Geographic Information Technology Training Alliance

Lesson :

Primary sources

Unit 3:

Field data acquisition

B-DC Lesson 2 / Unit 3

Claude Collet

Department of Geosciences - Geography

UNIVERSITAS FRIBURGENSIS

Content of Lesson

Unit 1: Digital sources

Unit 2: Remote sensing

Unit 3: **Field data acquisition**

Unit 4: Field survey techniques

Unit 2: Field data acquisition

1: Introduction

2: Planning

3: Data collection

1 Introduction

Objectives of the Unit

- **To understand the process of field data acquisition and its major steps**
- **To relate this process with:**
  - the previous one on Spatial modelling (module B-SM)
  - the following one on regionalisation and GDB setup (module B-AN)
- **To identify and illustrate concepts and techniques in volved in this process**

What is field data acquisition process ?

- **It is the action to measure (to observe) properties of a phenomenon from the reality (terrain)**
  - Properties are measured throughout space in order to describe the spatial distribution of phenomenon properties
  - The timespan covered is from the present to a future
- **Field measurement process implies:**
  - a sampling process for selecting field sites
  - A clearly defined planning of the acquisition process

Not isolate but a linked process

As data are sampled from the reality, one should keep in mind related processes that produce relevant and representative information:

- **Spatial modelling of the reality (B-SM):**
  - Information collected and structured into the GDB must satisfy the objectives of the reality model
- **Regionalisation of variables (B-AN):**
  - In order to describe the whole surface of the study area, sampled data should be regionalised. This process requires specific properties from the sample to be satisfy, during the data acquisition phase

From the field to the geographical database (GDB)

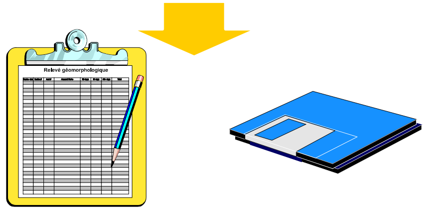
1. Field data acquisition

Reality

Location sampling
Selection of sampling sites

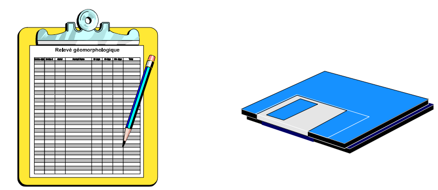
Data collection
Measurement + Storage

Spatial Data



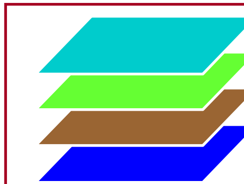
2. Regionalisation

Spatial Data



Regionalisation process
From point data to spatial unit properties

GDB
geographical database



The environment of field data acquisition

• **Information dimensions:**	***Thematic, spatial, temporal***
• **Spatial distribution of properties:**	***Continuous, discontinuous (discrete)***
• **Acquisition mode:**	***Sampling***
• **Instrumentation:**	***Man driven, automatic***
• **Acquisition system:**	***Fixed network, mobile system***
• **Information content:**	***Measured, interpreted***
	Nominal, ordinal, cardinal level

2 Planning field data acquisition

From phenomenon behavior to location properties

How to measure phenomenon behavior according to the defined model of reality?

- **In the thematic dimension:**
  - What characteristic(s) of the phenomenon should be retained in the data?
- **In the spatial dimension:**
  - Where to locate measurement sites in order to express the spatial distribution of phenomenon properties?
- **In the temporal dimension:**
  - When measurements should be acquired?

The keyword is Representativity

The collected data sample should represent the real spatial distribution of properties

- **In the thematic dimension:**
  - Measured properties should contain the whole range of diversity present in the study area during the considered period of time
- **In the spatial dimension:**
  - Sites of measurement should be located throughout the study area, even outside
- **In the temporal dimension:**
  - Repeatition of measurements should cover the whole considered period of time

Thematic representativity (1)

Data expressing phenomenon behaviour

- **Objective:**
  - to express relevant characteristics of a phenomenon through measured variable(s)
- **Example:**
  - Major characteristics of a broad phenomenon such as climate can be expressed by a set of following variables:
    - wind parameters
    - air temperature, humidity and pressure
    - …
    - precipitation

Thematic representativity (2)

Conditions for representativity

- **Selection of relevant instruments:**
  - What potential instruments can provide the most reliable information about selected phenomenon characteristics?
- **Covering the whole range of possible properties:**
  - Regionalisation methods assume all possible properties to be present in the data sample:
    - For discrete distribution, unsampled categories will be ignored
    - For continuous distribution, properties out of the measured range cannot be infered so simply

Spatial representativity

From point sample data to the surface

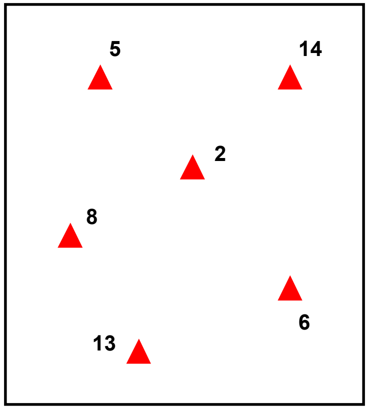
- **Assumption:**
  - Most field data are collected from point locations (sites of measurement, stations, …)
- **Infering properties outside sites of measurement:**
  - Regionalisation is in charge of spatial inference:
    - For predefined objects, their property is the “overall” property measured at their sampled locations
    - For resulting objects, property of any location in the study area is estimated from an interpolation process

Spatial representativity: for predefined objects (1)

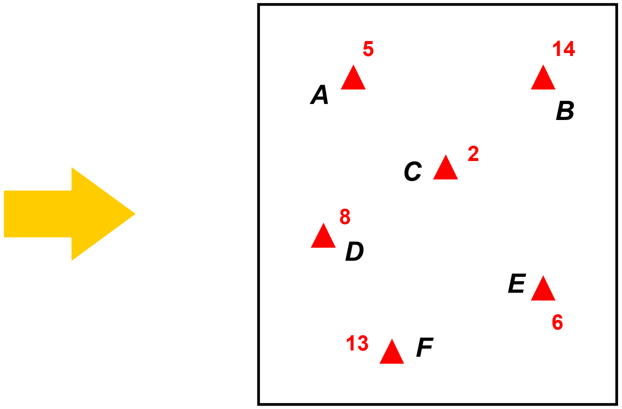
From point sample to point features

- **Usualy properties of point features are acquired with a sample of points at the same location**
- **There is no regionalisation process as the measurement property matches the location of the feature to be characterized**

n : Measured property



Location of acquisition points



Properties assigned to point features

Id : Feature identification

n : Assigned property

Spatial representativity: for predefined objects (2)

From point sample to linear features

- Several point locations should be selected along each linear feature in order to estimate its "overall" property
- The regionalisation process combines point measurements to estimate the property of each linear feature

Location of acquisition points

Properties assigned to linear features

Spatial representativity: for predefined objects (3)

From point sample to areal features

- Several point locations should be selected within each areal feature in order to estimate its "overall" property
- The regionalisation process combines point measurements to estimate the global property of each areal feature

n : Measured property

Id : Feature identification

? : Assigned property after regionalisation

Location of acquisition points

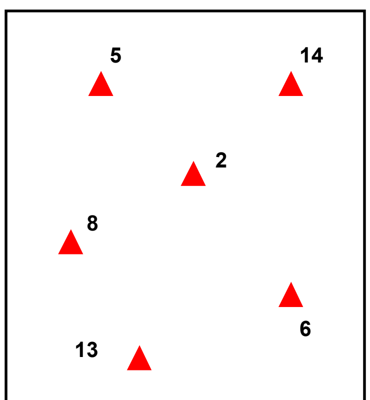
Properties assigned to areal features

Spatial representativity: for resulting objects (1)

From point sample to areal features
for discrete spatial distribution

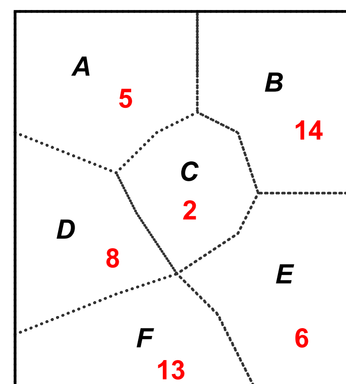
- **Numerous point locations should be selected within the study area in order to infer the distribution of properties in the study area**
- **The regionalisation process generates areal features with homogeneous property (space partitioning or zoning)**

n : Measured property



Location of acquisition points

Zoning by inference



Properties assigned to areal features

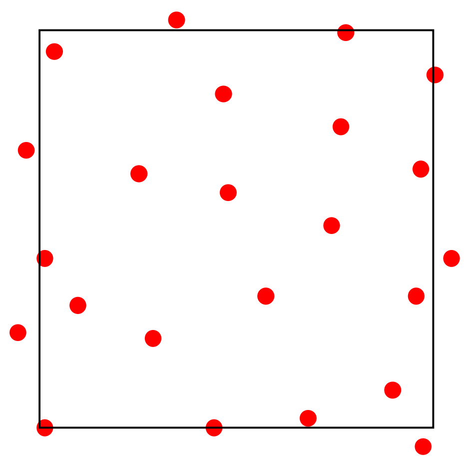
Id : Feature identification of the resulting object

n : Assigned property

Spatial representativity: for resulting objects (2)

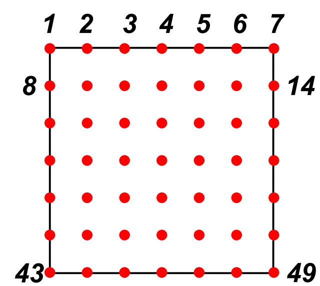
From point sample to continuous surface for continuous spatial distribution

- **Numerous point locations should be selected inside and outside the study area in order to infer the distribution of properties in the study area**
- **The regionalisation process generates point or areal features based on spatial dependancy of properties (spatial interpolation)**



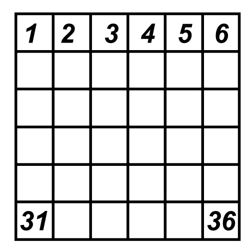
Location of acquisition points

Interpolation by inference



Regularily distributed point features

Id : Feature identification of the resulting object



Regularily distributed areal features (cells)

Id : Feature identification of the resulting object

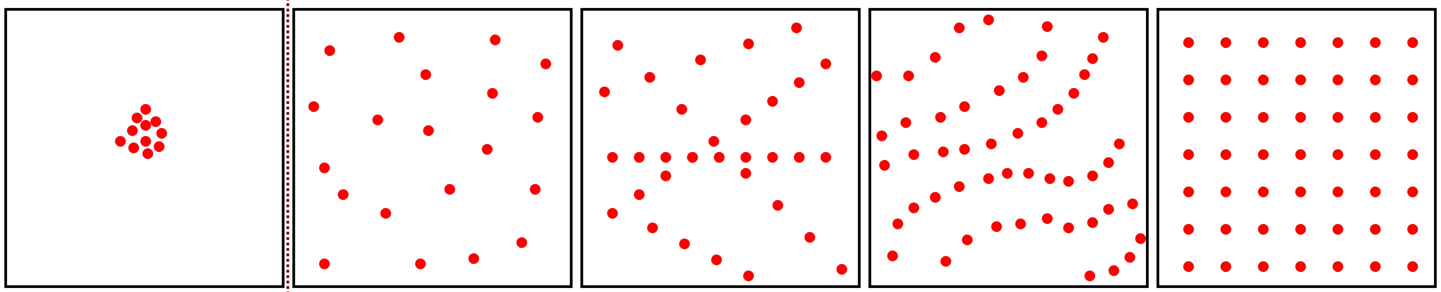
Properties assigned to point or areal features

How to assess spatial representativity ?

- **For predefined spatial features:**
  - sufficient point samples should be distributed inside the area or along the linear feature in order to estimate its « overall » (global) property such as the central tendency, often accompanied with the dispersion value
- **For resulting spatial features:**
  - sufficient point samples should be distributed inside and outside the study area in order to estimate a property for any location, through an interpolation process

Spatial distributions of points

R=0 | R=1 | R=2.1419



clustered | **random** | **systematic** *(profiles)* | **systematic** *(contour lines)* | **systematic** *regular*

Representativity of the spatial distribution of the phenomenon ?

No | ***Yes***

How to describe the spatial distribution ?

The nearest neighbour index (R index) compares an observed point distribution with a theoretical random distribution. It is a ratio

- **R index :**

$$R = \frac{\text{Mean distance between pairs of points in the sample}}{\text{Mean distance between pairs for a random distribution}} = \frac{\bar{d}}{\bar{d}_{al}}$$

- **Mean distance for the observed distribution (sample):**

$$\bar{d} = \frac{1}{n}\sum_i d_i$$

d_i = distance from point i to its nearest neighbour

- **Mean distance for a random distribution :**

$$\bar{d}_{al} = 0.5\sqrt{n/s}$$

s : surface de la zone d'étude
n : nombre de points de mesure

Comments on the use of R index

- **Interpretation of the R index value :**
  - see previous illustration for the different distributions and their related R value

  R = 1 describes a random distribution, $0 \geq R \leq 2.1419$

  - R value can be statistically tested against a random or regular distribution
- **How to modify R value from the observed distribution:**
  - adding few well located points inside the study area can change a clustered non representative distribution into a random and therefore representative distribution
  - reducing the study area surface can achieve similar effect

Spatial sampling techniques

How to produce a spatially representative sample ?

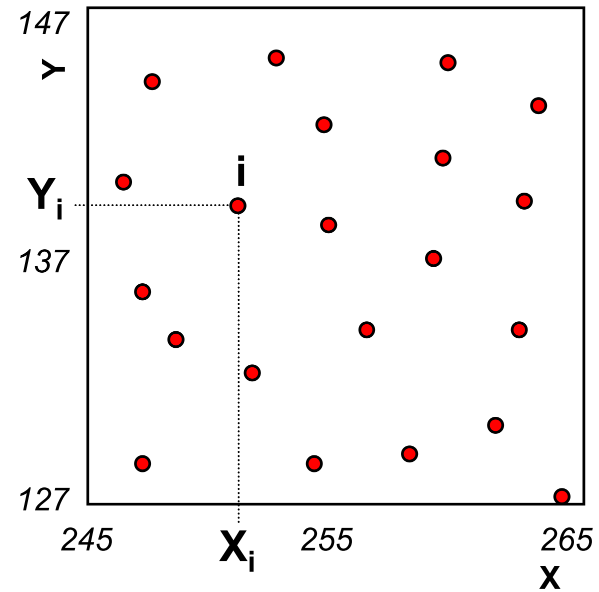
- **The type of spatial distribution to produce is influenced by:**
  - the instrumentation used for field data acquisition
  - the phenomenon characteristics to be investigated (variable)
  - the terrain accessibility and the resources available
- **Several sampling techniques are offered:**
  - Simple sampling: random or systematic
  - Stratified sampling: random or systematic

Simple random sampling

Technique

- Define the number of sites **n** to sample
- Define x_i, y_i coordinates for each site using a random number table or a random generator

Example of a simple random sample



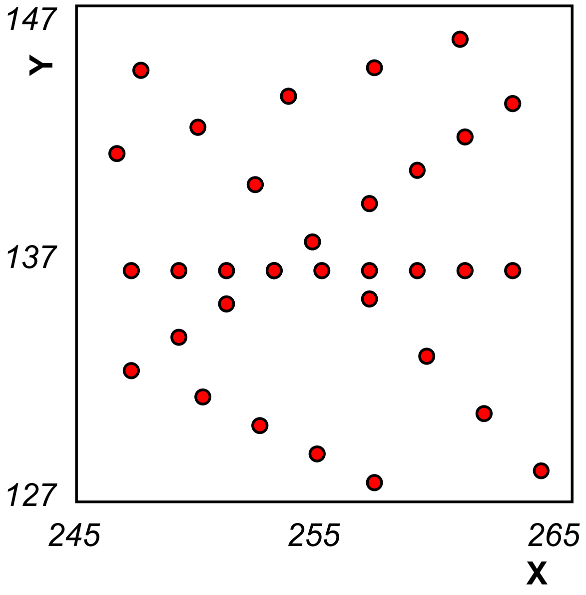
Simple systematic sampling (1): transverse

Technique

Some instrumentation requires to measure along profiles (transverses), such as geo-electric measurements

- Locate profiles in different directions, covering the whole study area
- Generally measurement sites are regularly distributed along profiles

Example of a simple systematic sample with transverses (profiles)

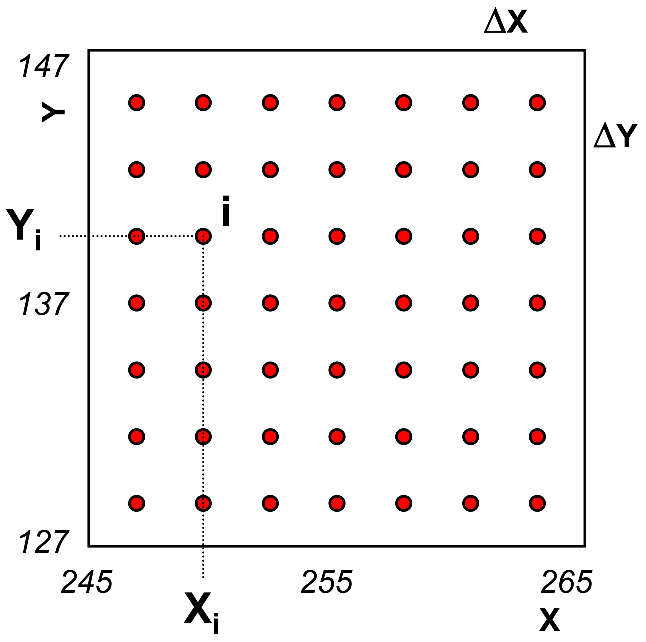


Simple systematic sampling (2): regular

Technique

- Define the the **mesh intervals** ΔX and ΔY (they can be equal)
- Locate the **first site i** randomly (it becomes the mesh origin)
- Other sites are then **arranged regularly** according to the mesh definition

Example of a simple systematic sample with regular mesh



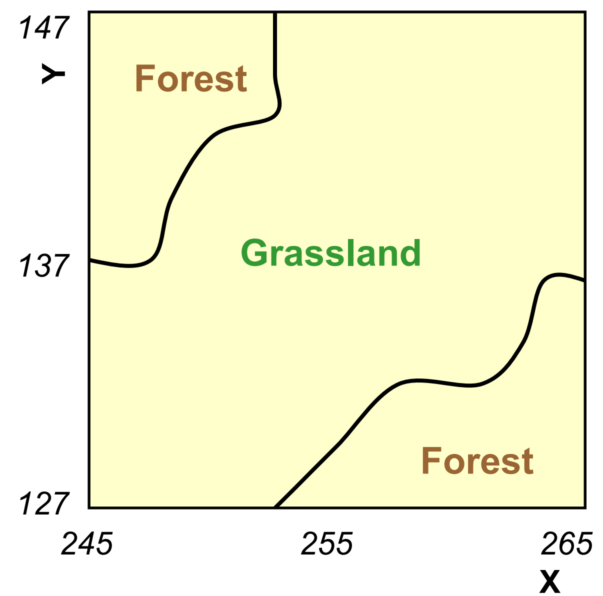
Stratified sampling (step 1): definition of strata

Step 1: definition of strata producing a set of regions

Technique

- **The study area is segmented into a set of regions in order to locate a defined number of sites with respect to each stratum or thematic category**
- **Examples:**
- **to draw an equal number of sites for the categories forest and grassland**
- **to draw the same number of sites for the 3 objects (regions)**

Example of a stratification scheme based on landcover type

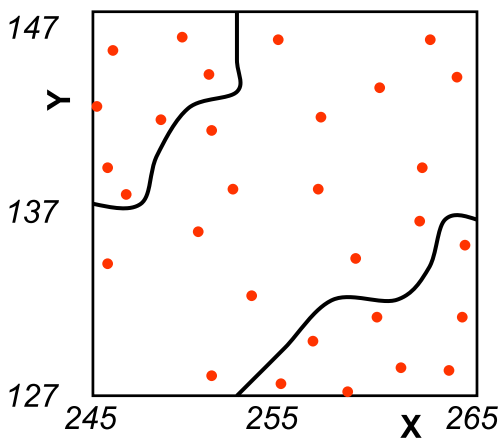


Stratified sampling (step 2): sampling within region

Step 2: selection of a sampling technique to apply in each region

Stratified random sampling

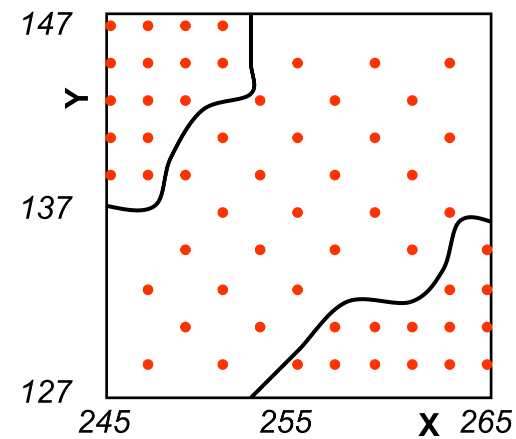
Equal representation



The 2 strata (forest et grassland) are equally represented with 15 observations randomly distributed: 7+8 and 15

Stratified systematic sampling

Equal representation



The 2 strata (forest et grassland) are equally represented with 29 observations regularly distributed: 15+14 and 29

Temporal representativity

When properties should be acquired ?

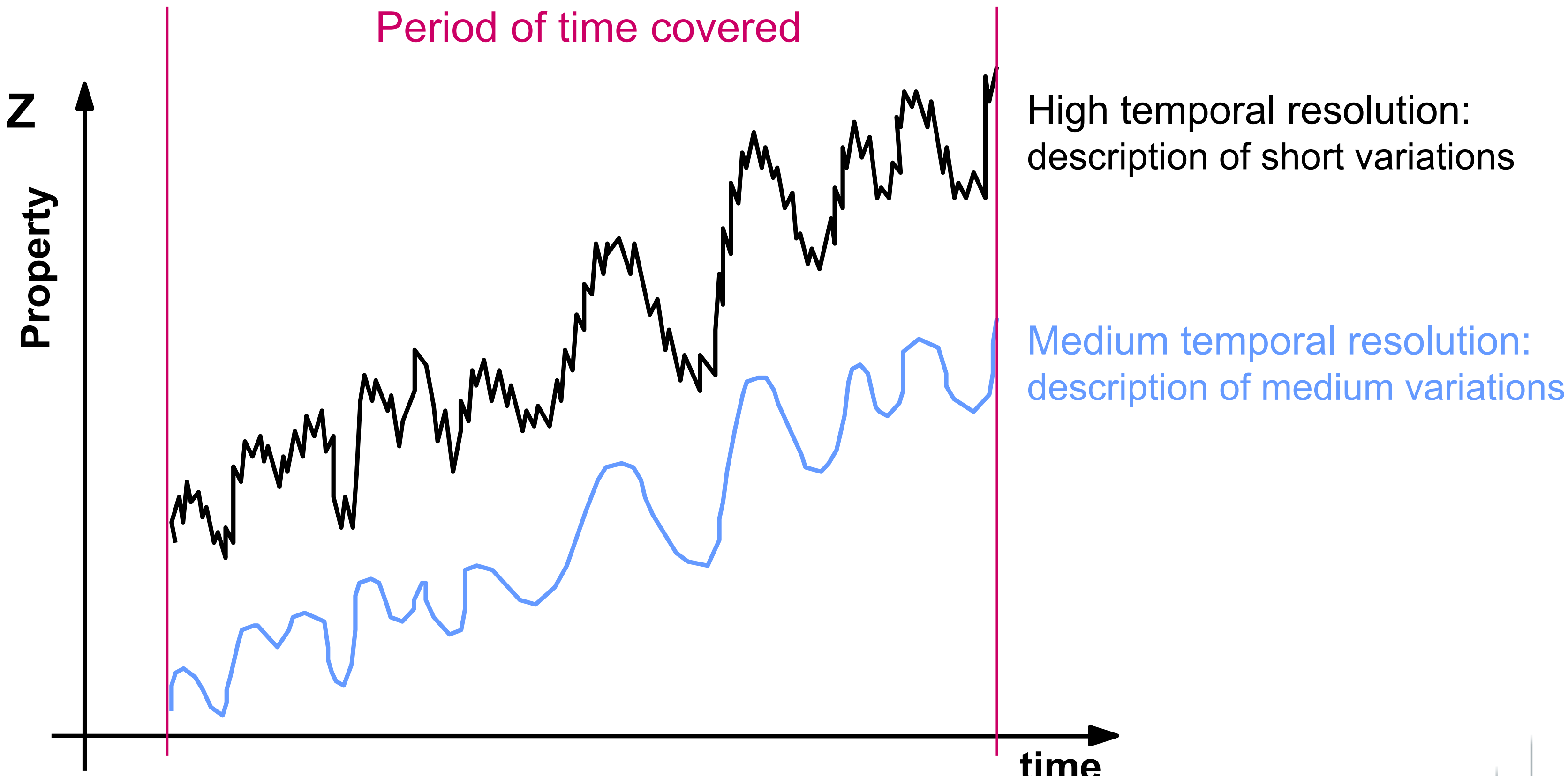
- **Selection of the time frame:**

  According to the model of reality requirements:

  - The period of time to be covered
  - The relevant temporal resolution corresponding to the characteristics of the phenomenon to be described (short, medium or long term variations)
  - The synchronicity of measurement throughout space (linked to the temporal variability of the phenomenon)

Not only properties change through time, but consequently spatial features too!

Properties change through time



3 Data collection

Acquisition systems

They are means to measure properties from the reality through space and time

- **They should provide information about:**
  - thematic dimension: property
  - spatial dimension: location
  - temporal dimension: time of measurement
- **Their components, made of different technologies, are:**
  - the sensor that captures properties
  - the storage device
- **They influence the acquisition process:**
  - sampling technique, production of data in digital form, ...

Components of an acquisition system

Reality (terrain)

Sensor module

Senses and translates measured properties as a:

Physical signal

Electrical signal

Chemical signal

Storage module

Stores measured signal under following form:

Analog:
- alpha-numerical text
- graphical
- electrical

Digital (media):
- magnetic
- optical
- solid state memory

The diversity of acquisition systems ranges from a human observer with field pad to an automatic station with wireless data communication

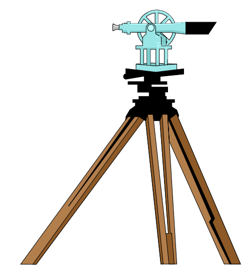
Sensor modules

Sensing devices

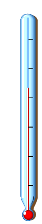
- **For the thematic dimension:**
  - observer, surveyor
  - thermometer, ph-meter
- **For the spatial dimension:**
  - theodolite, GPS
- **For the temporal dimension:**
  - clock, timer

Examples

Observer



Theodolite



Thermometer

Sonde

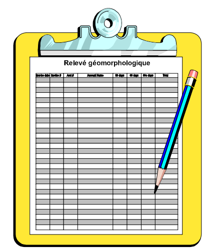
Storage modules

Storage devices

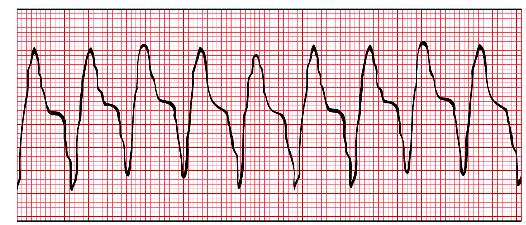
Different technologies to store data

- **Form:**
  - text: alpha-numerical
  - graphic
  - electric
  - digital
- **Medium:**
  - paper, film
  - magnetic, optic, solid state

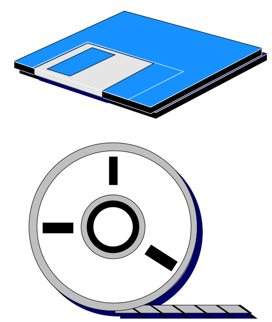
Examples



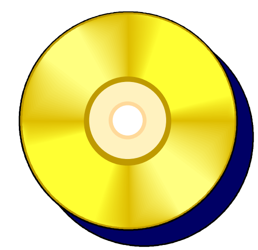
Field notes



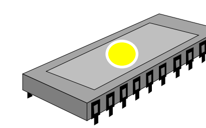
Analog graphic



Magnetic media



optic medium



Solid state

Field data in digital form

In order to be processed and structured into the GDB, field data should be set in digital form

- **From analog to digital form**

  When data are stored in an analog form, the conversion into a digital form requires one of already described digitizing process:

  - Manual entry with the keyboard
  - Semi-automatic digitizing with a digitizing table or on-screen digitizing
  - Automatic digitizing with a scanner or a camera