Methods

Handling Spatial Data

HANDLING SPATIAL DATA MANUALLY

- 1940s or 1950s.
- Maps could be found in households, shops and businesses.
- many applications: identifying areas, routing delivery, planning, targeting services.

HANDLING SPATIAL DATA MANUALLY

- The concepts of sieve mapping were introduced.
- Each map layer is placed in light table and areas of interest.
- In the early days, colour pens and shading patterns to represent different values.

HANDLING SPATIAL DATA MANUALLY

- The techniques of sieve mapping were used by planning agencies, nature conservation agencies.
- Land capability maps show the best or most appropriate crop type such as slope, climate and wetness.

HANDLING SPATIAL DATA MANUALLY

- Manual sieve mapping is inaccurate.
- Route finding, for planning deliveries or an emergency evacuation from a disaster zone.

HANDLING SPATIAL DATA MANUALLY

- The string is used to trace the route and converted to distance using the map scale.
- Inaccurate, difficult to repeat, and time-consuming.

HANDLING SPATIAL DATA MANUALLY

- Calculating areas from maps is even more difficult
- Tracing areas onto graph paper and then ‘counting squares’.
- An innovative technique is used by the US Bureau of Census.

HANDLING SPATIAL DATA MANUALLY

- Manual techniques were in widespread use until computers
- Slow speed, difficulties in extracting, inaccuracies, and the inflexibility.
- Other problems associated with the media.
- Stretch and shrink), inconsistencies, destroyed, considerable space and difficult to transport.

HANDLING SPATIAL DATA MANUALLY

- Other spatial data, printed census
- Birth and death records, marriage certificates and other documents
- Inconsistencies and changes

THE DEVELOPMENT OF COMPUTER METHODS FOR HANDLING SPATIAL DATA

- Graphics technology,
- Data access and storage,
- Digitizing,
- Programming and interfaces; and
- Developments.

THE DEVELOPMENT OF COMPUTER METHODS FOR HANDLING SPATIAL DATA

- Areas of prime importance
- Hardware developments
- Areas of cartography

THE DEVELOPMENT OF COMPUTER METHODS FOR HANDLING SPATIAL DATA

- Harvard Graphics Laboratory played a role in the development of GIS(1991).
- In 1960s little commercial development.
- (CAD) and (AM/FM).

The systems approach

- Problem solving and decision making.
- Macroscopic view (1971).

The systems approach

- Boulding (1956), Von Bertalanffy (1968), Churchman (1968) and Ackoff (1971).
- Systems theory is the conceptual approach

The systems approach

- (Reeve, 1997)
- (Goodchild, 1995).
- (Burrough, 1986; Aronoff, 1989);

The systems approach

- (Ward, 1975) and (Huggett, 1980).
- Methodological framework.
- Two approaches, hard and soft systems analysis

Computer cartography

- Cartography(Mather, 1991).
- Atlas (Bickmore and Shaw, 1963)
- SYMAP (Goodchild, 1988).
- Process of map production (Tomlinson, 1988).

Spatial statistics

- Spatial statistics and geographical analysis methods (Tomlinson, 1988).
- Middle of the twentieth century.
- During 1960s huge increase in the application

AM/FM and CAD

- Automated mapping / facilities management and computer-aided design are two areas of technology
- 1970s.
- (Reina, 1997).
- (Henry and Pugh, 1997).
- GIS/CAD hybrid packages

THE DEVELOPMENT OF GIS

- (Tomlinson, 1988).
- (Tomlinson, 1990).
- First GIS appeared in 1964.
- The Canadian Geographic Information System (CGIS).

THE DEVELOPMENT OF CGIS

- The land inventory also classified.
- (Crain, 1985).
- CGIS is 1964, until 1971.
- (Symington, 1968).
- Canada Land Data System.
- It was the first general-purpose GIS

THE DEVELOPMENT OF CGIS

- The first GIS to employ the data structure(1975)
- (Crain, 1985).
- (Heywood, 1990)

THE DEVELOPMENT OF GBF-DIME

- The development of the GBF-DIME late 1960s was a major step forward in GIS data models.
- some of the major developments:
- 1965
- 1966
- 1967
- 1970s
- DIME assisted efficient digitizing and

The development of ESRI

- 1969 ESRI
- 1970s ESRI
- 1981 ARC/INFO GIS
- 1980s ESRI
- 1986 PC ARC/INFO

The development of ESRI

- 1991 ARCVIEW
- 1992
- 1995 SDE
- 1996 ARC/INFO

The development of ESRI

- 1997 ESRI
- 1998
- 1999 ArcInfo
- 2000 Geography Network

The development of ESRI

- 2001 ESRI.
- 2002
- 2003
- 2004
- 2005

The development of GIS in 1970's

- GIMMS, MAPICS and SURFACE II.
- In 1971.
- (Rhind, 1987).
- (Tomlinson, 1990).
- (Tomlinson, 1988).

The development of GIS in 1970’s

- (Coppock and Rhind, 1991).
- (Tomlinson, 1990).
- (Taylor, 1991a)
- Late 1980s and 1990s.
- Mid-1980s (Rhind, 1987).

The development of GIS in 1980's

- (Maguire, 1989).
- 1986.
- (Taylor, 1990).
- (Martin, 1991).
- (Goodchild, 1988).
- (Goodchild and Rhind, 1990; Masser, 1990).

The development of GIS in 1980’s

- Burrough (1986) and Peuquet and Marble (1990);
- (Unwin *et al., 1990*)
- (Goodchild and Kemp, 1990).

The development of GIS in 1990's

- CORINE
- The main achievements were:
- the standardization and methods;
- the demonstration of the feasibility; and
- the development of similar activities.

DATA QUALITY ISSUES

Data quality and errors – Sources of error in GIS – Finding and modeling errors of GIS – Managing GIS error

INTRODUCTION

- Quality is a word in a dictionary define that ‘degree of excellence’.
- In GIS, *data quality is used to give an indication* of how good data are.

INTRODUCTION

- In many industries, quality control is everything.
- GIS users strive for quality products from their systems.

INTRODUCTION

- Errors can enter a GIS in many different ways.
- For example:
- A mistake in digitizing a field boundary.
- A mistake in the transcription of a grid reference.
- Misinterpretation of satellite imagery.

INTRODUCTION

- But what is a high-quality product?
- What exactly are good quality data?
- How can we describe and recognize poor quality output?

INTRODUCTION

- Two issues are:
- First, the terminology, and
- Second, the sources, propagation and management.
- Describing data problems in GIS is difficult since many of the words used are also common in everyday language.
- Words such as quality, accuracy and error not only mean different things to different people but also have precise technical definitions.

INTRODUCTION

- The terms used for data error and quality are introduced to solve problems.
- The types and sources of errors to recognize and deal with problems.

INTRODUCTION

- Techniques for modelling and managing errors are.
- A few data sets are error-free.
- GIS users should document the limitations.

INTRODUCTION

- Earlier in the book we compared the fuel in a car to data in a GIS.
- In the same way that a poor quality fuel may cause problems with the running of the car, poor quality data will introduce errors into your GIS.

DESCRIBING DATA QUALITY AND ERRORS

- In GIS, *data quality is used to give an indication* of good data.

DESCRIBING DATA QUALITY AND ERRORS

- Issues such as *error, accuracy, precision and bias can help to assess the quality*.
- In addition, the *resolution and generalization*.

DESCRIBING DATA QUALITY AND ERRORS

- A systematic error would have occurred.

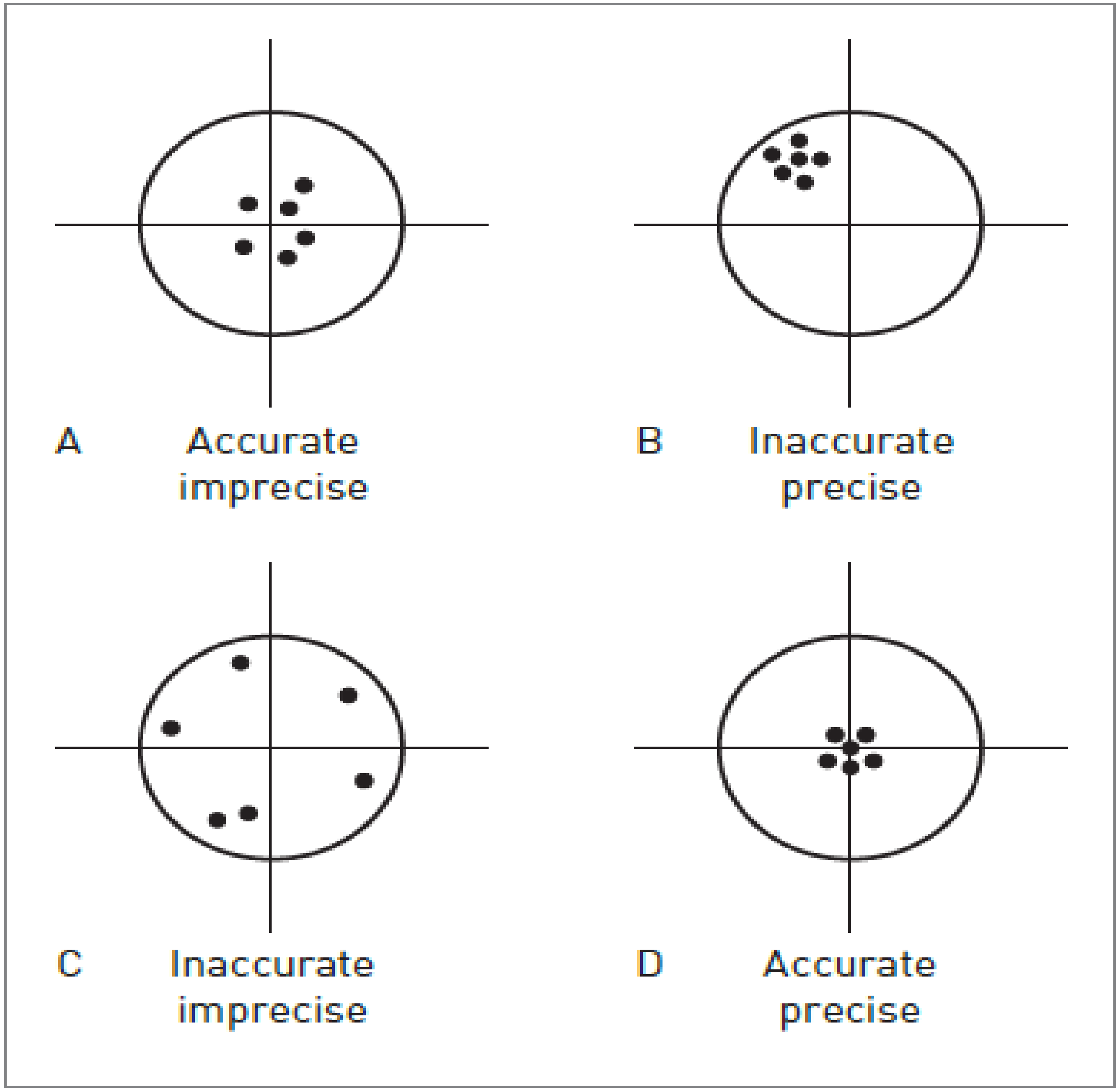
DESCRIBING DATA QUALITY AND ERRORS

- Accuracy of reality.

DESCRIBING DATA QUALITY AND ERRORS

- Precision is the recorded.

The difference between accuracy and precision is Shown:



DESCRIBING DATA QUALITY AND ERRORS

- Bias in GIS data is the systematic variation of data

Resolution and generalization of raster datasets

(a) 25m resolution vegetation map

(b) 5m resolution colour aerial photograph

DESCRIBING DATA QUALITY AND ERRORS

- Generalization is the process of simplifying the complexities.

DESCRIBING DATA QUALITY AND ERRORS

- *Completeness, compatibility, consistency and applicability are introduced here.*
- A complete will cover period of interest.

DESCRIBING DATA QUALITY AND ERRORS

- Maps containing data measured in different scales of measurement cannot be combined easily.

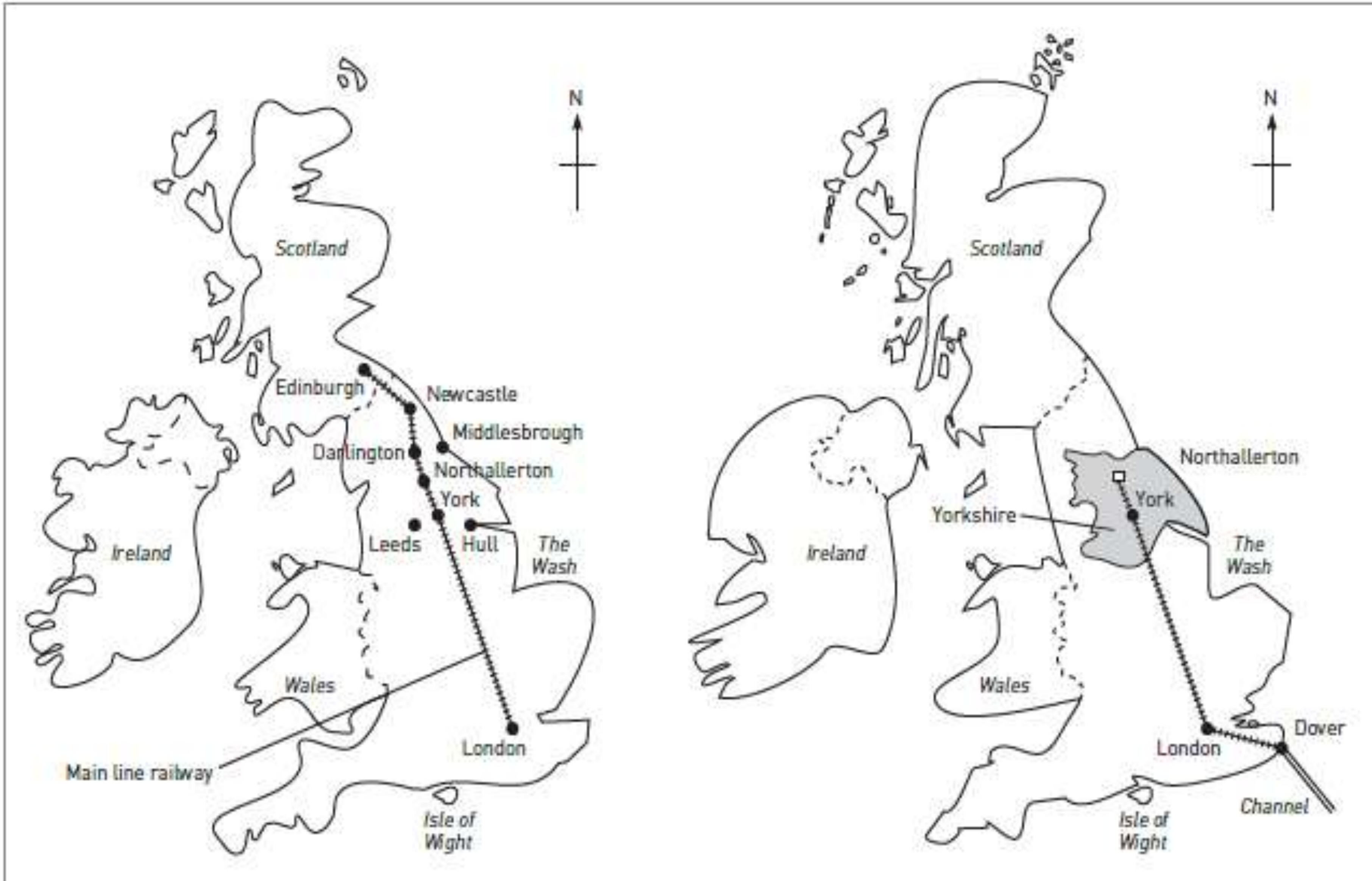
DESCRIBING DATA QUALITY AND ERRORS

- Consistency applies not only to separate data sets but also within individual data sets.

SOURCES OF ERROR IN GIS

- Spatial and attribute errors can occur at any stage

MENTAL MAPS


N
Scotland
Edinburgh
Newcastle
Middlesbrough
Darlington
Northallerton
York
Leeds
Hull
The Wash
Ireland
Wales
London
Main line railway
Isle of Wight
N
Scotland
Northallerton
York
Yorkshire
Ireland
The Wash
Wales
Dover
London
Channel
Isle of Wight

Errors arising from our understanding and modelling of reality

- Our perception of reality influences our definition of reality.

Errors arising from our understanding and modelling of reality

- Consider the problem of defining a common geographical feature such as a mountain in a GIS.

Finsteraarhorn, Switzerland (4273m)

The Finsteraarhorn, Switzerland (4273 m)

Terrain model of Mount Everest (8849 m)

Terrain model of Mount Everest and its surrounding area based on photogrammetric survey data. 3D vizualization of Mount Everest (*Source*: Martin Sauerbier/Institute of Geodesy and Photogrammery)

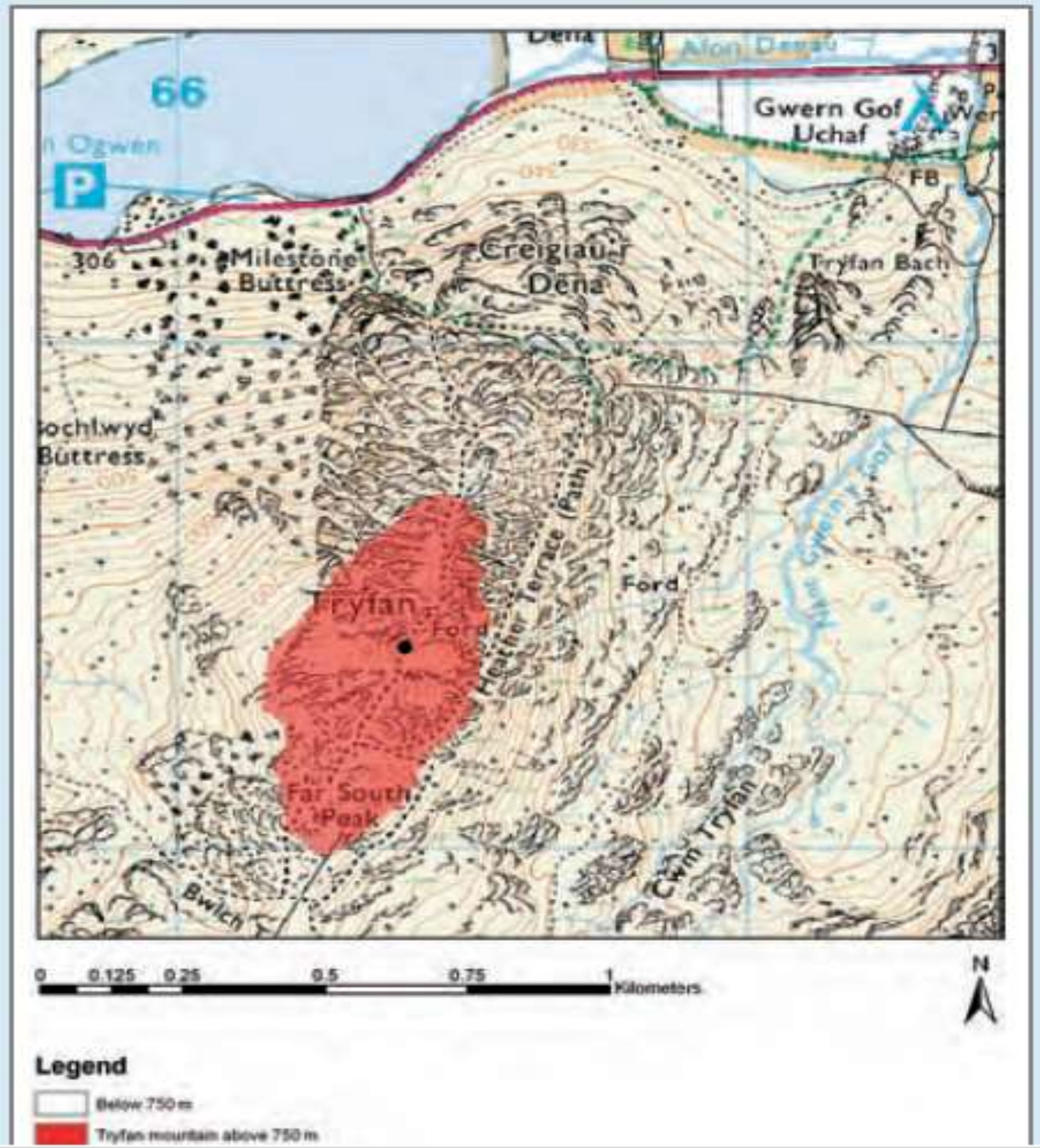
Errors arising from our understanding and modelling of reality

- The problem of identifying mountains.

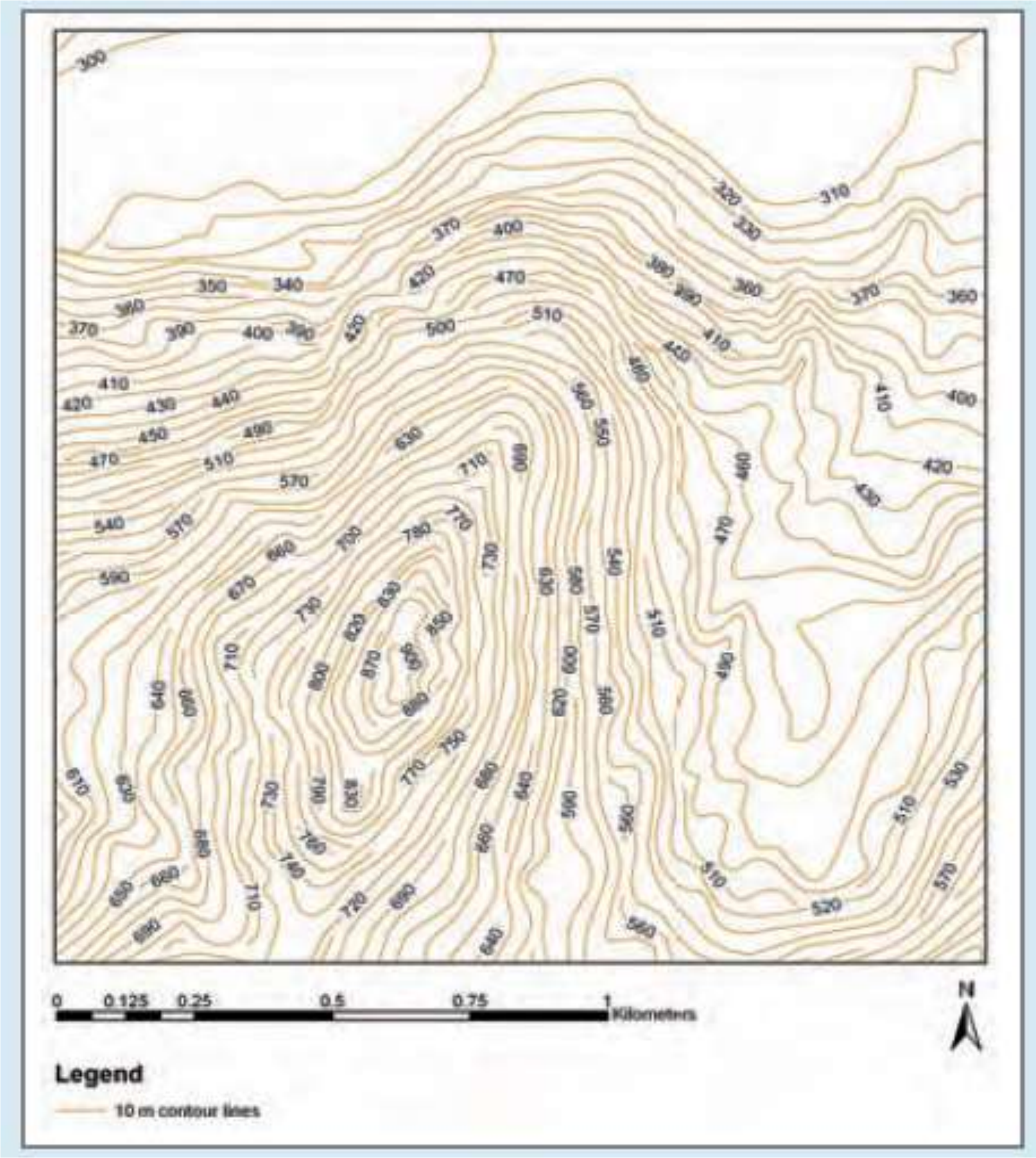
Visualize the problem for Tryfan; a mountain in north Wales.

Photograph of Tryfan

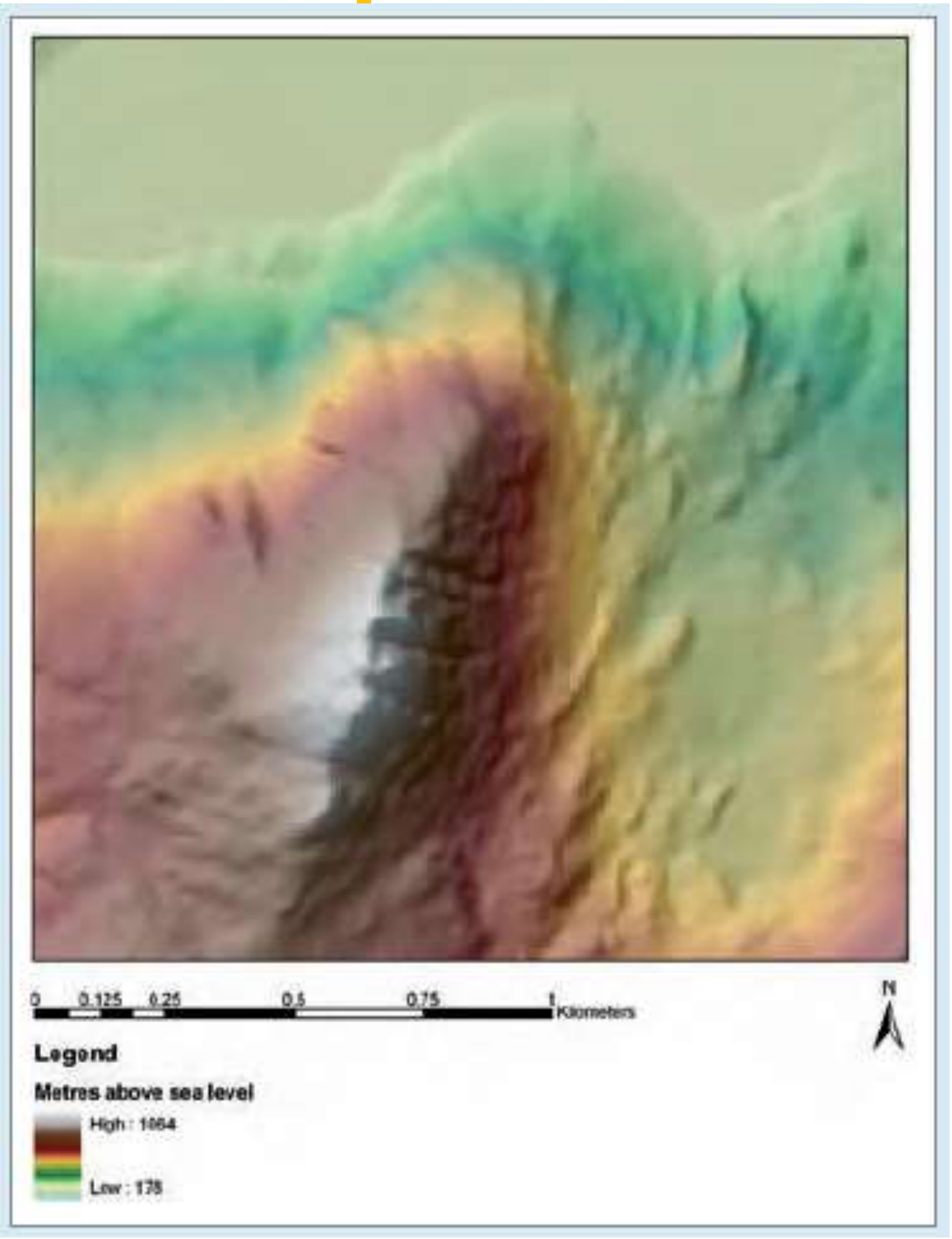
Ordnance Survey 1:50,000 topographic map showing area above 750m



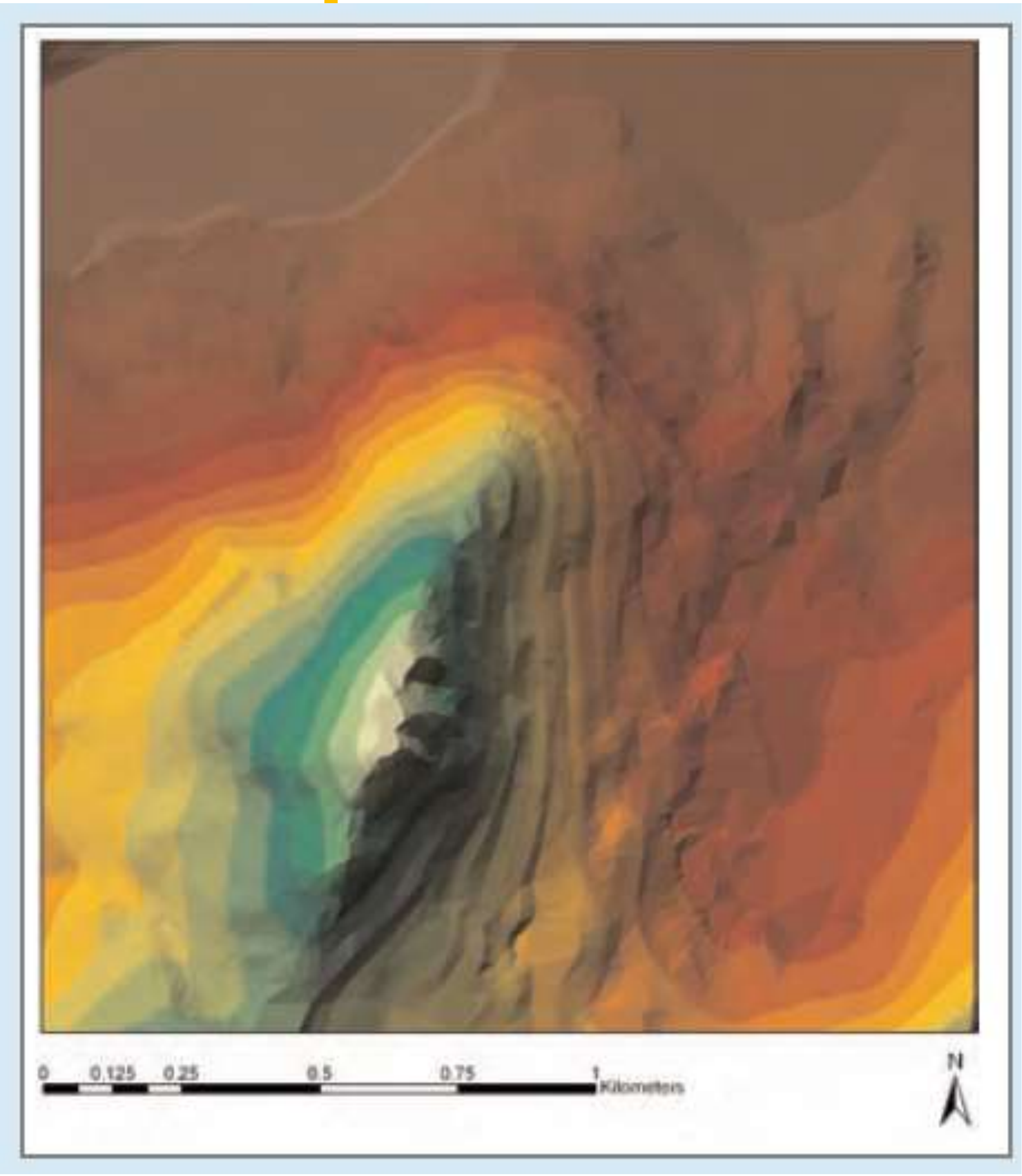
Contour map



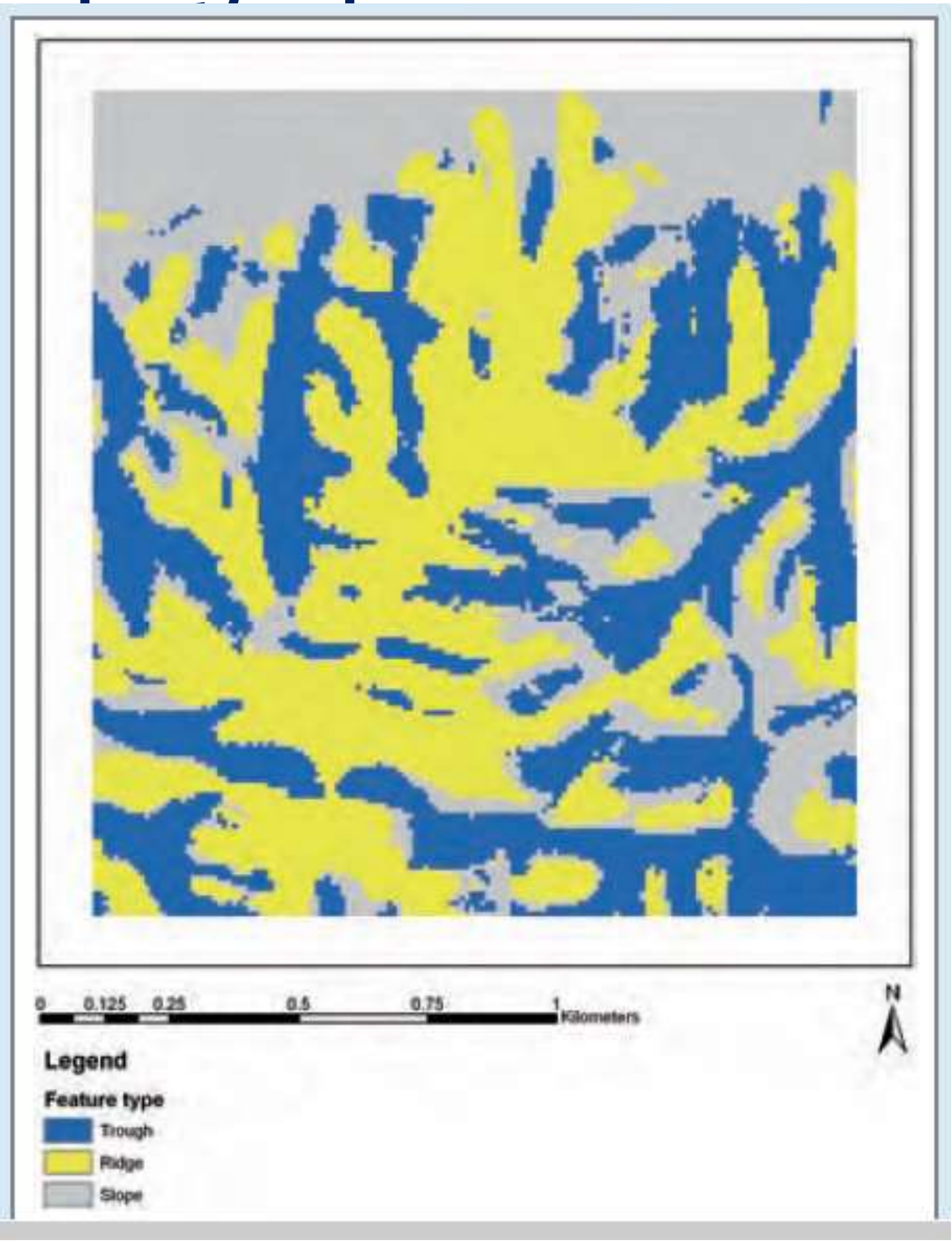
DEM Representation



TIN Representation



Topographic features



Errors in source data for GIS

- **All sources of spatial** and attribute data are likely to include errors.

Errors in source data for GIS

- Cartographic data sources error are:
  - CONTINUOUS DATA
  - FUZZY BOUNDARIES
  - MAP SCALE
  - MAP MEASUREMENTS

DATA QUALITY ISSUES

Data quality and errors – Sources of error in GIS – Finding and modeling errors of GIS – Managing GIS error

INTRODUCTION

- Quality is a word in a dictionary define that ‘degree of excellence’.
- In GIS, *data quality is used to give an indication* of how good data are.

INTRODUCTION

- In many industries, quality control is everything.
- GIS users strive for quality products from their systems.

INTRODUCTION

- Errors can enter a GIS in many different ways.
- For example:
- A mistake in digitizing a field boundary.
- A mistake in the transcription of a grid reference.
- Misinterpretation of satellite imagery.

INTRODUCTION

- Two issues are:
- First, the terminology, and
- Second, the sources, propagation and management.
- Words such as quality, accuracy and error not only mean different things to different people but also have precise technical definitions.

INTRODUCTION

- data error and quality are introduced to solve problems.
- sources of errors to recognize and deal with problems.

DESCRIBING DATA QUALITY AND ERRORS

- In GIS, *data quality is used to give an indication* of good data.

DESCRIBING DATA QUALITY AND ERRORS

- Issues such as *error, accuracy, precision and bias can help to assess the quality*.
- In addition, the *resolution and generalization*.

DESCRIBING DATA QUALITY AND ERRORS

- A systematic error would have occurred.

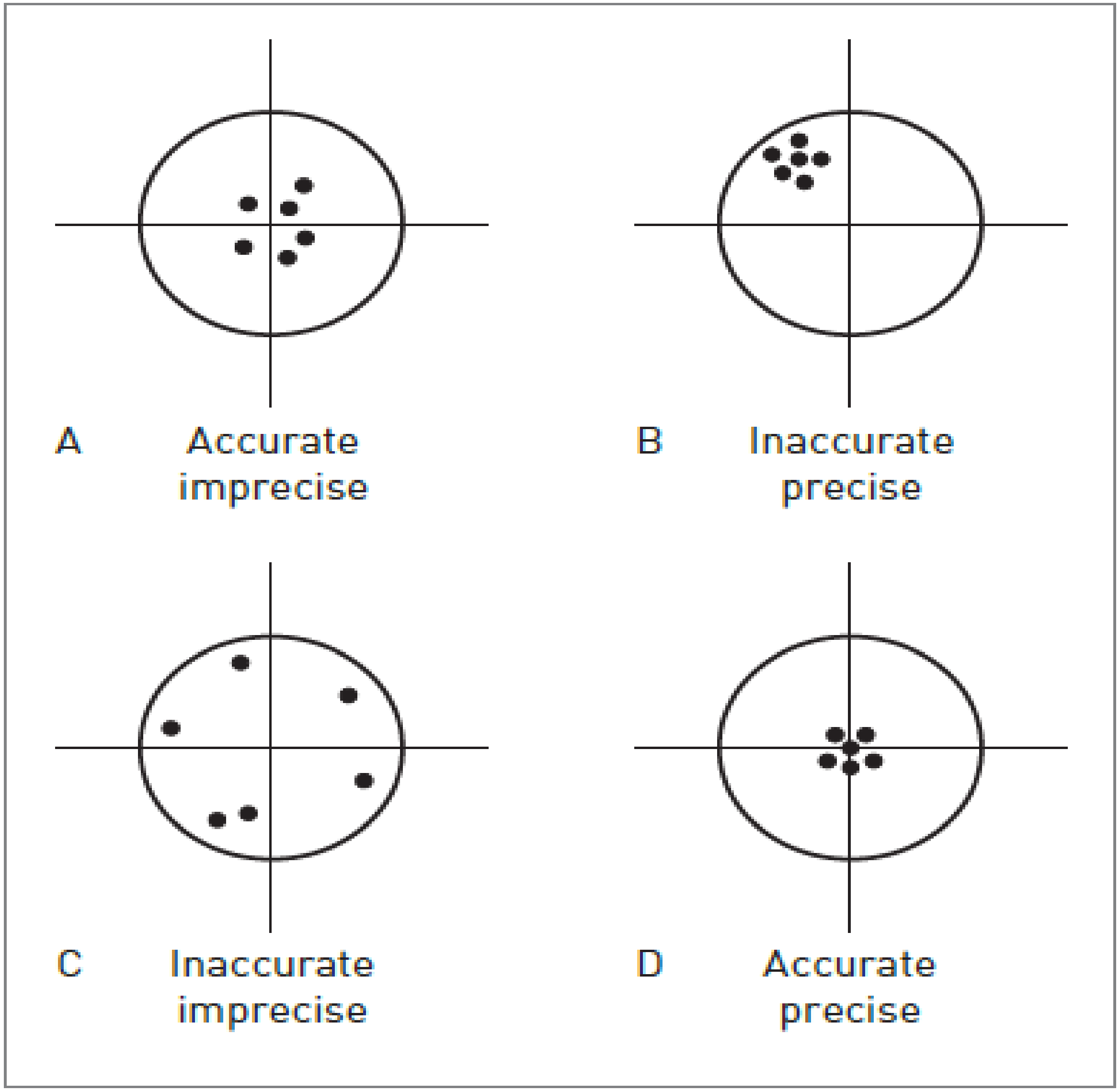
DESCRIBING DATA QUALITY AND ERRORS

- Accuracy of reality.

DESCRIBING DATA QUALITY AND ERRORS

- Precision is the recorded.

The difference between accuracy and precision is Shown:



DESCRIBING DATA QUALITY AND ERRORS

- Bias in GIS data is the systematic variation of data

Resolution and generalization of raster datasets

(a) 25 m resolution vegetation map

(b) 5 m resolution colour aerial photograph

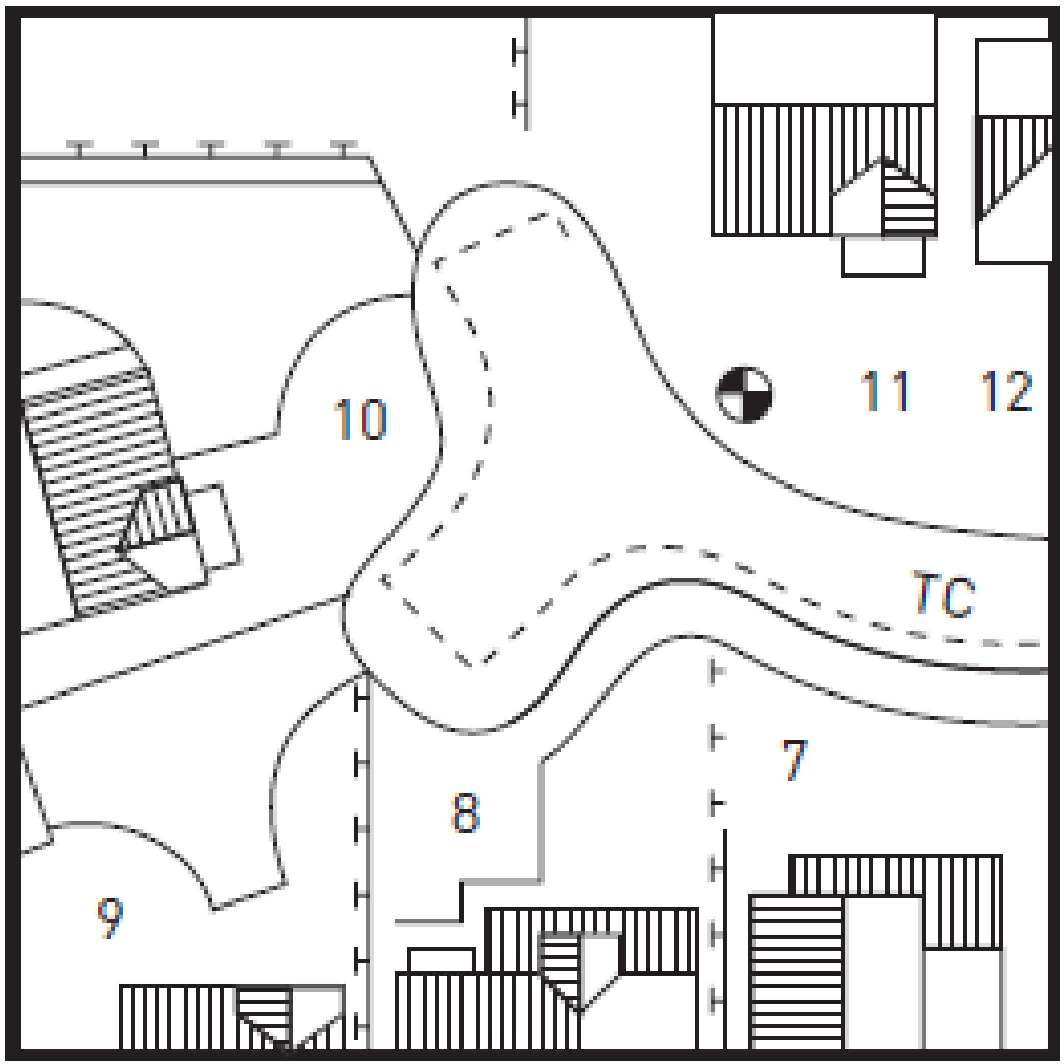
DESCRIBING DATA QUALITY AND ERRORS

- Generalization is the process of simplifying the complexities.

DESCRIBING DATA QUALITY AND ERRORS

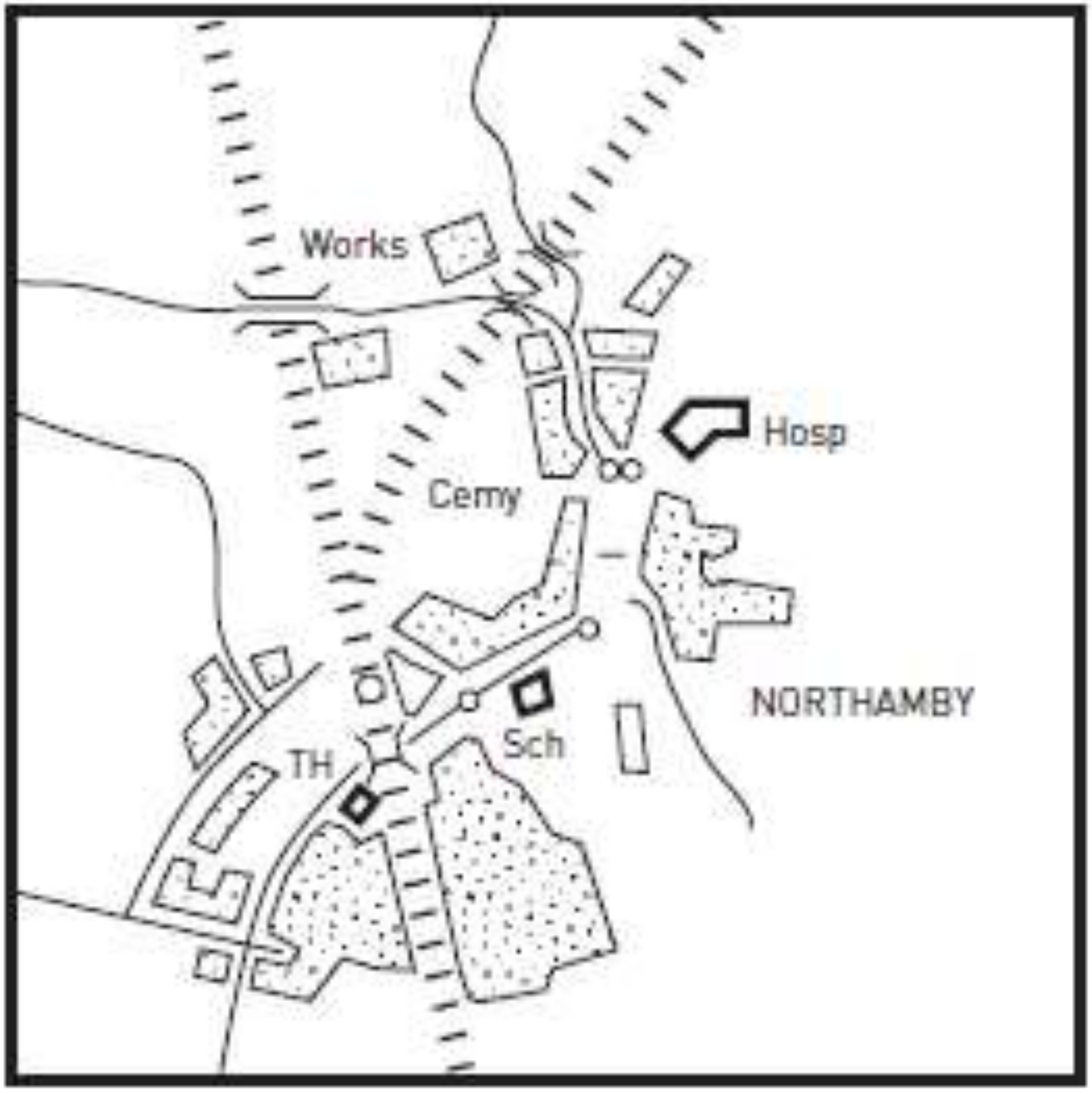
- *Completeness, compatibility, consistency and applicability are introduced here.*
- A complete will cover period of interest.

Scale-related generalization



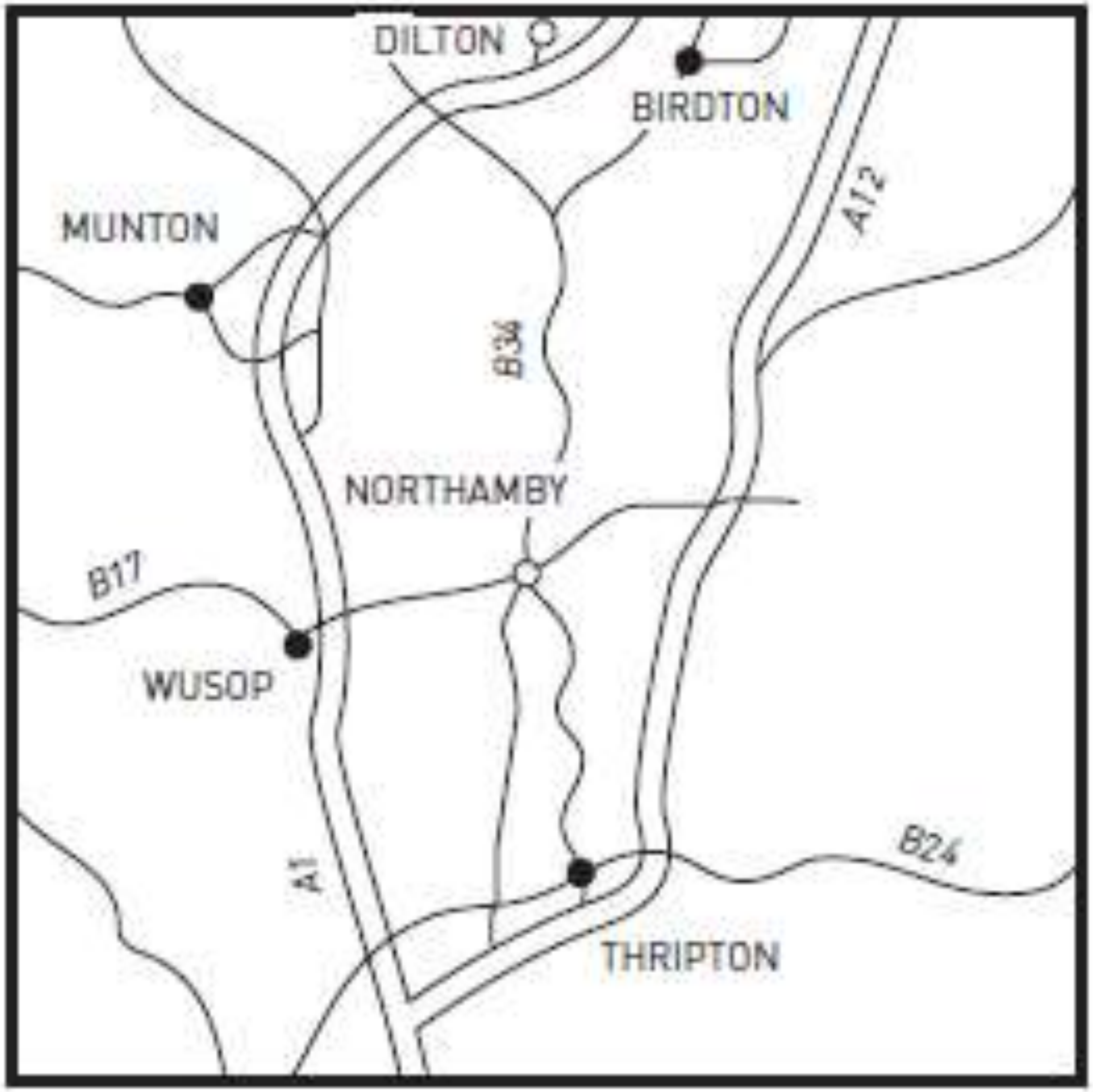
1:10,000

Scale-related generalization



1:50,000

Scale-related generalization



1:500,000

DESCRIBING DATA QUALITY AND ERRORS

- Maps containing data measured in different scales of measurement cannot be combined easily.

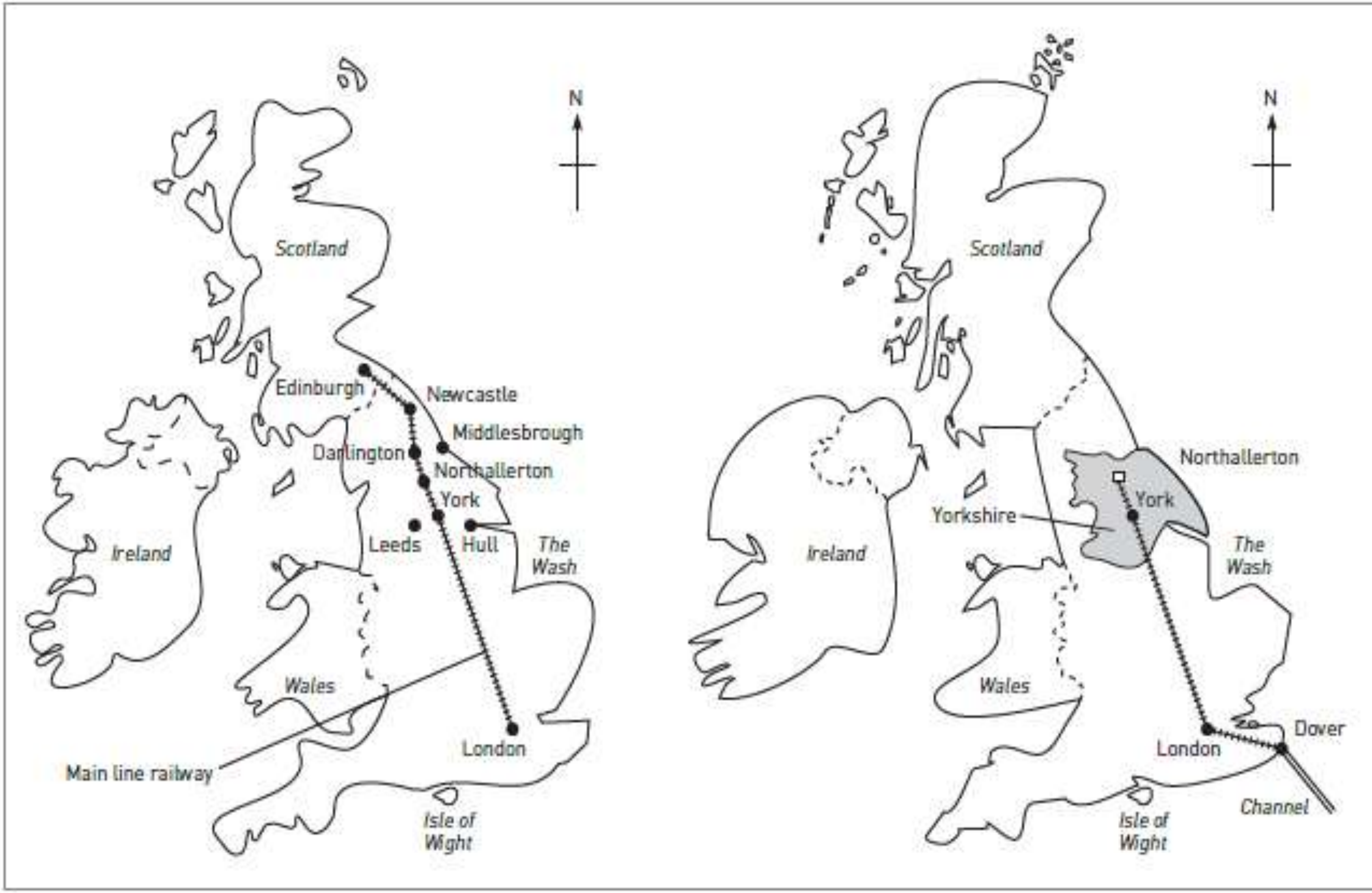
DESCRIBING DATA QUALITY AND ERRORS

- Consistency applies not only to separate data sets but also within individual data sets.

SOURCES OF ERROR IN GIS

- Spatial and attribute errors can occur at any stage

MENTAL MAPS


N
Scotland
Edinburgh
Newcastle
Middlesbrough
Darlington
Northallerton
York
Leeds
Hull
The Wash
Ireland
Wales
London
Main line railway
Isle of Wight
N
Scotland
Northallerton
York
Yorkshire
The Wash
Ireland
Wales
London
Dover
Channel
Isle of Wight

Errors arising from our understanding and modelling of reality

- Our perception of reality influences our definition of reality.

Errors arising from our understanding and modelling of reality

- Consider the problem of defining a common geographical feature such as a mountain in a GIS.

Finsteraarhorn, Switzerland (4273m)

The Finsteraarhorn, Switzerland (4273 m)

Terrain model of Mount Everest (8849 m)

Terrain model of Mount Everest and its surrounding area based on photogrammetric survey data. 3D vizualization of Mount Everest (*Source*: Martin Sauerbier/Institute of Geodesy and Photogrammery)

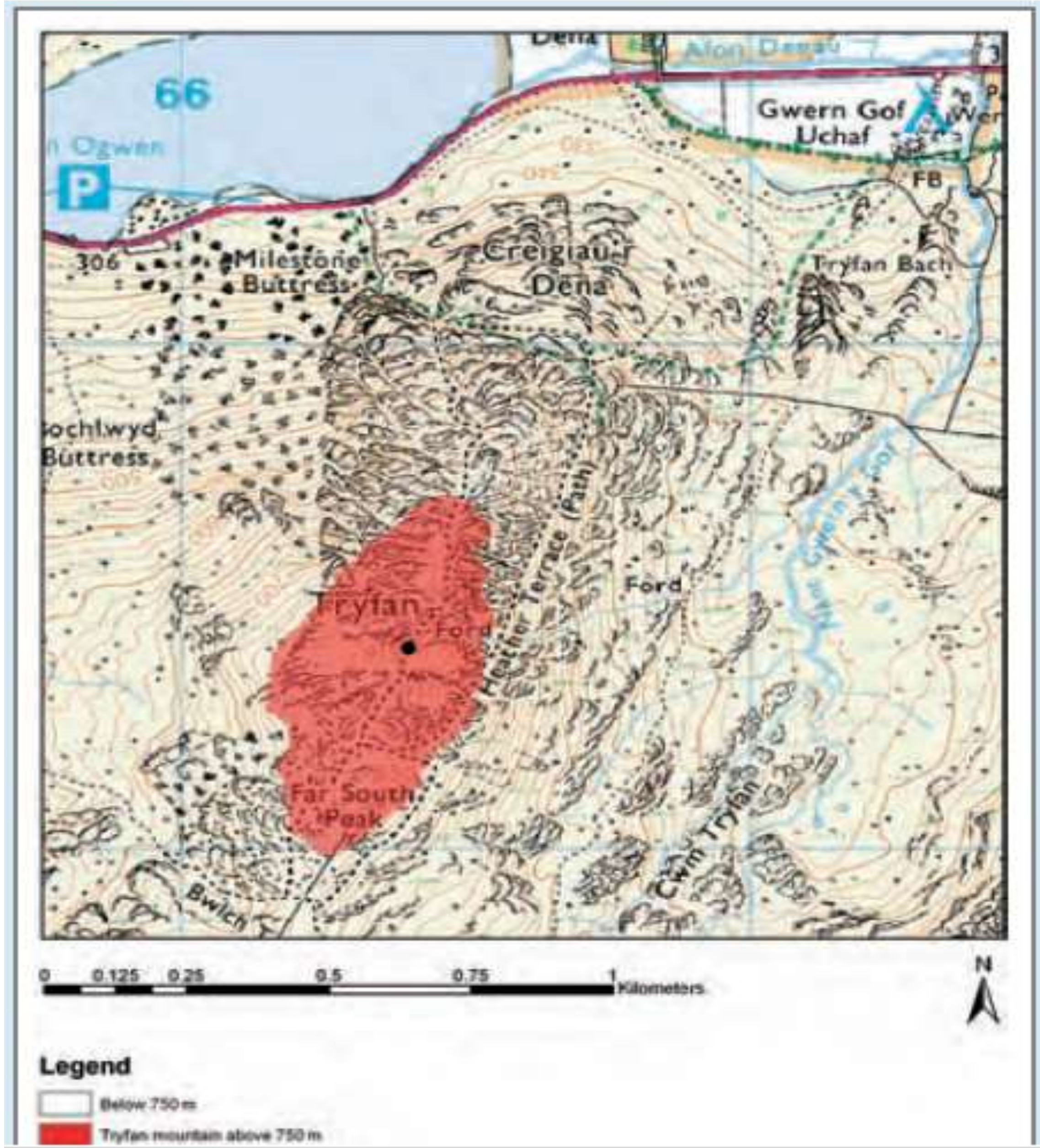
Errors arising from our understanding and modelling of reality

- The problem of identifying mountains.

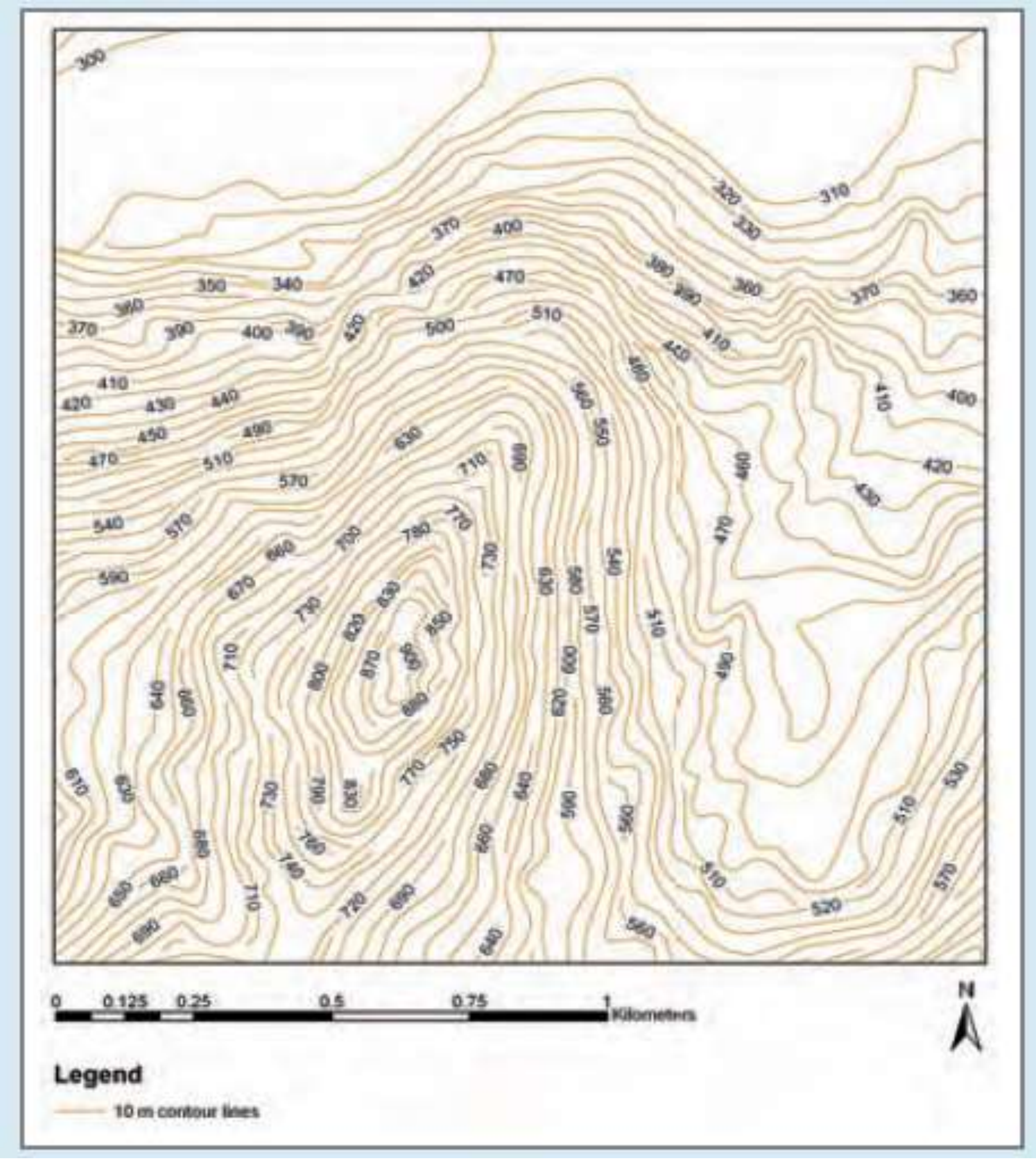
Visualize the problem for Tryfan; a mountain in north Wales.

Photograph of Tryfan

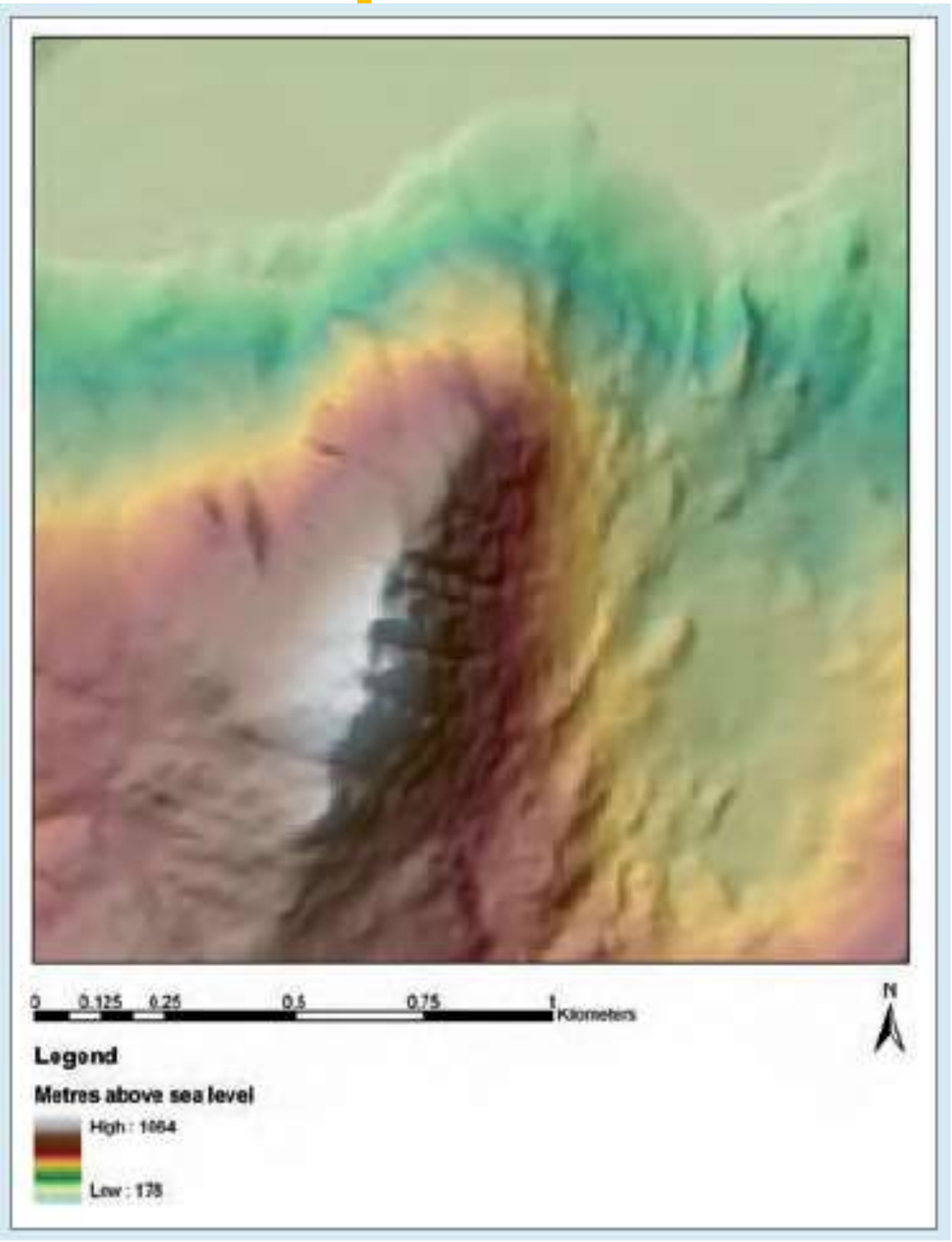
Ordnance Survey 1:50,000 topographic map showing area above 750m



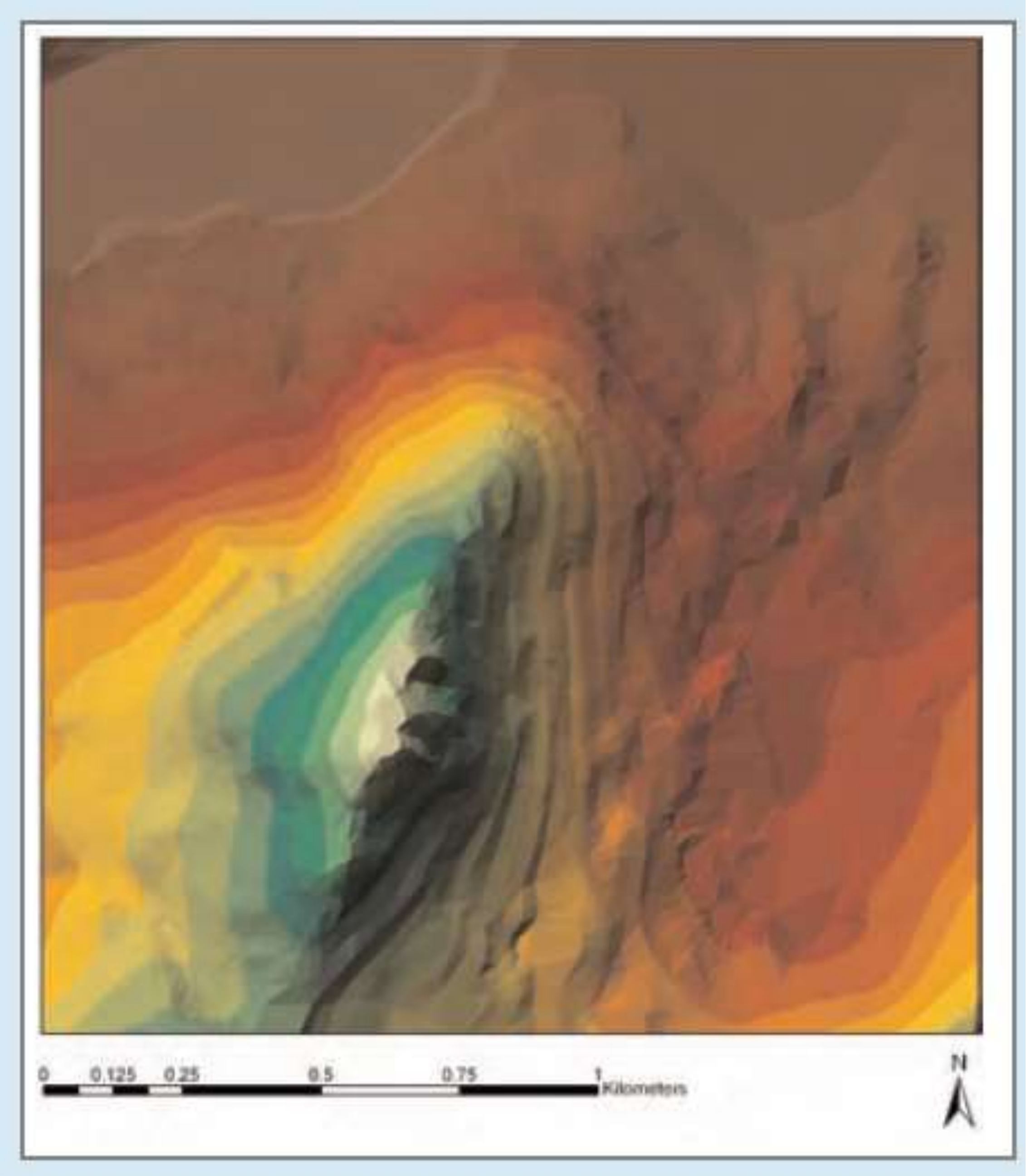
Contour map



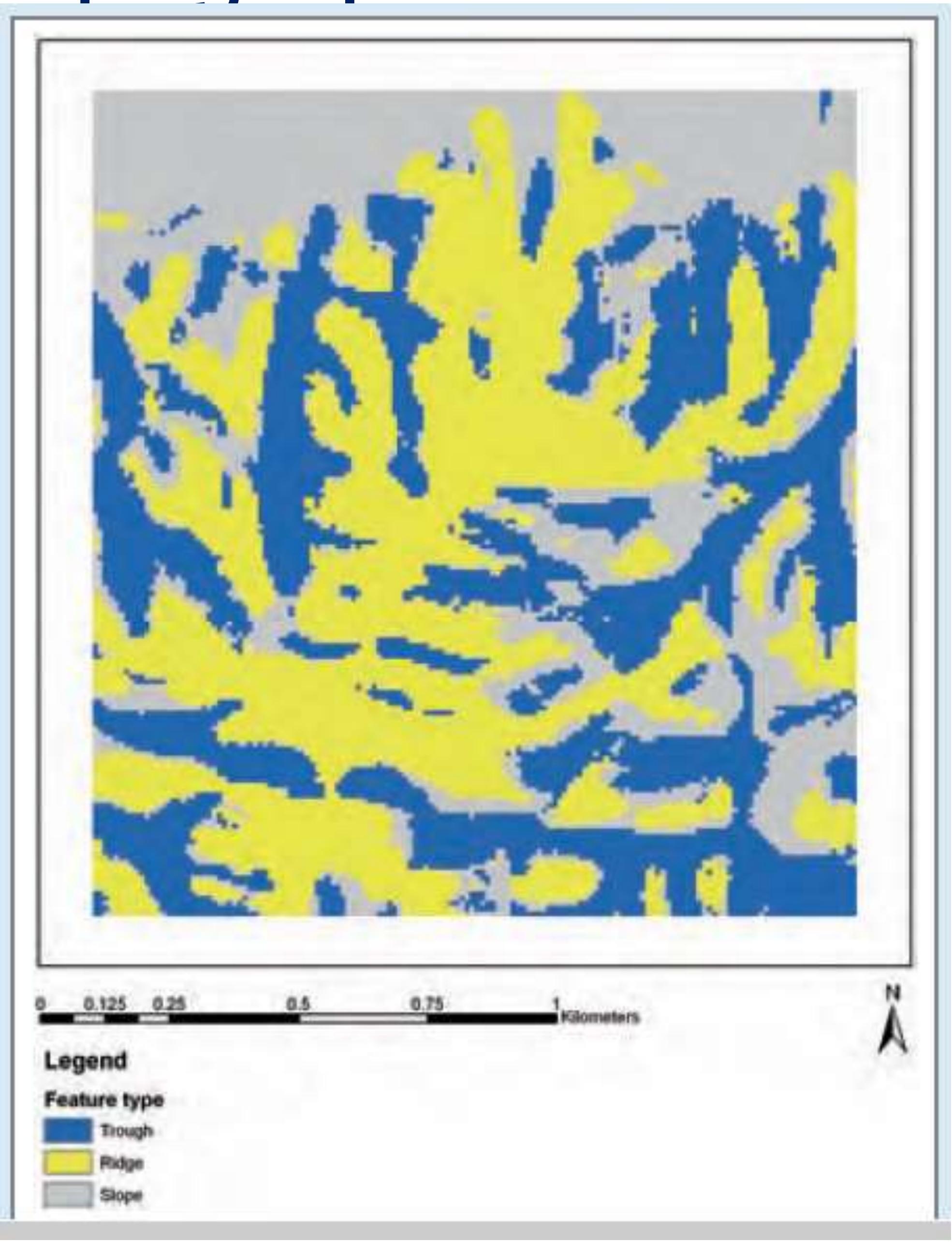
DEM Representation



TIN Representation



Topographic features



Errors in source data for GIS

- **All sources of spatial** and attribute data are likely to include errors.

Errors in source data for GIS

- Cartographic data sources error are:
  - CONTINUOUS DATA
  - FUZZY BOUNDARIES
  - MAP SCALE
  - MAP MEASUREMENTS

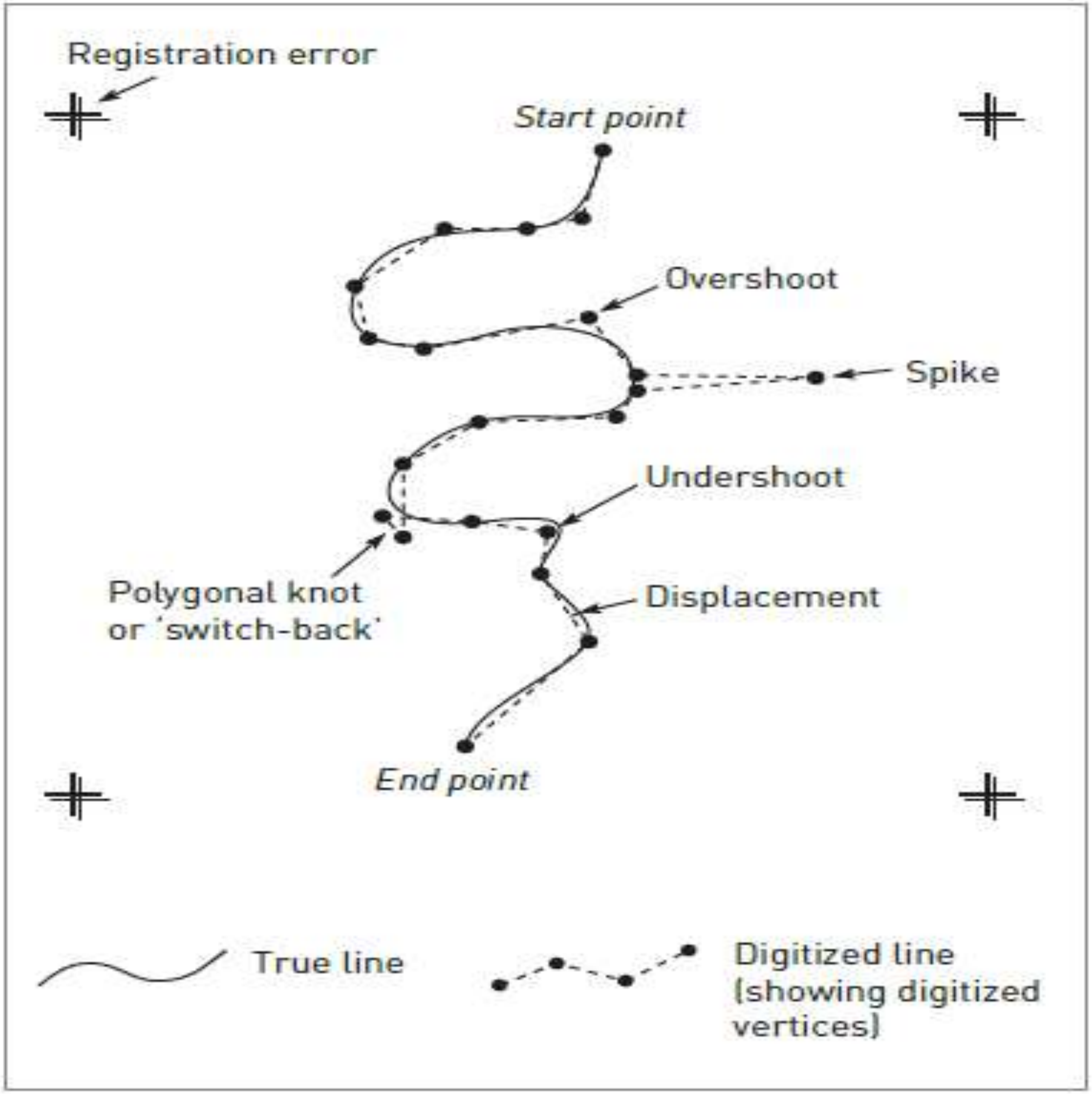
Errors in data encoding

- Data encoding is the process by which data are transferred from some non-GIS source.

Errors in data encoding

- Two main types:
- source map error and
- operational error.
- Categorizes human digitizing error into two types:
  - **psychological and**
  - **physiological.**

Operational errors



Operational errors

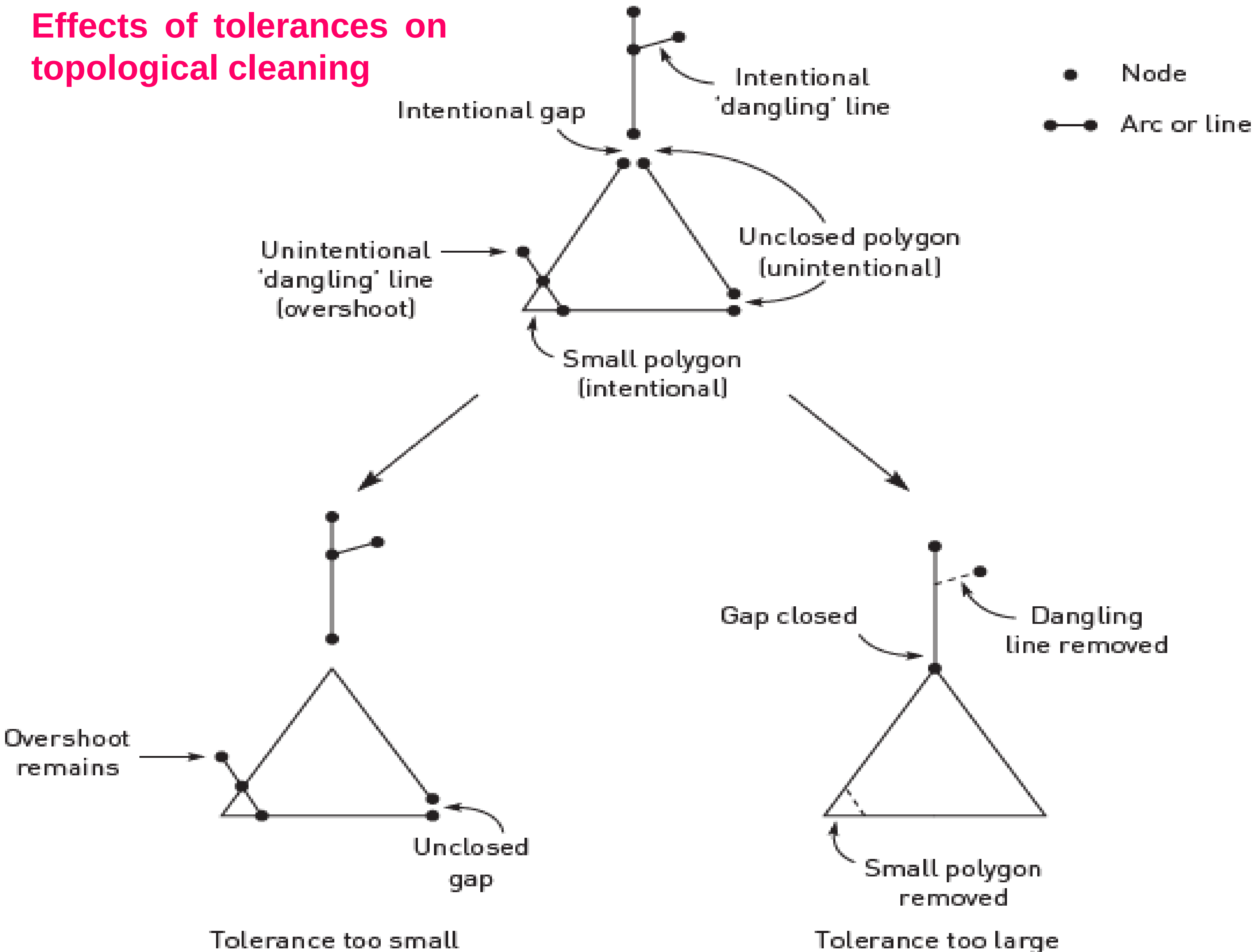
- 3 *Line thickness.*
- 4 *Method of digitizing. Two methods* :
  - point mode and
  - stream mode.

Errors in data editing and conversion

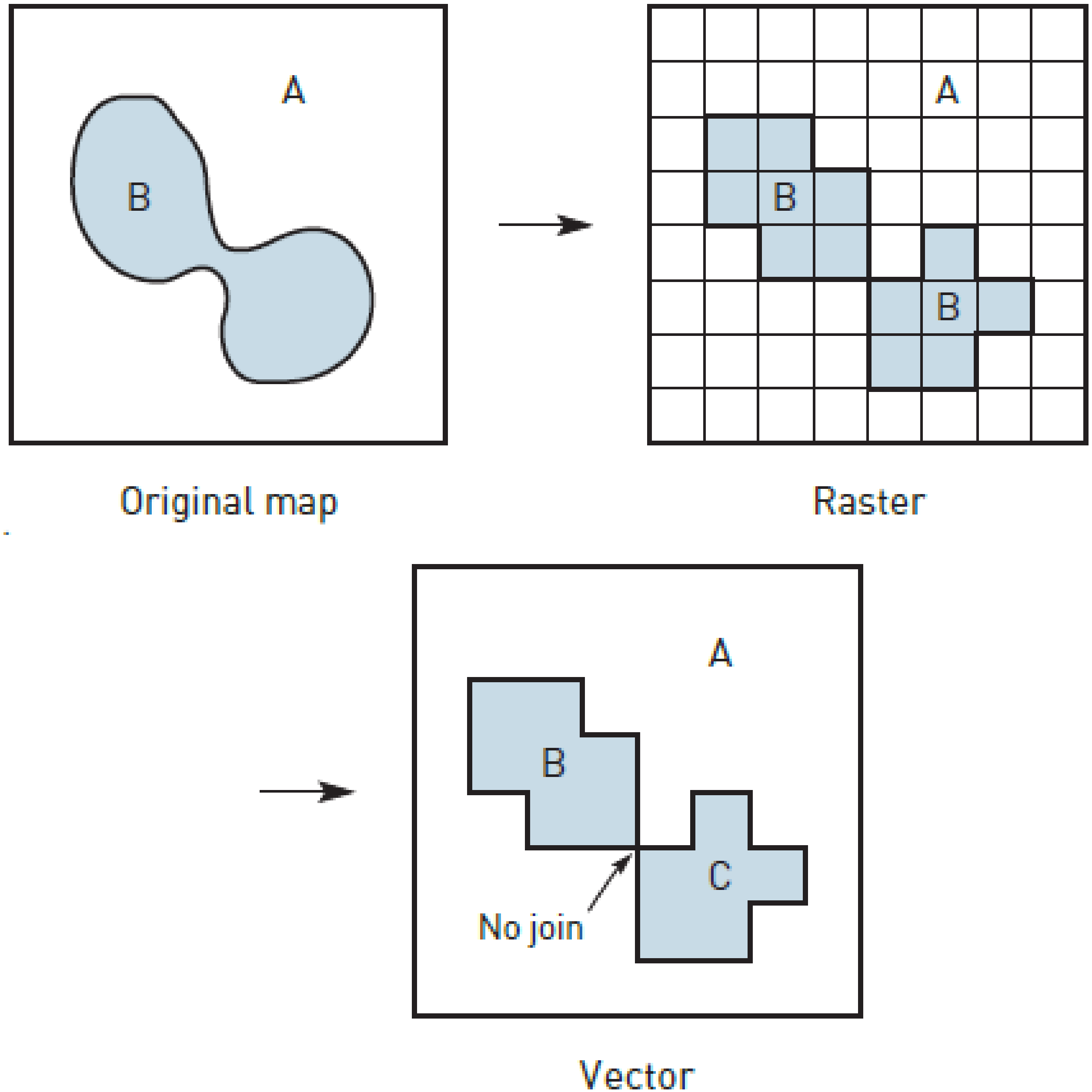
- After data encoding is complete, cleaning and editing.

Topological errors in vector GIS

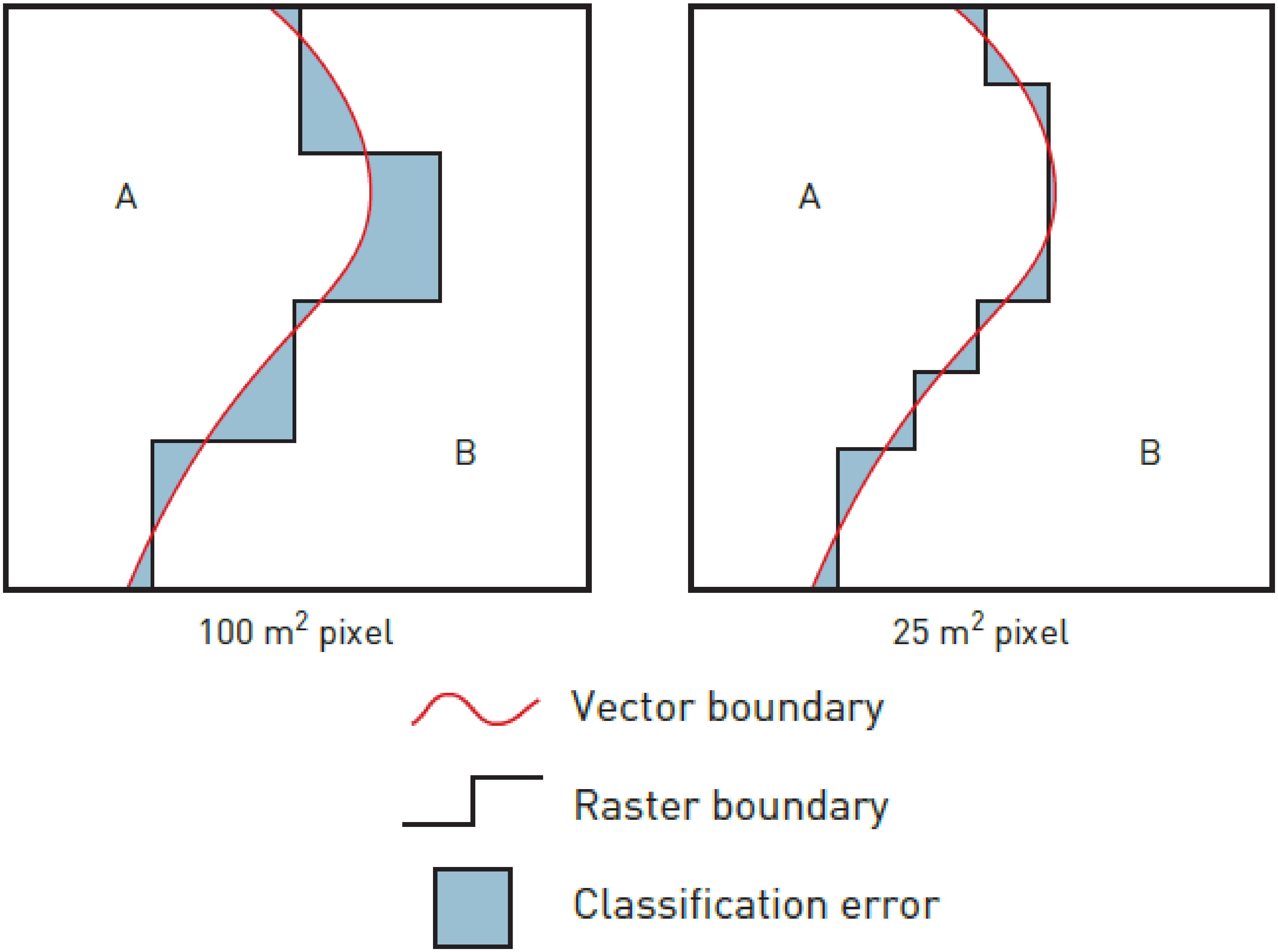
Effects of tolerances on topological cleaning



Topological ambiguities in raster to vector conversion



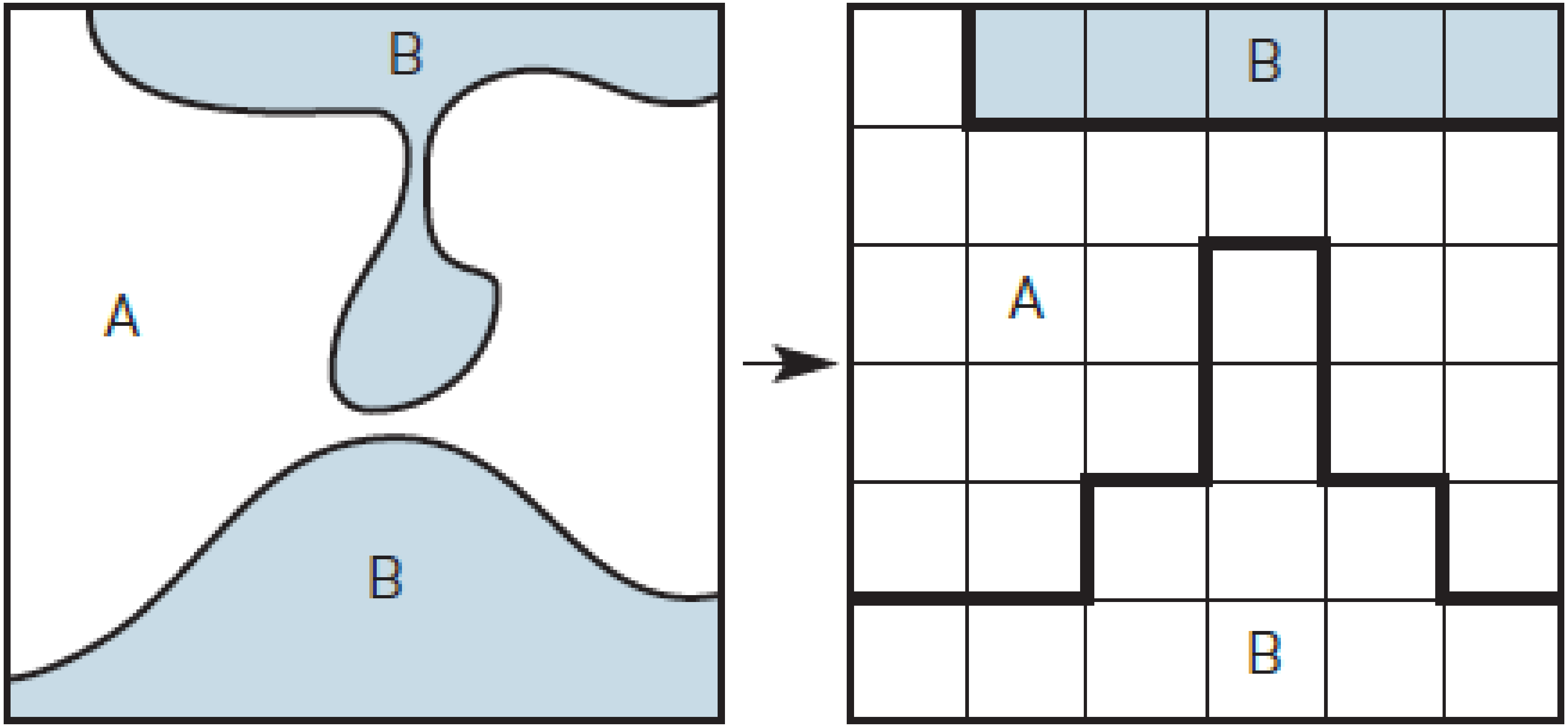
Vector to raster classification error



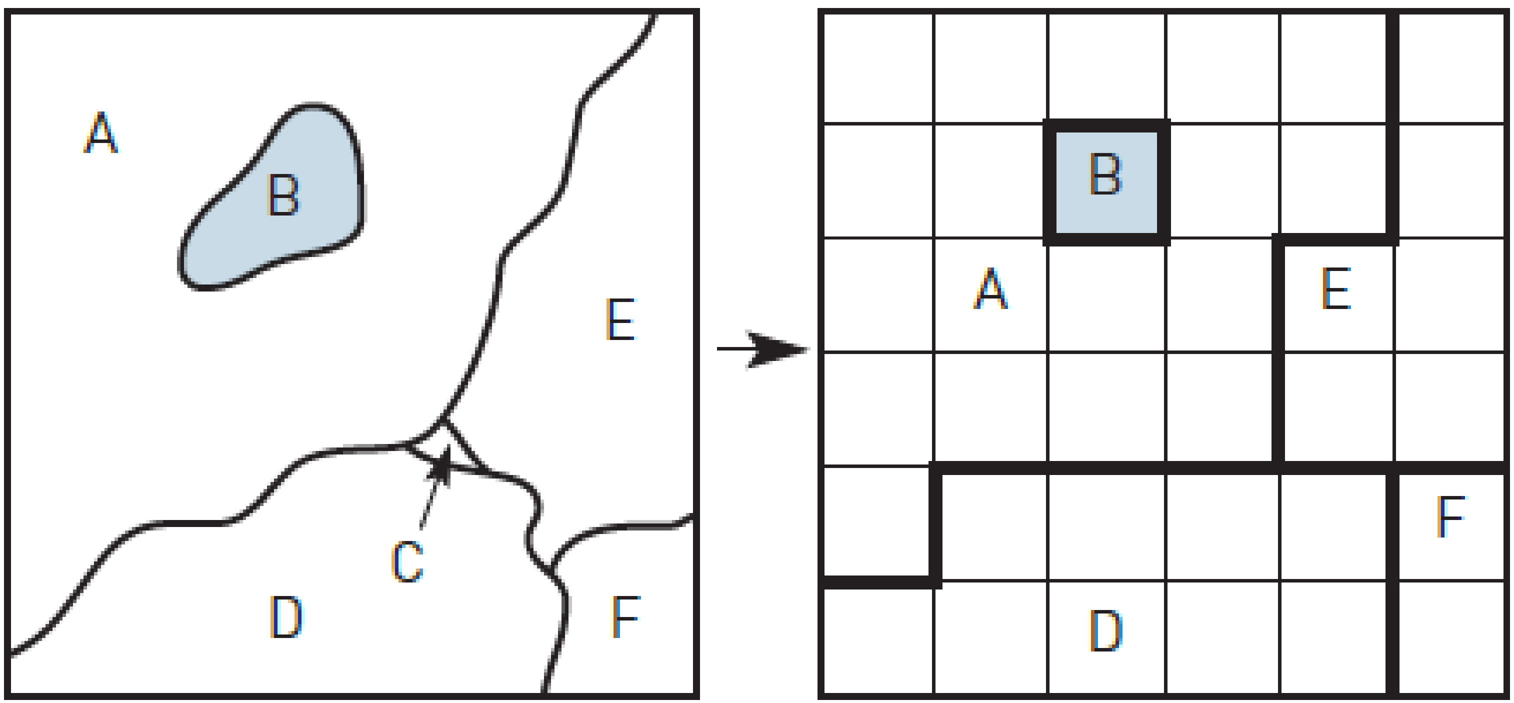
Rasterization errors

- Vector to raster conversion.
- For example:
- 1 *Topological errors.*
- *2 Loss of small polygons.*
- 3 *Effects of grid orientation.*
- 4 *Variations in grid origin and datum.*

Loss of connectivity and creation of false connectivity



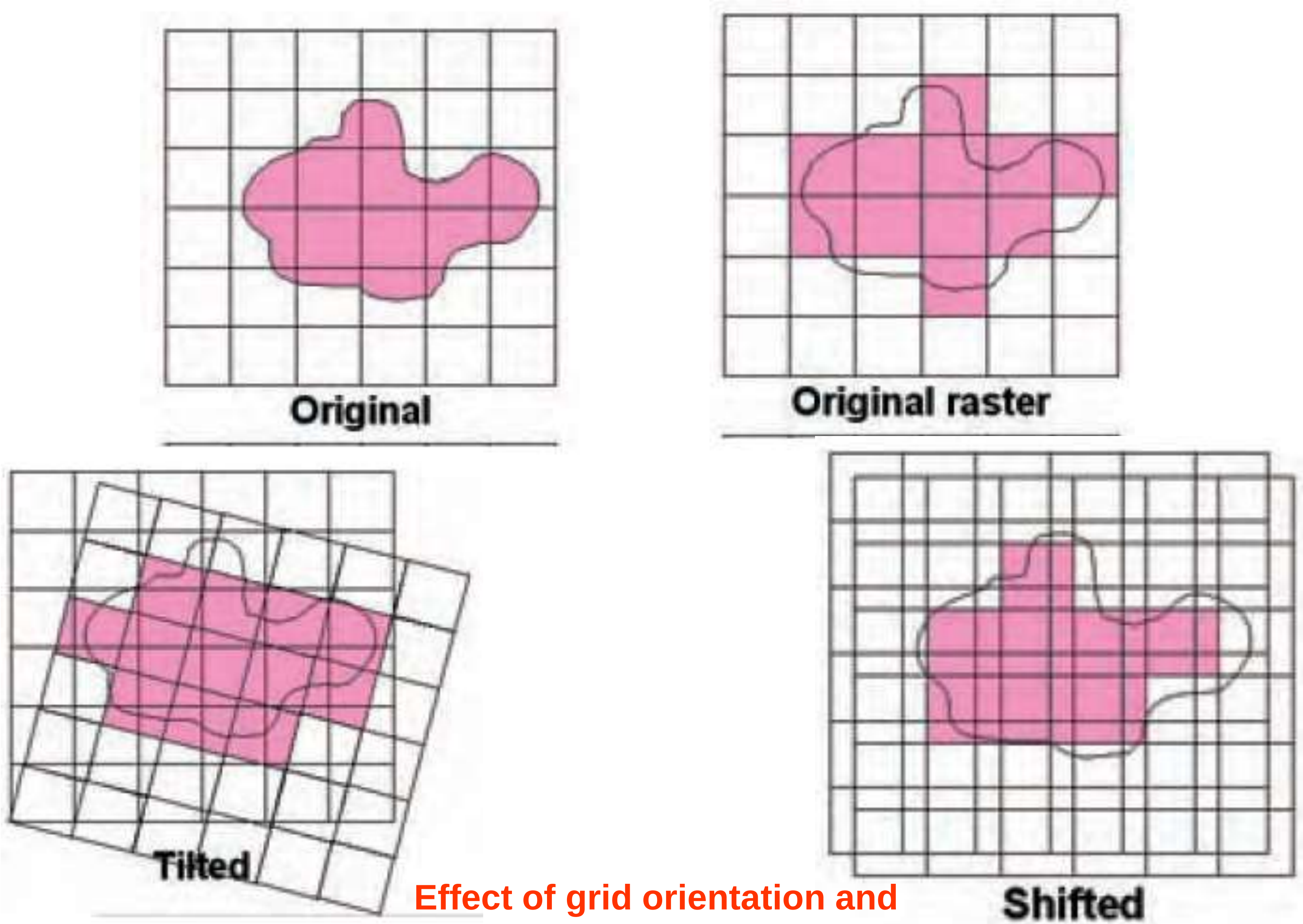
Loss of Information



(b)

What happened to 'C'?

Effect of grid orientation



Effect of grid orientation and origin on rasterization

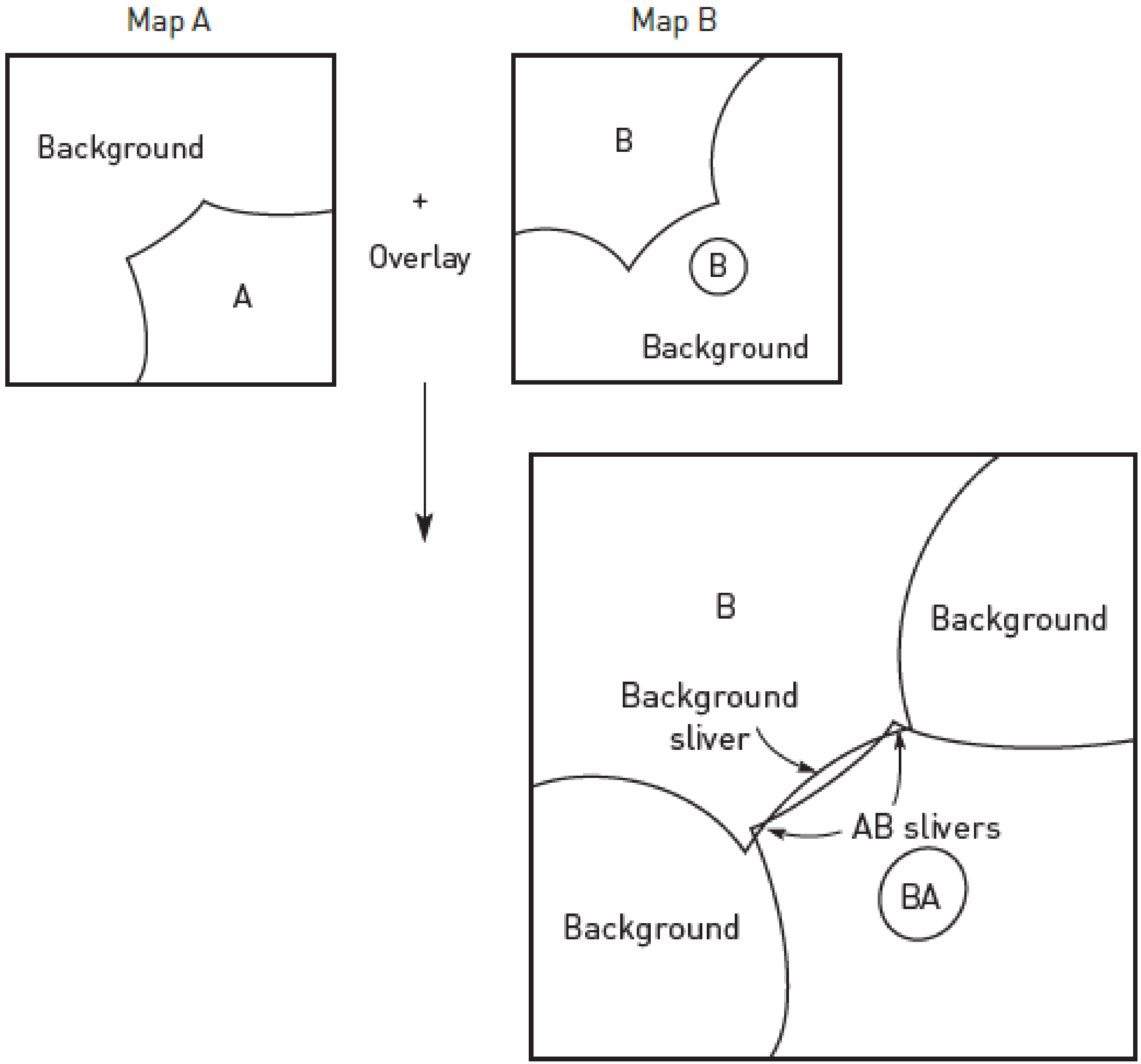
Errors in data processing and analysis

- Errors may be introduced during the manipulation and analysis of the GIS database.
- For example: Are the data suitable for this analysis? Are they in a suitable format? Are the data sets compatible?
- Are the data relevant? Will the output mean anything? Is the proposed technique appropriate to the desired output?

Errors in data processing and analysis

- Classification errors also affect raster data.
- Classified satellite images provide a reflectance value
- Raster maps of environmental variables, present the reflectance values.

Generation of sliver polygons



Composite (A + B) with sliver polygons