

Slides by [Wheaton et al. \(2009-2014\)](#) are licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](#)

Watershed Sciences 4930 & 6920

GEOGRAPHIC INFORMATION SYSTEMS

UNCERTAINTY IN GIS...

& WATS 6915 WRAP UP

Joe Wheaton

HOUSEKEEPING



No Class Today... Another PODCAST

Joe Wheaton

Jan 31 at 9:36am

Good morning folks.

I might have misled some of you in class on Tuesday, when I said 'see you Tuesday'. I didn't look at the schedule to discern that today (Tuesday the 31st) was also a Podcast. So I should have said, 'see you next Thursday, the 2nd of February'. My apologies for the confusion.

As a reminder, by Thursday I expect that you will have watched the following Podcasts since our TLS lecture on Tuesday:

- [Raster Imagery, Data Sources & Intro to Remote Sensing](#) - Follow up to Tuesday, Jan 24th ([Topic ref](#))
- [Vector Data, Editing & Attributing Data + Metadata](#) - Thursday, Jan 26th Materials ([Topic ref](#))
- [Geoprocessing, Model Building & Scripting](#) - Tuesday, Jan 31st Materials ([Topic ref](#))

I am not assigning any reading to accompany these Podcasts, but you are welcome to reference your book for helpful follow ups to these podcast topics:

- Chapter 6 & 7 - For follow up to raster imagery and data sources
- Chapter 4 - For follow up to editing

The podcast for today on [Geoprocessing](#) (an ESRI term) is really important. We won't do any labs in this class that specifically require you to use geoprocessing, batch processing or scripting (to make your life easier and make tedious tasks go quicker). However, the real key to advanced geospatial analyses is knowing how to bolt multiple processes together to represent a workflow (i.e. model building in ESRI), and scripting. Chris Garrard offers an excellent series of Python Geoprocessing course in the fall (WILD 6900), that I encourage you to take as a follow up. However, you can (optionally) try out some of these skills on future labs to see how well you can put the basic concepts into practice to make the labs go quicker and more efficiently for you.

Please post any questions or comments about any of the podcasts to our discussion board: [General Class, Lecture or GIS Questions](#).

On Thursday, we will review these podcasts and provide a summary of the GIS Fundamentals we've covered thus far. We'll also discuss uncertainty in GIS and how to work meaningfully with it.

Best Wishes,

Joe



HOUSEKEEPING



Website permissions

Shannon Wing Belmont

Feb 2 at 5:25am

Hey everyone,

Please double check your websites and verify that you have shared permissions for all the bits & pieces.

I made around 25 permission requests for lab 2. Just started lab 3 and right off found blocked permissions.

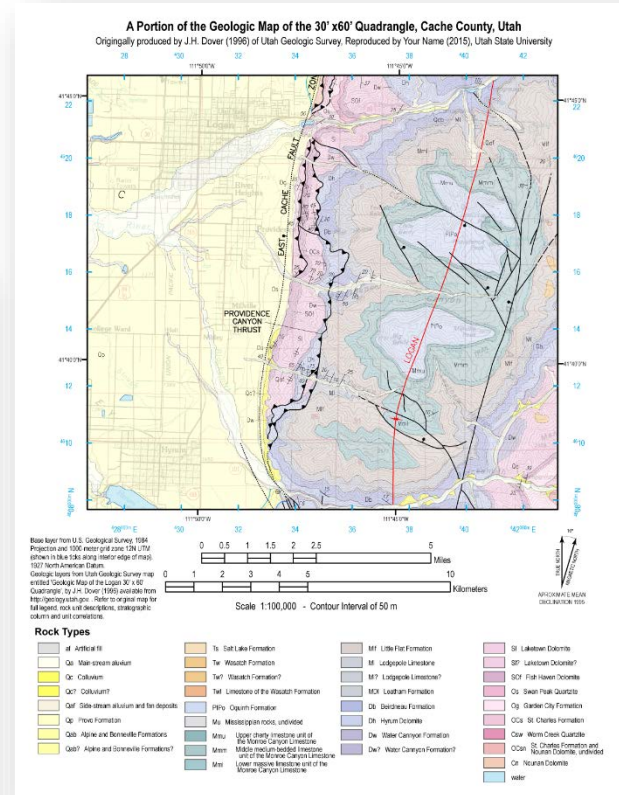
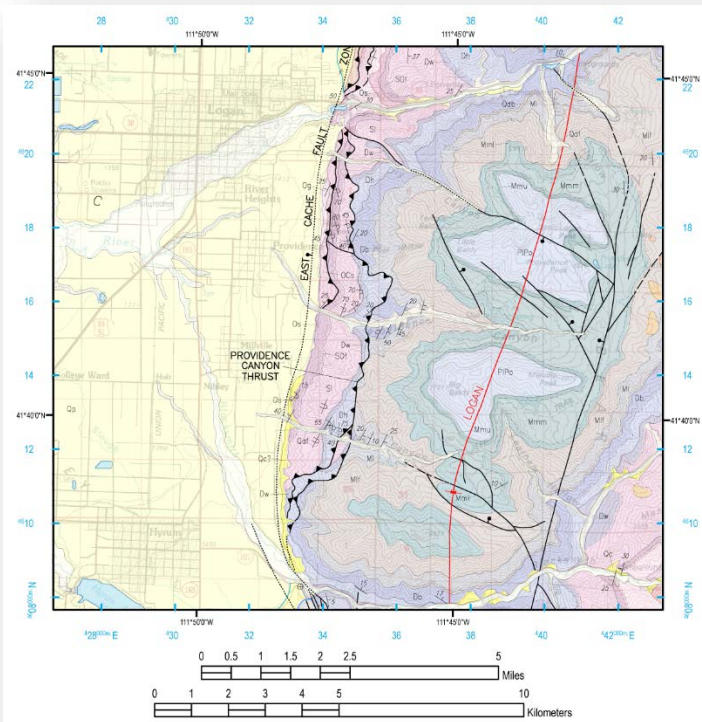
SMrS = Spare My remaining Sanity



:)

HOUSEKEEPING

- Lab 3
- Any questions?



THIS WEEK'S LAB

Lab 4: Digitizing, Editing, Sharing Data & Intro to Geoprocessing

- Edit existing vector data layers
- Create new vector data layers and edit them (e.g. digitizing)
- Edit and create tabular data
- Share vector and tabular data with other GIS users
- Learn how to use Model Builder & Geoprocessing

[Assignments > Labs \(4930/6920/6915\) >](#)

Lab 04 - Digitizing, Editing, Sharing Data & Intro to Geoprocessing

Contents

- 1 Introduction to Lab
 - 1.1 Background
 - 1.2 Lab Objectives
- 2 Lab Tasks Overview
 - 2.1 IMPORTANT
 - 2.2 Task 1 - Create Campground Design Map
 - 2.3 Task 2 - Share GIS Data for Campground Design in Different Formats
 - 2.4 Data for Lab
- 3 What to Submit
- 4 Additional Resources
 - 4.1 Web Links:

Key Links

[Corresponding Lecture Topic\(s\)](#)
[Week 03 & 04: Working with Data in GIS](#)
[See all Topics](#)

Lab Due Date:
Submitted before start of next week's lab on [Canvas](#)
[See Calendar](#)

Relevant Discussion(s):

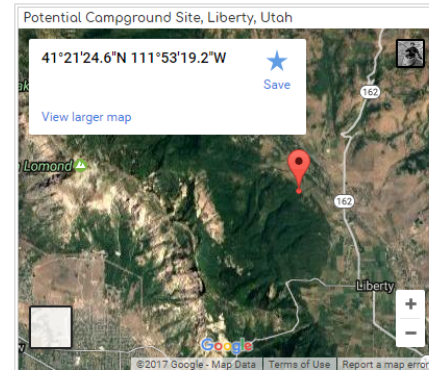
- [Lab 04](#)
- [Week 03 & 04](#)

Introduction to Lab

Background

The Forest Service has funds to create a new campground. The site will be located in northern Utah, on the east side of the Wasatch Range, north of Liberty. You have been tasked with creating a map to present a campground of your own design, reporting some development specifics, and making your plan (a map) and GIS files available via your website.

The manipulation and creation of vector data (points, polylines and polygons) is a fundamental part of your GIS toolkit. In ArcGIS, the manipulation, modification, drawing, and digitizing of vector data is referred to as editing. We need to build proficiency in these tasks before we can do vector-based spatial analyses (i.e. [Lab 05](#) & [Week 05](#)).



Lab Objectives

The purpose of this lab is to build your proficiency in working with vector data sets. Namely, you will learn how to:

- Edit existing vector data layers
- Create new vector data layers and edit them (e.g. digitizing)
- Edit and create tabular data
- Share vector and tabular data with other GIS users
- Learn how to use Model Builder & Geoprocessing

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:

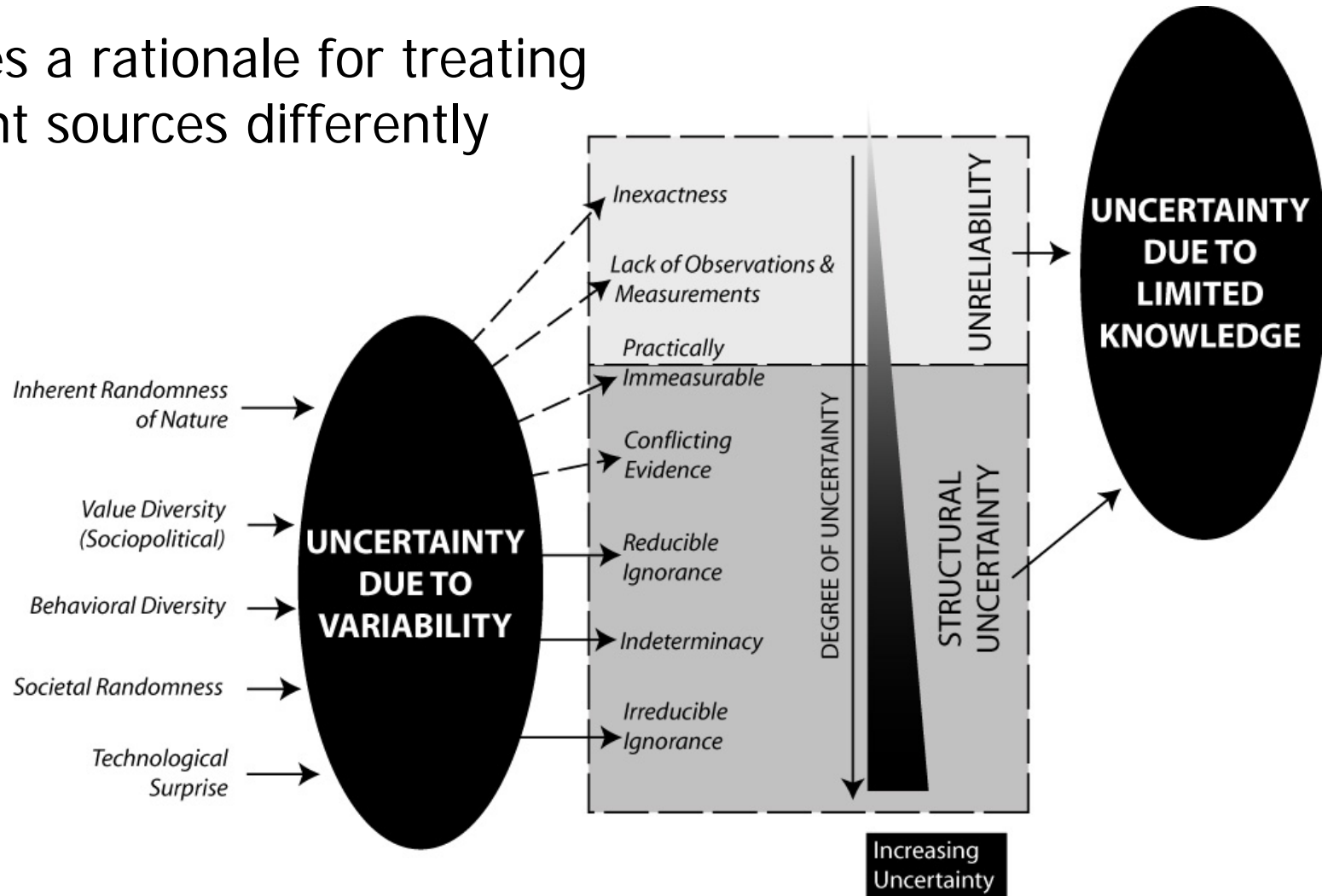
- Sign of weakness
- Like saying you don't know anything
- Confusing

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



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

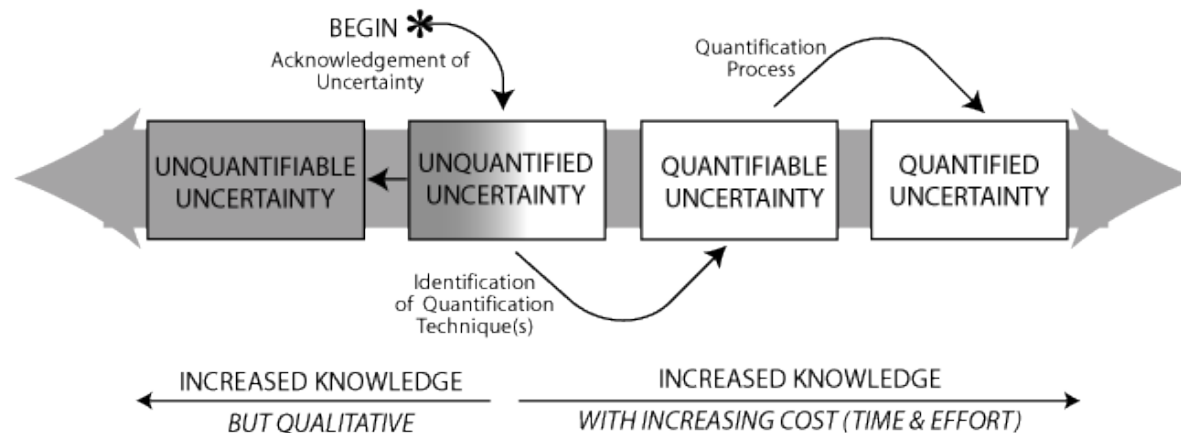


Figure from Wheaton et al. (2008)

TODAY'S PLAN...

I. Uncertainty & Error

II. GIS Errors

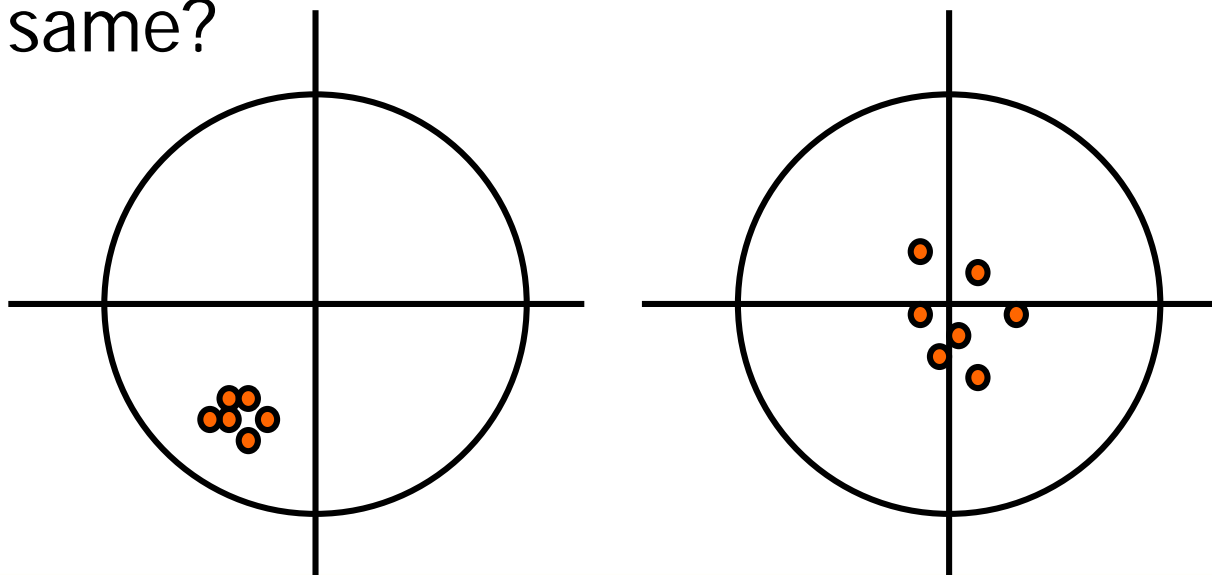
III. Error Propagation

IV. All Bad?

V. Summary of GIS Fundamentals

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 or blunders
- Are they same?



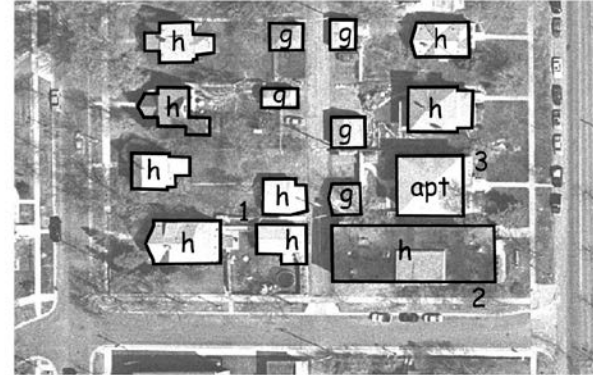
OTHER WAYS TO BE WRONG...

- Spatial data accuracy issues:

a) Positional accuracy



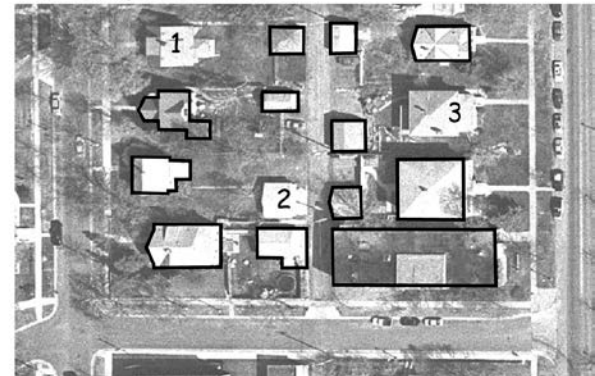
b) Attribute accuracy



c) Logical consistency



d) Completeness



FOUR END MEMBERS

- Positional accuracy of intersection of two freeways

high average accuracy,
high precision



low average accuracy,
high precision



high average accuracy,
low precision



low average accuracy,
low precision



PRECISION OR RESOLUTION

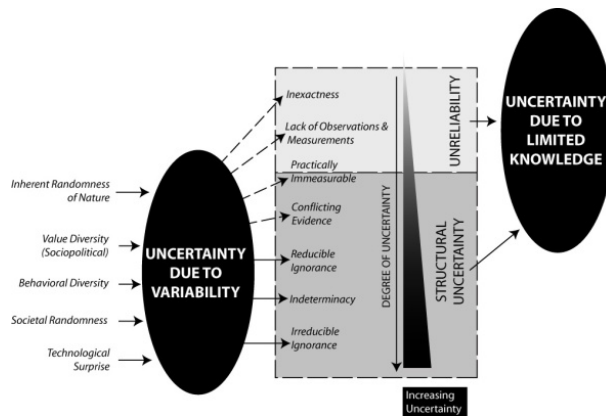
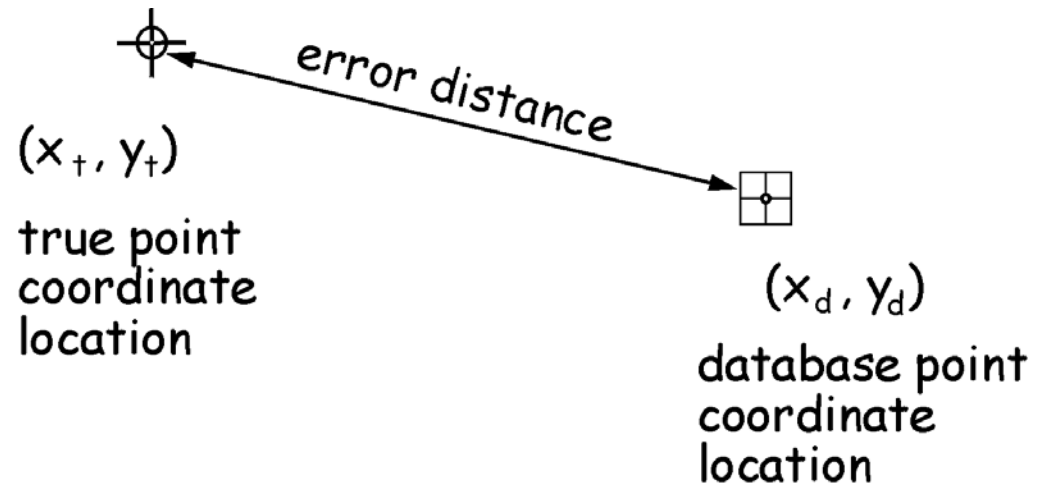
NOT THE SAME AS ACCURACY!

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

HOW POSITIONAL ACCURACY IS CALCULATED

- All you need is measured coordinates and 'true' coordinates
- The lower the error distance, the more accurate...

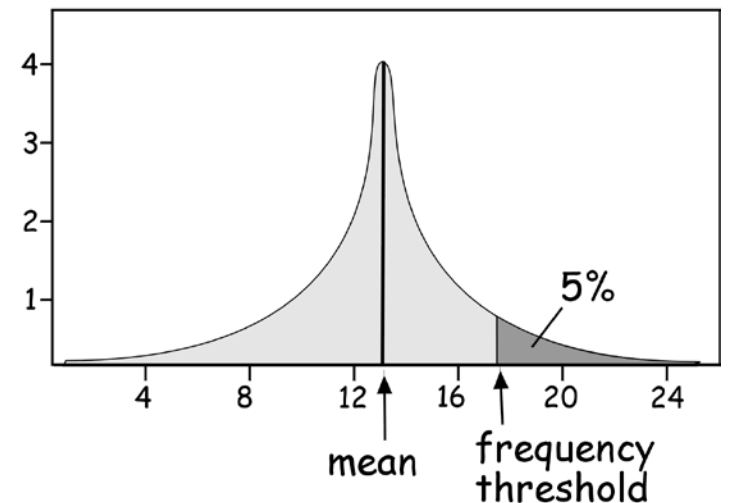
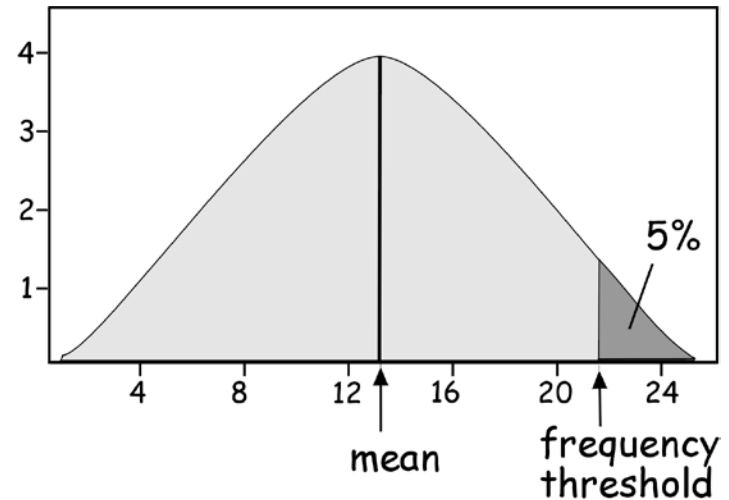


$$\text{error distance} = \sqrt{(x_t - x_d)^2 + (y_t - y_d)^2}$$

From Chapter 14 of Bolstad (2008)

IMPLICATION OF ERROR DISTRIBUTIONS

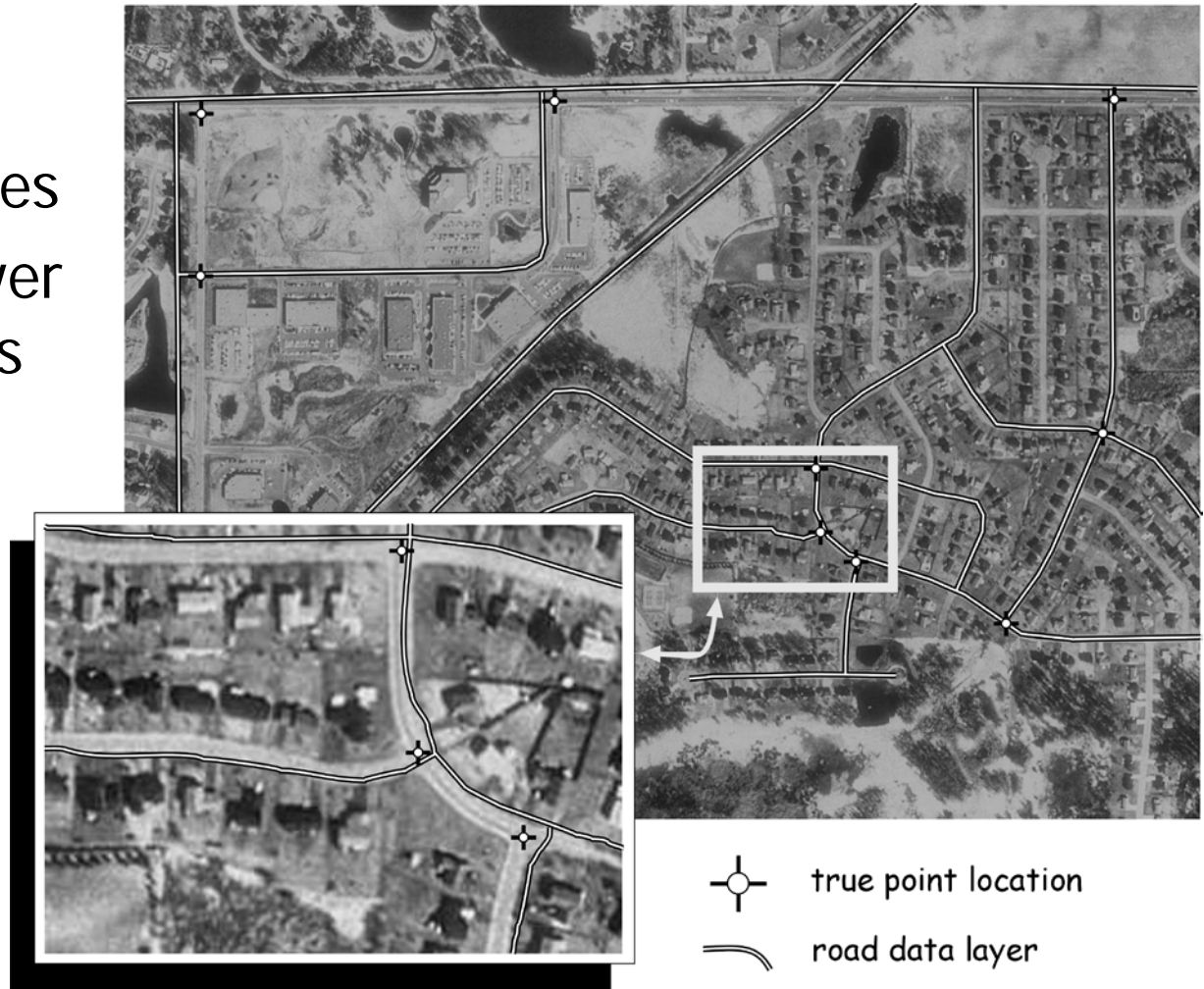
- How would I get a plot like this?
- If we take 95% of the error...
- With same mean, but different distributions, implications are quite different...



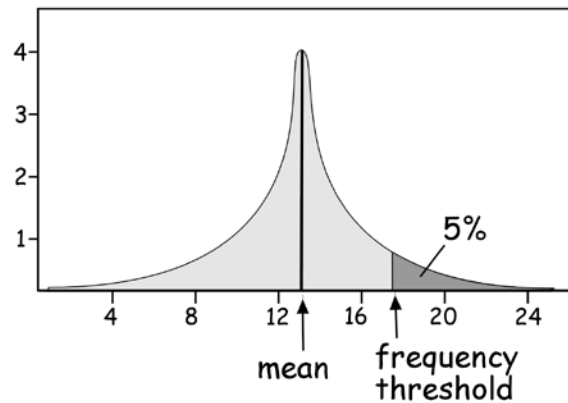
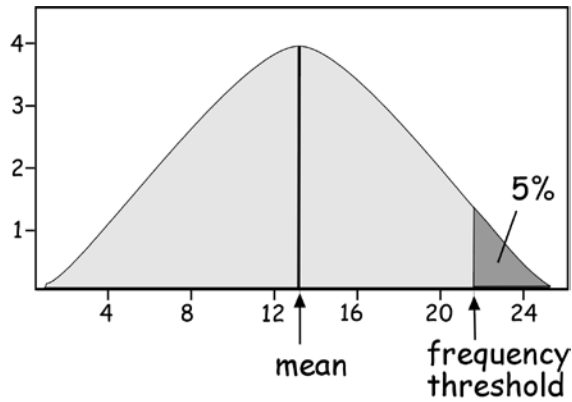
positional error

HOW TO CALCULATE THOSE POSITONAL ERRORS

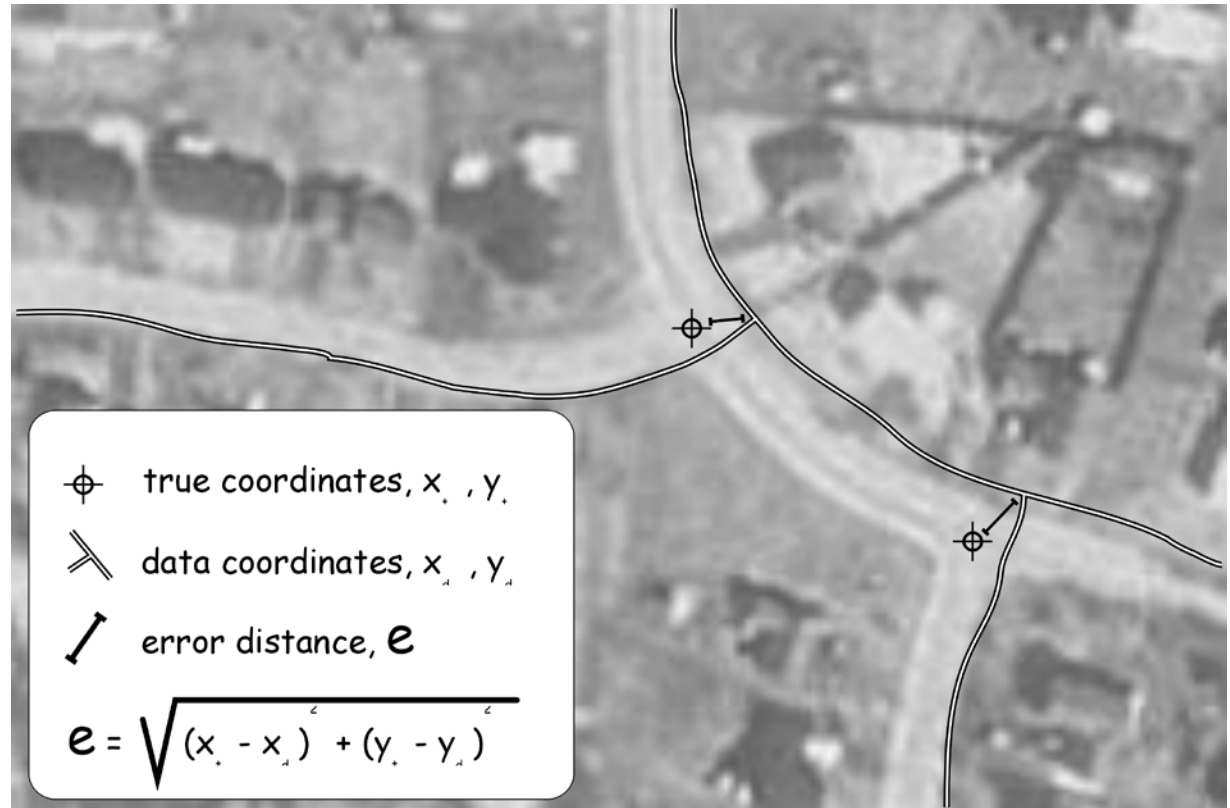
- Find, define or assume *true* values
- Find values of layer to calculate errors for
- Create error field
- Plug and chug
- THIS IS NOT TECHNICALLY CORRECT



A CLOSER LOOK



positional error



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

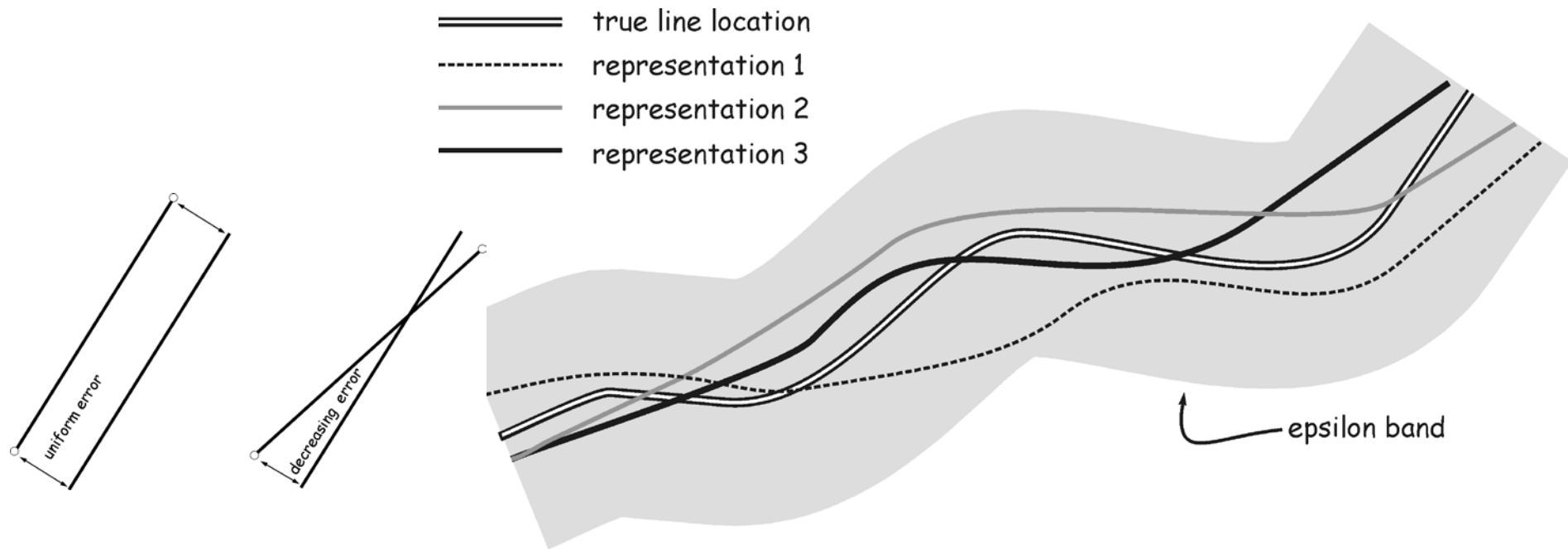
PUT IT ALL TOGETHER...

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

ID	x (true)	x (data)	x differ- ence	(x differ- ence) ²	y (true)	y (data)	y differ- ence	(y differ- ence) 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)

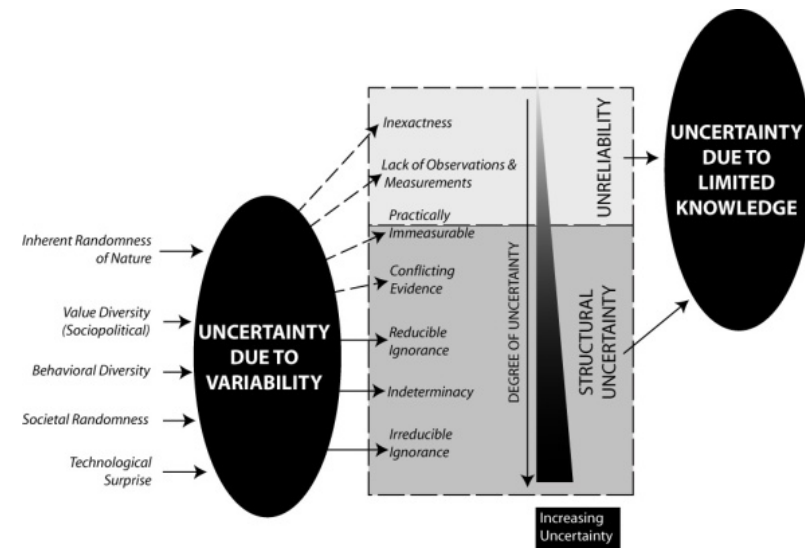
MEASUREMENT OF POSITIONAL ACCURACY

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

- $RMSE = \sqrt{\frac{e_1^2 + e_2^2 + e_3^2 + \dots + e_n^2}{n-1}}$ where e_i is the distance (horizontally or vertically) between the true location of point i on the ground, and its location represented in the GIS.
- 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

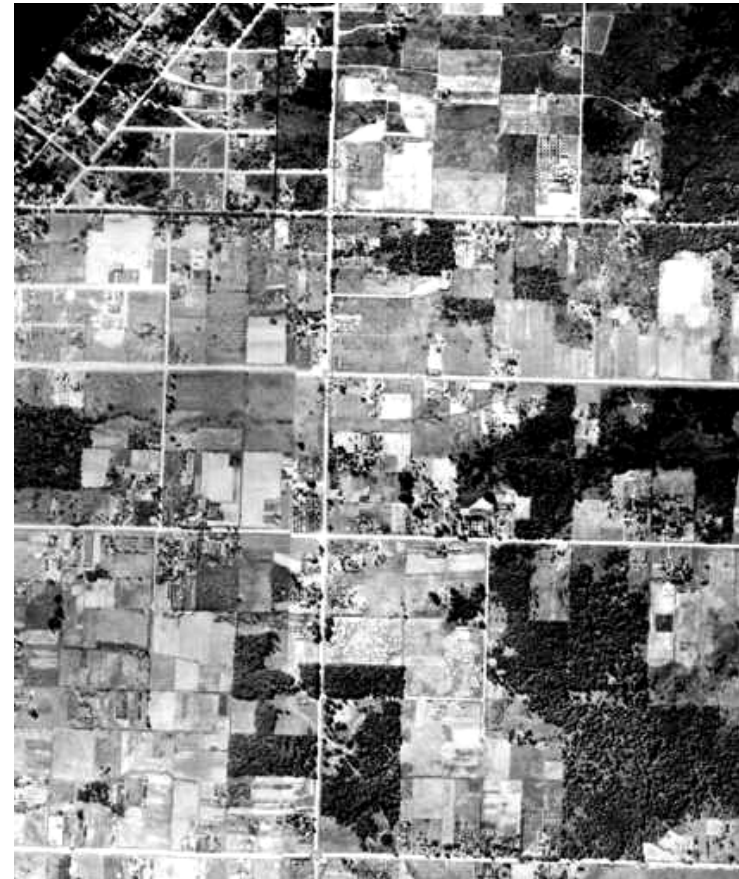


ERROR – OUT OF DATE

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



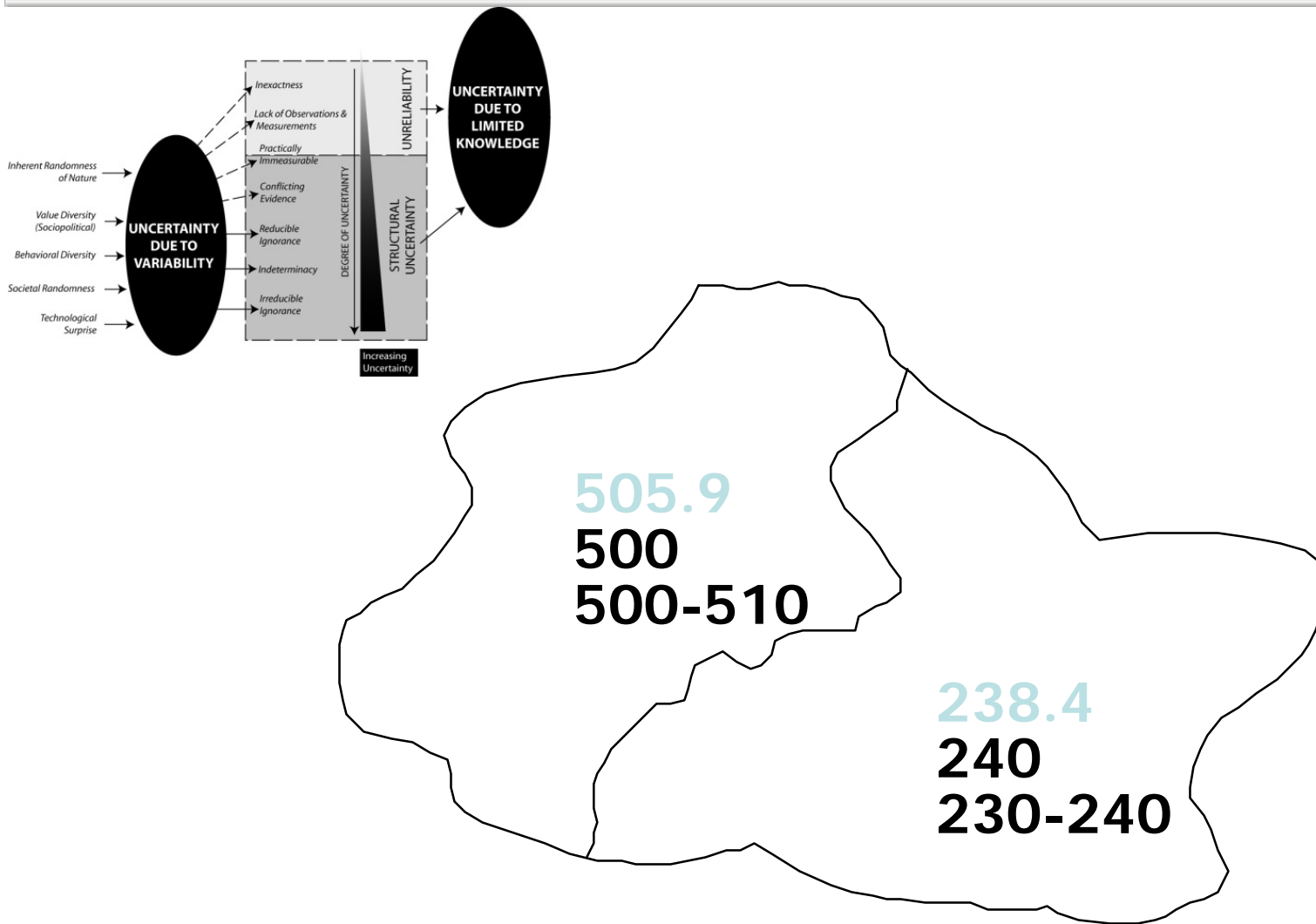
1997



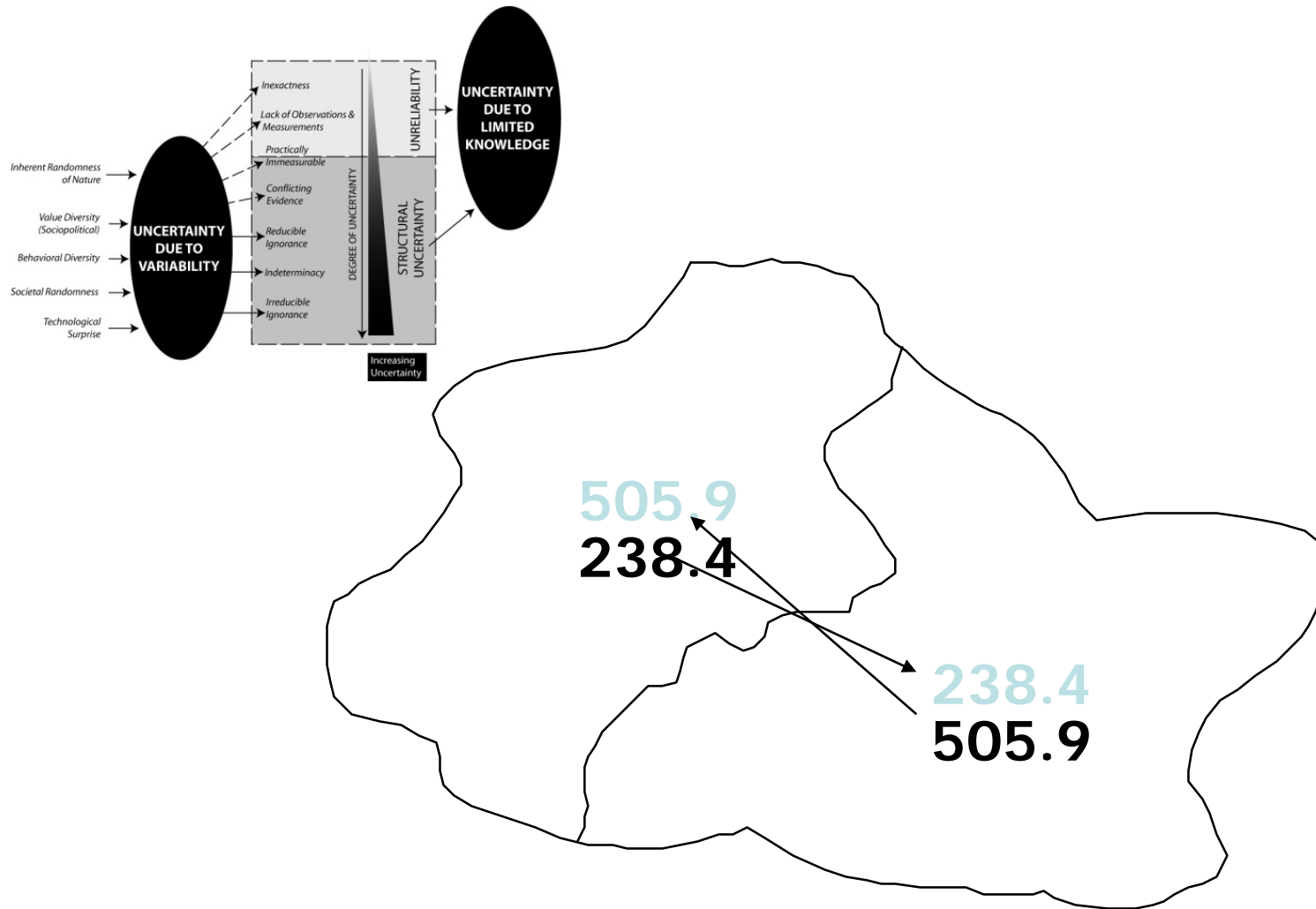
1936

From Chapter 14 of Bolstad (2008)

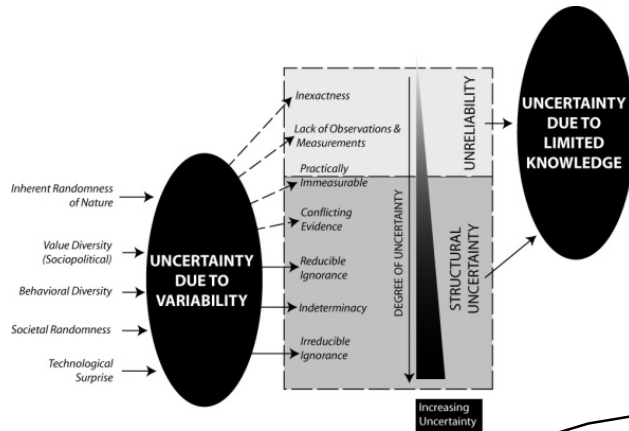
IMPRECISE AND VAGUE



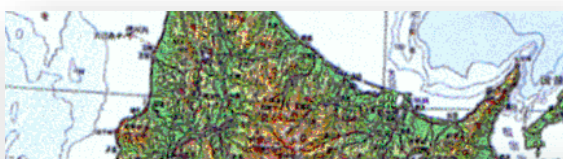
MIXED UP



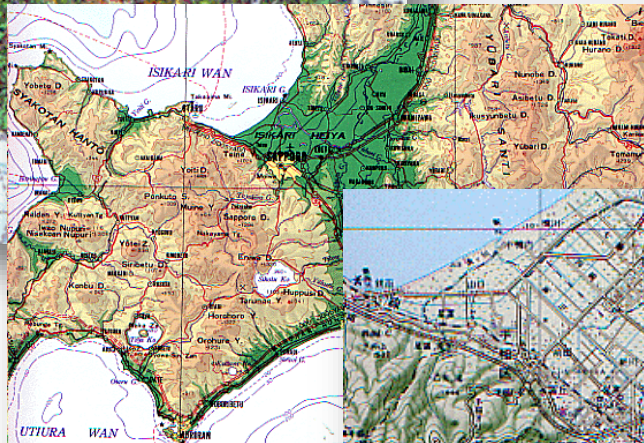
JUST WRONG



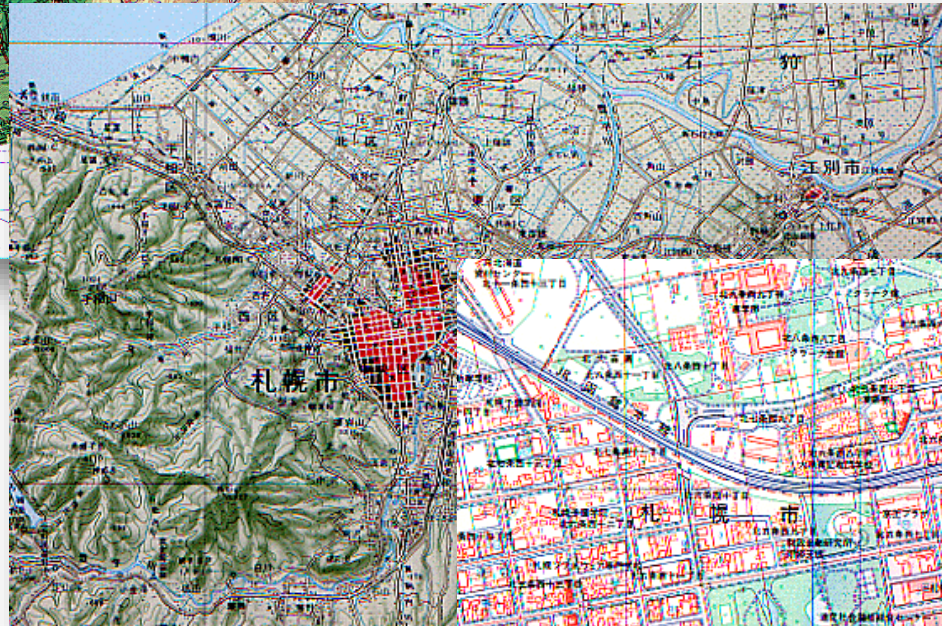
GENERALIZATION



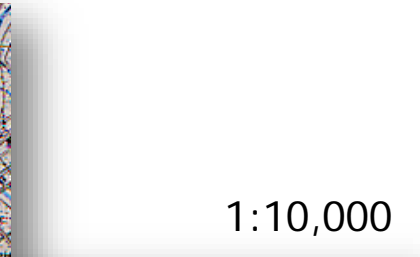
1:3M



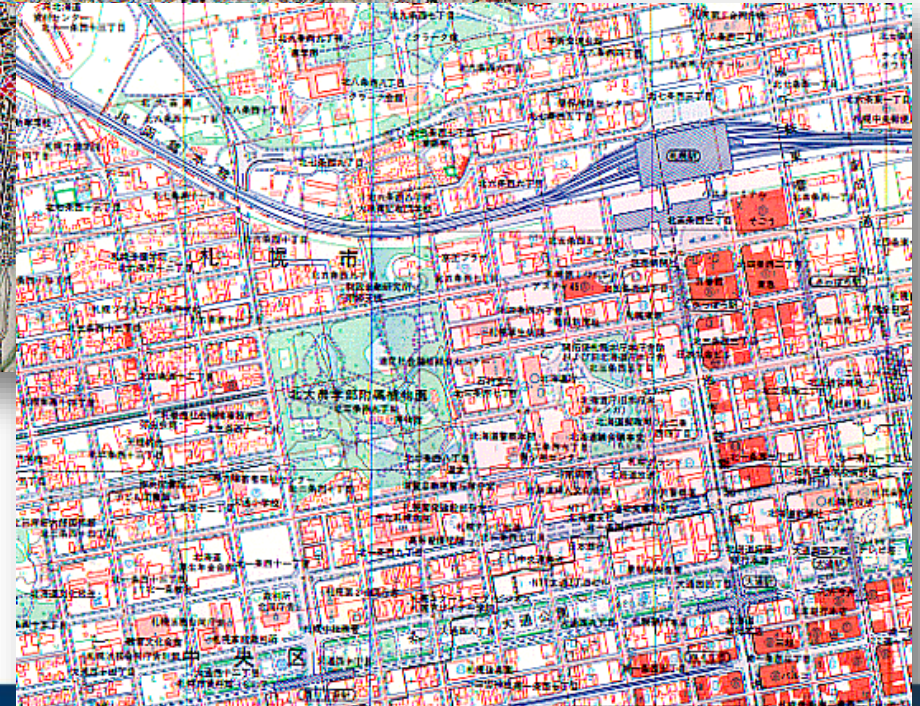
1:500,000



1:25,000



1:10,000



City of Sapporo, Japan

UNCERTAINTY IN ANALYSIS

- Just because 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?
 - Don't underestimate the value of a laugh-test.
- Functional REDUNDANCY:
 - There is almost always another (often faster) way of performing any analysis
 - Should produce the *same* result... try it?

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



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



DEM DIFFERENCING

RASTER CALCULATOR....

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

NEW DEM

OLD DEM

= DoD

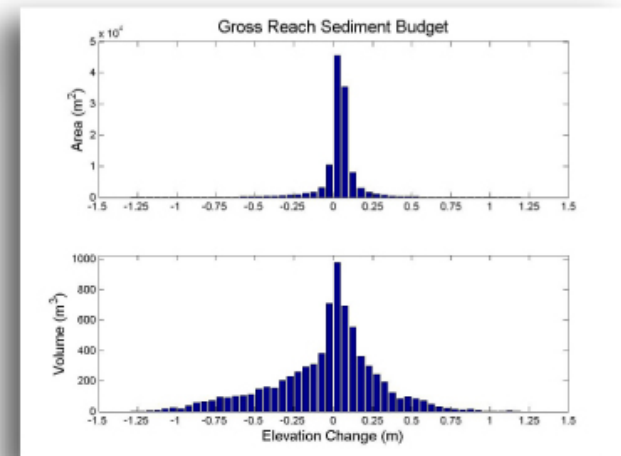
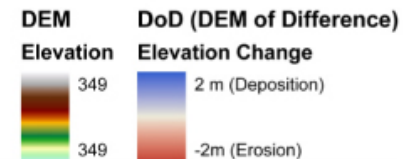
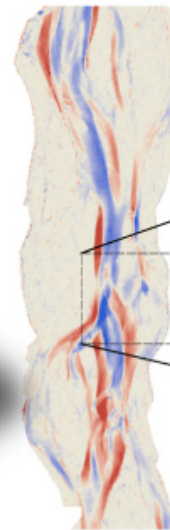
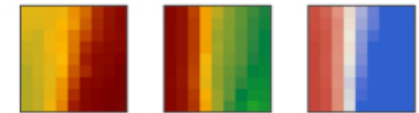
CONSERVATION OF MASS
VOLUMETRIC

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

\uparrow Volumetric rate of bed material transport
 \uparrow Porosity of bed material

McLean & Church (1988) – Water Resources Research

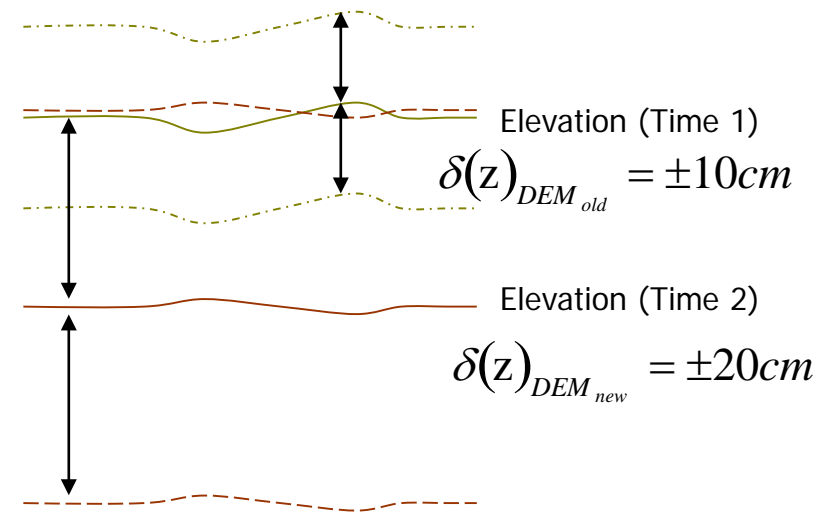
$$\text{NEW DEM} - \text{OLD DEM} = \text{DoD}$$



© Wheaton 2008

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



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

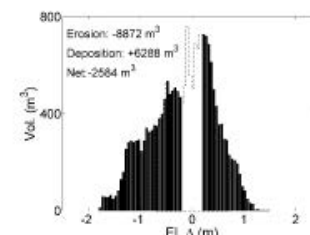
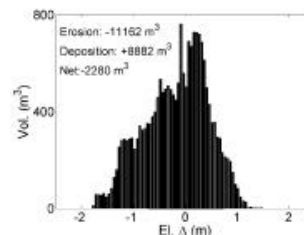
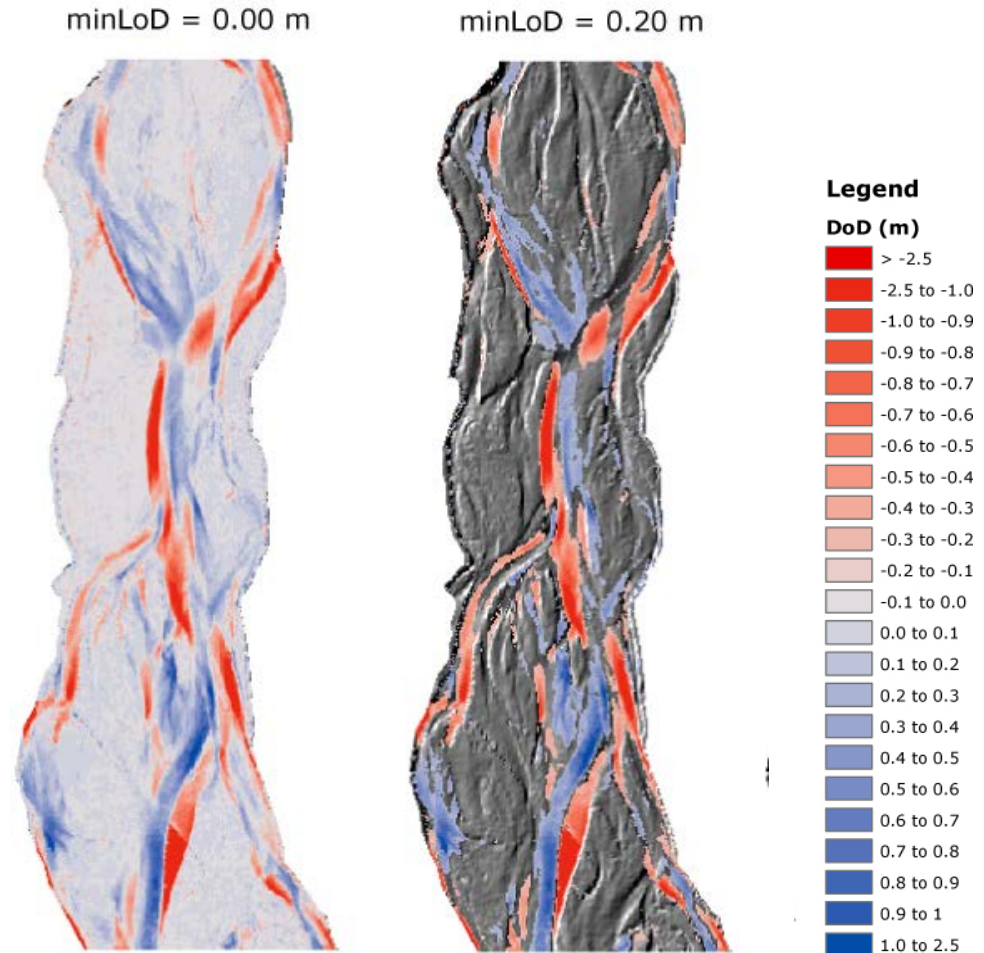
e.g. $\delta(z) = \sqrt{(10)^2 + (20)^2} = 22.36$
 $22.36 \text{ cm} \approx 8.8 \text{ in}$

See

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

HOW DOES A minLoD GET APPLIED?

- You take original DoD, and remove all changes $\leq \text{minLoD}$
- For example ± 20 cm
- How would you do that?
- What is the assumption here?



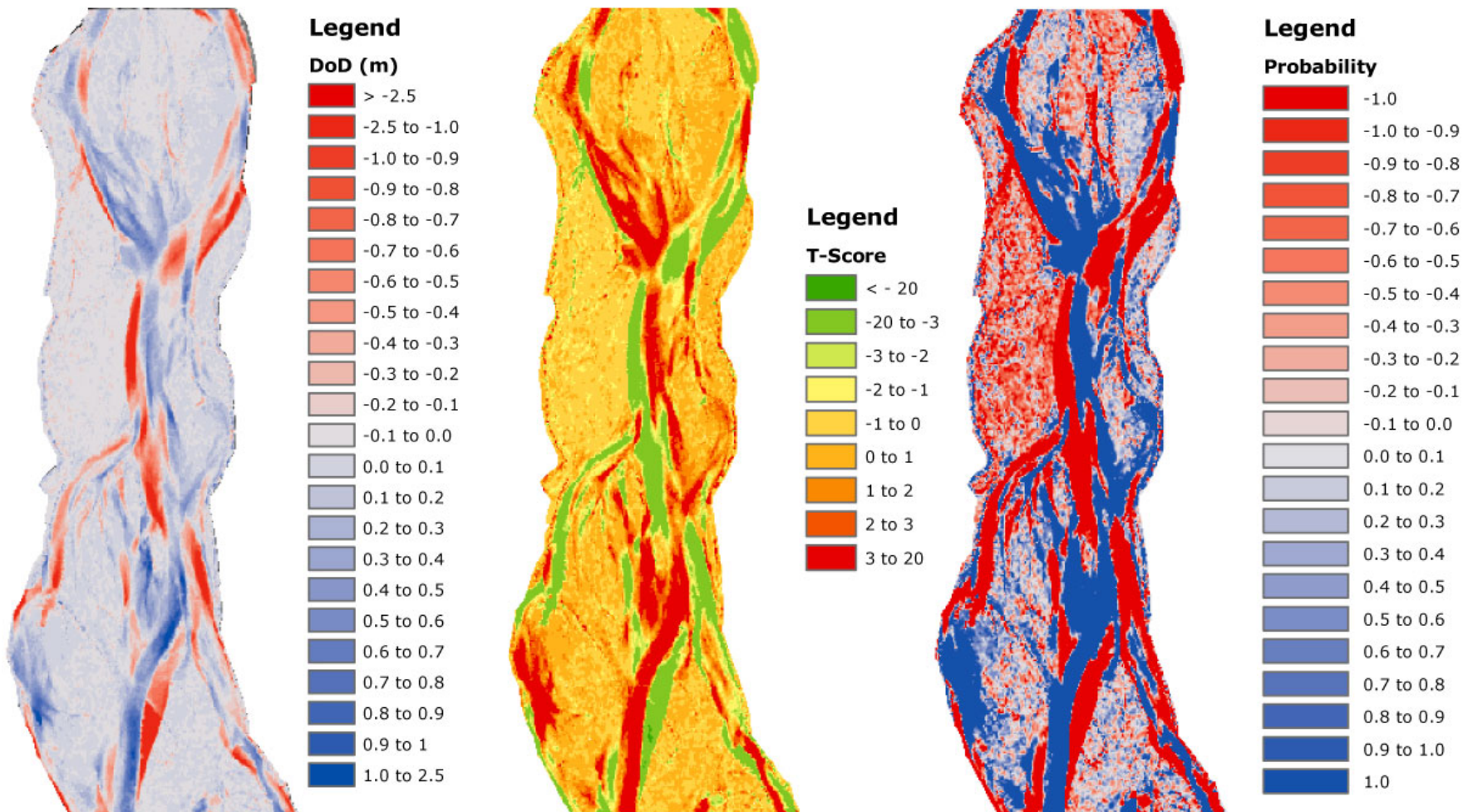
HOW COULD I REPRESENT AS PROBABILITY?

- Using inferential statistics, we'll calculate a t-score
- σ_{DoD} is the characteristic uncertainty
 - In this case $\sigma_{DoD} = \min LoD$
- Just the ratio of actual change to $\min LoD$ change
- Assuming two-tailed test, t is significant at:
 - 68% confidence limit when $t = 1$
 - 95% confidence limit when $t = 1.96$

$$t = \frac{|z_{DEM_{new}} - z_{DEM_{old}}|}{\sigma_{DoD}}$$

PROBABILITY THAT CHANGE IS REAL

Original DoD → Propagated DoD Uncertainty → Calculated T-Score → Converted Probability

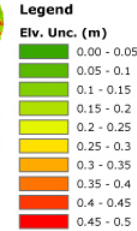
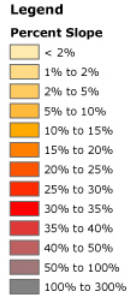
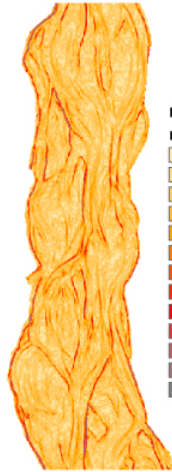
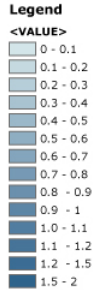
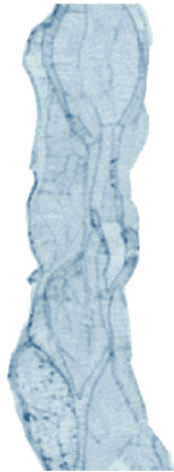


© Wheaton (2008)

Even when \min LoD is spatially constant,
probability varies in space... why?

APPLY FIS ON CELL BY CELL BASIS

2006



FIS Input 1
(Point Density)

+

FIS Input 2
(Slope)

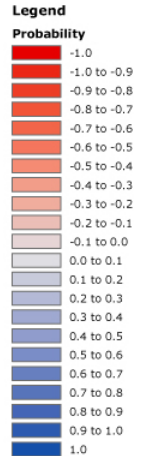
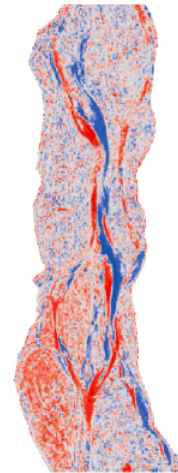
+

FIS Input 3
(GPS Quality)

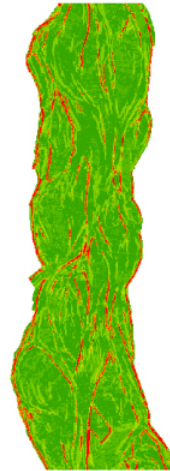
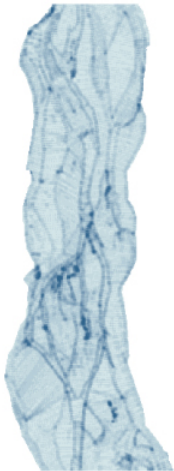
=

FIS Surface
(El. Unc. (m))

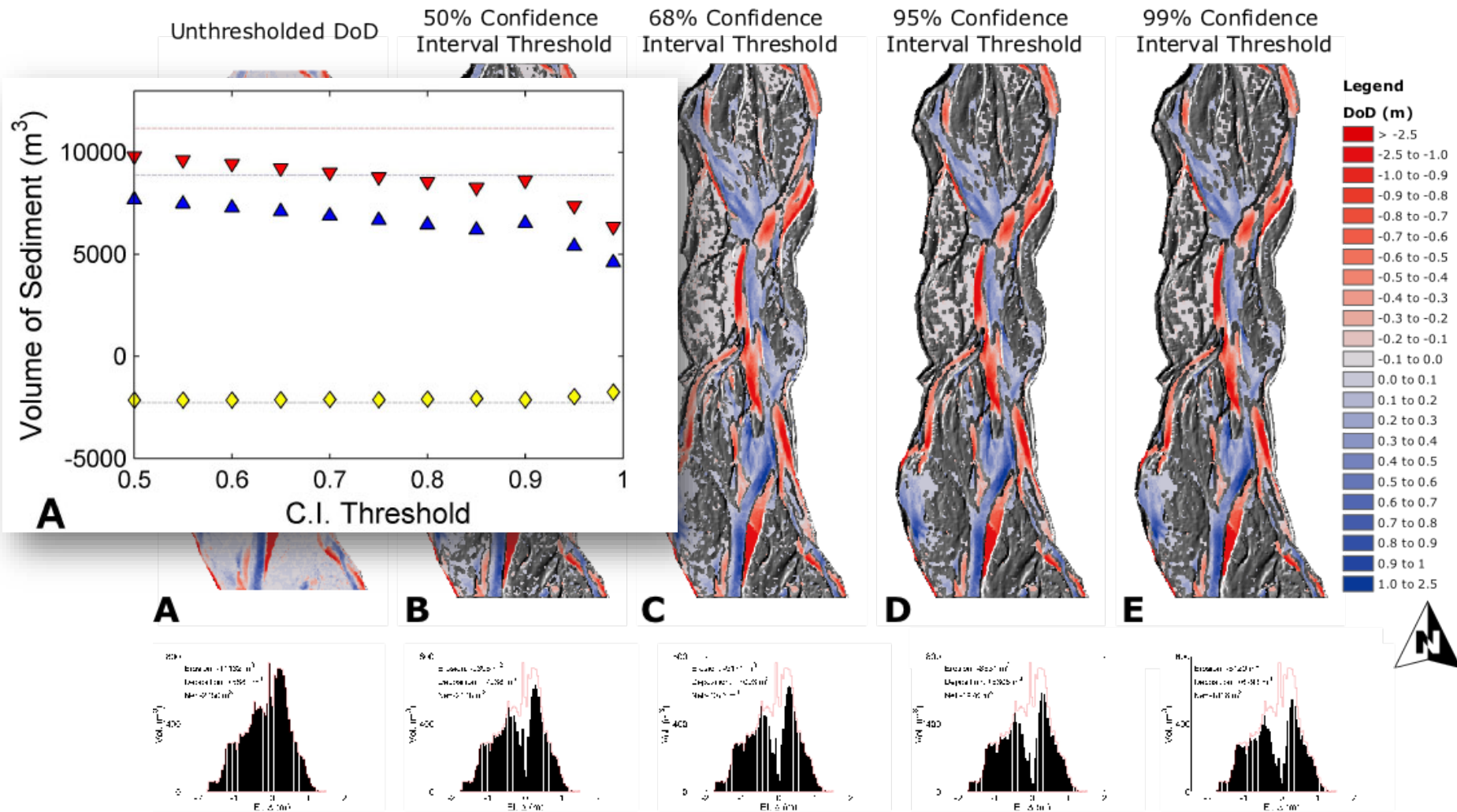
Both FIS Surfaces
Combined to DoD
Probability



2005



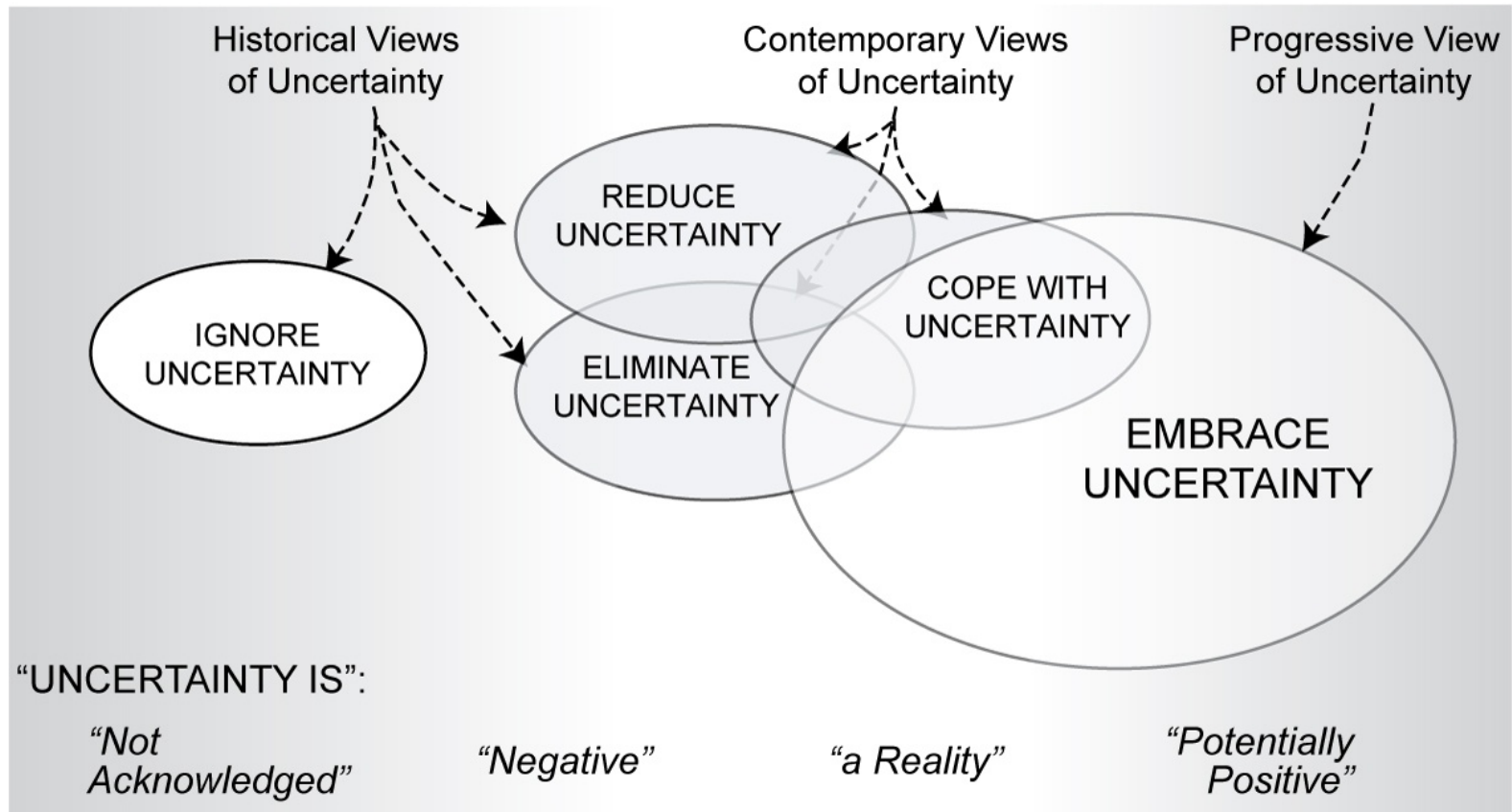
SENSITIVITY OF THRESHOLD?



TODAY'S PLAN...

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

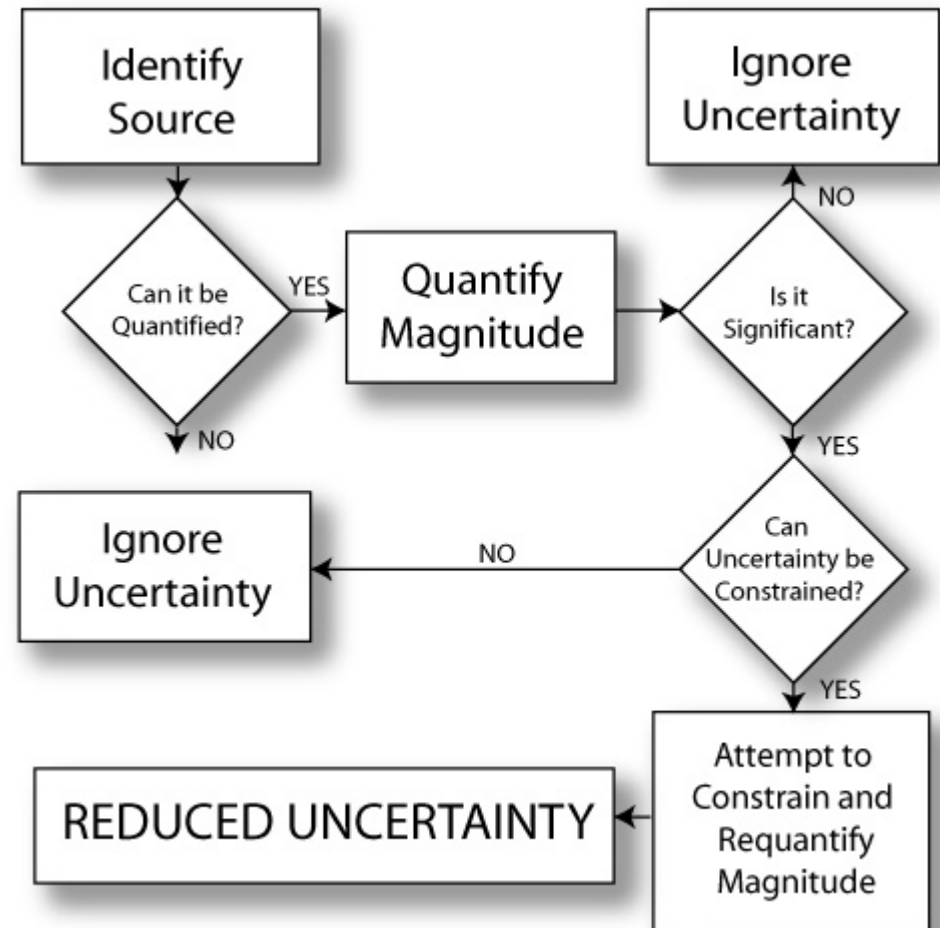
PHILOSOPHICAL ATTITUDES TO UNCERTAINTY



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

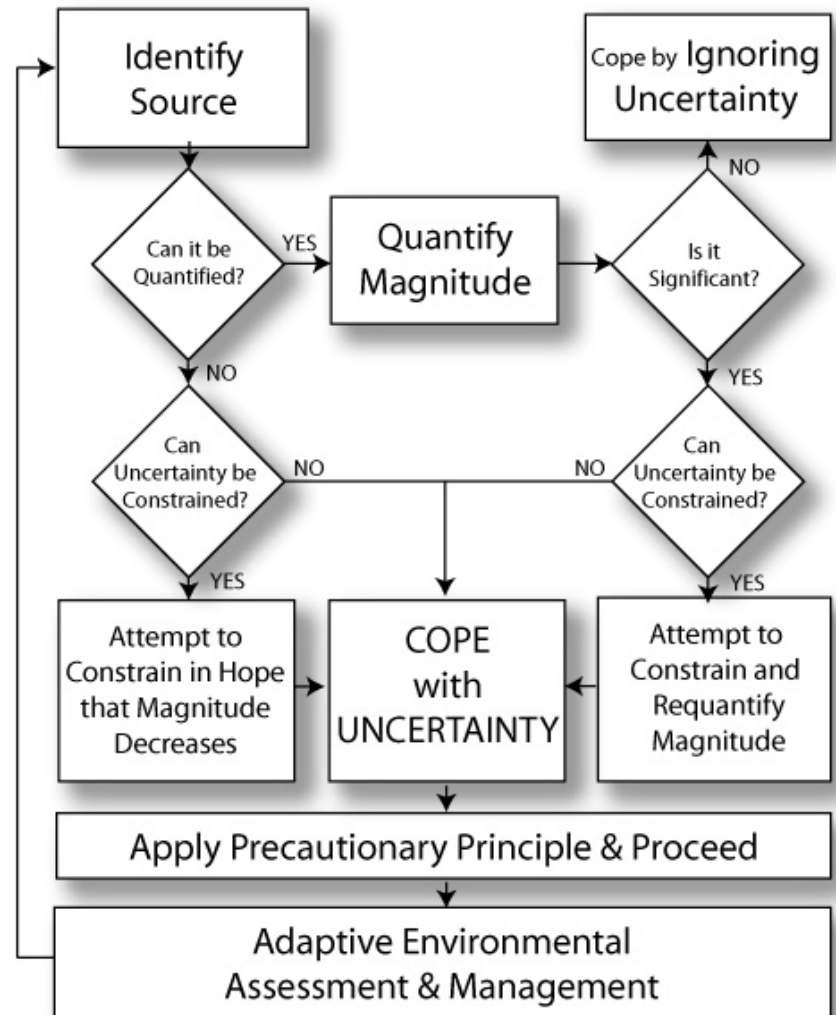
REDUCE UNCERTAINTY

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



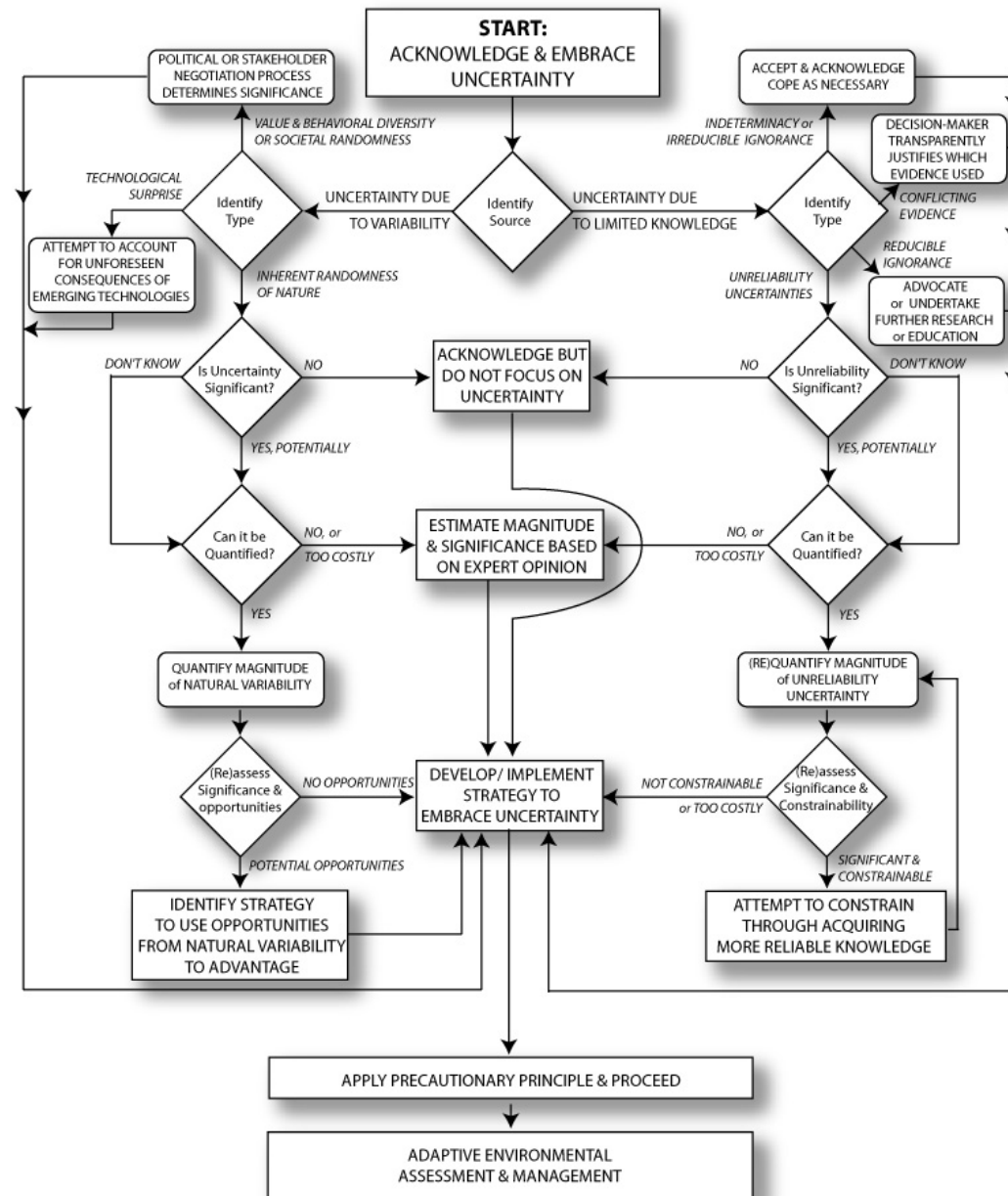
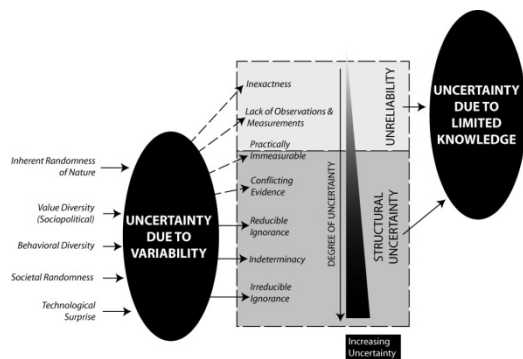
COPE WITH UNCERTAINTY

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



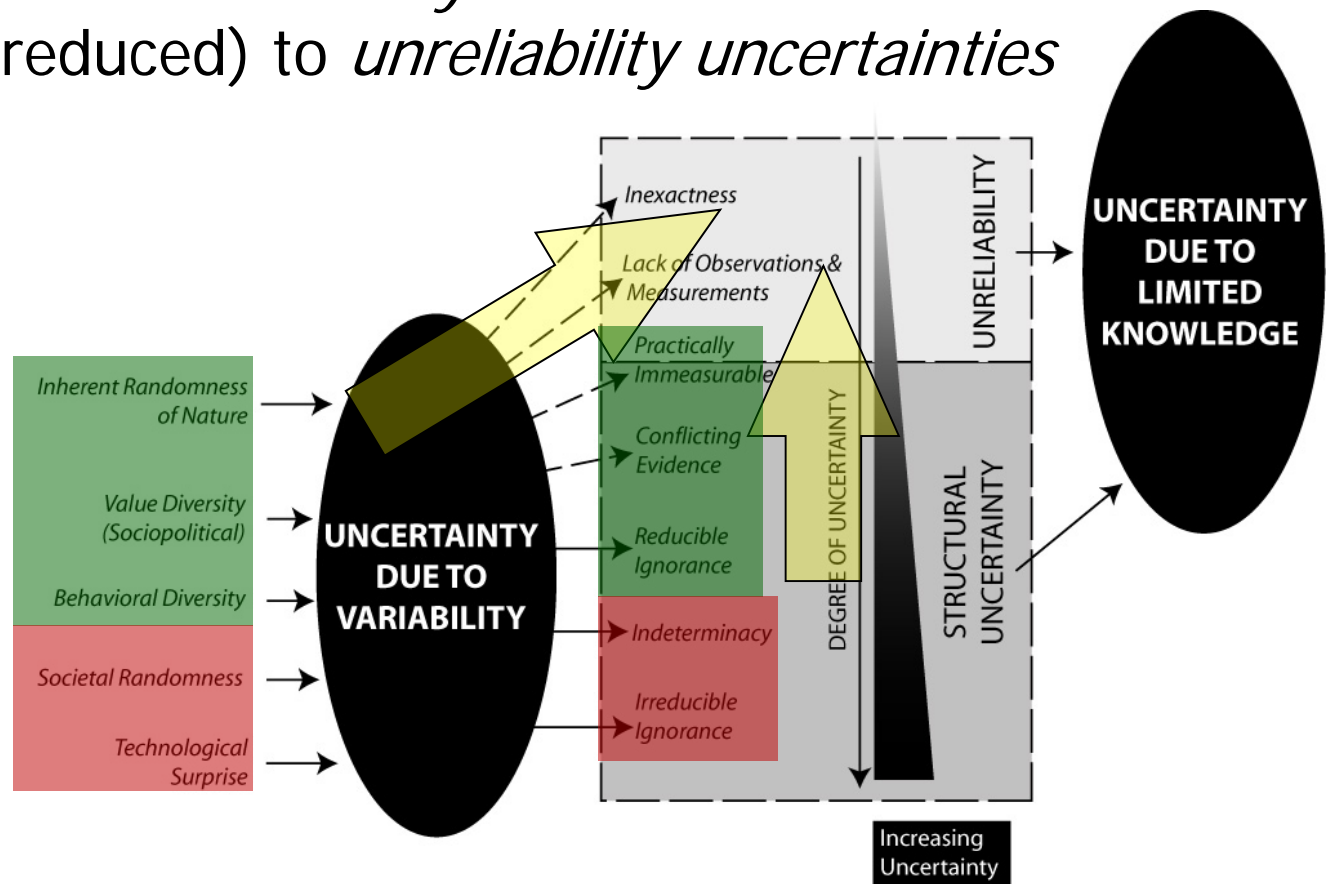
EMBRACE UNCERTAINTY

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



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*



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 THE 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



FOR THE WATS
4930/6920 FOLKS...



WHAT FUNDAMENTALS?

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



WHAT YOU SHOULD HAVE GOTTEN... (so far)

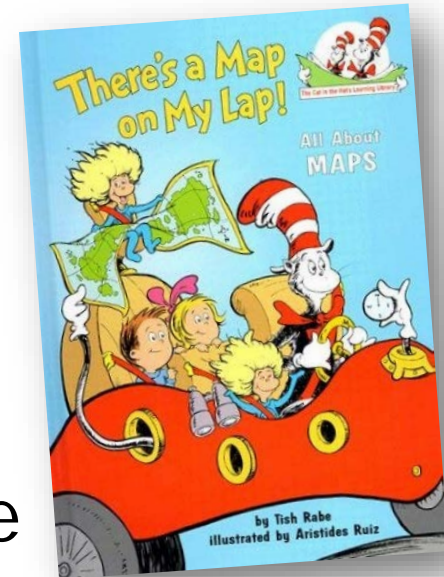
The above learning outcomes apply to the courses as follows:

Learning Outcome:	WATS 4930/6920	WATS 4931/6921	WATS 6915
1 - GIS Theory	Core	NA	Core
2 - Proficiency 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 & tachometric 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 Python 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 Python 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 Toolkit
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

4930, 6920, 6915	WEEK 4	<u>Working with Data in GIS - End of WATS 6915 Lectures</u>	Jan 27* & Jan 29	3. <u>Reproducing Maps - Geologic Map</u>
4930, 6920, 6915	WEEK 5	<u>Vector Analyses</u>	Feb 3* & 5*	4. <u>Digitizing & Editing & Sharing Data</u> - Last WATS 6915 Lab
4930 & 6920	WEEK 6	<u>Raster Analyses</u>	Feb 10 & 12*	5. <u>Vector Analysis</u>
4930 & 6920	WEEK 7	<u>Raster Analyses</u>	Feb 19	6. <u>Working w/ DEMs</u>
4930 & 6920	WEEK 8	<u>GIS Modeling</u>	Feb 24 & 26*	7. <u>Building DEMs</u>
4930 & 6920	WEEK 9	<u>GIS Modeling</u>	Mar 3 & 5	8. <u>Morphometric Analyses</u> or <u>Habitat Modelling</u>
Spring Break – March 9 -13				(Catch Up)
4930 & 6920	WEEK 10	<u>Collecting Your Own Data & Synthesis</u>	Mar 17 & 19	9. <u>Blimp & Georeferencing Lab</u>
End of WATS 4930/6920				
4931 &				