

Re-Thinking Best Practices in Cartographic Data Capture and Data Modeling

Barbara P. Buttenfield

University of Colorado-Boulder USA
babs@colorado.edu

Charlie Frye

Cartographic Research Manager, ESRI USA
cfrye@esri.com

Data modeling support by Kiyoshi Yamashita, CU-Boulder

Funding from ESRI 2003B6345, B6347

Data Capture Defines Fitness for Use

- **A single data representation cannot serve:**

- all audiences
- all types of use, or
- all mapping scales

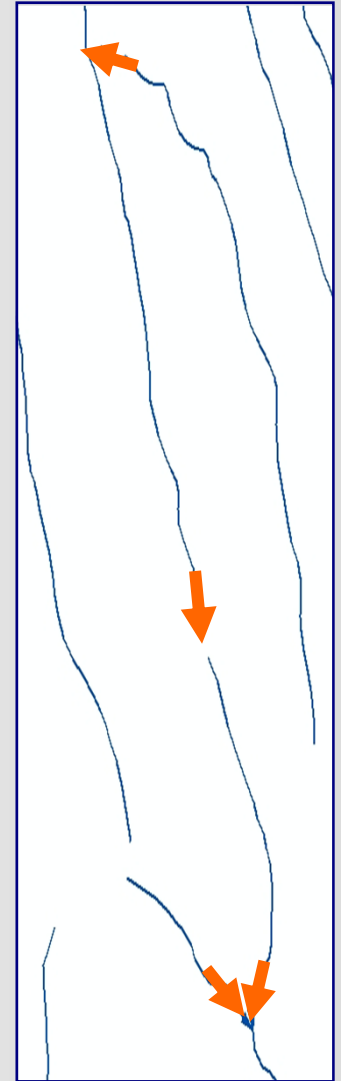
without data modeling, alternative symbolization, or both.

- **Local data managers do not plan to transform the data for specific map scales or uses when they capture data.**

- **Regardless of data precision or accuracy, limits are imposed upon data by practices used for capture.**

Data Capture Defines Fitness for Use

- Digitize “from” / “to” nodes in uniform direction
- Avoid problems mapping hydrology or flow accumulation
- “Informed” data capture and modeling
 - Capture for particular purposes that are loosely or tightly coupled to a particular resolution and mapping scale
 - Specify intended use *in advance of data capture* to minimize subsequent processing problems.



Outline of the paper

- **Basic factors constraining cartographic data capture and data modeling**
 - Cartographic project design
 - Business project management
- **Case studies demonstrate where “informed” data capture can reduce or obviate intensive data modeling**
 - Guidelines for best practice data capture for general use
 - Implications for data sharing
 - Informing data partners

Multiple Use Concept in Cartography Sherman and Tobler (1957)

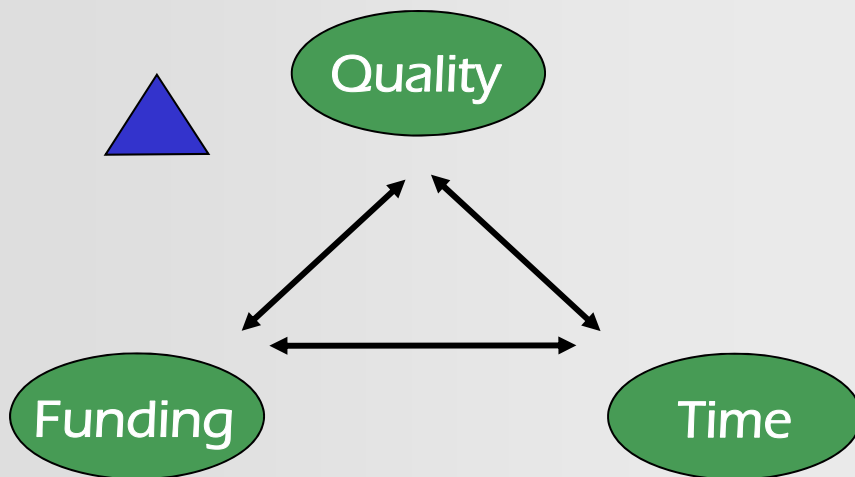
- When data capture is planned to maximize the uses to which it can be applied, overall production costs and problems will be minimized.
- Analogy of building a hydroelectric dam:
 - ... to serve purposes of flood control, river navigability, irrigation and electric power.
 - Applying the *multiple use concept* satisfies the need for optimum allocation of the diverse but finite resources.

Multiple Use Concept in Cartography Sherman and Tobler (1957)

- **The same concept can be applied to cartography with similar results.**
 - **Separate captured map elements for later “selective reassembly” depending on mapping scale or purpose**
 - **Separated elements can be re-symbolized, masked, re-scaled, or integrated with new data**
 - **Even if initial capture tasks take longer, labor costs overall should be lower than for initiating new data compilations for each mapping project**

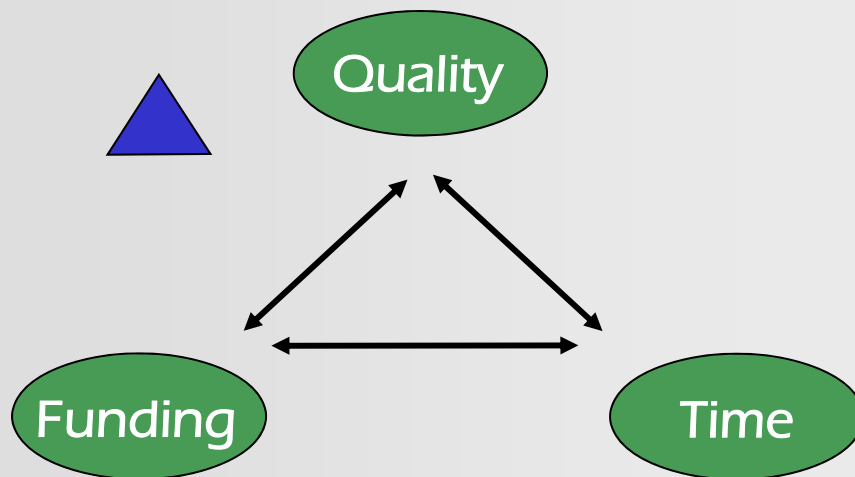
Business Project Management: “The Iron Triangle”

- **Basic Factors Constraining Data Capture**
 - On Time
 - High Quality
 - On Budget



Basic Factors Constraining Data Capture: “The Iron Triangle”

- The Iron Triangle is Dynamic
 - Constraints interact
 - Priorities re - balance resources



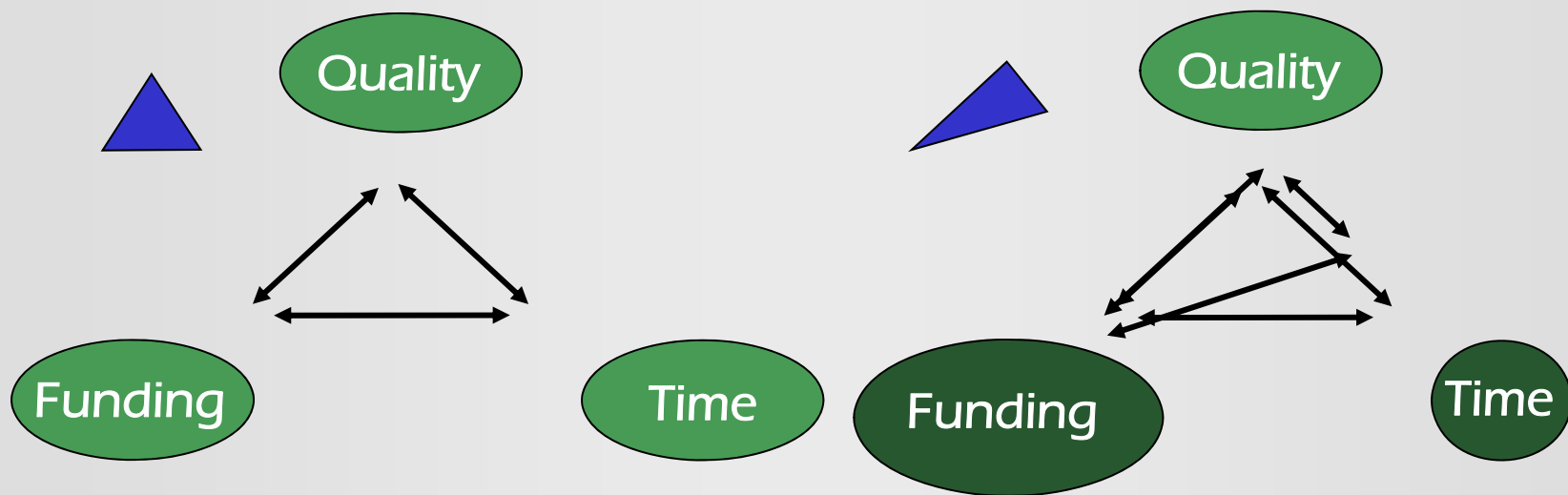
Length of each side =
amount of funding, time
or quality that can be
achieved

Length of any side can vary,
BUT

Sum of all lengths remains
constant

Basic Factors Constraining Data Capture: “The Iron Triangle”

- **Example: short production time is required; maintain high quality**
 - More labor to complete work on time
 - Preserve high quality product only at higher labor cost



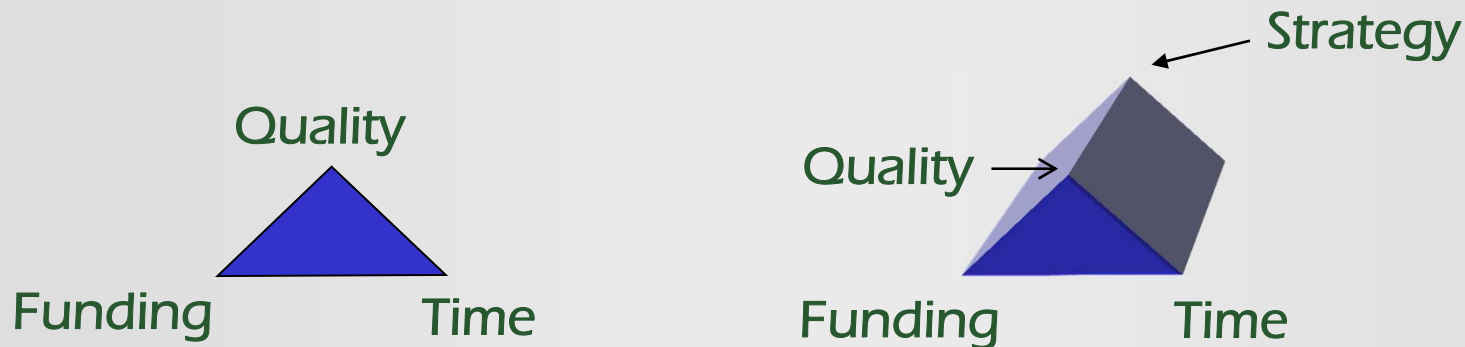
4-10 August 2007

23rd ICC

Moscow, Russia

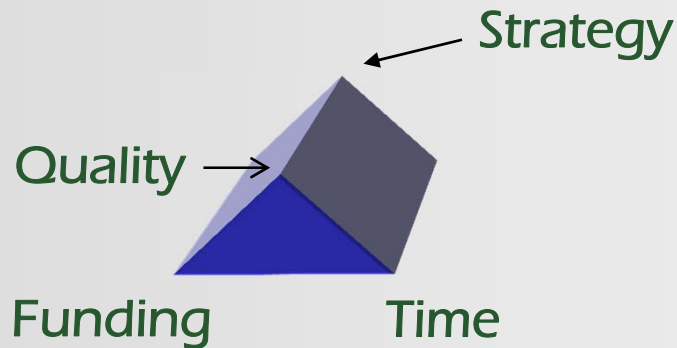
Are Three Constraints Sufficient?: Extending “The Iron Triangle”

- One more (important!) constraint = STRATEGY
 - What is the purpose of the product?
 - What is its intended fitness for use?



- When goal is general purpose topographic map capture and modeling, strategy means flexibility.

Flexibility, Data Capture and Fitness for Use



- Flexibility means “Capture topographic data with an eye to general use.”

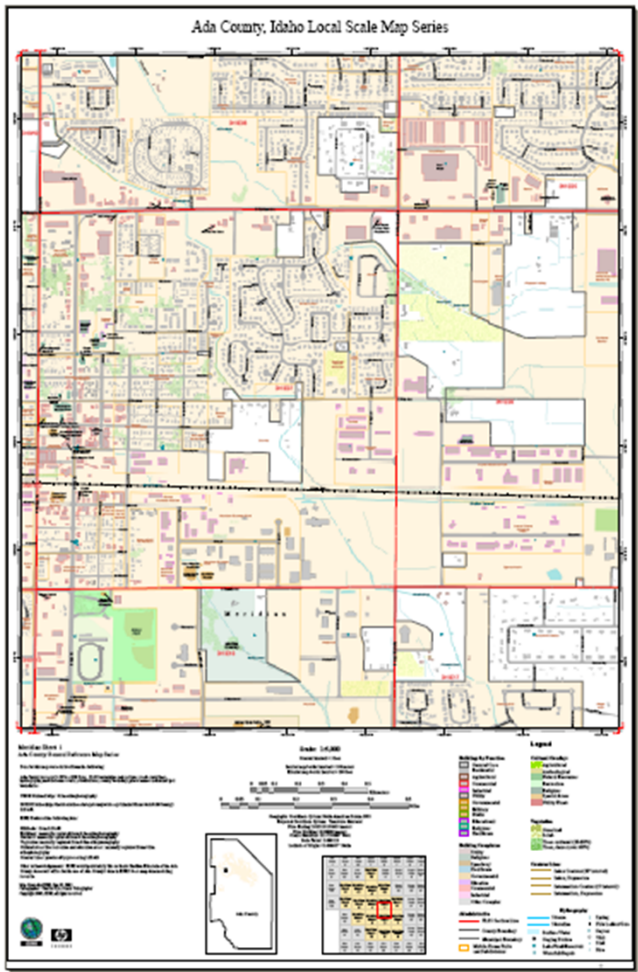
- Opt for:
 - Multiple uses
 - Multiple scales

- Capture features that can be used to derive other features, within existing constraints

Case Studies from Ada County, Idaho (USA)

- **Objective: build a multi-scale, multi-purpose topographic database for a county government office in fast-growing region (Boise Idaho)**
 - Arid montane region – water an important feature
 - State capitol – transportation and infrastructure priority
- **Begin with data already captured at 1:5,000 and generate another database at 1:25,000**

Ada County, Idaho



1:5,000
Community scale base map

4-10 August 2007

23rd ICC

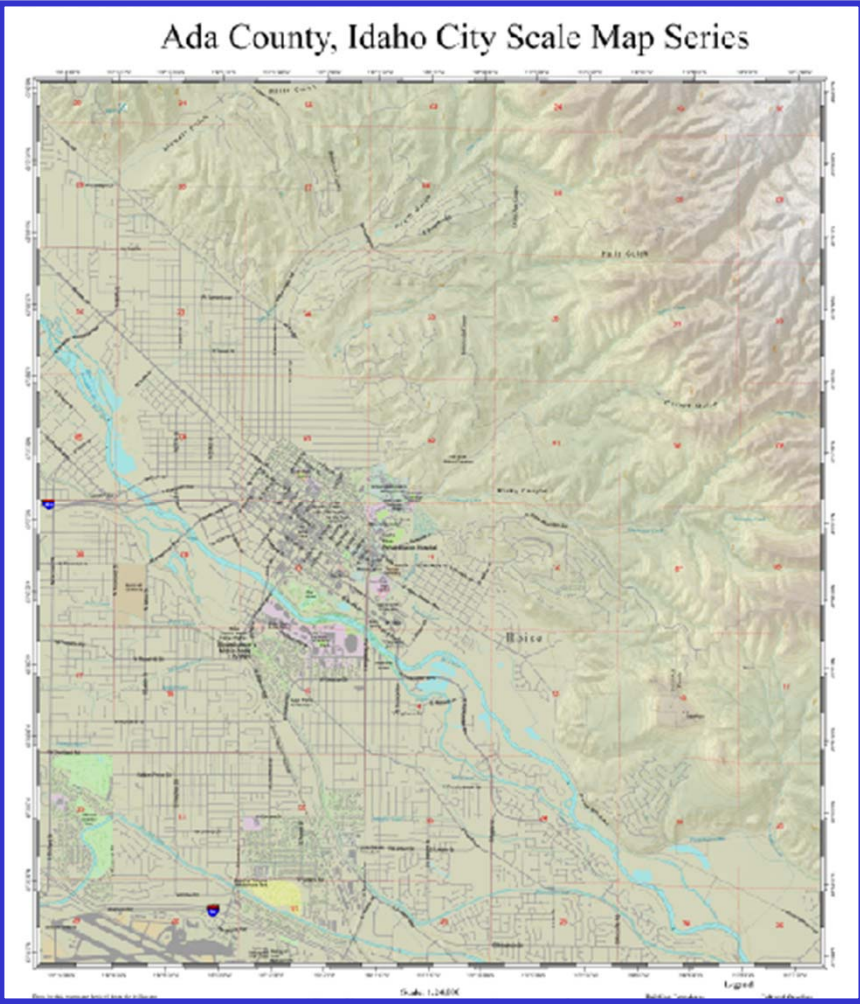
Moscow, Russia

Ada County, Idaho



4-10 August 2007

23rd ICC



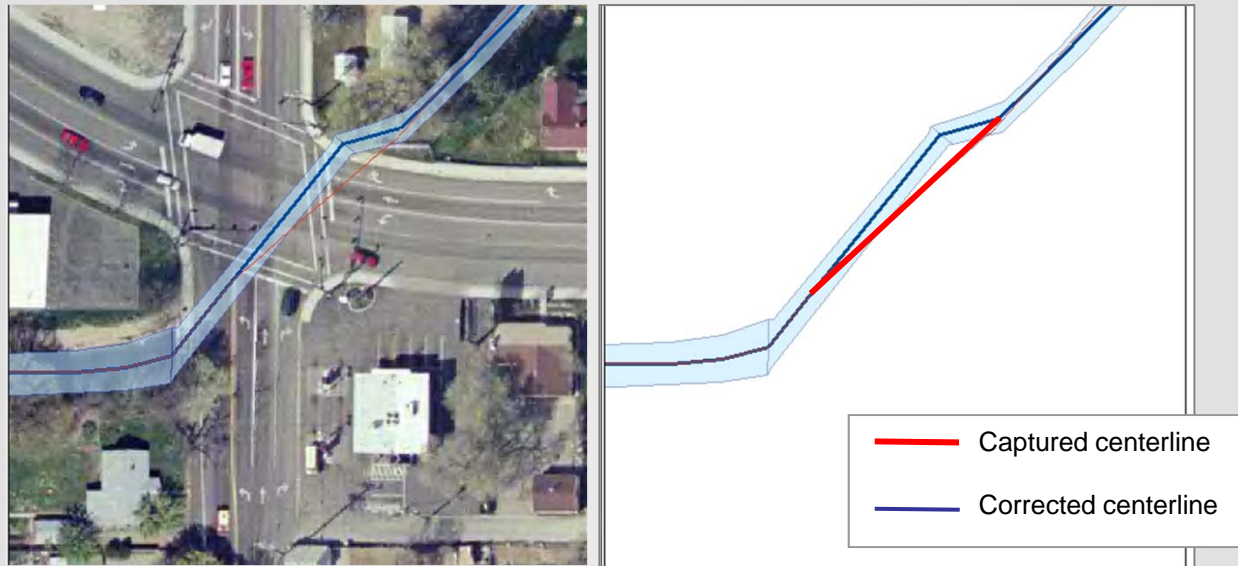
1 : 25,000
City scale base map

Case Studies

- **All concern problems arising in modeling scale change from 1:5,000 to 1:25,000**
 - Demonstrate how data capture limits fitness for use (in this case, suitability for multi-scale mapping)
- **All concern hydrography**
 - High sensitivity to scale change
 - High priority for topographic mapping in this region
- **Each case study will conclude with a guideline for establishing best practices in data capture**

Case Study 1

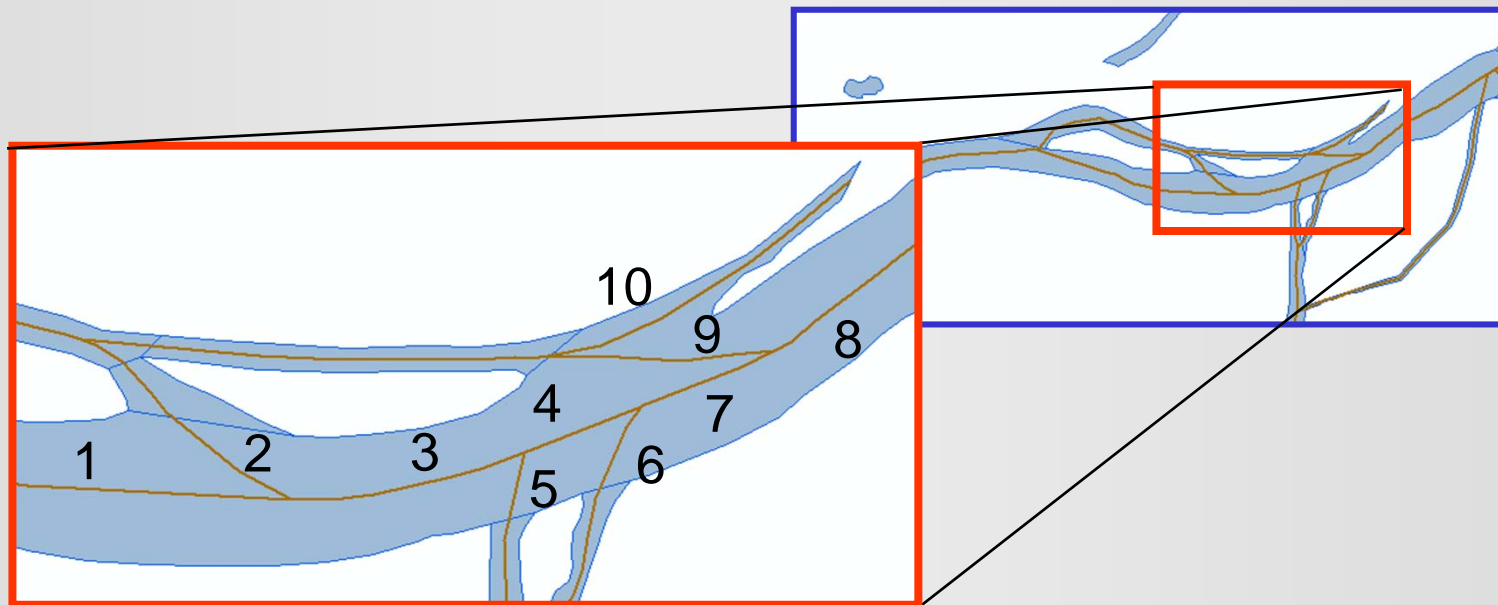
- Problems arise when centerlines are captured solely from imagery.



- **Best Practice: capture stream centerlines to lie inside captured stream polygons – base capture on imagery and previously collected hydrography.**

Case Study 2

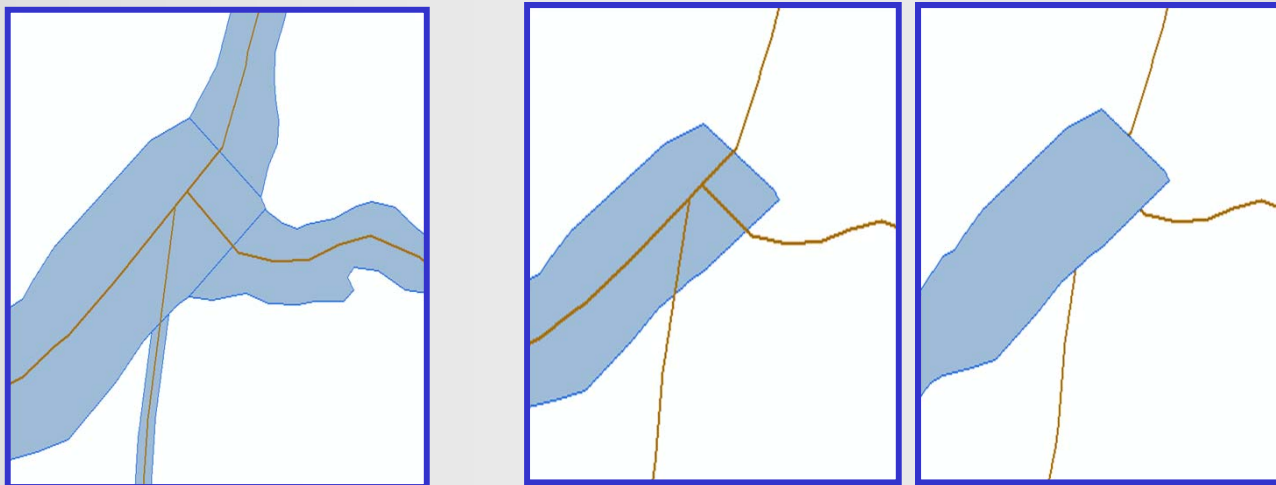
- Multiple centerlines inside stream polygon boundary



- Best Practice: capture polygons on basis of width. Attribute channel widths to centerlines for flexible selection and elimination with scale change.

Case Study 2: more

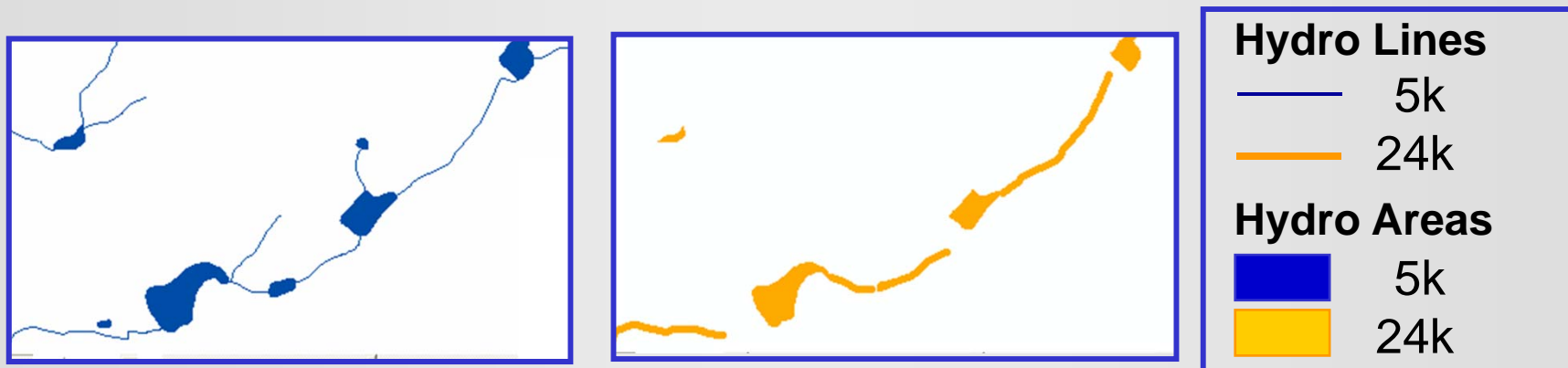
- **Centerline nodes do not correspond precisely with polygon boundaries.**
 - When centerlines are substituted for smaller channels, stream confluences appear illogical.
 - Problem not entirely corrected by modified drawing order.



- **Best Practice: capture no more centerlines than polygons. Don't use straight closure lines.**

Case Study 3

- Stream segments do not begin and end at adjacent hydrographic area extents.
 - Selection queries on stream width damage topology
 - Insert stream order as explicit attribute. Query on both.
 - Else code stream segments by stream “notoriety”.



- Best Practice: correct stream topology at the time of data capture.

Summary 1

- **Re-think best practices for data capture**
 - **Insure multiple use mapping at multiple mapping scales**
- **Data modeling informs data capture**
 - **Minimize data production costs**
 - **Maximize use for multiple products at multiple scales**

Summary 2

- **Strategic planning in advance of data capture:**
 - **Balances the data production workflow**
 - Where to concentrate labor, skill and time
 - **Formalizes a set of requirements for regional and local project managers**
 - **Maintains consistent quality for distributed workflows**
 - Data sharing

- **Best practice at the time of data capture:**
 - **Reduces subsequent processing**
 - **Sustains the goal of capturing data for general uses across a wide range of mapping scales**