Getting the Map into the Computer: GIS Data Development

- 4.1 Analog-to-Digital Maps
- 4.2 Finding Existing Map Data
- 4.3 Digitizing and Scanning
- 4.4 Field and Image Data
- 4.5 Data Entry
- 4.6 Editing and Validation

GIS maps are digital not analog

- Maps have a **communications function** but...
- A map has a **storage function** for spatial data
- Somehow, the visually "stored" data must **get digital**
- Real and Virtual maps

GIS Data Conversion

- Traditionally **most of the cost** of a GIS project
- This is a **one-time cost**
- Depends on **reuse**
- Requires maintenance

GIS Data Development Approaches

- Research for existing data
- If analog maps exist, **convert analog maps to virtual maps** (GIS data)
- If no analog maps, aerial photography, remote sensing, GPS and ground survey usually are obtained.

Finding Existing Map Data

- Map libraries
- Reference books
- State and local agencies
- Federal agencies
- Commercial data suppliers e.g. GDT, Thompson, ETAK

Existing Map Data

- Existing map data can be found through a **map library**, via network searches, or on media such as CD-ROM and disk.
- Many major data providers make their data available via the **World Wide Web**, a network of file servers available over the Internet.
- GIS vendors package data with products.

Commercial vendors



Commercial vendors

- www.navteq.com
- www.teleatalas.com
- www.geographynetwork.com
- www.google.com
- www.mapquest.com

Federal Data Agencies

- USGS
- NOAA
- Census Bureau
- NIMA
- EPA
- many more...

National Spatial Data Infrastructure



National Spatial Data Clearinghouse

📥 Netscape: MetaStar Gatew	vay – Test Search Page						•
File Edit View Go Com	imunicator						Help
Back Forward Delead Llaws	🧀 👫 🎯	Coouvitur Otan					N
Back Polward Reload Home	: Search Guide Finit	2.184/servl	t/FGDCServlet				
				ANIE			4
	nywnere 🗆			AND			
	Title 🗆			AND			
	Edition 🗖			AND			
A	Abstract 🗖			AND			
A	nywhere 🗖			AND	-		
<u>Step 3 - Numeric Query Ter</u>	<u>ms (Optional):</u>						
Field	Relation	n	Criter	ia		Boolean	
Scale of Source \Box	Equal To					AND =	1
Cloud Cover 📼	Equal To						
Step 4 - Select Databases:							
						T	
Africa Data Disseminati	on Service I Data Clearinghour	20			lŕ	ì	
Australia – ERIN Enviror	nmental Data Direc	tory					
Australia – ERIN Index							
Australia – Victorian Spatial Data Directory							
BTS National Transporta	ation Atlas Database	ogator Online					
Biological Resource Maps, Costa Rica							
California CERES/CGIA	Test Server						
Caribbean Environment	Programme	loaringhouse					
CIESIN/EPA – Federal F	Facilities Public Heal	lth Assessme	nt Data Access Sv	stem Catal	og		
CIESIN/USDA – Global Environmental Change Data Assessment and Integration Catalog							
CIESIN/EPA – Great Lal	kes Environmental I	nformation Sy	stem Catalog				
CIESIN/NASA – Socioed	conomic Data and A	applications C	enter (SEDAC)		7	×	5
						U 🛞 🖉	i de 🍫

USGS: National Mapping

Netscape: U.S. Geological Survey Home Page		
File Edit View Go Communicator	Help	
🗳 🐳 强 🏦 🚈 🛋 🛋	Stop	
😻 Bookmarks 🕭 Location: http://www.usgs.gov/	🔽 🕼 What's Related	
🥒 🗶 The Weather Channel – Santa Barbara, CA 🥒 Santa Barbara MTD Winche	ester Canyon 🥒 HotBot .	
Science for a changing world Home Biology Geology Mapping Water Products Glossary Site Map Searce	Netscape: EPOS Data Center, Sieux Falls, SD	
	File Edit View Go Communicator	Help
U.S. Geological Survey (<u>Text Version</u>)	Beckmarks & Lossian: http://edowny.or. usos.gov/	
About Us Contact Us Search Clear	👻 🐨 🕄 🏦 🧟 🗟 🕸 📽 🕄 🏭 Back Finnerri Relaad Hone Search Netscape Print Security Shap Shap	
Rowes Our Tanice	🖉 Google 🥒 "Mekyl Homepege" 🥒 Santa Bartsara MTD 🥒 UCSB Dept of Geograpy Discussion Board	
WHAT S REW Congress USGS Information by State EDUCATION Other Government Floods JOBS Scientists Maps LIBRARY Teachers and Students Volcances PARTNERSHIPS More topics	Earth Resources Observation Systems (EROS) Data Center STE SUNCI Products Solence NASA DAAS Solence National Archive About ERDS	
USGS Headlines Explore Our Products and Data Forecasting Tomorrow for Better Decisions Today: FY 2001 Ask USGS - ESIC Atlas of Antarctic Research CINDI - Natural Disaster Info Transition of the test of the test of the test of test of the test of t	Actial Mag. Execution Satellite Land Cover	
The Future of Planet Earth, and the Challenges We Face in the Coming Millennium – A Message from the Director <u>Text Audio</u> <u>Audio</u>	Featured Site The USSS Obdal Visialization Yawar <u>Nitro Motors and any</u> is a spick and any search and orderisod for saleshed sale He data. The vision-alogs agentized by access to all available degree brocks images from the Landard 7 (2104) and Landard 45 (110)	
Wigging United Factories Carlor Applications Now Mapmoer USGS Research on Historic Streamflows Maps La Niña's Impacts Minerals Publications & Data West Nile Virus Update – Wildlife Plays Key Role National Attas USCS Ta Ca approce Mosting on Large Scale Consequention NBII - Biological Data	Set suis and other may and them the USSIS ERO'S Data Center machines	
be held June 9–12, 2000, in Missoula, Montana Water Use 100%	Edit UE 05: HOME Electory Gard agy Masping Water Sm. Add Construction of Management Address of Suprem Add GB Planae and Management Address of Suprem Add GB Planae and Management Address of Suprem Add GB Planae and Construct Address of Suprem Address of Suprem Add GB Planae and Construct Address of Suprem Address of Suprem Add GB Planae and Construct Address of Suprem Address of Suprem Add GB Construct Construct Address of Suprem Address of S	
	2° 1076 11 56 42 59 62	2

U.S. Bureau of the Census



The Decennial National Census

- •This is a **survey based, national-scale** collection of data, conducted by the U.S. Census Bureau (a Federal agency) every 10 years
- •This provides a snapshot of a wide range of **socio**economic data for the entire nation, that can be used in two ways:
 - •Comparisons can be made to conditions as captured by a previous census so that **change in time** can be studied <u>AND</u>

•The data can be analyzed **spatially**, using levels of geography from local to national scales

Census 2000 Short Form and Long Form



Census Data as Spatial Data

- •The information collected on the surveys form the **attribute part** of the data
- •Individual surveys are aggregated spatially to **geographic units**, at the various levels of census geography, that uses a nested scheme that builds higher-level units out of lower ones

Small-Area Geography Overview

Census Small-Area Geography



Geographic Products: The TIGER Data Base

> Topologically Integrated Geographic Encoding & Referencing

The source of **ALL** census geographic products

Basic TIGER/Line File Topology



One census block: 3 GT-polygons 1 point landmark (school) 1 area landmark (park)

NOAA Weather and other data



David Tenenbaum - EEOS 265 - UMB Fall 2008

Eros Data Center

- Distributed active archive center
- Sioux Falls,
 SD
- Operated by USGS



US GeoData ftp access to DEM DLG GNIS **GIRAS** etc.

Location: http://edcwww.cr.usgs.gov/doc/edchome/ndcdb/ndcdb.html

What's New?
What's Cool?
Destinations
Net Search
People
Software

US

GSU.S. Geological Survey - National Mapping Information - EROS Data Center

US GeoData

NOTICE: Effective July 1, 1996, the U.S. Geological Survey will no longer provide the Digital Line Graph (DLG) data in the Standard format. The Optional (native) format will continue to be available. The newer Spatial Data Transfer Standard (SDTS) format availability will continue to increase as data are converted and/or produced in this format.

FTP File Access

During the upcoming year, the EROS Data Center will be providing file transfer protocol (FTP) access to a variety of USGS digital data sets. The following is a list of each data set, a link to **data examples**, a hyperlink to **user guide** information, and a list of the methods available for transferring the data using FTP. You can download files by connecting to the data set's anonymous FTP account and selecting files from the existing directory structure (**FTP from account**), by selecting from a list of states that will guide you to the appropriate files (**FTP sorted by state**), or by using the graphic option which allows you to select files using an index map (**Graphics FTP**). Not all FTP options are available for each data set. The **user guide** provides detailed information about the data set including distribution formats.

If you use **FTP from account** make sure to review the **README** (or 00README) file located in the first directory you encounter.

<u>Public Domain Software</u> for use with DEM, DLG and LULC data <u>Spatial Data Transfer Standard (SDTS)</u> information.

1:250,000-Scale Digital Elevation Model (DEM)



Geographic Na	mes Informatio	n System Query H	Results	NTV CAL	TEODNIA SCU	20
	OF CALIFORT	IIA, SANTA BAR	CBARA COU	NTI, CAI	IFORNIA SCH	50.
Feature Name	: University Of (California At Sant	a Barbara			
Feature T	ype: schoo					
State 	Ca	ounty				
Californi	a Sá	anta Barbar	a			
Latitude	Longitude	USGS 7.5	' x 7.5'	Мар		
342458N	1195053W	Goleta				
<u>Show Feature L</u> site is busy, try t	<u>ocation</u> using map his <u>alternate map</u>	s produced from th <u>site</u> .	e U.S. Census	Bureau's T	ïger Map Server. If t	his
Find the Waters Watershed site.	<u>hed</u> for this featur	e using the U.S. En	vironmental Pro	otection Ag	ency's <u>Surf Your</u>	
	ry Form 🗖 Ma	nning Information	Home Page			
URL: <i>http//www</i> Maintainer: gnis_	-nmd.usgs.gov/c _manager@usgs.g	gi-bin/wow/GNIS ov	veb.DisplayCı	nty		
Last modified: 2	4SEP1996 rlb					

GNIS Feature locations

David Tenenbaum - EEOS 265 - UMB Fall 2008

Location: http://www.init.com/action.com/action/						
What's New?	What's Cool?	Destinations	Net Search	People	Software	

U.S. Geological Survey - National Mapping Information - EROS Data Center

Examples of 1:250,000-Scale Land Use and Land Cover (LULC) data.



Spokane, WA. Land Use/Land Cover, Census, and Political Units



Spokane, WA. Land Use/Land Cover, Hydrologic and Political Units

Go to Previous Page. Go to EDC Home Page

<URL:http://edcwww.cr.usgs.gov/doc/edchome/ndcdb/exampleslulc.html> Page owner: <edcweb@edcwww.cr.usgs.gov> Last modified: 31 December 1994.

🗾 🕺 🛛 Document: Done

⊠?

GIRAS Land Use and Land Cover Data

GEOCODING

- Geocoding is the conversion of spatial information into digital form
- Geocoding involves capturing the map, and sometimes also capturing the attributes
- Often involves address matching

Address Matching

- Most GISs contain capability for address matching
- **Start with** 100 Morrissey Boulevard, Boston, MA 02125
- End with coordinates (lat-long for UMB)
- May need to interpolate along blocks
- Street number range, **left and right side** e.g. 101-199

TIGER/Line Address Range Basics



The complete chain has two address ranges: left: odd-numbered right: even-numbered.

Potential address ranges along a complete chain have values that encompass the addresses of existing structures, as well as those not yet built

GEOCODING LEAVES A "STAMP" ON DATA

- The **method** of geocoding can influence the **structure and error** associated with the spatial information which results
- **Examples**: scanning (raster), digitizing (vector)

Geocoding methods for maps

- Digitizing
- Scanning
- Field data collection

Digitizing

- Captures map data by tracing lines from a map by hand
- Uses a cursor and an electronically-sensitive **tablet**
- Result is a **string of points** with (x, y) values

The Digitizing Tablet



- Digitizer cursor transmits a pulse from an electromagnetic coil under the view lens.
- 2. Pulse is picked up by nearest grid wires under tablet surface.
- 3. Result is sent to computer after conversion to x and y units.

Digitizing

- Stable base map
- Fix to tablet
- Digitize control
- Determine coordinate transformation
- Trace features
- Proof plot
- Edit
- Clean and build

Digitizing

- Cursor data entry
- Secondary tablet (menu/template)
- Voice command entry
- Point select
- Stream mode
- Distance mode

Selecting points to digitize



Some common digitizing errors

- Slivers
- Duplicate lines
- Duplicate nodes
- Unended lines
- Gaps
- Zingers

Scanning

- Places a **map on a glass plate**, and passes a **light beam** over it
- Measures the **reflected light** intensity
- Result is a **grid of pixels**
- **Image size** and **resolution** are important
- Features can "drop out"

Scanning

- **Types** of scanners:
 - Flat bed vs. Drum
- Characteristics
 - DPI & File size





Scanning example

- 15 x 15 cm (3.6 x 3.6 km)
- grid is 0.25 mm
- ground equivalent is 6 m
- 600 x 600 pixels
- one byte per color (0-255)
- 1.08 MB



David Tenenbaum - EEOS 265 - UMB Fall 2008

Field data collection





David Tenenbaum – EEOS 265 – UMB Fall 2008

Global Positioning System (GPS)

- A space-based 3-dimensional **measurement and positioning system** that operates using radio signals from satellites orbiting the earth
- Created and maintained by the US Dept. of Defense and the US Air Force
- The system as a whole consists of three **segments**:
 - satellites (space segment)
 - -receivers (user segment)
 - ground stations (control segment)
- Note: Russia and a European consortium are implementing similar systems.

GPS – Space Segment (Satellites)

- 24 NAVSTAR satellites in the GPS constellation
 - orbit the Earth every 12 hours
 - ~11,000 miles altitude (a very **high orbit**)
 - positioned in 6 orbital planes (4 per plane)
 - orbital period & planes are designed to keep 4-6
 satellites above the horizon at any time everywhere on the planet
 - controlled and monitored by five ground stations around the globe



20,200 km Altitudes, 55 Degree Inclination

GPS – User Segment (Receivers)

- Ground-based devices that can read and **interpret the radio signals** from several of the NAVSTAR satellites at once
- Use **timing** of radio signals to calculate the receiver's **position** on the Earth's surface
- Calculations result in varying **degrees of accuracy** that depend on:
 - quality of the receiver
 - user operation of the receiver
 - local & atmospheric conditions
 - current status of system



GPS – Control Segment (Ground Stations)

Peter H. Dana 5/27/95



Global Positioning System (GPS) Master Control and Monitor Station Network

- Five control stations
 - master station at Falcon (Schriever) AFB, Colorado
 - **monitor** satellite orbits & clocks
 - broadcast orbital data and clock corrections to satellites

GPS - How Does it Work?

- GPS allows us to determine a position by calculating the **distance** between a receiver and multiple satellites
 - Distance is determined by **timing** how long it takes the signal to travel from satellite to receiver
 - Radio signals travel at **speed of light**: 186,000 mi / sec
 - Satellites and receivers generate **exactly the same signal** at exactly the same time
 - Signal travel time = delay of satellite signal relative to the receiver signal



GPS - Satellite Signals

- Satellites have accurate **atomic clocks** and all 24 satellites are transmitting the same time signal at the same time
- The satellite signals contains information that includes
 - Satellite number
 - Time of transmission
- Receivers use an **almanac** that includes
 - The **position** of all satellites every second
 - This is updated monthly from control stations
- The satellite signal is received, compared with the receiver's internal clock, and used to calculate the **distance** from that satellite
- **Trilateration** (similar to triangulation) is used to determine location from multiple satellite signals

GPS - Trilateration



Start by determining the distance between a single GPS satellite and your position (**a sphere**)

1 satellite = sphere

Adding a second distance measurement to another satellite narrows down your possible positions to **a circle** where the spheres intersect



GPS - Trilateration Cont.

Adding a third satellite narrows down the position to **two points** where the three sphere intersect, and usually only one point is a **'reasonable'** answer



3 satellites = 2 points



The intersection of four spheres occurs at **one point**, but the 4th measurement is not needed, and is used for **timing** purposes instead

GPS - Using the 4th Signal

- How do we know that satellites and receivers generate the **same signal** at the **same time**?
 - The satellites have **atomic clocks**, so we know they are accurate
 - But **receivers do not** -- so can we ensure they are exactly accurate? No! If the receiver's timing is off, the location in 3-D space will be off slightly...
- So: We use the 4th satellite to **resolve any signal timing** error instead by
 - determining a **correction factor** using the 4th satellite
 - (like solving multiple equations ... there will only be one solution that satisfies all equations)

- Satellite errors
 - satellite position / orbit error
 - satellite clock error
- Atmospheric errors
 - Speed of electromagnetic waves in the atmosphere
 - Path taken by the signal
- Multi-path distortion errors
- Receiver errors
- (Selective availability)

- Satellite Errors
 - Although the satellites are in high orbits to minimize their deviations, sometimes there is a slight 'wobble' due to local gravitational forces



- While the atomic clocks used in the satellites are extremely accurate (and quadruple redundant), sometimes clock errors can occur
- These can contribute up to 1-5 **meters** of error



- Atmospheric Delays/Bending
 - The speed of light is only precisely 186,000 miles per second in a vacuum, and is slightly slower in the atmosphere, varying by composition
 - The signal can be **bent** as it moves through the atmosphere (sphere size based on a straight path)
- Up to 30m of error



- Multi Path Interference
 - The signal can bounce off of buildings, trees, etc.
 and this again distorts the time and distance between the receiver and the satellite
- Up to 1m of error



- Receiver Errors (Timing/Rounding)
 - Satellites have quadruple redundant atomic clocks that are accurate to nanoseconds (about \$800,000 of clock hardware on each satellite), e.g. "the time is 2:02:01.23456789012"
 - Receivers are powered by 4 AA batteries worth about \$2.99, generate their clock signal with an oscillating crystal that is sensitive to battery current, e.g. "the time is 2:02:01.2345"
- Up to 10 meters of error







Poor

• Satellite Coverage in Sky

– Position Dilution of Precision (PDOP)



The Role of Error

- **Enforcement** for map data is usually by **using topology**
- Map and attribute data errors are the **data producer's responsibility**, but the GIS user must understand error
- Accuracy and precision of map and attribute data in a GIS affect all other operations, especially when maps are compared across scales

Precision and Accuracy

- When describing error we need to distinguish between **two characteristics**:
 - Accuracy refers to the amount of distortion from the true value in a measurement
 - Precision refers to the variation among repeated measurements, and also to the amount of detail in the reporting of a measurement

Precision and Accuracy

•These related concepts are often confused:

- •**Precision** refers to the exactness associated with a measurement (i.e. closely clustered)
- •Accuracy refers to the extent of systematic bias in the measurement process (i.e. centered on the middle)



David Tenenbaum - EEOS 265 - UMB Fall 2008

Getting the Map into the Computer: GIS Data Development

- 4.1 Analog-to-Digital Maps
- 4.2 Finding Existing Map Data
- 4.3 Digitizing and Scanning
- 4.4 Field and Image Data
- 4.5 Data Entry
- 4.6 Editing and Validation

Next Topic:

What is where?

David Tenenbaum - EEOS 265 - UMB Fall 2008