Watershed Sciences 4930 & 6920
GEOGRAPHIC INFORMATION SYSTEMS

WEEK TWELVE –

UNCERTAINTY IN GIS

Joe Wheaton

HOUSEKEEPING

- Labs
  - Any questions/concerns
  - Interactive Maps
- Office hours today 1-3
- WATS 6915 – This is it!
  - Drop dead date Labs 1 -4: April 20th
- WATS 4931/6921
  - Project Proposals due tomorrow!
- WATS 4930/6920
  - Start on Spatial Analysis Next Week

TODAY'S PLAN… UNCERTAINTY IN GIS

I. **Uncertainty & Error**
II. GIS Errors
III. Error Propagation
IV. All Bad?
V. Summary of GIS Fundamentals

"But there are also unknown unknowns: the ones we don't know we don't know."
-Donald Rumsfeld

"It's not the things you don't know that matter, it's the things you know that ain't so."
- Will Rogers

UNCERTAINTY…

Lack of sureness about something… NOT a lack of knowledge.

To the general public and decision makers:

- Sign of weakness
- Like saying you don't know anything
- Confusing

To you and I (scientists):

- A statement of knowledge
- Useful information
- Full-employment act

MORE CONSTRUCTIVE DEFINITION

- Considered in terms of sources
- Provides a rationale for treating different sources differently

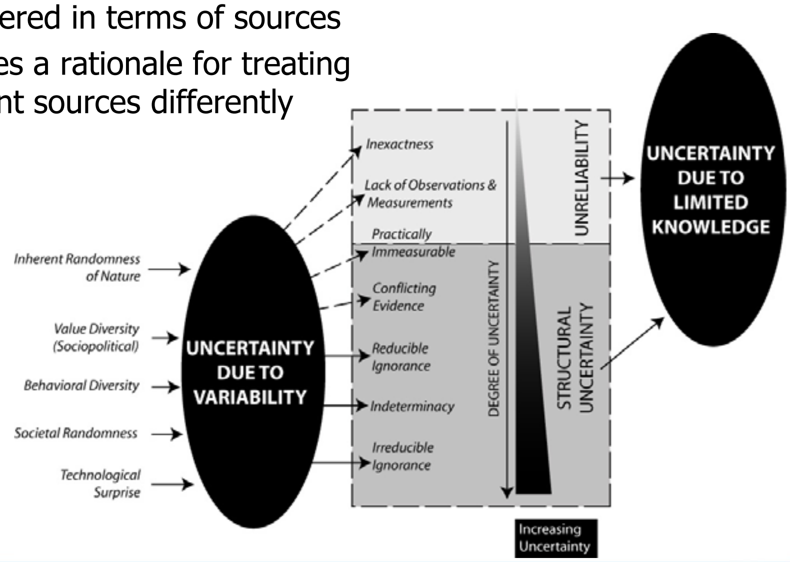


Figure Adapted from Van Asselt and Rotmans (2002): http://dx.doi.org/10.1023/A:1015783803445

UNCERTAINTY REVISED

- Uncertainty does not equate to a lack of knowledge
- A statement of uncertainty is not a sign of weakness… it is useful information
- ‘What in life is worth having that you didn’t have to take a risk to get?” – Mike Clark

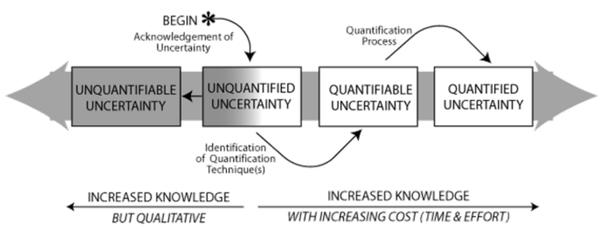


Figure from Wheaton et al. (2008)

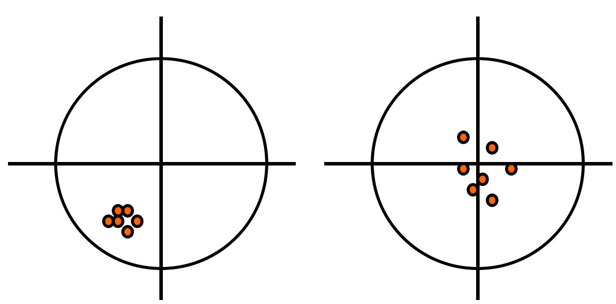
ERROR AND UNCERTAINTY

- Error
  - wrong or mistaken
  - degree of inaccuracy in a calculation
    - e.g. 2% error (calculated by difference between ***known*** & measured values)
- Uncertainty
  - lack of knowledge about level of error
  - unreliable

TODAY'S PLAN…

I. Uncertainty & Error
II. GIS Errors
III. Error Propagation
IV. All Bad?
V. Summary of GIS Fundamentals

ACCURACY AND PRECISION



The term **precision** is often used to refer to the repeatability of measurements. In both diagrams six measurements have been taken of the same position, represented by the center of the circle. On the left, successive measurements have similar values (they are *precise*), but show a bias away from the correct value (they are *inaccurate*). On the right, precision is lower but **accuracy** is higher.

GIS DATA ACCURACY

- *Accuracy* is how close an observation (or GIS data layer) is to the *truth*
- *Error* is the measure of how far a measure or observation deviates from the truth
- Many different ways to have errors
- Are they same?

UtahStateUniversity WATS 4930

ACCURACY

Positional Accuracy
- Spatial: horizontal and vertical: difference from true location and height
- Temporal: Difference from actual time and/or date

Attribute Accuracy or Consistency (a feature is what it purports to be)
- A railroad is a railroad, and not a road
- A soil sample agrees with the type mapped

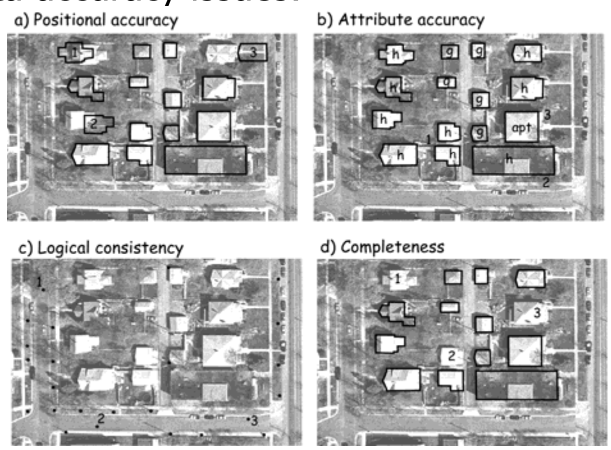
Completeness--the *reliability* concept
- Are all instances of a feature the GIS/map claims to include, in fact, there?
- Partially the criteria for including features: when does a road become a track?

Logical Consistency: The presence of contradictory relationships in the database
- Attribute
  - Some crimes recorded at place of occurrence, others at place where report taken
  - Data for one country is for 2000, for another its for 2001
- Lineage
  - Annual data series not taken on same day/month etc.
  - Data uses different source or estimation technique for different years)

UtahStateUniversity WATS 4930

OTHER WAYS TO BE WRONG…

- Spatial data accuracy issues:



From Chapter 14 of Bolstad (2008)

PRECISION OR RESOLUTION

<u>NOT</u> THE SAME AS ACCURACY!

Precision: the exactness of measurement or description

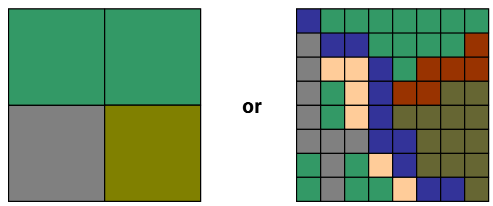
- the "size" of the "smallest" feature that can be displayed, recognized, or described

- For raster data, it is the size of the pixel (resolution)
- For vector point data, it is the point density

- resolution and positional accuracy
  - you can see a feature (resolution), but it may not be in the right place (accuracy)
  - higher accuracy generally costs *much* more to obtain than higher resolution

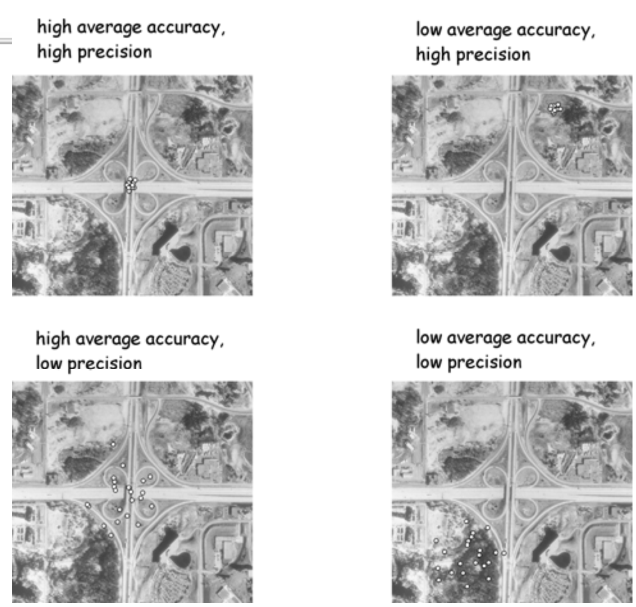
SPATIAL RESOLUTION

<u>Spatial Resolution</u>: The size of an object that can be resolved based on pixel size. Could also refer to the scale at which the data was collected.



FOUR END MEMBERS

- Positional accuracy of intersection of two freeways



From Chapter 14 of Bolstad (2008)

TEMPORAL RESOLUTION

Temporal Resolution: (also known as the repeat cycle) describes the frequency with which event occur or data is modeled or collected.

Every second

Every hour

Every day

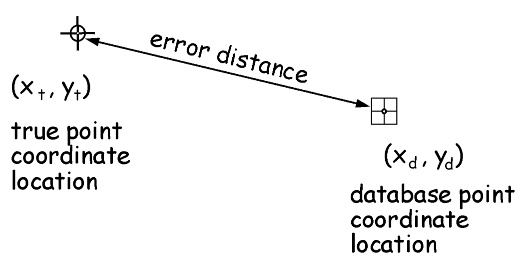
Every week

…

High and low temporal resolution depends upon the event occurring.

HOW POSITIONAL ACCURACY IS CALCULATED

- All you need is measured coordinates and ‘true’ coordinates
- The lower the error distance, the more accurate…

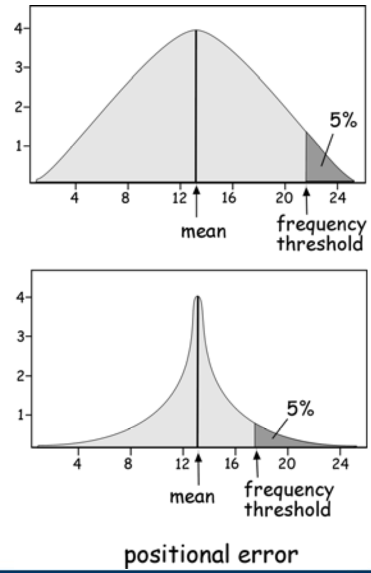


$$\text{error distance} = \sqrt{(x_t - x_d)^2 + (y_t - y_d)^2}$$

From Chapter 14 of Bolstad (2008)

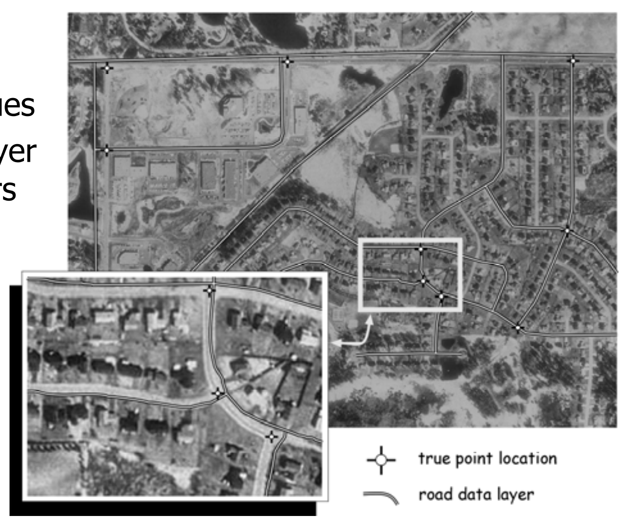
IMPLICATION OF ERROR DISTRIBUTIONS

- How would I get a plot like this?
- If we take 95% of the error…
- With same mean, but different distributions, implications are quite different…

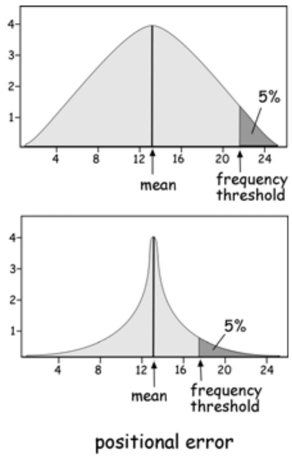


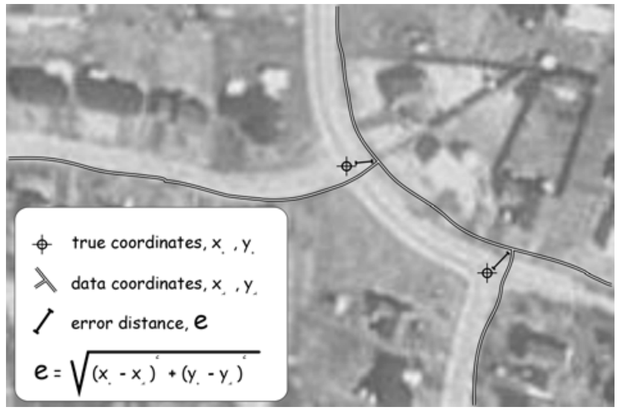
HOW TO CALCULATE THOSE POSITOINAL ERRORS

- Find, define or assume *true* values
- Find values of layer to calculate errors for
- Create error field
- Plug and chug
- THIS IS NOT TECHNICALLY CORRECT



A CLOSER LOOK





- 95% of the data… depends on distribution shape…

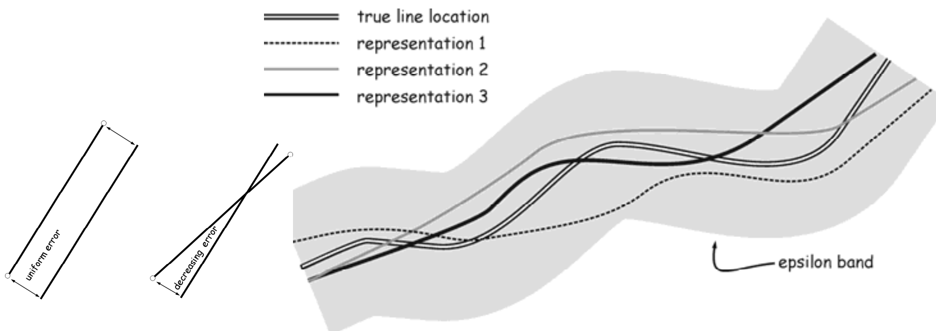
PUT IT ALL TOGETHER…

- Simple excel or field calculator exercise?
- How would you do it?

ID	x (true)	x (data)	x difference	(x difference)^2	y (true)	y (data)	y difference	(y difference)^2	sum x diff^2 + y diff^2
1	12	10	2	4	288	292	-4	16	20
2	18	22	-4	16	234	228	6	36	52
3	7	12	-5	25	265	266	-1	1	26
4	34	34	0	0	243	240	3	9	9
5	15	19	-4	16	291	287	4	16	32
6	33	24	9	81	211	215	-4	16	97
7	28	29	-1	1	267	271	-4	16	17
8	7	12	-5	25	273	268	5	25	50
9	45	44	1	1	245	244	1	1	2
10	110	99	11	121	221	225	-4	16	137
11	54	65	-11	121	212	208	4	16	137
12	87	93	-6	36	284	278	6	36	72
13	23	22	1	1	261	259	2	4	5
14	19	24	-5	25	230	235	-5	25	50
15	76	80	-4	16	255	260	-5	25	41
16	97	108	-11	121	201	204	-3	9	130
17	38	43	-5	25	290	288	2	4	29
18	65	72	-7	49	277	282	-5	25	74
19	85	78	7	49	205	201	4	16	65
20	39	44	-5	25	282	278	4	16	41
21	94	90	4	16	246	251	-5	25	41
22	64	56	8	64	233	227	6	36	100
								Sum	1227
								Average	55.8
								RMSE	7.5
								NSSDA	12.9

WHAT ABOUT POSITIONAL ACCURACY OF SHAPES AS OPPOSED TO VERTICIES?

- Compare true line location to various representations of actual to define epsilon band…



From Chapter 14 of Bolstad (2008)

MEASUREMENT OF POSITIONAL ACCURACY

- Usually measured by root mean square error*: the square root of the average squared errors*

- $RMSE = \sqrt{\frac{{e_1}^2+{e_2}^2+{e_3}^2+\cdots+{e_n}^2}{n-1}}$ where e_i is the distance (horizontally or vertically)between the tue location of point $_i$ on the ground, and its location represented in the GIS.
- Usually expressed as a probability that no more than P% of points will be further than S distance from their true location.
- Loosely we say that the RMSE tells us *how far recorded points in the GIS are from their true location on the ground, on average*.
- More correctly, based on the normal distribution of errors, 68% of points will be RMSE distance or less from their true location, 95% will be no more than twice this distance, providing the errors are random and not systematic (i.e., the mean of the errors is zero)

DIGITIZATION ERRORS

- Manual digitizing
  - significant source of positional error (roads, streams, polygons)
- Source map error
  - scale related generalization
  - line thickness
- Operator error
  - under/overshoot
  - time related boredom factor

LINEAGE.... REMEMBER METADATA?

- Identifies the *original sources* from which the data was derived
- Details the *processing steps* that led to its creation
- Rough outline of the route the data has gone to reach its current form
- Both impact its accuracy
- Both should be in the metadata, and are required by the *Content Standard for Metadata*

CURRENCY: *IS MY DATA "UP-TO-DATE"?*

- Data is always relative to a specific point in time
  - there are important applications for historical data (e.g. analyzing trends), so don't trash old data

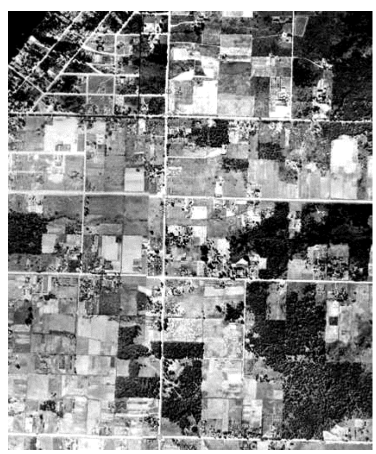
- Currency is not really an independent quality dimension; it is simply a factor contributing to lack of accuracy regarding
  - *consistency:* some GIS features do not match those in the real world today
  - *completeness:* some real world features are missing from the GIS database

Many organizations spend substantial amounts acquiring a data set without giving any thought to how it will be maintained!

ERROR – OUT OF DATE

- Belvue Washington… At one time it was 'right'

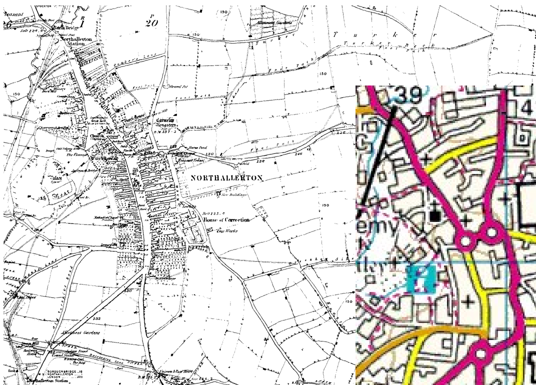
1997



1936

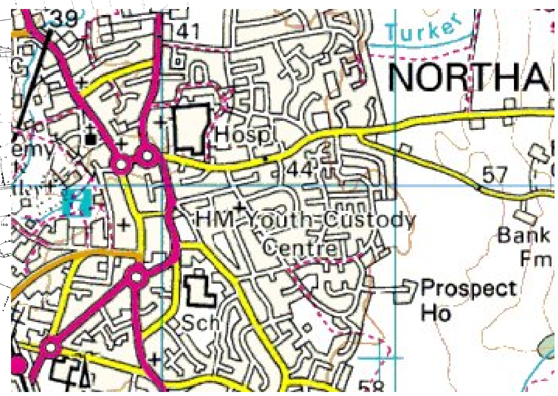
From Chapter 14 of Bolstad (2008)

DATA CURRENCY



Northallerton circa 1867

Northallerton circa 1999



ATTRIBUTE UNCERTAINTY

- Uncertainty regarding characteristics (descriptors, attributes, etc.) of geographical entities
- Types:
  - imprecise (numeric), vague or ambiguous (descriptive)
  - mixed up
  - plain wrong!
- Sources:
  - source document
  - misinterpretation (human error)
  - database error

IMPRECISE AND VAGUE

UtahStateUniversity

WATS 4930

MIXED UP

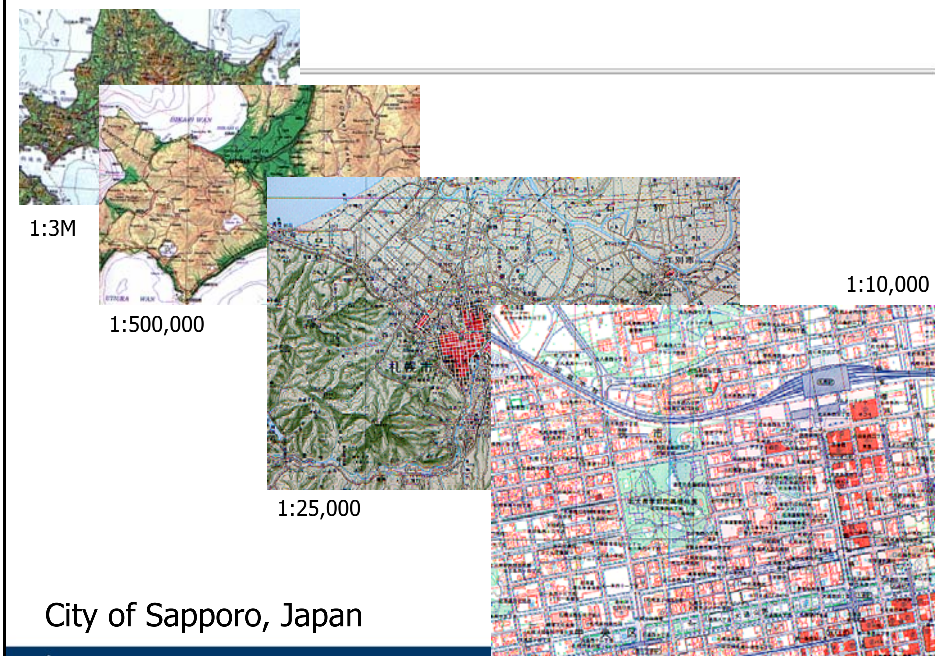
UtahStateUniversity

WATS 4930

JUST WRONG

GENERALIZATION

- Scale-related cartographic generalization
  - simplification of reality by cartographer to meet restrictions of:
    - map scale and physical size
    - effective communication and message
  - can result in:
    - reduction, alteration, omission and simplification of map elements
    - passed on to GIS through digitizing



City of Sapporo, Japan

UtahStateUniversity

WATS 4930

UNCERTAINTY IN ANALYSIS

- Just cause you think it will work does not guarantee success—Always LOOK at the results of your analysis!
  - What would a certain combination of inputs result in?
  - How is that likely to change across all inputs?
- Functional REDUNDANCY:
  - There is almost always another (often faster) way of performing any analysis
  - Should produce the *same* result… try it?

UtahStateUniversity

WATS 4930

HANDLING ERROR AND UNCERTAINTY

- Must learn to cope with error and uncertainty in GIS applications
  - minimize risk of erroneous results (QA/QC)
  - minimize risk to life/property/environment

- More research is needed:
  - mathematical models
  - procedures for handling data error and propagation
  - empirical investigation of data error and effects
  - procedures for using output data uncertainty estimates
  - incorporation as standard GIS tools

UtahStateUniversity WATS 4930

DATA QUALITY: *HOW GOOD IS YOUR DATA?*

- Scale
  - Can be an output issue; at what scale do I wish to display?
  - Analyses are only as good as the coarsest input
- Precision or Resolution
  - the exactness of measurement or description
  - Determined by input; can output at lower (but not higher) resolution
- Accuracy
  - the degree of correspondence between data and the real world
  - Fundamentally controlled by the quality of the input
- Lineage
  - The original *sources* for the data and the *processing steps* it has undergone
- Currency
  - the degree to which data represents the world at the present moment in time
- Documentation or Metadata
  - data about data: recording all of the above
- Standards
  - Common or "agreed-to" ways of doing things
  - Data built to standards is more valuable since it's more easily shareable

ERROR HANDLING 101

- Awareness
  - knowledge of types, sources and effects
- Minimization
  - use of best available data
  - correct choices of data model/method
- Communication
  - to end user via metadata, honest and thorough reporting of uncertainties

TODAY'S PLAN…

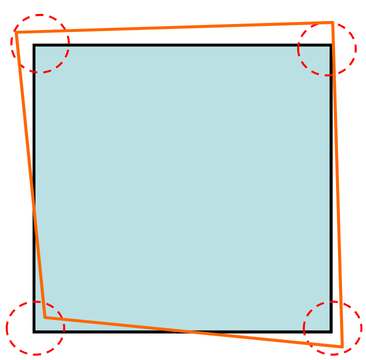
I. Uncertainty & Error
II. GIS Errors
III. Error Propagation
IV. All Bad?
V. Summary of GIS Fundamentals

UtahStateUniversity WATS 4930

ERROR PROPAGATION

- Methods for assessing the effects of known degrees of error in a model's inputs
  - Producing measures of confidence in model outputs
  - Normally by simulation

U3: UNCERTAINTY OF ANALYSIS
ERROR PROPAGATION



Error in the measurement of the area of a square 100 m on a side. Each of the four corner points has been surveyed; the errors are subject to bivariate Gaussian distributions with standard deviations in x and y of 1 m (dashed circles).

The red polygon shows one possible surveyed square (one *realization* of the error model).

In this case the measurement of area is subject to a standard deviation of 200 sq m; a result such as 10,014.603 is quite likely, though the true area is 10,000 sq m. In principle, the result of 10,014.603 should be rounded to the known accuracy and reported as as 10,000.

DEM DIFFERENCING

RASTER CALCULATOR….

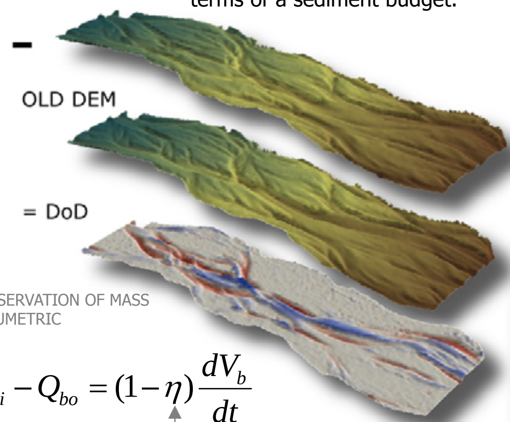
Simple method of quantifying spatial variations in change in storage terms of a sediment budget.

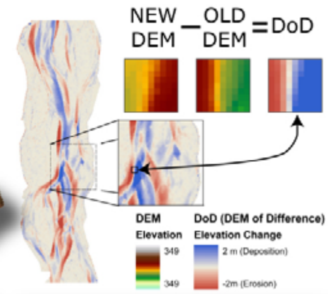
NEW DEM

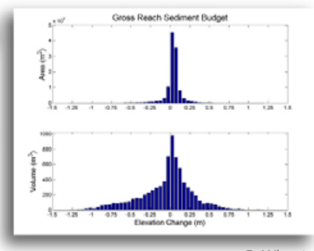
–

OLD DEM

= DoD







CONSERVATION OF MASS
VOLUMETRIC

$$Q_{bi} - Q_{bo} = (1-\eta)\frac{dV_b}{dt}$$

Volumetric rate of bed material transport

Porosity of bed material

Mclean & Church (1988) – Water Resources Research

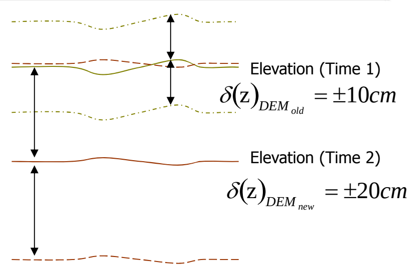
© Wheaton 2008

UtahStateUniversity

WATS 4930

MINIMUM LEVEL OF DETECTION

- Distinguish those changes that are real from noise
- Use standard Error Propagation
- Errors assumed to be spatially uniform, but can vary temporally



$$\delta(z) = \sqrt{\left(\delta(z)_{DEM_{old}}\right)^2 + \left(\delta(z)_{DEM_{new}}\right)^2}$$

e.g. $\delta(z) = \sqrt{(10)^2 + (20)^2} = 22.36$

22.36 cm ≈ 8.8 in

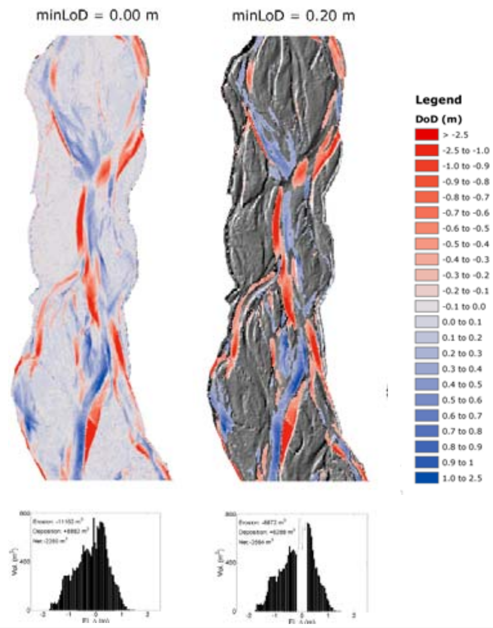
See
•Brasington et al (2000): *ESPL*
•Lane et al (2003): *ESPL*
•Brasington et al (2003): *Geomorphology*

UtahStateUniversity

WATS 4930

HOW DOES A $_{MIN}$LoD GET APPLIED?

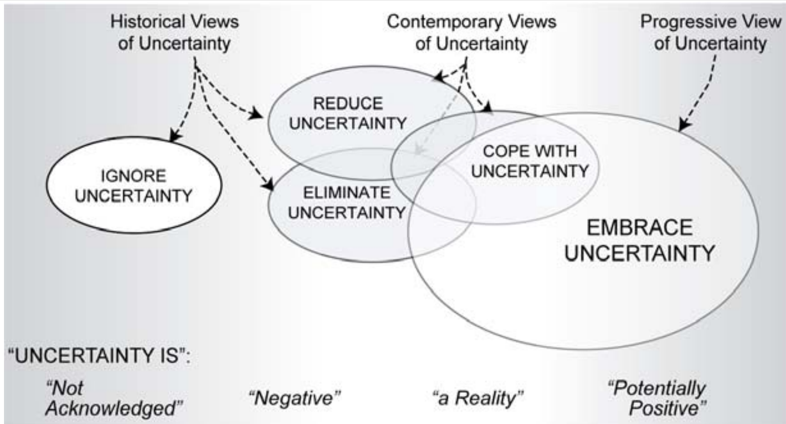
- You take original DoD, and remove all changes <= $_{min}$LoD
- For example +/- 20 cm
- How would you do that?
- What is the assumption here?



TODAY'S PLAN…

I. Uncertainty & Error
II. GIS Errors
III. Error Propagation
IV. All Bad?
V. Summary of GIS Fundamentals

PHILOSOPHICAL ATTITUDES TO UNCERTAINTY



These contrasting philosophical approaches to dealing with uncertainty are rarely explicitly identified.

(From Wheaton et al. 2008)

REDUCE UNCERTAINTY

- Uncertainty is a nuisance
- It should be constrained wherever possible
- Unquantifiable uncertainty difficult or impossible to constrain

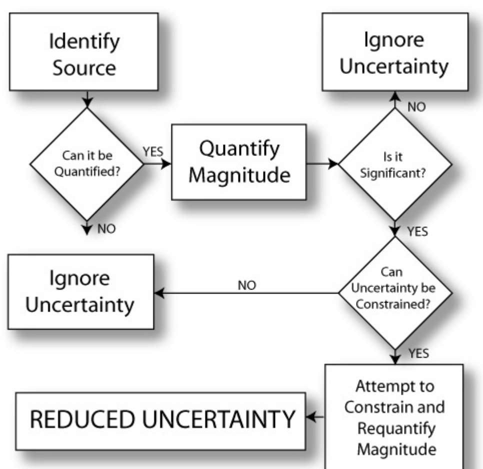


Figure from Wheaton (2004)

COPE WITH UNCERTAINTY

- Fuller appreciation of types of uncertainty
- Uncertainty still viewed as a nuisance
- Acceptance of uncertainty as a given
- Explicit link to adaptive management

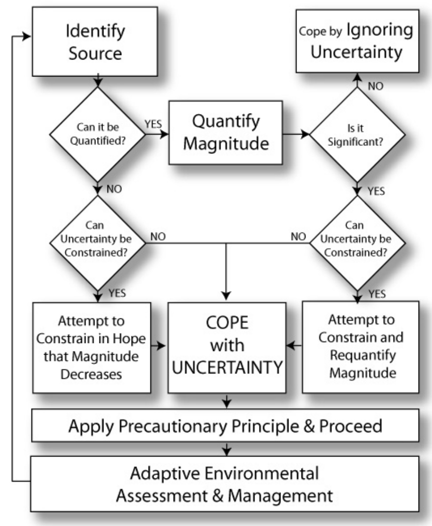
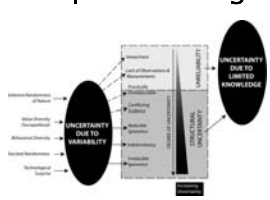


Figure from Wheaton (2004)

EMBRACE UNCERTAINTY

- Uncertainty seen as useful information
- Explicit recognition of uncertainty sources
- Use of natural variability as an opportunity
- Explicit linked to adaptive management



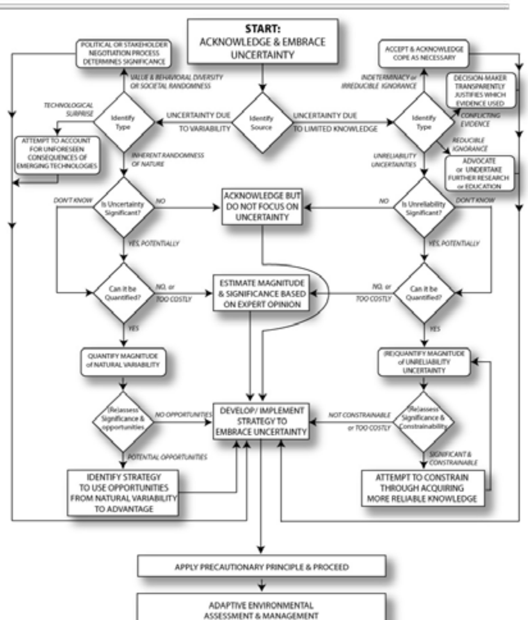


Figure from Wheaton (2004)

TRANSFORM UNCERTAINTY TYPES

- Central to embracing uncertainty
- Many examples of *structural uncertainties* & *uncertainties due to variability* can be transformed (and thereby reduced) to *unreliability uncertainties*

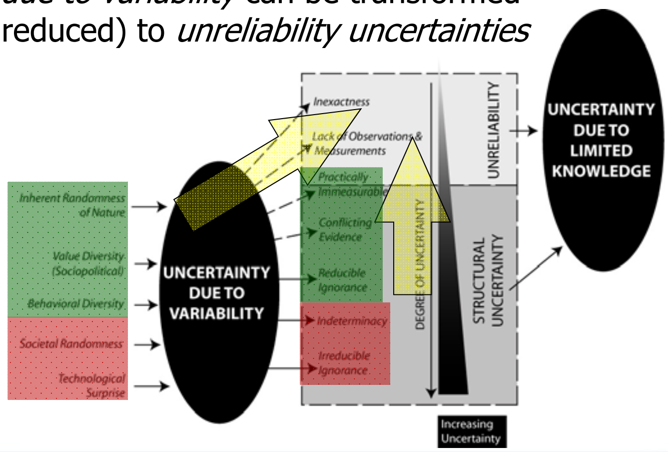
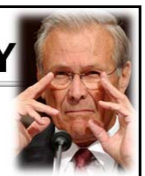


Figure Adapted from Van Asselt and Rotmans (2002): http://dx.doi.org/10.1023/A:1015783803445

HOW TO COMMUNICATE UNCERTAINTY WITHOUT SOUNDING LIKE A QUACK?



- Know the **audience** (general public vs. peers)
- Complete transparency of **source** and **type** of uncertainties
- Relate **significance** in terms of audience's criteria
- Clear identification of uncertainties leading to **risks** versus **opportunities** versus both
- Distinguish between **transformable** uncertainties & total unknowns (e.g. irreducible ignorance)
- Highlight **tradeoff** between cost of **knowing more** and taking **acceptable risks**

MIX OF COMMUNICATION OPTIONS

Method	Appropriate For
Qualitative Description	Unquantifiable and/or unquantified uncertainties
Probabilities	Expressions of confidence or likelihood
Measures of Variance	Uncertainties due to variability
Upper & Lower Limits (+/-)	Well constrained uncertainties due to inexactness
Fuzzy Numbers	Uncertainties due to vagueness and ambiguity
Scenarios & Conceptual Models or Simulation Models	Uncertainty about future (gets away from actual prediction)
Definition of Plausible Outcomes	Structural & Variability Uncertainties Leading to Predictive Uncertainty

TODAY'S PLAN…

I. Uncertainty & Error
II. GIS Errors
III. Error Propagation
IV. All Bad?
V. Summary of GIS Fundamentals

IF THIS FIRST FOUR WEEKS WAS SLOW

- Sorry….
- Repetition helps (even if a little boring)
- You forget…
- The fundamentals matter
- We'll pick up the pace now…
- Last six weeks will push you

UtahStateUniversity WATS 4930

WHAT FUNDAMENTALS?

- Introduction Review to GIS
  - Review of Maps (Cat in the Hat)
  - Intro WebGIS
- Abstracting World to Digitial Maps
  - Projections & Coordinate Systems
  - Data Types
- Data/Data/Data
  - Remote Sensing/Imagery Data Sources
  - Geoprocessing Intro
  - Editing & Attributing Data + Meta Data
  - Uncertainty in GIS

ESRI python

UtahStateUniversity WATS 4930

WHAT YOU SHOULD HAVE GOTTEN… (so far)

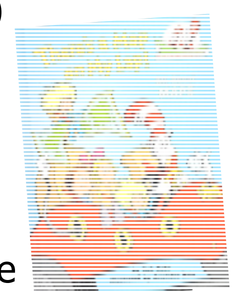
The above learning outcomes apply to the courses as follows:

Learning Outcome:	WATS 4930/6920	WATS 4931/6921	WATS 6915
1 - GIS Theory	Core	NA	Core
2 - Profiency in Spatial Analyses & Cartography	Core	Partial	Partial
3 - Self-Teaching & Troubleshooting	Core	Core	Partial
4 - Spatial Analysis in Research	NA	Core	NA
5 - Communicating with GIS	Core	Core	Core

- 1. **GIS Theory**: Understand the fundamental theory of Geographic Information Science behind Geographic Information Systems (GIS), and in so doing build an awareness of what GIS can and cannot be used for
- 5. **Communicating with GIS**: Become effective in building maps that can be shared with non-GIS users (e.g. PDF maps and interactive webGIS maps)

SO ALL OF YOU BETTER DAM WELL KNOW:

- How to make an effective map (6 C's)
- How to make an interactive map
- How to make a website
- Understand, read, convert coordinate systems and transform if necessary
- How to create, edit, query, manipulate and display vector data
- How to share GIS data
- ENOUGH GIS to be dangerous
- ENOUGH GIS to tell if someone else is dangerous



YOU ARE NOT DONE!!!!

- You should know how to teach yourself
  - GIS Help
  - Forums
  - Peers
  - ESRI Community
  - Self-Paced Courses
  - Follow up Courses

condensed format:

Taught Courses From ESRI ($$)

- ArcGIS Desktop I: Getting Started with GIS
- ArcGIS Desktop II: Tools & Functionality
- ArcGIS Desktop III: GIS Workflows & Analysis

Free Self-Paced Courses

- Getting Started in ArcGIS: webinar (9 hours)
- Using ArcMap in ArcGIS Desktop 10 webinar (3 hours)
- Other Free Training

Follow Up Courses

At Utah State University

There are too many courses to list, which employ GIS skills that you might learn in WATS 4930/6920 or an equivalent. However, here are a few follow ups that you might find useful:

Course	Title	Cr	Trm	Notes:
BIOL 4750/6750	Introduction to Computer Programming and Database Management for Ecologists	3?	Fa	See announcement here.
CEE 2240	Engineering Surveying	3	Sp, Su	Fundamentals of geomatics & tacheometric surveying.
ECE 5930	Small Satellite Imager Design	3	Sp.	If you want to learn more about the blimp platforms we covered in Lab 8, this is the class!
WATS 5300/ 6300	Remote Sensing of Land Surfaces	4	Sp.	Covers principles of remote sensing
WILD 5750/ 6750	Applied Remote Sensing		Fa	Learn image classification using Imagine
WILD 6900 (section 3)	GIS Programming with Phython I	1	Sp	This is a great follow up that focuses on geoprocessing and scripting in ArcGIS (1/2 Semester)
WILD 6900 (section 4)	GIS Programming with Phython II	1	Sp	This section focuses on Python scripting with OpenSource GIS libraries (1/2 semester)
WATS 6900	Restoration Monitoring: Geomorphic Change Detection	1	Su	This is an ICRRR short course I teach in Park City the week after finals. 3 days.
WATS 6900	River Bathymetry Toolkit	1	Su	This is a new 3 day short course on the River Bathymetry Tookit
Ecology Center	Landscape Genetics	?	Sp	Talk to Karen Mock for more information on this course

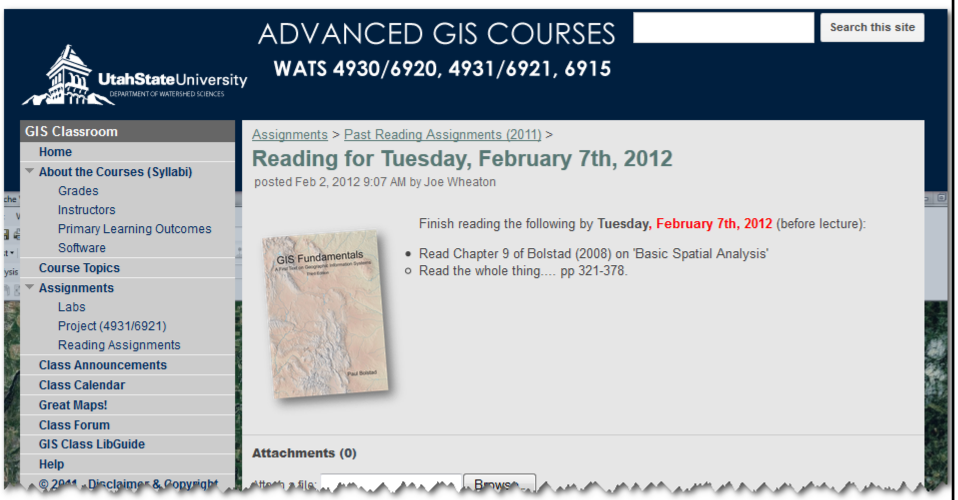
WATS 4930/6920... WHERE WE'RE GOING

- WATS 6915... welcome to tag along for any, all or none

Approximate Schedule

GIS Course(s)	Semester Week	Course Topics	Dates	Lab
4930 & 6920	WEEK 5	Vector Analyses	Feb 7 & 9	5. Vector Analysis
4930 & 6920	WEEK 6	Raster Analyses & Digital Elevation Modeling	Feb 14 & 16	6. Working w/ DEMs
4930 & 6920	WEEK 7	Raster Analyses & Uncertainty in GIS	Feb 21 & 24	7. Building DEMs
4930 & 6920	WEEK 8	GIS Modeling	Feb 28 & Mar 1	8. Morphometric Analyses
4930 & 6920	WEEK 9	Collecting Your Own Data	Mar 6 & 8	9. Georeferencing
Spring Break – March 12 -16				
4930 & 6920	WEEK 10	Collecting Your Own Data	Mar 20 & 22	10. Blimp Lab
End of WATS 4930/6920				

READING…


ADVANCED GIS COURSES
WATS 4930/6920, 4931/6921, 6915
Search this site
UtahStateUniversity
GIS Classroom
Home
About the Courses (Syllabi)
Grades
Instructors
Primary Learning Outcomes
Software
Course Topics
Assignments
Labs
Project (4931/6921)
Reading Assignments
Class Announcements
Class Calendar
Great Maps!
Class Forum
GIS Class LibGuide
Help
Assignments > Past Reading Assignments (2011) >
Reading for Tuesday, February 7th, 2012
posted Feb 2, 2012 9:07 AM by Joe Wheaton
GIS Fundamentals
Finish reading the following by Tuesday, February 7th, 2012 (before lecture):
Read Chapter 9 of Bolstad (2008) on 'Basic Spatial Analysis'
Read the whole thing…. pp 321-378.
Attachments (0)