user selects "pop" as the weighting variable these become population centroids for the output areas. Thus the application can be used to determine population centroids for any of the geocodes on the database. For example: calculate the population centroids of all MSA-CMSA's in Florida.

There are many potential applications for such correlation lists and we'll be creating some background papers and related documents describing these. But at this stage we are looking for early beta testers of the system who are probably already familiar with the use of such correspondence files. Some important notes and caveats about the current **\*\*very experimental\*\*** implementation at CIESIN: It can be very slow during the day. Especially if your request is for small-level output or you need to access large-state files. Be prepared to wait several minutes for each state selected for processing. If you're from Texas you can go to lunch and your output will be there when you get back. Its much better in off hrs. The default

form has the state of Alabama selected: turn this off before selecting your states (unless, of course, you want data for Alabama). Be sure you understand the note at the top form about doing multiple selects. You need to be able to select multiple items from the various select boxes. Yes, its possible to get data for the whole U.S. by selecting ALL the states but please don't. (Not yet, at least.) There is a small glitch in how the output files are presented. If you specify that you do not want a listing file it currently still shows you a menu with that output selection on it. Just ignore it. We have noticed that our version of MOSAIC does not handle checkboxes properly. But you can type a "0" over the "1" in the box to turn off the option ("uncheck"). Ugly. Most of our testing has been with IBM's WebExplorer (OS/2) and Netscape 1.2. There is an option on the form that allows you to specify a custom name for your listing and .csv output files: the default value is "geocorr.lst" and "geocorr.csv". Currently that option does not work at CIESIN; just accept the default.

Henk Meij has gone on vacation until March 4 so there is really none at CIESIN who can go in and fix things if they break (related to this specific application). So try not to break anything until he gets back or be patient. You can leave messages. If you only need data for Mo, Il, Ks, Co, Mi, Ct, Nv or Al then you can access an alternate site here in Missouri. The URL here is <http://www.oseda.missouri.edu/sedac-2.2/geocorr>. This is a very unreliable site, however, since it may be "improved" at any moment without notice. This is the actual construction site. The source of the MABLE database at this point is entirely the 1990 ZIP-Block equivalency files produced by the U.S. Census Bureau. Thus all of the geographic information is as defined on 1-1-90 except for ZIP codes which were assigned by a contractor and are as of approximately July, 1991. There

are a whole bunch of caveats about these ZIP codes.

We already have several enhancements under development and others that we are working on proposals for. One of the more useful and soon-to-be-implemented extensions will be the addition of 1990 PUMA codes to the MABLE database. This will allow easy correlating of the geographic areas on the 1990 Public Use Micro Sample files with all other geographies. We leave it as an exercise for the class to determine why this might be useful. We plan to create a series of help modules with hyper-links on the form to provide better explanation of what the options are about and what to watch out for. Also, with some kind of background, overview piece to explain what a correlation list really is.

Future possible enhancements include the addition of School District and Transportation Analysis Zone geography (based on TIGER 94). The ultimate extension would involve relating the census blocks to a regular square-kilometer (or smaller) grid that would allow creating lists relating political jurisdictions with regular grids. This would provide a tool for allowing data to be converted back and forth between these types of geographies, i.e. households and family income data for grid areas, and land-coverage data aggregated to place and/or ZIP codes. Since we have x-y coordinates on the block data, we could implement a selection criteria based on distance from one or more specified points. For now, you can ask for the coordinates on your .csv file and implement this yourself (if you know the secret formula.) We encourage comments and suggestions as long as they are not overly abusive. But we don't have a feedback button on the form yet so you'll have to use e-mail. Send it to me and/or to Henk Meij ([hmeij@ciesin.org](mailto:hmeij@ciesin.org)). John Blodgett, Urban Information Center / Office of Computing, University of

Missouri - St. Louis, 8001 Natural Bridge Rd.,  
St. Louis, Mo 63121-4499; Phone: (314)  
516-6014/6000 FAX: 516-6007, e-mail: c1921  
@umslvam.umsl.edu.

### III. GIS Outreach

A. Responses to February 26 announcement "GIS HELP sought from Evelyn Choury, MPH" are shown below. Special thanks are extended to **Bill Henriques**, **Brian Plikaytis**, and **Jerry Feder**. Please note the ATSDR GIS team that Bill is coordinating. Original request as follows: **Nancy Stroup** suggested I contact CDC/ATSDR GIS Users regarding a project we are undertaking. I am with the Coastal Health District office--six counties south of Savannah, GA to FL border. We have a GIS system (startup, at this point), and I am conceptualizing the rudimentary data base elements. We have 21 hazardous waste sites in our county (Glynn, major city of 16,000+ is Brunswick) including 3 on NPL, and want to do some morbidity/mortality mapping in relation to the sites. We just got another small grant from NACCHO to look at issues around the newest site: GA Pacific. Do you know anyone who has a model we can follow or modify? Many thanks. Tel.# 912-264-3907; FAX 912-262-2315.

1. From **Bill Henriques**, ATSDR: I called Evelyn and spoke with her concerning the recent resurgence of the GIS efforts here at ATSDR. I have recently been appointed as the GIS Coordinator as of March 1st. We have also been lucky to acquire **Virginia Lee**, Medical Officer and **Janet Heitgerd**, Demographer/Statistician for ATSDR Team GIS. Evelyn is very interested in working with us to use GIS technologies for her demographic work.

2. From **Brian D. Plikaytis**, NCID: There is an 11 disk CD ROM set which can be purchased from the Census Bureau for a

nominal charge that may be of interest to Evelyn. These are digitized maps of the U.S. that overlay the locations of EPA regulated sites. It uses a front-end program developed at the Census Bureau called LandView II. This is a geographic reference. like an atlas, that displays EPA-regulated sites, demographic and economic information from the 1990 census, and key geographic features of the U.S. This includes a subset of facilities, sites, and monitoring stations represented in five EPA data bases: Aerometric Information Retrieval System (AIRS), Biennial Reporting System (BRS), CERCL Information System (CERCLIS), Permit Compliance System (PCS), and the Toxic Release Inventory System (TRI).

LandView II also includes the complete Bureau of Census TIGER/Line '92 files. These files are data containing spatial characteristics of streets, rivers, railroads, and other landmarks. With this software, the user can view EPA and Bureau of Census information along with geographic boundaries including states, congressional districts, counties, cities, tribal lands, census tracts, census block groups, and metropolitan areas. This software will also access the much broader STF3A Census disks, if these are available. User-defined and created databases may also be used by this software. I have spoken to some of the programmers at the Bureau of Census who developed the product and they are more than eager to help people with customized applications. I am not tremendously well-versed in the use of this software, but I can attempt to demo it for anyone interested in stopping by.

3. From **Jerry Feder**, USGS (Email to Evelyn): My colleagues and I at the US Geological Survey have collected much water quality data in the area you are presently studying. This data includes: trace elements, organic contaminants, radioactive elements, herbicides and pesticides. We also have

extensive GIS coverages of your study area. If you think we could help you out or cooperate on a study please let me know. I am also FAXing you an article that appeared recently in

the Jacksonville, FL newspaper. Geochemically, the environment in NE Florida is similar to the coastal areas of Georgia. Please feel free to e-mail me at: glfeder@usgs.gov or FAX: 703-648-5274.

B. The following announcement, from **Ron Welch**, S. Dakota School of Mines and Technology, concerns the development of a GIS infectious disease and climatological database and seeks GIS Users input:

For an introduction, I am head of a large and active satellite remote sensing research group at South Dakota School of Mines and Technology. Together with Dr. Paul Epstein at Harvard Med School, Dr. Claire Parkinson at NASA Goddard Space Flight Center, and Dr. Vickie Connors at NASA Headquarters, we are in the process of constructing a global infectious disease database (disease, date and location) that will be populated/correlated with a wide range of ancillary data (ecosystem, terrain, %water, elevation, several vegetation indices, soil characteristics) and satellite derived climatological parameters (precipitation, max/min day/night temperatures, cloudiness, wind speed, winter/summer severity indices, El Nino/La Nina year, flood/drought occurrences, etc). Both a GIS system in ARC/INFO and a less sophisticated relational database (initially for Silicon Graphics workstations, but may be ported to X) are being constructed. These databases will be distributed freely on the internet, with the exception of restrictions that may be placed upon certain sensitive data categories by the CDC or state agencies.

At present we have the encouragement of Dr. Ruth Berkelman, Deputy Director,

National Center for Infectious Diseases (CDC/Atlanta) and a great deal of support from a large number of people at CDC (Atlanta and Ft. Collins). We have made formal requests for data from Dr. Duane Gubler, Director, Division of Vector-Borne Diseases (CDC, Ft. Collins), Dr. Denise Koo, Chief, Systems Operations and Information Branch (CDC/Atlanta) and to Dr. Ann Fleiger, Special Pathogens Division (CDC/Atlanta). In each case, the maximum assistance for data has been promised, consistent with restrictions for certain categories. It should be noted that data is only requested for dates and locations. No data entries will be made as a function of race, gender, age, economic status, etc. We acknowledge and appreciate the encouragement, active assistance, and additional contacts provided by Dr. Steven Ostroff, Dr. Marta Akers, Dr. Mike Cramer, Dr. Chet Moore, and Dr. Charles Croner. The purpose of this letter is not only to make your acquaintance with this database development project, but also to solicit additional data sources. We would appreciate receiving cases, statistics, etc, concerning all infectious diseases from 1982 to the present (and future) that can be entered into the database - electronic, hardcopy, or references submissions are useful. Please send data to: Dr. Ronald M. Welch, Institute of Atmospheric Science, South Dakota School of Mines and Technology Rapid City, SD 57701; (605) 394-2291, (605) 394-6061 fax, welch@cloud.ias.sdsmt.edu [Editor: I believe Dr. Welch's project will constitute a competitive research proposal and I advise anyone interested to call for more information]

C. From **Howard Campbell**, District Health Department: We at the District Health Department, Elizabeth City, North Carolina, need educational sessions on using census, mortality and morbidity data and reflecting this

in a GIS format at the county level to show health status and health needs. Can you direct me to any CDC/ATSDR GIS Users who can provide guidance? Thank you <hbc@nascar.pgcc.dst.nc.us>

D. The following is a summary announcement of the 1995 edition of the "Electronic Atlas Newsletter", a HARD-COPY, monthly GIS newsletter which has been published since January 1990 by Brian Matuschak. Free samples of the newsletter are available to interested parties who e-mail, write, call, or fax their physical mailing address to Electronic Atlas Enterprises. For info regarding the newsletter (ask for price) or consulting services, contact: Electronic Atlas Enterprises, latlas@accessone.com//Tel./Fax: (206) 525-7155//P.O. Box 75394, Seattle, WA 98125-0394.

"Year-End Wrap-Up" : The *January* issue featured a retrospective of the first five years of the newsletter. It included a cumulative index of the first 60 issues and a color cover featuring a SPOT image scene which displayed most of the city of Seattle. Also, the major policies of the newsletter were described. They pertain to the newsletter's relaxed writing style, its direction to educational resources for GIS, submissions for the subscriber's "GIS Networker(tm)" column, and software reviews. *February's* lead topic was a commentary, "Bringing GIS into the Mainstream: What Will It Take?" Suggestions offered to the GIS vendor community included making products available for Windows 95, increasing product visibility and training opportunities, adopting open standards for data transfer, and bundling of datasets along with software. A new column was started in this edition, "GIS Resources from the Internet." Part 1 of this column featured the USGS Node of the NSDI Clearinghouse and Oracle Corporation's home page. Finally, the conclusion of the

five-year retrospective of the newsletter related a history of the newsletter. In *March*, yet another feature for the newsletter was started. The first installment of "On-Line GIS Services" reviewed what was available from the GISnet BBS in Boulder, CO. The review was complemented by a "GIS Networker(tm)" column by the BBS's founder, Bill Thoen, and it elaborated on the need for BBS systems despite the accessibility of the Internet. *April's* edition featured a guest column about the relationship between virtual reality and GIS by Dr. Robert Jacobson, founder of Worldesign in Seattle. This issue also included Part 2 of "GIS Resources from the Internet" which described the new listserver address for GIS-L and mentioned the availability of declassified satellite photos. The *May* issue's focus was on precision requirements for building GIS landbases. The practice of beginning a GIS with a less precise landbase and then using rubbersheeting and conflation to a better landbase was examined. These approaches were not recommended for governmental agencies needing more accurate information in the future. In addition, this edition contained a review of chapters 17 and 20 of the Multipurpose Land Information Systems: The Guidebook and the third installment of "GIS Resources from the Internet." This installment described how to download ESRI's ArcView 1.0 from the Internet. The *June* edition was an extensive review of Caliper Corporation's Maptitude 3.0a for Windows. There were two color supplements which illustrated the "look and feel" of the software along with the depth of information found in its data bundle. The conclusion of the review stated the software was an "excellent value for businesses-cartographers-and governmental agencies as well."

*July's* issue was devoted to press release coverage from ESRI and Caliper Corporation.

To end the issue, there was an open response to an e-mail message received from an employee of SPOT Image Corporation. The traditional theme for the *August* newsletter is "Applying Cartographic Techniques to GIS." This year's emphasis was on working with color. This edition discussed basics of color reproduction, the need to calibrate color schemes for thematic maps to enhance a map's ability to be legibly printed on black-and-white printers, and "thematic overload," defined as a situation where too many separate layers are attempted to be displayed on a thematic map. *September's* emphasis was on the implications and pitfalls to upgrading to Windows 95. A list of enhancements from Windows 3.1 was listed in this issue along with description of potential hardware conflicts and other things a user should watch for. Two new utilities in Windows 95, Briefcase and Dial-up Networking, were discussed in relation to GIS data collection. Finally, a mention was made of a GPS add-in for Caliper's Maptitude. In *October*, ArcView 2.1 was reviewed. Enhancements from ArcView 2.0 and Avenue were examined. Part 4 of "GIS Resources from the Internet" discussed the PARC Map Viewer and its location on the World Wide Web. *November's* edition provided an introduction to the concept of Open GIS(tm) and to the Open GIS Consortium, Inc. (OGC). Open GIS promises to deliver transparent access to GIS datasets and applications across networks. This issue ended with an interview conducted with the vice president of OGC. Finally, the *December* edition featured "thumbnail" reviews of products from eight different GIS vendors. The products run on various versions of MS Windows and are either concerned with data or its acquisition or they are GIS software packages. The products include R2V 2.06 from Able Systems, Inc., the Data Automation Kit from ESRI, the GEONAME Digital Gazetteer 1.1, the 1990

U.S. Census Block Groups and Centroids CD's, MapInfo Professional for Windows 95 by MapInfo, the Mapping and GIS Office Suites by Intergraph, ER Mapper 5.1 from Earth Resources Mapping, and Atlas GIS and Atlas Select from Strategic Mapping Inc. (SMI). Some of these products will be reviewed in more detail in 1996's (Vol. 7's) editions of the Electronic Atlas Newsletter. [Editor: This is not intended as an endorsement of this commercial product; I agreed to post this announcement as an informational service to our readers]

#### IV. Special Field Report: GIS Project on Dengue

(New section-submissions are open to all)

**Amy Morrison**, PhD and **Paul Reiter**, DPhil, Dengue Branch, Centers for Disease Control and Prevention, and **Marilyn Santiago**, M.S., Caribbean District, U.S. Geological Survey San Juan, Puerto Rico

At the Dengue Branch of the CDC in San Juan Puerto Rico, we have been studying the geographic and temporal distribution of dengue cases using a Geographic Information System (GIS). This project has been a collaborative effort between the CDC and Puerto Rico Department of Health who manage Puerto Rico's dengue surveillance system, and the U.S. Geological survey who have developed a GIS for the island. A single dengue epidemic occurring during the later half of 1991 in municipality Florida, Puerto Rico was studied. Florida, a small rural community of about 8,700 people located in north central Puerto Rico, is the smallest municipality in both area and population but it had the highest incidence rate for dengue during 1991. Dengue cases from the municipality (377) were identified from a laboratory-based dengue surveillance system

and georeferenced by their residential addresses. The exact latitude and longitude of each household was determined by plotting its position on a digital zoning map, using a geopositioning system (GPS), or measuring the distances between the residence and a known location on a U.S.G.S. topographic map. The locations of the cases and all relevant demographic information were converted into ARC/INFO<sup>™</sup> GIS software format.

Weekly plots of dengue cases occurring between 1991-1992 were generated and showed clustering of cases within housing tracts separated by steep hills (eroded karst topography). The spread of DEN was extremely rapid; nearly all of the neighborhoods effected in the outbreak had at least one reported case within a month of the first reported cases. Statistical confirmation of these observations are currently underway, using second-order analysis and local statistics ( $G_i$  and  $G_i^*$ ).

For descriptive purposes, the assumption was made that the nearest geographic case would have the highest probability of being the source of infection for subsequent cases. A data matrix of all possible case pairs occurring during the outbreak was generated using the pointdistance command in ARC/INFO. This matrix included the date of onset of symptoms, distance and time interval between each case. For each case, the paired case separated by the shortest distance was extracted to form a new matrix including only "nearest case pairs". If only case pairs occurring 0-30 days apart were considered--since it is highly unlikely a case occurring >30 days earlier could have been the source of infection of the second--54% of the case pairs were separated by  $\leq 10$  days. The results suggest that a significant percentage of DEN transmission occurred rapidly, or that large numbers of people are infected simultaneously. We speculate that this pattern suggests that multiple blood feeding is a more

common source of virus transmission than serial transmission between person-mosquito-person. This interpretation was further supported by examining the fifty-one houses where multiple cases occurred ( $n=123$  cases). Of the 2nd to 6th cases within a house ( $n=72$ ), 37 (51%) occurred within 5 days of each other.

Additionally, data from this study indicates that 80% of nearest cases occur within 500 m of each other independent of the time interval examined; for the 0-30 day interval 56% of the nearest case pair were located within 100 m, 78% within 200 m, 84% within 300 m, 87% with 400 m and 90% within 500 m of each other. These distances are consistent with published reports on the flight range of *Ae. aegyptii* Puerto Rico (Reiter et al. 1995). It is possible that the small size of the study area (26 km<sup>2</sup>) may have restricted virus dispersal. Interpretation of these results has been confounded by a lack of data on the movement of people in this area which could also account for the patterns observed. These results indicate that control practices which react to individual dengue cases are unlikely to be effective, because they do not account for the rapid temporal and spatial dissemination of the virus.

Also of interest to newsletter readers is some discussion of the technical problems we experienced. Most of our problems arose from faulty assumptions about what data would be available before we started our project. We had hoped to use U.S. Census Bureau Tiger Files to address reference our cases. We quickly learned that for Puerto Rico this would not be possible. As an alternative we obtained digital zoning maps from the Puerto Rico Planning Board. These maps were out-dated and did not cover the entire municipality. In addition, we later learned that these maps were digitized from aerial photographs without any photo rectification. Thus, instead of using already available data as we had originally hoped, we

were forced to develop our own base maps. Because address georeferencing was not an option we were forced to locate each residence in the field and determine its location on our base maps. Florida, has at least three different address systems (they periodically reassign addresses) and we soon discovered that mailing addresses may have no relation to a residential address. Fortunately, the small size of the municipality saved us since most residents seemed to know everybody in town. We question the ability to carry out this type of project for a larger area, say the entire island of Puerto Rico. Based on this experience we chose to carry out a similar type of study in the municipality of Ponce. This time we checked the quality of address data and for the availability of base maps ahead of time only to have terrible problems converting existing maps of the area from MAP GRAFIX to ARC/INFO. It looks as if we have solved our problem, but we have been working on it for 5 months. We are still working with Tiger Files for the area in order to have denominator data. We have

discovered so many inconsistencies with the TIGER file information and the house counts we have made we are unsure how useful this information will be.

Just a few words of warning for anyone starting out. Never assume that base maps or data are available or of good quality until you have them safely in your GIS and be sure that you have thought carefully about how you will georeference your information (cases). I'm sure these types of problems may be old news for most, but I hope our experience may help. [Ref: Paul I. Reiter, et al., Dispersal of *Ae aegypti* in an urban area after blood feeding as demonstrated by rubidium-marked eggs, *Am J Trop Med. Hyg.* 1995;52(2):177-179]

#### V. GIS Activities at CDC/ATSDR

(Editor: please continue to submit any omissions, corrections or changes as they occur)

#### 1995 Survey of CDC/ATSDR GIS Projects

Disease, environmental condition or related topic	Time period	Geographic scale
polio	1995	national, China
malaria	1993-present	villages, western Kenya
Guinea Worm	1990-present	villages, 15 countries, sub-Saharan Africa
onchocerciasis	1988-present	villages in and around known endemic areas, Guatemala, Mexico, Colombia, Venezuela, Brazil, Ecuador
dengue fever	1989-present	municipalities, Puerto Rico

Disease, environmental condition or related topic	Time period	Geographic scale
plague	1994-present	national, western U.S. counties
Asian Tiger mosquito, aedes albopictus	1994-present	national, U.S.
arboviral encephalitis 1994	Alabama, Florida, Georgia	
Lyme disease	1994-present	Baltimore county, Maryland
Lyme disease	1995	national, U.S. counties
coccidioidomycosis	1994-present	Ventura county, California
respiratory diseases	1994-present	national, U.S. counties
health status indicators	1993-present	Dade county, Florida
violent and firearm-related injuries	1994-present	census tracts, DeKalb county, Georgia
well water co	1994-present	N.Dakota, S.Dakota, Iowa, Minnesota, Wisconsin, Kansas, Nebraska, Illinois, Missouri
ntamination		
public health assessment, contaminant monitoring	1994-present	Weldon Spring, Missouri; Oak Ridge, Tennessee; Kennecott Utah Copper Corp., Utah; Aberdeen Pesticides, North Carolina
environmental justice, demographics component	1994-1995	census blocks, 1200 NPL sites
demographic updating and mapping	1995	selected NPL sites
primary brain cancers	1994-present	census blocks; Virginia, Florida, California, Massachusetts, New York, Pennsylvania
cancer mortality and birth outcome	1995	census division, 8 counties, Washington state

Disease, environmental condition or related topic	Time period	Geographic scale
health statistics reviews	1994-present	variable NPL sites (census tract, municipality)
release of 1,3 Butadiene	1994-present	Robstown, Nueces county, Texas; point, plume, census areas, ZIP Code
groundwater flow and chemical transport	1994-1995	Groton, Massachusetts and Elkhart, Indiana
vital events	1993-present	Georgia; county, health district, perinatal regions, block
violent crimes	1994-present	convenience stores, Virginia; selected counties, census tracts and block groups
birth defects	1995	five county metro Atlanta; street address and ZIP Code
primary and secondary syphilis	1990-present	national, U.S. counties
spatial analytic methods (cancer rates)	1995	Ohio and Idaho
geocoding, National Health and Nutrition Examination Survey	1994-present	national, U.S.

Source: Survey of CDC/ATSDR GIS Users Group, 1995. Compiled by Charles M. Croner, Office of Research and Methodology, National Center for Health Statistics.

#### VI. Public Health GIS Literature

(This will become a new section to include past or current citations submitted by GIS Users--note below examples; for the next edition, please submit **1995-1996** citations only)

Glass, G., Schwartz, B., Morgan, J., et.al. Environmental risk factors for Lyme disease identified with Geographic Information Systems. *Am J Public Health* 1995;85:944-948.

Waller, LA. Geographic information systems and environmental health. *Health and*

*Environmental Digest* 1996(Feb);9(10):85-8. [Also note related commentary: Marbury M. GIS: New tool or new toy? *Health and Environmental Digest* 1996(Feb);9(10):88-9.]

#### VII. Attachments

A. Excerpts of interviews with James Plasker, Associate Chief of the National Mapping Division, USGS, and Joel Morrison, Chief of the Geography Division, Bureau of the Census, appearing in the *ACSM Bulletin*, January/February (159), 1996:

James Plasker (USGS): It has been a very turbulent year in the geospatial data community across the board and definitely at USGS. We began the calendar year after the November 1994 election with the Contract with America and some of the legislation that was derived from that including the proposal to abolish the Geological Survey.

I think we were lucky to be one of the early agencies targeted. It probably put us in the position of being able to voice our concerns clearly without a lot of other static on the airwaves. And, because the USGS has a very long history and a very broad mission, we have a lot of supporters in all sectors...

We started this year with a 100 percent budget cut--in other words, being abolished. By the time we got to the Appropriations Committee hearing, they were saying that we would survive, but perhaps take a 20 percent cut. That was later reduced to perhaps a 10 percent cut. By the time we got through the appropriations process and both Houses of Congress had marked up on the appropriations bill, we were actually level-funded across the agency, and in fact the mapping area had a slight increase. So we went from a minus 100 percent to a minus 20 to maybe a plus 1 or 2. In today's environment, that's pretty good. But it was a long, emotional roller coaster, and the ride won't be over until the appropriations bill is signed by President Clinton.

Joel Morrison (Census): First of all I must state that with all the political infighting in Washington today, including the recent furloughs and threatened elimination of the Commerce Department, Congress, by itself, cannot cancel the Census because it is mandated in the Constitution--Article I, Section II: "There will be every 10 years, a total numeration of people." The primary purpose for the enumeration is that we have to redistrict 435

legislative seats in the House of Representatives. So Congress has a direct vested interest in the Census.

As far as geography and mapping in the census, the big event of the '80s leading to the '90s census was the construction of the TIGER files. I think of the TIGER files as the prototype of the NSDI (National Spatial Data Infrastructure). From a use point of view--everybody uses them.

During this decade and leading to Census 2000, we have to update and modernize TIGER. We have to try to more accurately position the street network, that is, increase the positional accuracy of it. We have to identify new buildings, streets, and attributes that were not there in 1990. And, the big thing of this decade is the creation of a master address file (MAF), which will geoposition all the viable addressable housing units in the country, and tie them to TIGER. Eventually the MAF will include businesses as well as residences.

Leading to the 1990 census, the Bureau also had to create a list of addresses of all the habitable units--in excess of 100 million. Congress, last year, passed a bill that allows the Postal Service [USPS] to share its delivery sequence file (DSF) with the Census Bureau. It required an Act of Congress to share and maintain a file which both need, but which was previously prevented by law.

More than 92 million "city-style" address (with complete street names and numbers) were contained in the DSF [but ungeocoded]. We geocoded 76 million on the first pass, which really is quite good. It gives us a 70+ million base, with some 50 million more to worry about between now and 2000. Only 20-some million of these addresses are city-style, however. Most of them are in rural areas.

Once we have the MAF on file, it will allow us to do continuous measurement. The Census Bureau continuous measurement

program will take, for example, a sample of a quarter million people each month during the next decade, and ask a series of questions. These surveys will be on a cycle such that estimates of needed parameters about our ever-changing American community are continuously being statistically updated throughout the decade.

B. Excerpts from the Public Health Service, NATIONAL COMMITTEE ON VITAL AND HEALTH STATISTICS, Subcommittee on Mental Health Statistics, September 13, 1995, Washington, D.C.:

UPDATE ON MENTAL HEALTH CONTENT IN THE 1996 NHIS: The core, or basic module, of the NHIS [National Health Interview Survey] will be in the field every year. It has three parts: a Family Core, with questions for every member of the household; a Child Core to be administered to a sample child, and an Adult Core for a sample adult. The survey will use pure random selection. (Dr. Manderscheid remarked that this could be augmented by oversampling people in the family who have particular conditions).

Ms. Simpson reiterated that the periodic module will be administered along with the basic module, probably every three years. The former is to take 20 minutes and the latter, 40 minutes. Several participants predicted a discontinuation of the traditional practice of expanding data collection to accommodate the agencies helping to pay for the HIS. Dr. Manderscheid suggested that the Subcommittee hear a presentation on the integration of NHIS and the NMES. The latter will be moved to NCHS as part of the Departmental reorganization, probably in 1996. He agreed with Dr. Zill that information from the periodic module can be used to enhance the NMES sample...

Dr. Wan asked about geographic

identifiers (zip code or county code), noting the Committee's longstanding interest in identifying community resources and assessing the general pathology in a given area. Ms. Simpson said that sensitive data will be geocoded, and Dr. Williams wondered if the geocoded data would be publicly available to authorized researchers. Ms. Simpson said she would contact Dr. Wan with the names of researchers who have been developing local small area estimates...

He [Dr. Zill] asked about questions on program participation, and was told that the survey will get information on SSI, SSDI, AFDC, food stamps, and possibly workman's comp, as well as health insurance coverage. Dr. Manderscheid described discussions about the possibility of having HCFA link the data files with Medicare and Medicaid and/or Social Security records, using the usual HCFA procedures. It may be possible to get access to a public use tape without identifiers. No plans have been made for these linkages.

Dr. Zill noted that the Subcommittee has agreed to write a letter to the Secretary supporting the inclusion of adult and child mental health items in the HIS. Dr. Manderscheid said the Center for Mental Health Services has appreciated working on this project with NCHS and is pleased with the outcomes of the joint effort.

C. Excerpts from the Public Health Service, NATIONAL COMMITTEE ON VITAL AND HEALTH STATISTICS, Subcommittee on Health Statistics for Minority and other Special Populations, December 7, 1995, Washington, D.C.:

She [Dr. Carter-Pokras] next discussed aspects of the review of OMB Directive 15 categories and a possible multiracial category. A project to get information on multi-racial and/or Hispanic mothers who have had babies during the last three years is having difficulty

getting enough subjects, and the contractor has advertised for participants. Project Race has changed its stance from favoring a stand-alone multiracial category to one with a follow-on check-off question.

OMB predicts a decision in mid-1997, and in January 1996 will issue a report on the findings of the Current Population Survey Supplement. One finding of that survey was that the addition of a stand-alone multiracial category would reduce the population counts of American Indians and Alaska natives by about 25 percent. A combined format for race/Hispanic ethnicity would reduce the Hispanic population count by 20 to 30 percent. The Hispanic advocacy groups (e.g., La Raza) that had advocated such a change revised their positions when they learned of these findings. Dr. Carter-Pokras noted that OMB will have a difficult time balancing the political concerns of advocacy groups and research findings. For example, when Hispanics were asked for their preference, they favored the combined format (not knowing its potential impact on population counts). The CPS also found that the term "African American" is still not the preferred term.

The National Content Survey is asking for comments on the race/ethnicity targeted test, and Dr. Carter-Pokras said she would get copies to Subcommittee members, along with the Federal Register notice and the directory of minority health data sets funded by AHCPR and ASPE. She noted that there is wide variation in the level of detail gathered on social and cultural data. There are plans to bring to the attention of the HHS Data Council the fact that some legislators are talking about limiting eligibility for some programs to people who have been U.S. citizens for a certain number of years. This information is not currently available in HHS statistics, and it might be needed to monitor the effect of eligibility changes. Dr. Williams

suggested that this distasteful prospect might best be faced in the unlikely event that it becomes a reality.

Returning to Directive 15, he asked if anyone in the Office of Minority Health was monitoring state implementation of requirements for the provision of multiracial status, such as in Michigan. Dr. Carter-Pokras said an intern in her office is doing so, contacting data system managers in every state. Dr. Williams asked that she be invited to make a presentation at the Subcommittee's February/March meeting (prior to full NCVHS meeting)

Ms. Golden noted that the Department of Education surveyed states on their race/ethnicity information practices. She speculated that there is quite a gap between data collection requirements and actual implementation, especially on the part of local health departments with limited resources.

Turning to another issue, Dr. Williams noted concerns about the National Health Interview Survey and other surveys, particularly in that data are not available to characterize some of the immigrant populations, notably Hispanics. NCVHS has endorsed a Subcommittee recommendation encouraging the Department to collect data on sociocultural factors, and Dr. Lee has written Ms. Jones expressing interest in discussing this matter. In preparation for such a discussion, the Subcommittee needs to review the redesign for the NHIS to see what questions have been retained, added, and dropped. Dr. Carter-Pokras suggested that the full Committee have a briefing on the NHIS and the NHANES as well as the Lee Report. She noted that the planning funds for the next NHANES have not been appropriated.

The group then reviewed the NHIS questions in the light of the Subcommittee's priorities. Dr. Parsons promised to update the Subcommittee by e-mail on the final version. It

was noted that respondents can check all that apply in terms of race/ethnicity, and then choose which group best represents them. Ms. Golden noted concerns about the fact that people born in Puerto Rico are identified as foreign-born. The following topics are addressed by current questions: national origin, ancestry, place of birth, year of migration if foreign born, language used at home, and how well English is spoken. Dr. Carter-Pokras observed that the Subcommittee had recommend an indication of the language of the interview if other than English, and this has been dropped. Dr. Parsons said there are no plans to translate the NHIS into Spanish, contrary to another Subcommittee recommendation.

Asked about the status of the NHIS redesign, Dr. Parsons said that the version going into the field in mid-1996 is final. A CAPI version will be instituted in 1997, and some changes may be incorporated in that version until the end of 1996.

Ms. Aldridge noted the absence of questions relevant to understanding instructions and gaining access to care: ability to read English and understand it, and frequency of visiting one's native country for the purpose of getting medical care or prescription drugs.

Overall, the group declared this version of HIS to be an improvement over the previous one. Dr. Parsons explained that the redesigners intended more far-reaching revisions, but were

limited by space and time constraints.

**SUBCOMMITTEE ANNUAL REPORT AND WORK PLAN:** Regarding the work plan, Dr. Williams agreed with Ms. Golden that the Subcommittee should monitor the Bureau of the Census's plans for the year 2000. He noted the pressure on the Hill to abandon the long form, despite the fact that every item on it was mandated directly or indirectly by Congressional legislation. The Bureau has said it will spend less time and energy reaching every U.S. household, but will instead do a sample survey of non-respondents and then project the estimate. This will save considerable money, and result in better data. The final census number will be based on both survey results and actual responses.

Dr. Williams noted that the Subcommittee had never followed up on its interest in requesting that NCHS do a report on Asian Americans, accumulating several years of HIS data. Also, the Subcommittee has discussed recommending that *Health U.S.* or another NCHS publication present data cross-tabulated by race or SES. These issues should be revisited, and may be addressed in a preliminary way in the discussion with NCHS staff planned for the February/March meeting. Steve Botman, John Barlow or Azil Bradi were suggested as sources of information, including on confidentiality.

**Chuck Croner, Editor, GIS NEWS AND INFORMATION, Office of Research and Methodology, National Center for Health Statistics**

**Spring '96... stay in GIS touch**