AESB2440: Geostatistics & Remote Sensing

Lecture 3: GNSS + measurement errors

Thursday, April 23, 2015

Roderik Lindenbergh

Dept. of Geoscience & Remote Sensing

Delft University of Technology

Lecture topics

GNSS

- Global Navigation Satellite System
- Global Positioning System
- Measurement errors

Error types

- Systematic errors
- Random errors

Univariate statistics

- Histogram
- Moments,
- Expectation
- Mean
- Variance
- Standard Deviation
- Quantiles

Interpolation

- Interpolation and extrapolation;
- deterministic vs. stochastic interpolation;
- Distance
- Metric
- Spatial Continuity
- Inverse Distance Interpolation

Dept. of Geoscience & Remote Sensing

A. GNSS

Source http://delta.tudelft.nl/article/students-watch-iceland-grow/27016

GNSS Positioning

GNSS: Global Navigation Satellite System

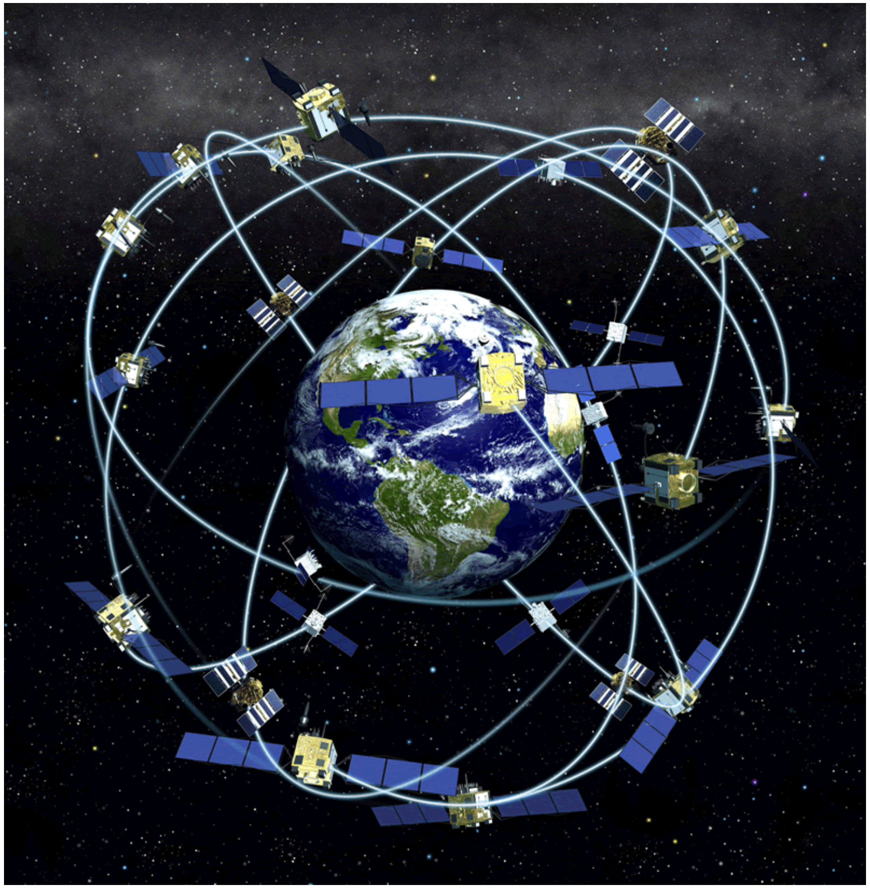
Different systems:

Fully operational

- US, NASA: GPS (32)
- Russia: Glonass (24)

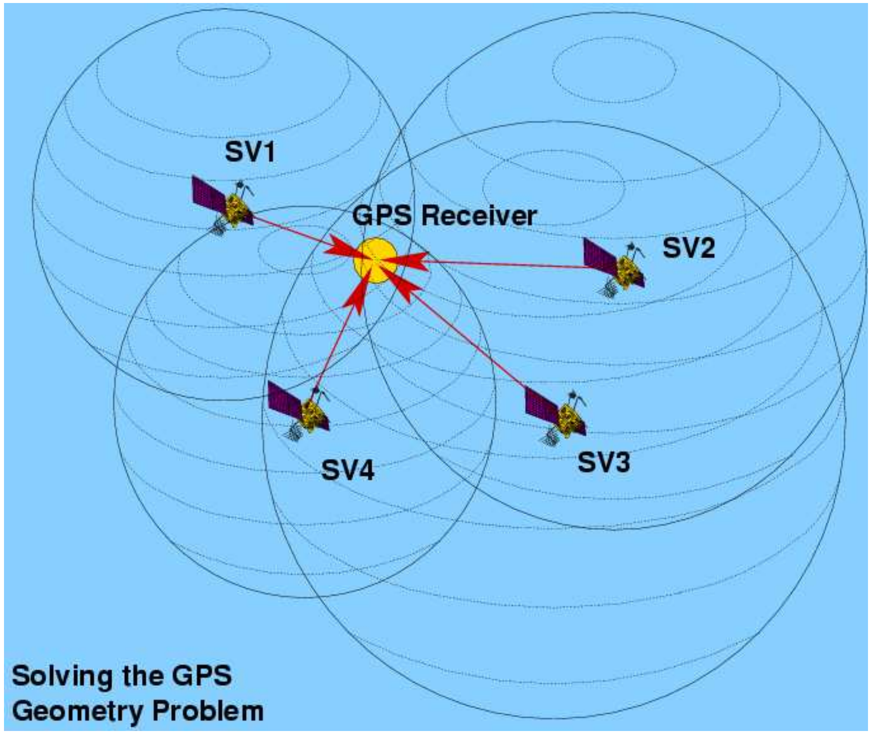
Under development

- China: Beidou (15)
- EU, ESA: Galileo (4)
- India: IRNSS (3)



GNSS Positioning

A) GNSS satellites transmits codes at constant intervals containing time and position at transmission



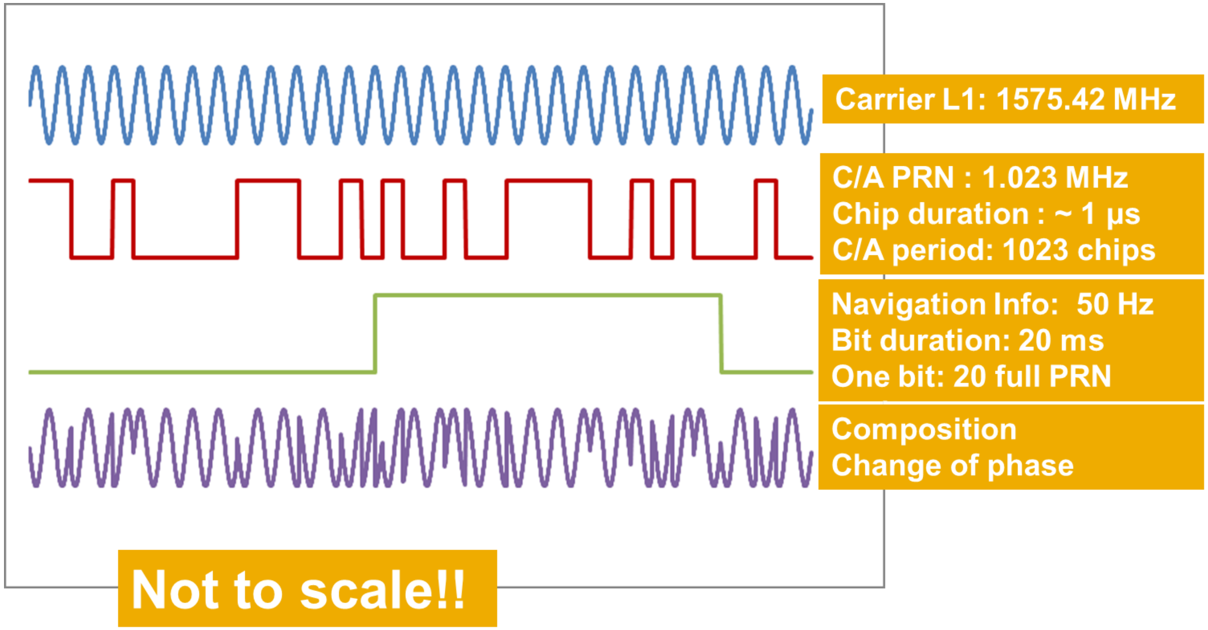
B) A GNSS receiver (e.g. in a car) measures the different arrival times from the satellites in sight.

The further the satellite, the longer the travel time

C) The time of flights are converted to distances.
In general four such distances are enough to pinpoint the receiver.

Question. Why?

GNSS signal



Carrier: Radio frequency sinusoidal signal at a given frequency.

Ranging code: Sequence of zeroes and ones, enabling the receiver to determine travel time of radio signal from satellite to receiver.
(Called: Pseudo-Random Noise (PRN) sequences).

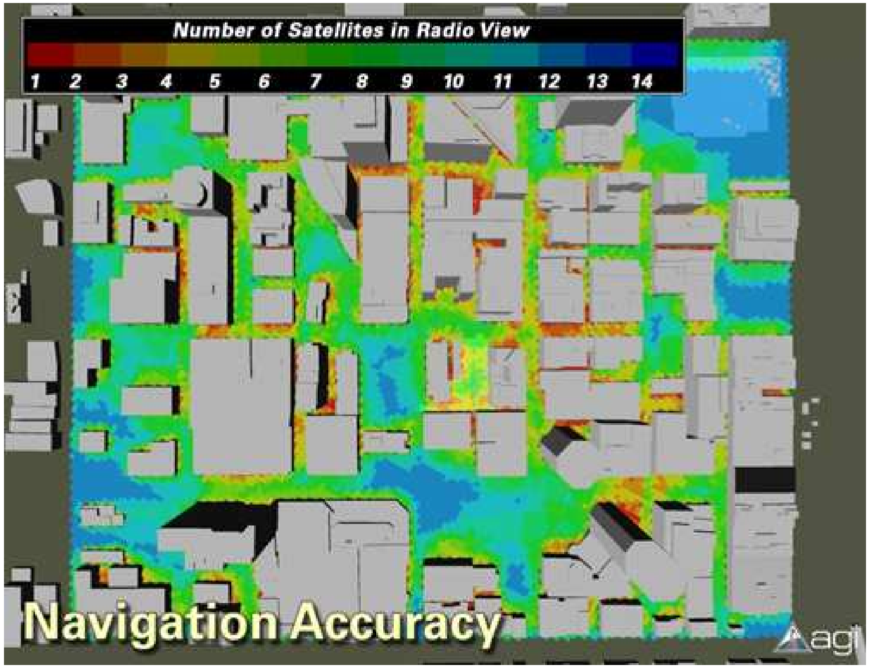
Navigation data: Message providing information on e.g. satellite orbit, clock bias parameters and satellite health status

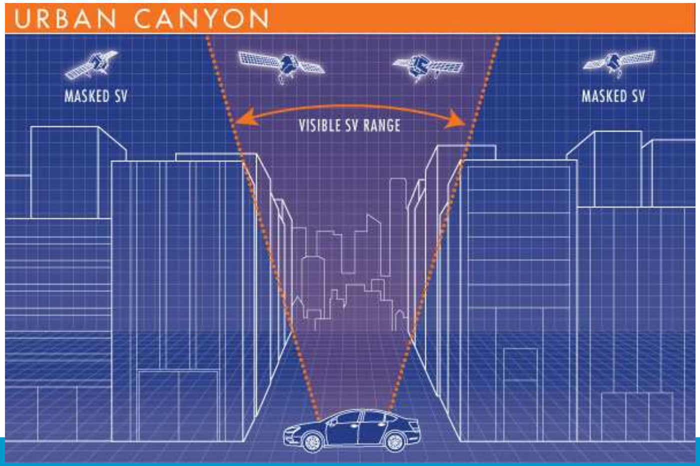
Source http://www.navipedia.net/index.php/GNSS_signal

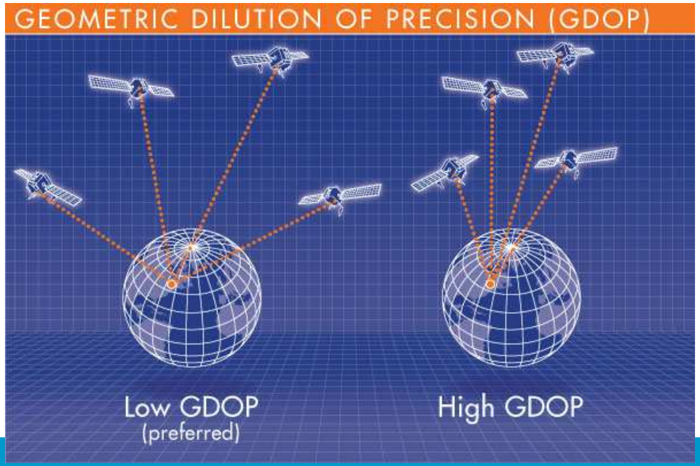
GNSS error sources

Error sources & Propagation

- Clock errors
- Orbital errors
- Atmospheric effects
- Geometric Dilution of Precision (GDOP) and Visibility
- Multipath (bouncing of the signal near the receiver)







Final Quality GNSS

Tom Tom

- Horizontal Precision: $\sigma_H < 5m$
- Vertical Precision: $\sigma_V \approx 5m$

Survey quality GPS

- Horizontal Precision: $\sigma_H \approx 1cm$
- Vertical Precision: $\sigma_V \approx 2cm$

Smartphone GPS

- Mekelpark experiment

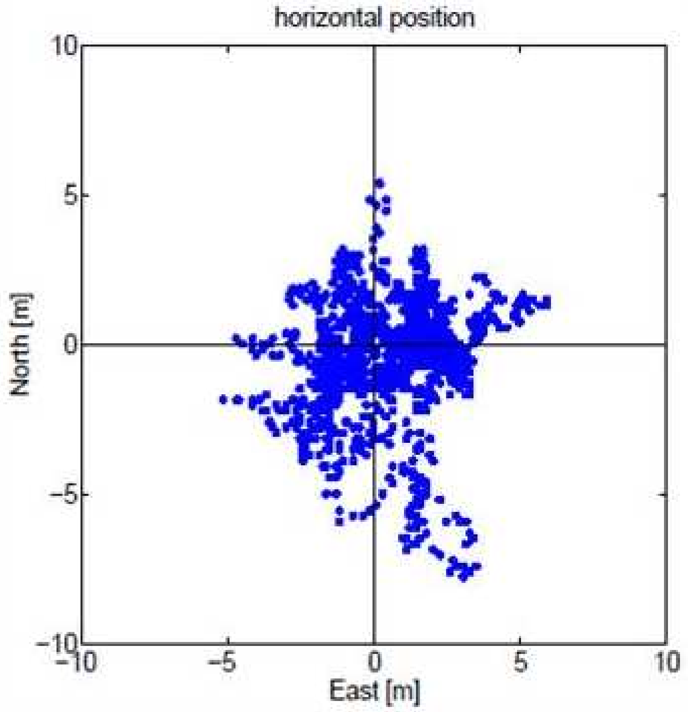
Question: what is precision?

Question: what is the difference between horizontal and vertical precision?

GNSS noise example

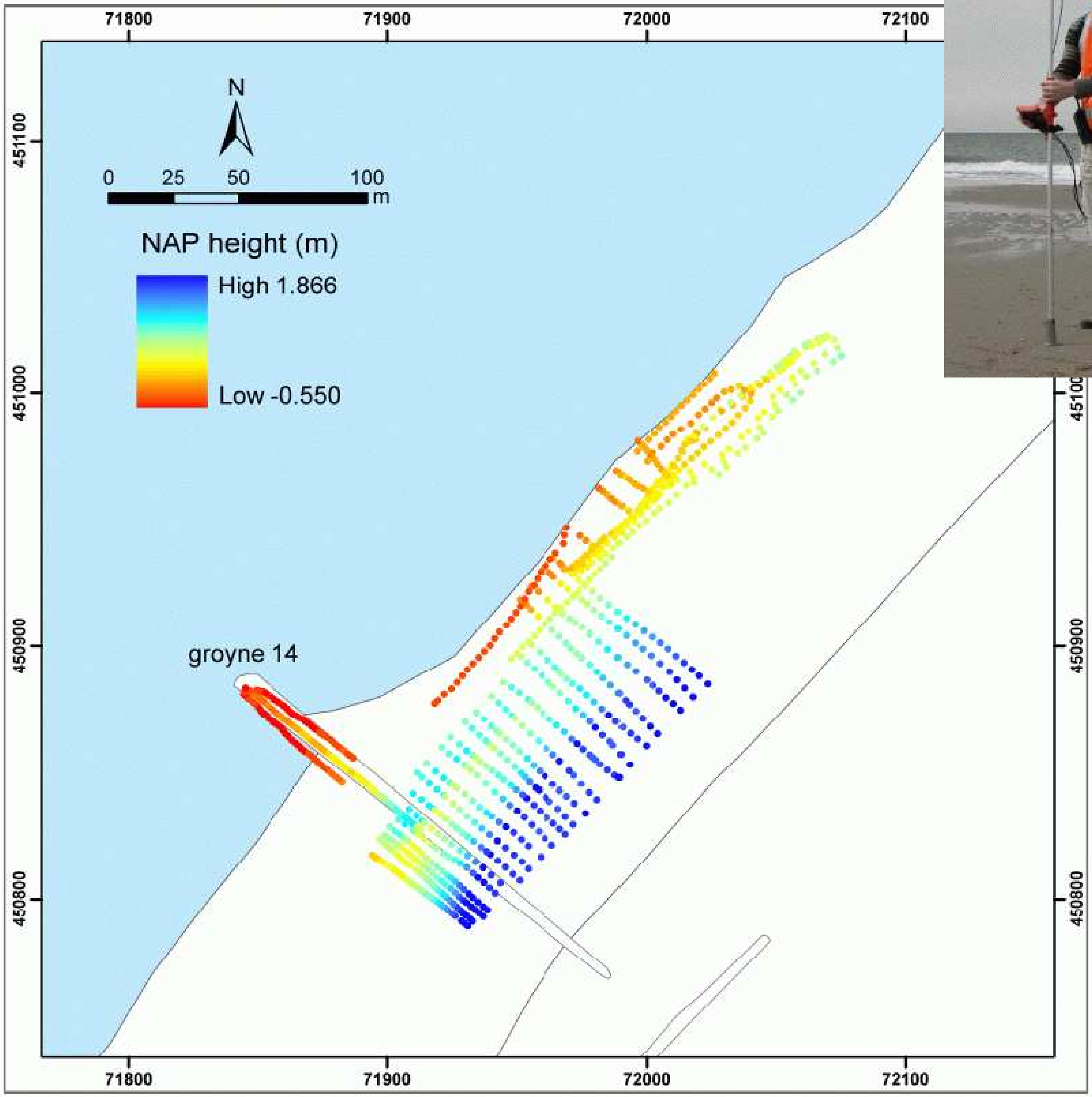
14.5 hours of GPS standalone positioning (at 30 seconds interval)

- with a Garmin GPS76 handheld receiver
- using an external antenna
- on a favorable location
- July 2002



Question. What do the measurements say on the quality of this device?

GNSS profiles



GNSS applications

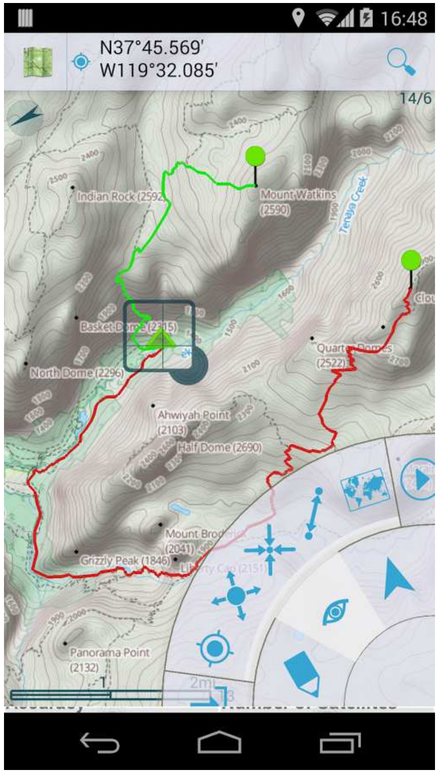
Ground Control Points:
Get coordinates for photos, field work, photogrammetry, terrestrial laser scanning, surveying

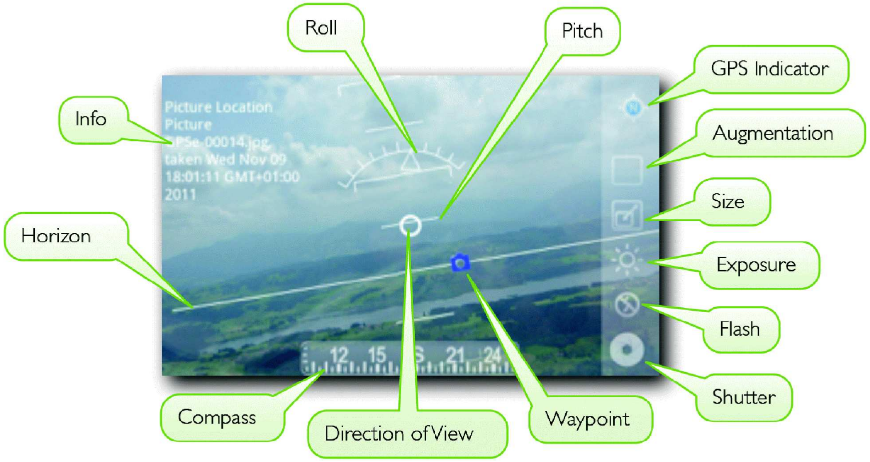
Navigation;
Positioning of cars, boats, airplanes and satellites

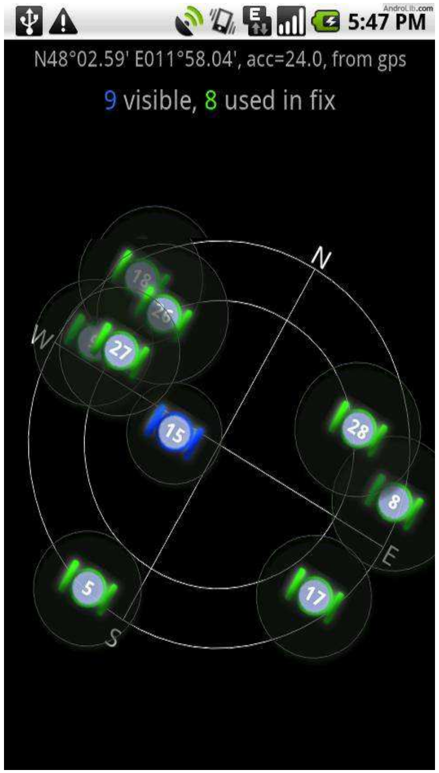
Mobile mapping
Combine (mobile) sensor acquisition with sensor location:

- Satellite remote sensing
- Airborne laser scanning
- Mobile mapping:
  e.g. using cameras or laser

GNSS app: GPS essentials







Home page: `http://www.gpsessentials.com/`

Manual: `http://www.mictale.com/gpsessentials/download/GPSEssentialsManual.pdf`

B. Error analysis

Source http://uk.smartnet-eu.com/news$_$116.htm?id=5047

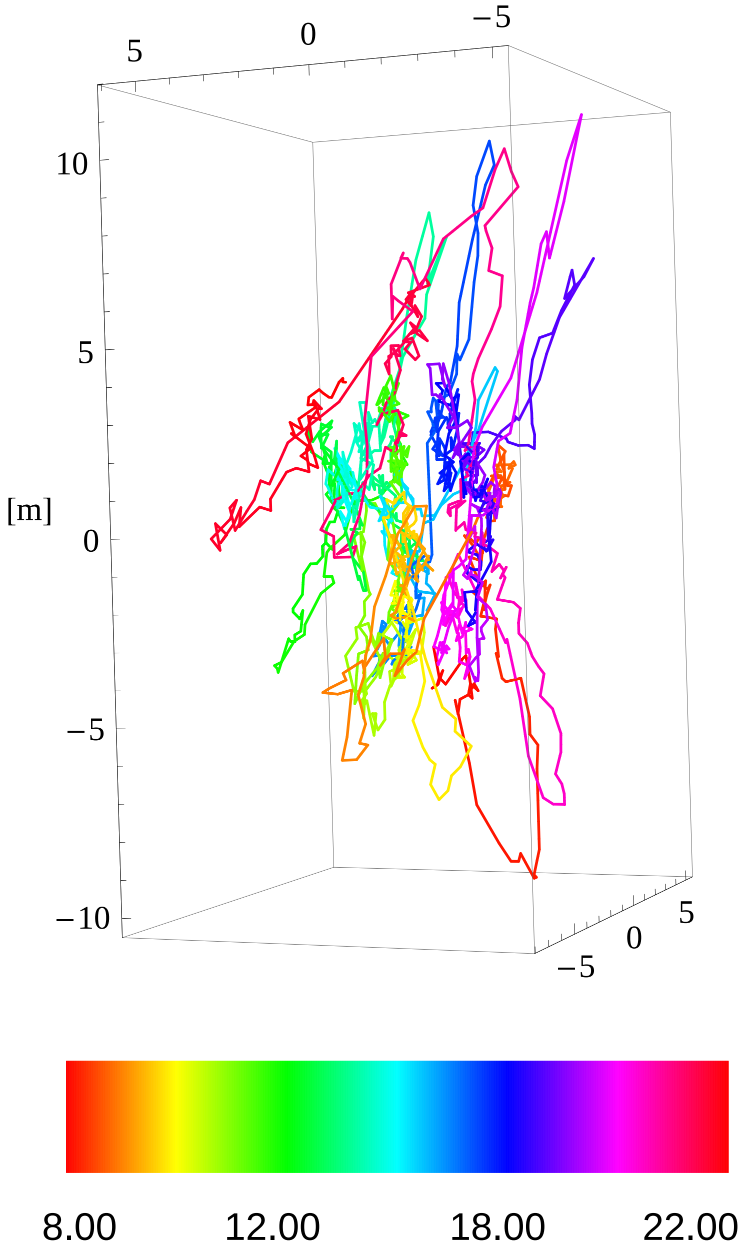
GPS data set

Recall: 14.5 hours of GPS standalone positioning (at 30 seconds interval)

Question: what is the best estimate of the position given the measurements?

Question: what type of errors are in play?

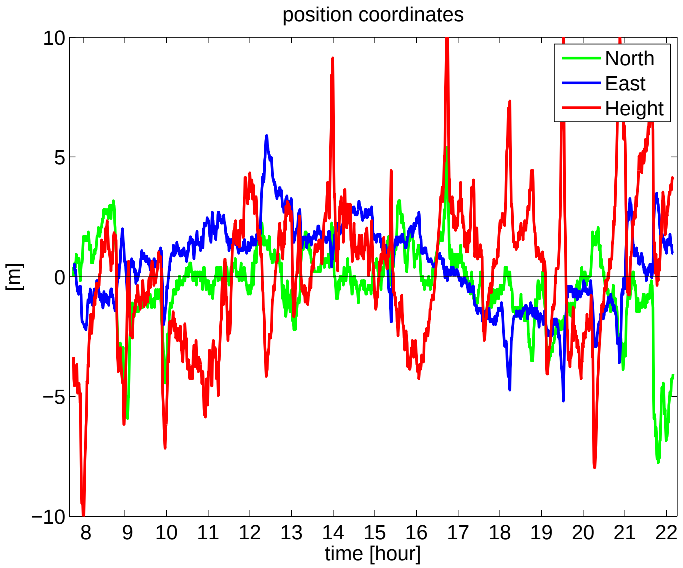
Question: what is the quality of this best estimation?



Solution over time

Estimated position per coordinate axis according to GPS device

$\Rightarrow$



Reference (true) position:

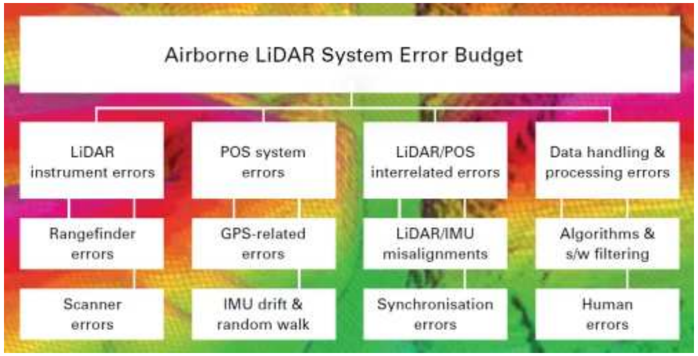
Latitude	51.98608984 [deg]
Longitude	4.38776741 [deg]
Height	74.2620 [m]

Error := true position - estimated position

Problem: in general the ground true (position) is not known

Systematic and random errors

Error budget: List of all errors, preferrably decomposed according to their relative contribution to the total error



Systematic errors: errors caused by systematic effects.

Example: a systematic offset or bias due to e.g. a wrong coordinate transformation.

Random errors: errors that are described by a random variable (and therefore follow some distribution)

Example: if you measure the length of the room 10 times with a laser ranger, you may find 10 slightly different outcomes

Source https://www.e-education.psu.edu/geog481/l3$_$p8.html

Accuracy and precision

Measurement errors equation:

$$\underline{y} = x + \theta + \underline{e}, \text{ with:}$$

$\underline{y}$ - observable (random variable)

x - unknown parameter of interest

θ - bias (some constant offset)

$\underline{e}$ - random measurement error (random variable)

Accuracy: Mean deviation of the measured values from the true value

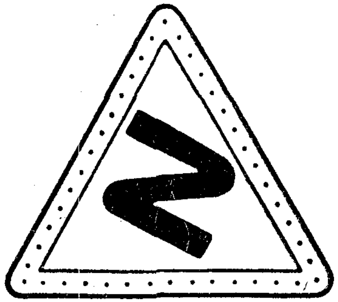
Precision: Spread of the measured values around the mean of the measured values

Measurement quality

High precision, low accuracy

High accuracy, low precision.

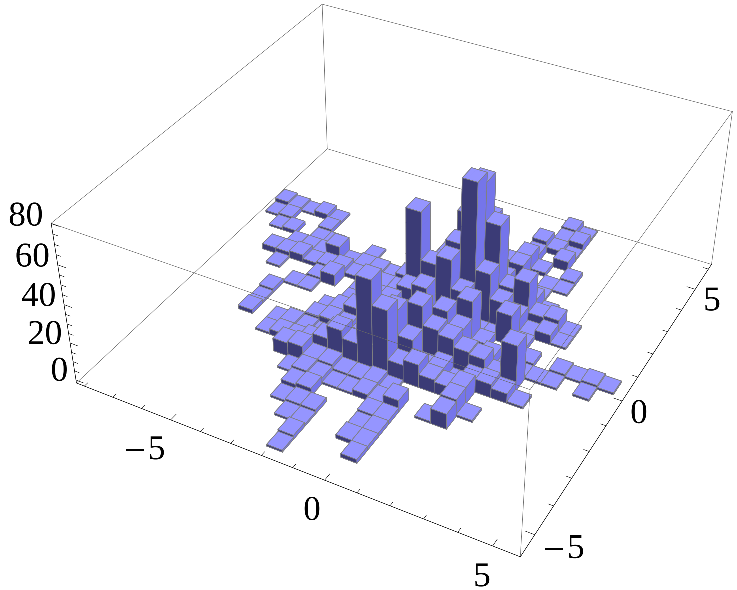
Remark. People often use these terms in a sloppy way. Always make clear what you and others mean by these terms when you discuss the quality of your results.



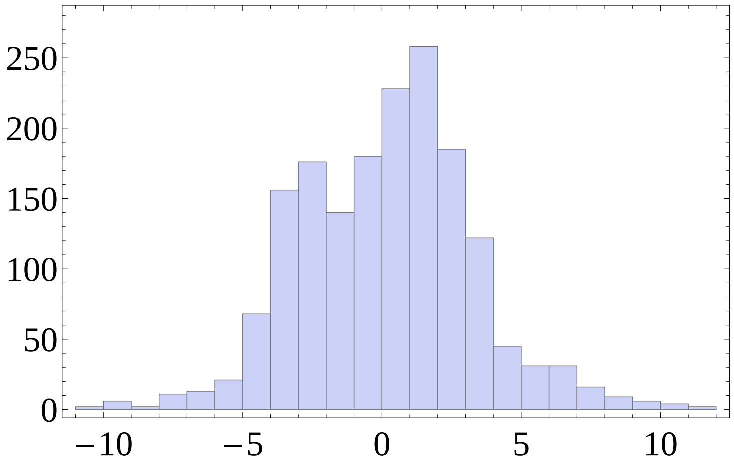
Source http://en.wikipedia.org/wiki/Accuracy$_$and$_$precision

Histograms

1. Divide the range of (many) measurements of the same random variable into bins of equal width
2. Plot the number of measurements per bin



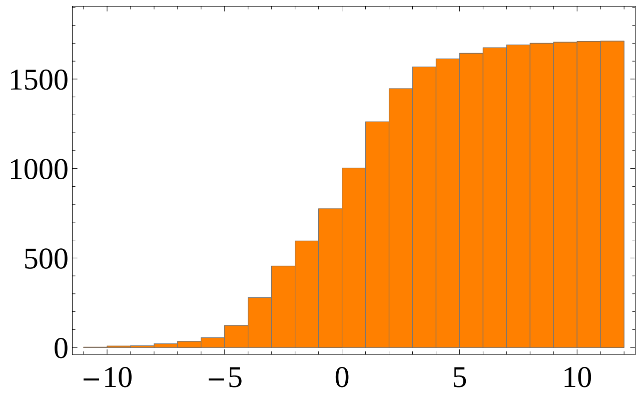
Histogram of horizontal (xy) GPS errors



Histogram of elevation (z) GPS errors

Remark. Rule of thumb: use $\sqrt{n}$ bins, given n measurements.

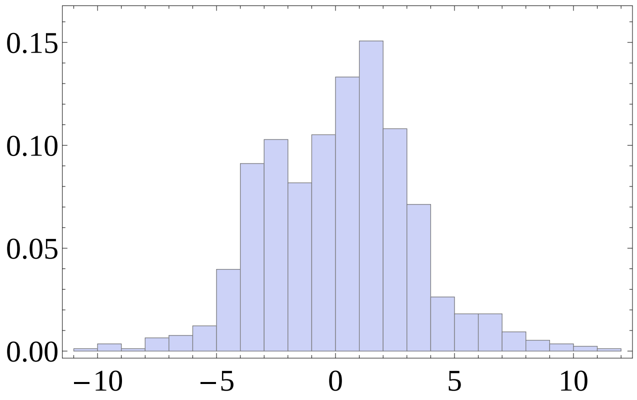
Back to distributions



Relation?

⟵

⟶



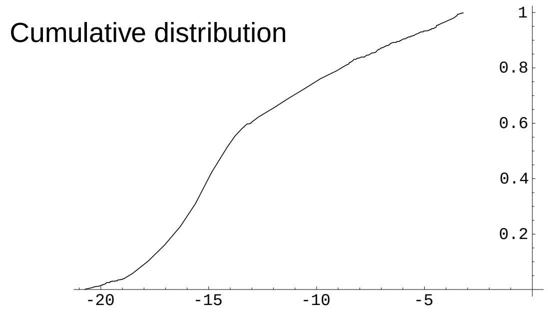
Question: how should the histogram be changed to serve as an emperical probability density function?

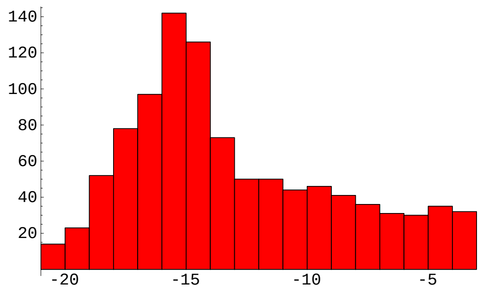
Question: What is the relation between ECDF & histogram?

Dept. of Geoscience & Remote Sensing

TU Delft

Histogram vs. Distribution





Every random variable $Z(\mathbf{p})$ is fully characterized by its non-decreasing cumulative distribution function $F : I\!R \rightarrow [0,1]$ such that

$$F(-\infty) = 0, \qquad F(\infty) = 1, \qquad F(z) = P(Z \leq z)$$

The probability density function $f(z)$ is a function with $f(z) \geq 0$ and $\int_{-\infty}^{\infty} f(z)dz = 1$ such that $f = F'$, that is:

$$F(z) = \int_{-\infty}^{x} f(\xi)d\xi$$

Mean and median

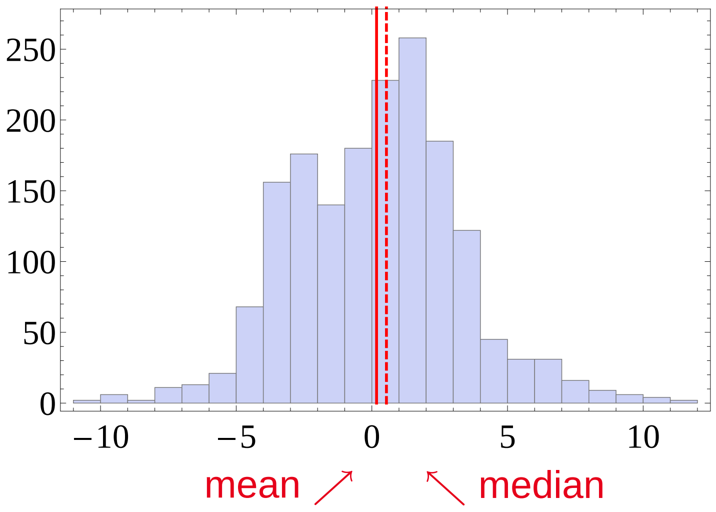
Running example: z errors of the GPS data set.

Given is a univariate dataset $S = \{x_1, x_2, \ldots, x_n\}$.

The sample mean $\overline{x}$ of S is defined as $\overline{x} := \frac{x_1 + x_2 + \ldots x_n}{n}$.
Example. $\overline{z}_{GPS} = 0.17$.

The sample median is the number in the middle if S is in ascending order.
Example. med$z_{GPS} = 0.54$.

Question:
when is the median really different from the mean?



Dept. of Geoscience & Remote Sensing

TU Delft

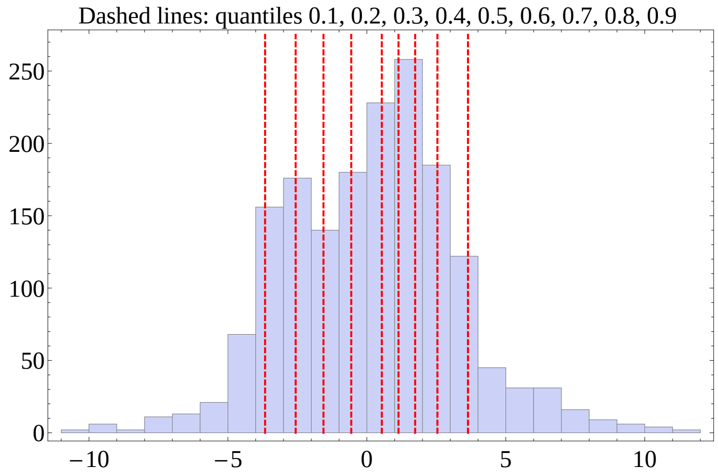
Quantiles

The p-th empirical quantile of S is a number $q(p)$ from S such that a proportion p of S is smaller then $q(p)$

Question: What is, in general, $q(0)$, $q(1)$ and $q(.5)$?

Example. $q_{GPS}(0) = -10.062$, $q_{GPS}(1) = 11.538$

Question: What is a robust alternative for the maximum and minimum of a large data set?



Dept. of Geoscience & Remote Sensing

Spread of a data set

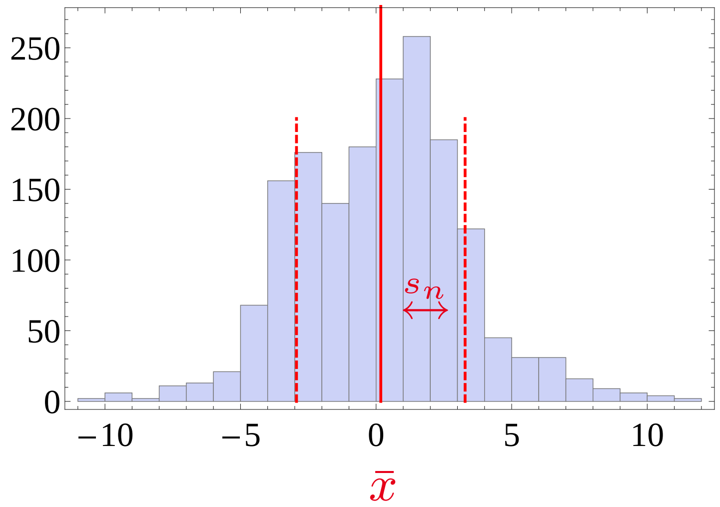
The sample variance of a data set is given by

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Technical remark. Division by $(n-1)$ ensures that s_n^2 is an unbiased estimator of the variance.

The sample standard deviation is defined as $s_n := \sqrt{s_n^2}$.

Question. Why is s_n always positive?



Dept. of Geoscience & Remote Sensing

Mean squared error

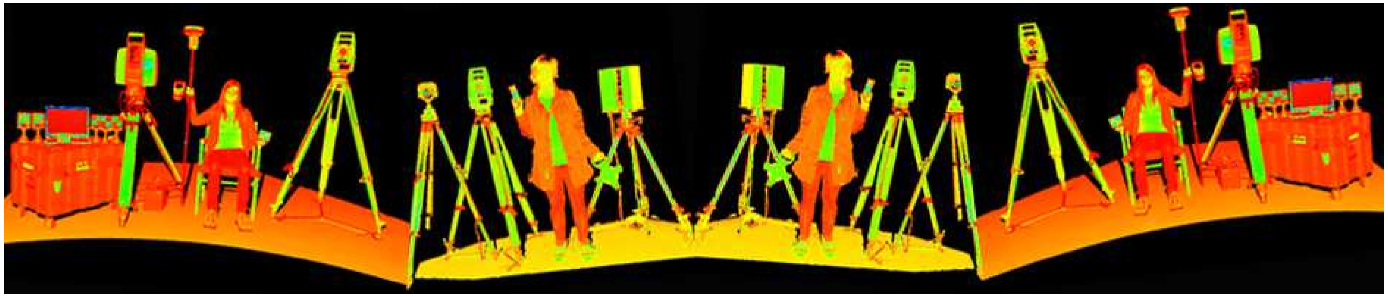
If the true value of an attribute is known, it is possible to assess the

Mean squared error:

$$\text{MSE} := s_n^2 + \theta^2, \quad \text{with,}$$

s^2	–	the sample variance
θ	–	the bias

Remark. The MSE is for example used in reporting on a calibration procedure



Source http://blog.hig.no/laserscanner/calibration-of-long-range-laser-scanner/

Dept. of Geoscience & Remote Sensing

Example, mean squared error

Before: mean of the z errors of the GPS data set: 0.17 [m]

1. Interpret this mean as a bias w.r.t. the real elevation, so $\theta = 0.17[m]$.
2. Centralize the z errors of the GPS data by substracting this bias
3. Determine the variance of the centralized data set: $sC_n^2 = 9.67$
4. So, MSE $:= sC_n^2 + \theta^2 \; = \; 9.67 + 0.03 \; = \; 9.70$

Question. Does the variance of a univariate dataset change if we subtract its mean?

Question. Is the bias in this case actually significant?

Expectation vs. mean

Let $f(z)$ denote a continuous probability density function.

Expected value or first moment versus Experimental mean

$$E\{Z\} = \int_{z \in \mathbb{R}} z f(z) dz = \mu \qquad \leftrightarrow \qquad \overline{z} = \frac{1}{n-1} \sum_{i=1}^{n} z_i$$

The mean minimizes the sum of square distances.

Question. Which distances?

Question. What is the discrete version of the expectation?

Variance and moments

The variance is the second moment about the mean (or 2nd central moment) and gives info on the spread around the mean.

Theoretical versus Experimental variance:

$$\begin{aligned} \text{var}(Z) &= E\{(Z - E\{Z\})^2\} = E\{(Z-\mu)^2\} \\ &= E\{Z^2\} - 2\mu E\{z\} + \mu^2 = E\{Z^2\} - \mu^2 \\ &= E\{Z^2\} - (E\{Z\})^2 = \sigma^2 \\ &\updownarrow \\ \sigma^2 &= \frac{1}{n}\sum_{i=1}^{n}(z_i - \overline{z})^2 \end{aligned}$$

Dept. of Geoscience & Remote Sensing

TU Delft

Higher central moments

The k-th central moment is defined as

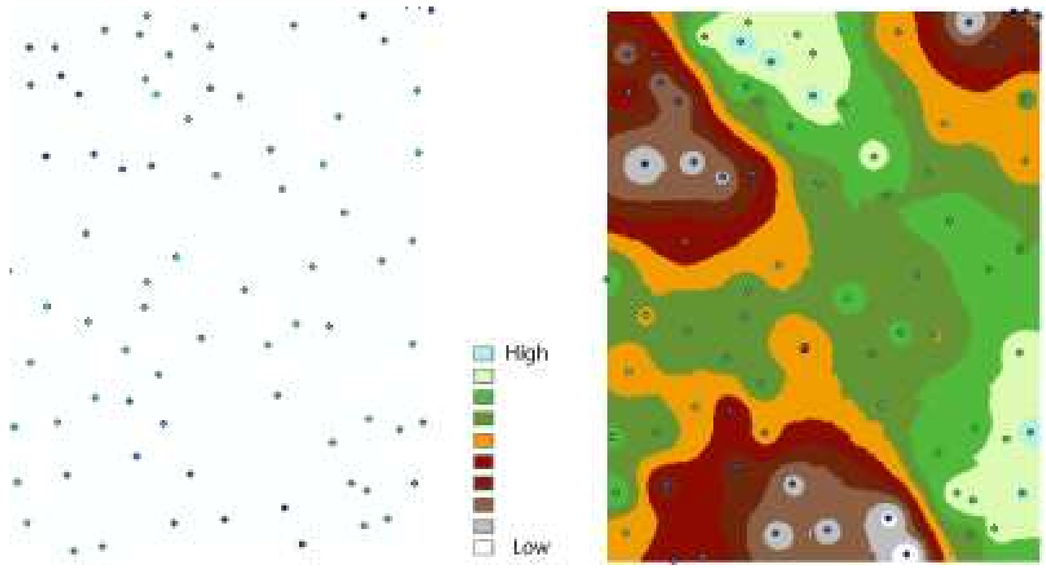
$$\int_{-\infty}^{\infty} (z-\mu)^k f(z) dz$$

Question: What tells the third central moment us, skewness?

Question: And what the fourth, kurtosis?

Dept. of Geoscience & Remote Sensing

C. Deterministic Interpolation



Source http://resources.esri.com/help/9.3/ArcGISDesktop/com/Gp_ToolRef/geoprocessing_with_3d_analyst/understanding_ras

The problem of Interpolation

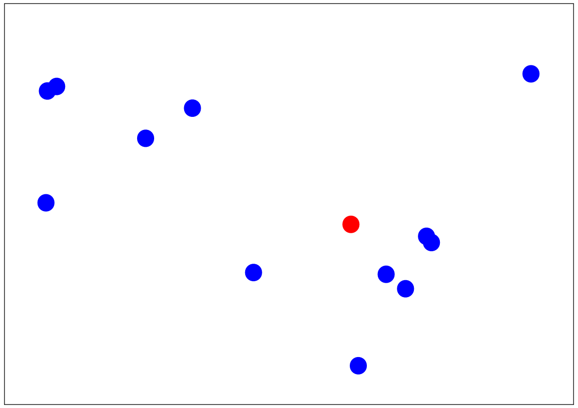
Given are height values

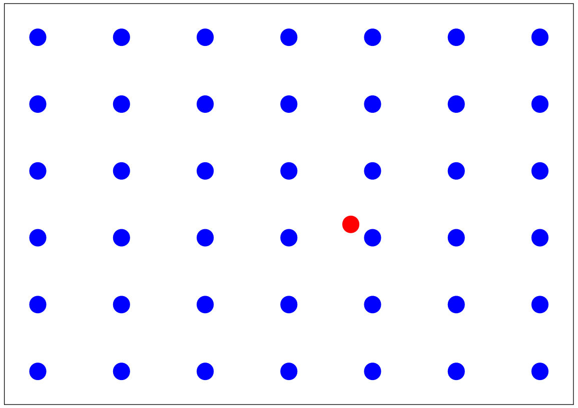
$$h_1, h_2, \ldots, h_n$$

at known (spatial) locations

$$p_1, p_2, \ldots, p_n.$$

in some domain D. What is the height h_0 at location p_0?





Dept. of Geoscience & Remote Sensing

TU Delft

Examples and problem specification

# Points	Technique
7	Leveling
668	GPS
2261	Single beam echo sounding
2 959 473	Laser altimetry data
25 million	Terrestrial laser scanning

- What part of the data should we use to estimate a height?
- What weight should we give to the different observations we use?
- How should we assess the quality of the observations?
- How should we assess the quality of the height prediction?

Related

- What dimension are we working in?

Interpolation 1D

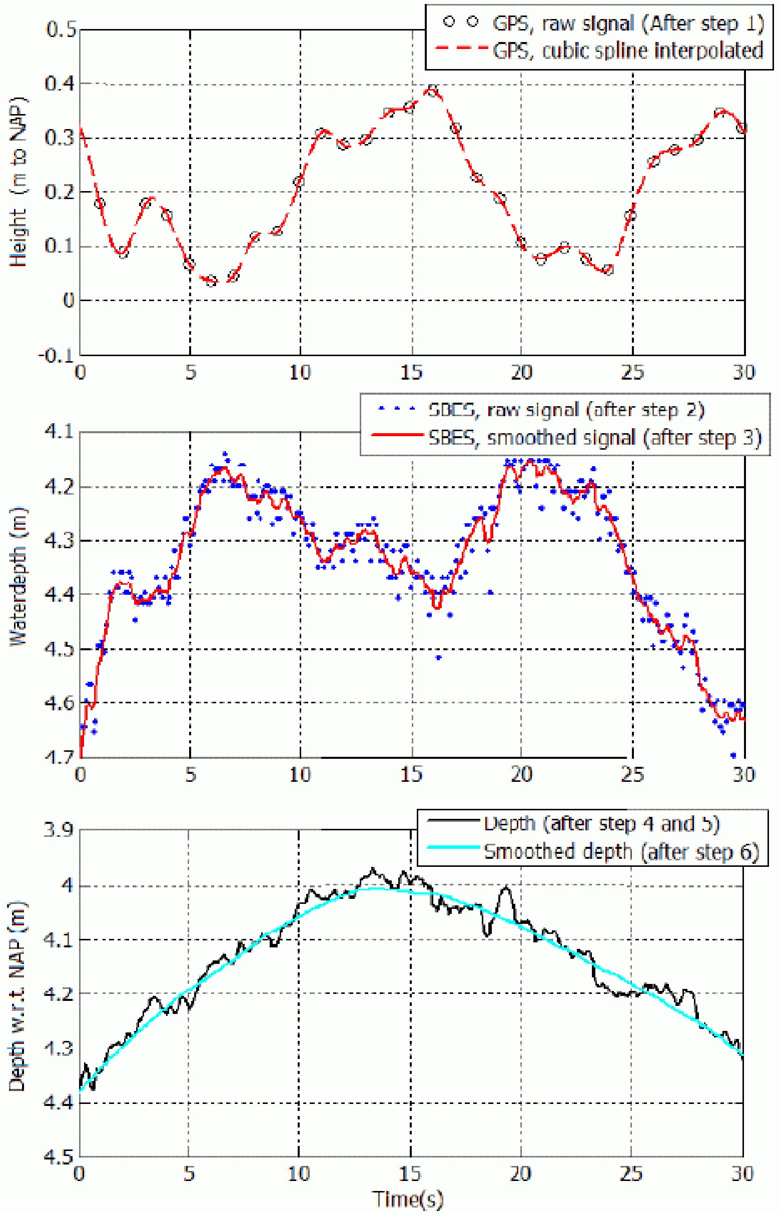
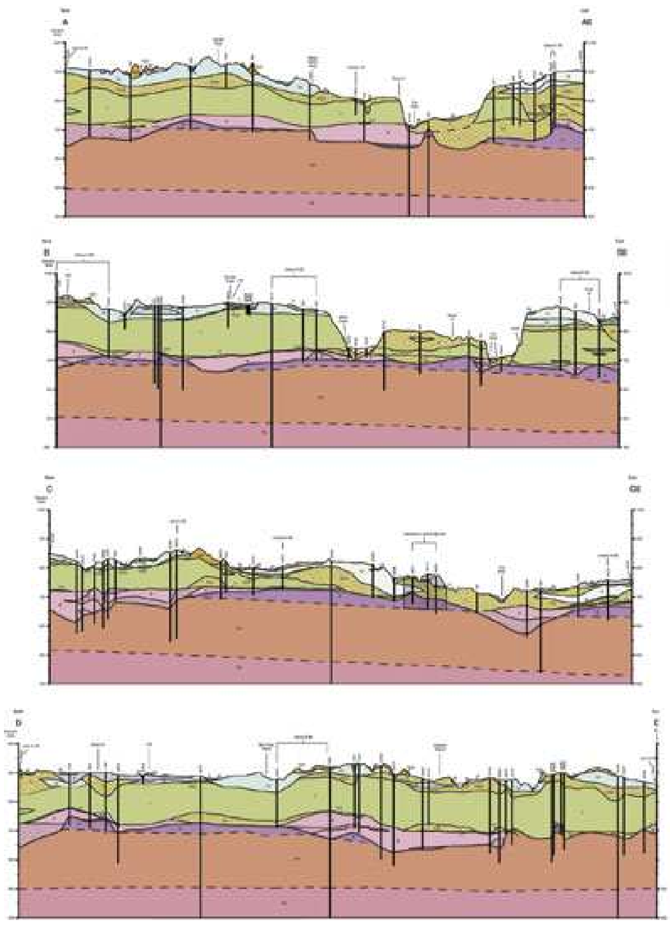


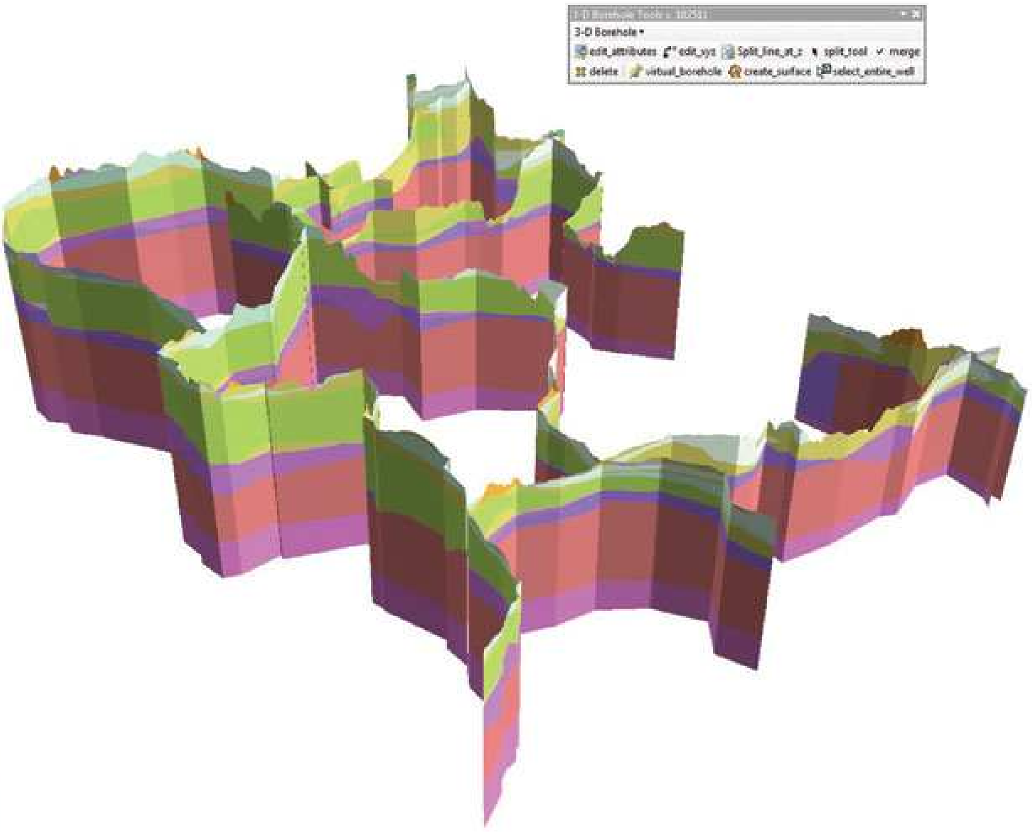
Figure 3: PWC platform elevation (top), Echosounder depth values (middle) and resulting seafloor topography (down)

Applications:

- Time series synchronization
- Noise removal
- ...

Interpolation 3D





http://www.esri.com/news/arcuser/0312/modeling-the-terrain-below.html

More examples: http://inside.mines.edu/ dhale/research.html

Deterministic interpolation.

Given: heights $h_1, \dots, h_n$ at positions $p_1, \dots, p_n$.
Wanted: height h_0 at position p_0.

Deterministic methods:

$$h_0 = f(p_0; (p_1, h_1); \dots; (p_n, h_n))$$

Deterministic: h_0 is completely determined given the parameters it depends on, which are here the $(p_1, h_1); \dots; (p_n, h_n)$ and the p_0.

Write:

$$h_0 = \mathbf{w} \cdot \mathbf{h} = w_1 \cdot h_1 + \cdots + w_n \cdot h_n$$

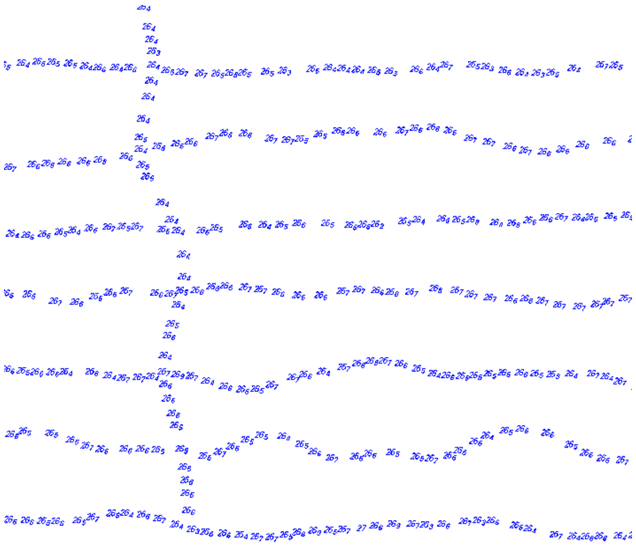
Running question: what are the weights, i.e. what are the w_i?

Interpolation Method 1: Arithmetic mean. Weights?

What is a good interpolation method?

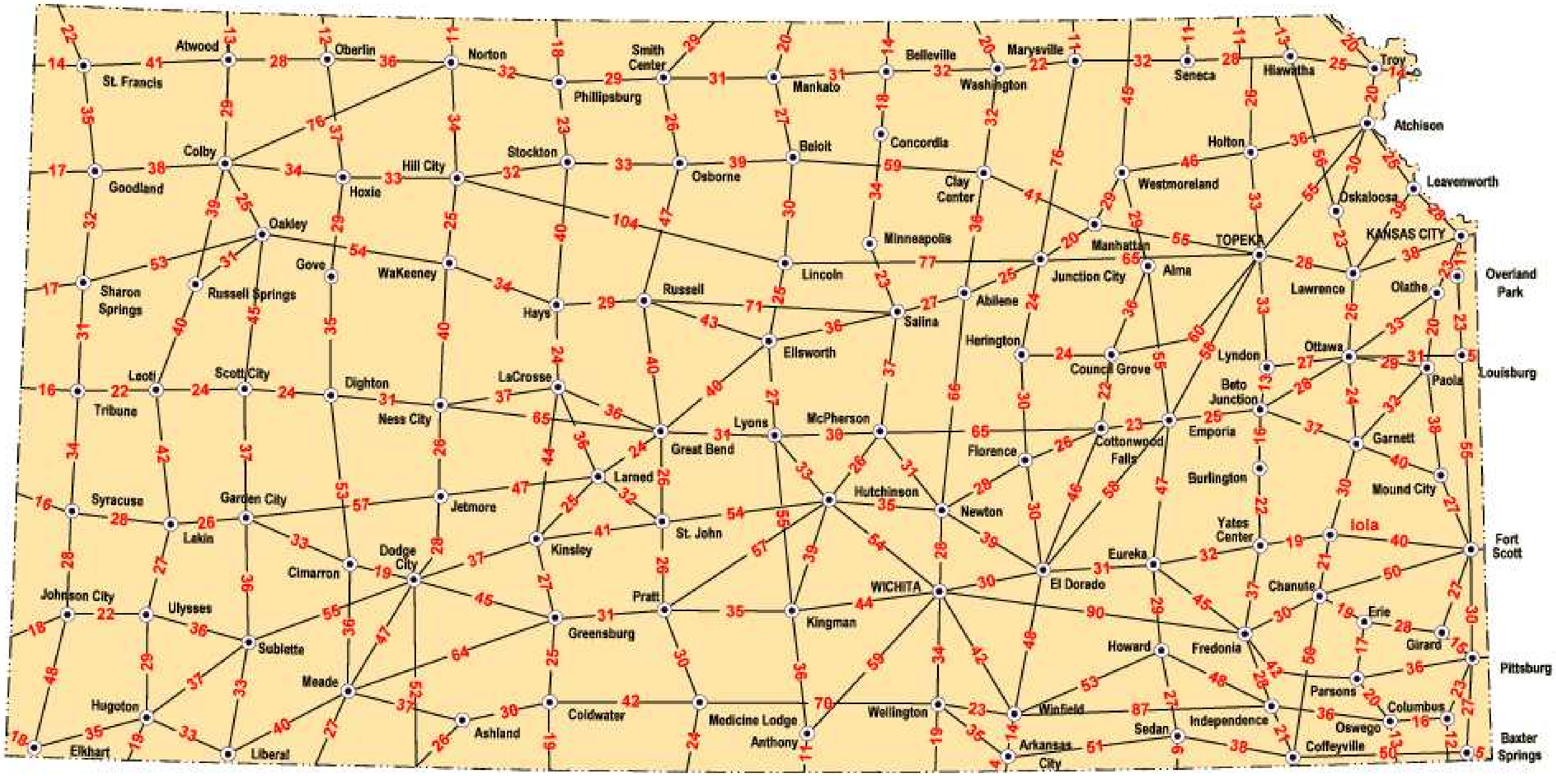
Question. What are possible criteria?

- ...
- ...
- ...
- ..



Single Beam Echo Sounder data

What is distance?

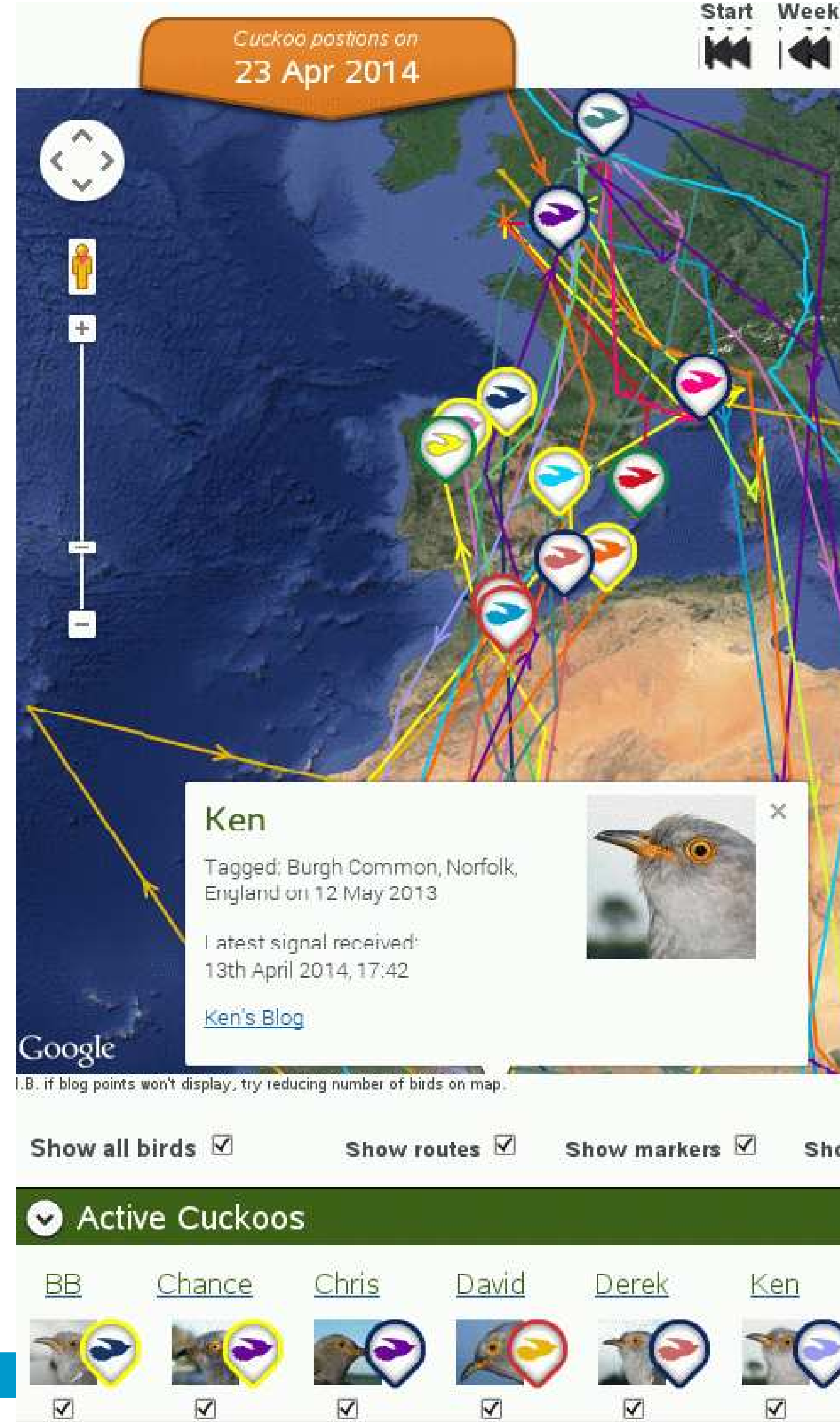


As the bird flies

Great Circle:
circle on a sphere centered at the center of the sphere.

Shortest path between two points on a sphere follows great circle

Question. Good notion of distance on Earth?



http://www.bto.org/science/migration/tracking-studies/cuckoo-tracking

Mathematical definition of distance

A distance map, d, is a map operating on a set X

$$\begin{aligned} d : X \times X &\rightarrow \mathbb{R}_{\geq 0} \\ (x,y) &\mapsto d(x,y) \end{aligned}$$

such that for each pair of points $x, y \in X$:

Symmetry	$d(x,y) = d(y,x)$
Positivity	$d(x,y) \geq 0$ and $d(x,y) = 0 \Leftrightarrow x = y$
Triangle Inequality	$d(x,z) \leq d(x,y) + d(y,z)$

Example.?? $X = \mathbb{R}$ and $d(x,y) = |x-y|$.
Example.?? $X = \mathbb{R}$ and $d(x,y) = |x| - |y|$.
Example.!! Euclidean distance $X = \mathbb{R}^n$ and

$$d(\mathbf{x},\mathbf{y}) = \sqrt{(x_1-y_1)^2 + \ldots (x_n-y_n)^2}$$

Example.!! Distance along great circles

Dept. of Geoscience & Remote Sensing

Inverse distance interpolation

Let $d_i = d(p_0, p_i)$ be the distance between points p_0 and $p_i, i = 1 \dots n$. The inverse distance estimation of the height at location p_0 is given by

$$\hat{h}_0 = \frac{1}{\sum_{i=1}^{n}(1/d_i)}(\frac{h_1}{d_1} + \cdots + \frac{h_n}{d_n})$$

Exercise: Given the locations $p_1 = (0, 1)$, $p_2 = (2, 0)$, $p_3 = (4, 3)$. What are the weights for each location for an estimation at $(0, 0)$?

Answer: $\mathbf{w} \sim (0.59, 0.29, 0.12)$

Question: why is the sum of the weights equal to 1?
(And, moreover, for each weight w_i holds that $0 \leq w_i \leq 1$).

Dept. of Geoscience & Remote Sensing

Power p inv. distance interpolation

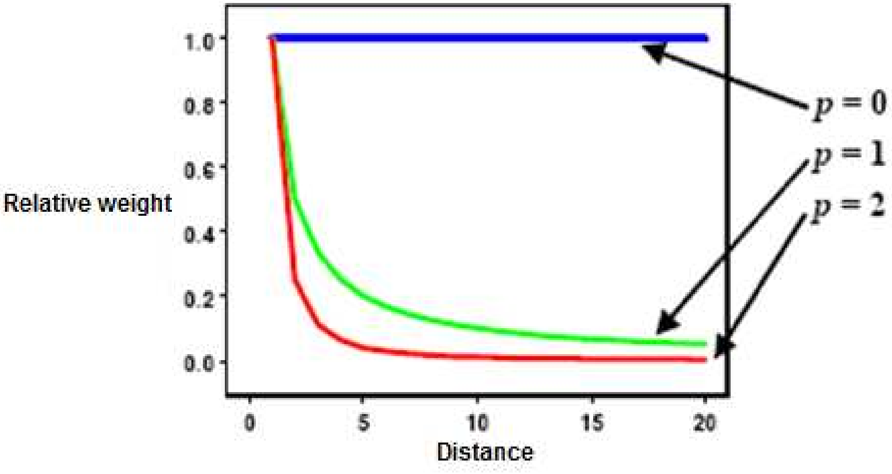
Inverse distance interpolation of power p:

$$\hat{h}_0 = \frac{1}{\sum_{i=1}^{n}(1/d_i^p)}(\frac{h_1}{d_1^p} + \cdots + \frac{h_n}{d_n^p})$$

Limit cases: $p = 0$; $p = \infty$.

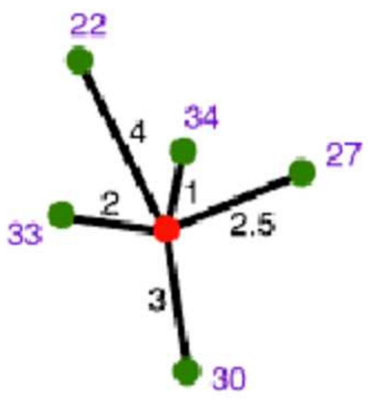
Question: how are the weights divided over the points in the two limit cases?

Question: to what two methods do these two limit cases correspond to?



http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00310000002m000000.htm

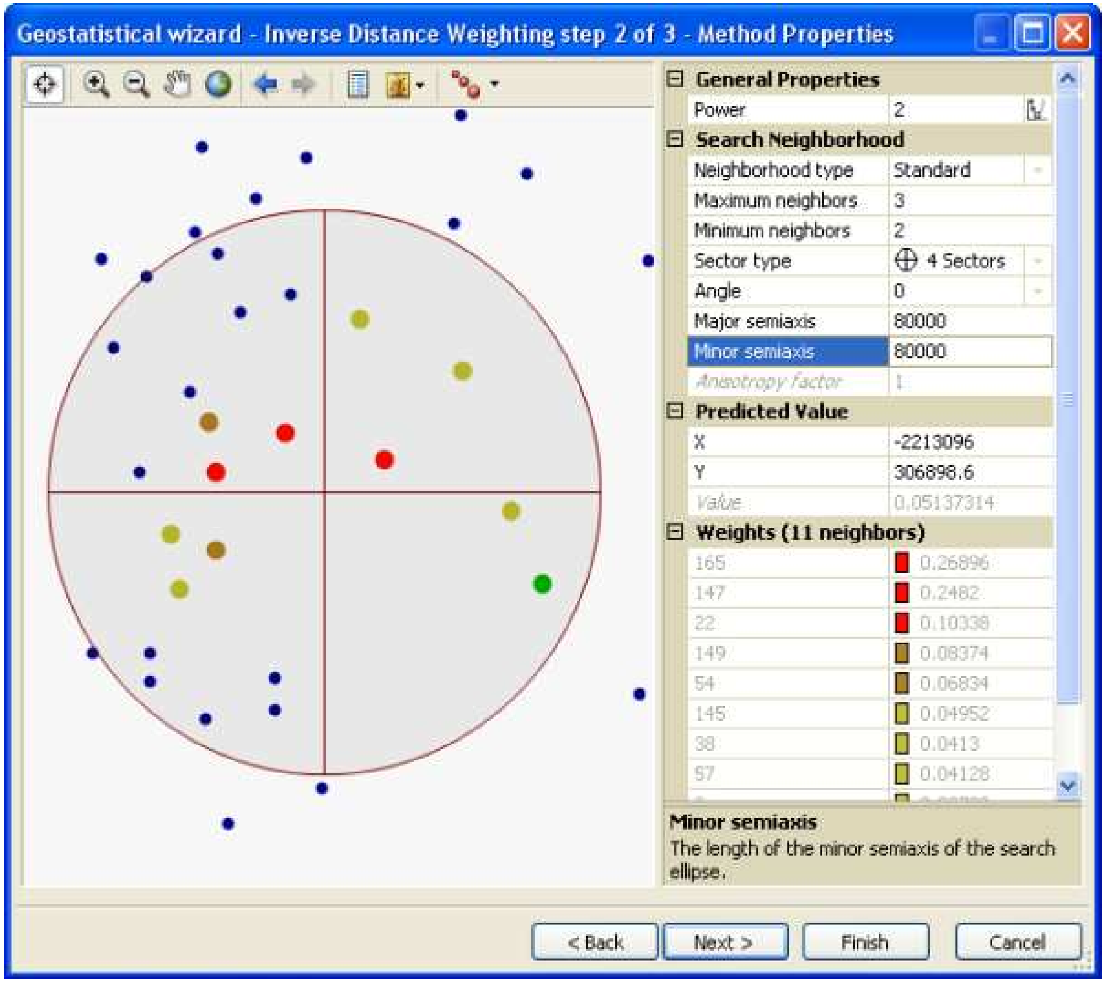
Example, Inverse Distance



$$Z(x) = \frac{\Sigma w_i z_i}{\Sigma w_i} = \frac{\frac{34}{1^2} + \frac{33}{2^2} + \frac{27}{2.5^2} + \frac{30}{3^2} + \frac{22}{4^2}}{\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{2.5^2} + \frac{1}{3^2} + \frac{1}{4^2}} = \mathbf{32.38}$$

Question.
(Dis)advantages, Inverse Distance?

Weight example, $p = 2$, ArcGIS



http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00310000002m000000.htm

Properties, inverse distance interpolation

- Sum of weights is 1.
- Results in a smooth map; results stay in between extreme values
- Needs all n observations

Question How to avoid using all observations?

- No automatic quality description of the result
- Quality of observations is not taken into account

Conclusions

New sensor principle: GNSS

- Position from travel times to multiple satellites
- Typically sparse spatial observations
- High temporal resolution (real-time updates)

Measurement errors

- Systematic and random errors
- Distributions, histograms and quantiles
- Statistics may be sensitive to outliers

Determistic Interpolation

- Key: distance to observations
- Different notions of distance exist
- Deterministic: uncertainty not incorporated and not reported
- Inverse distance: simple, intuitive method

Exercises

See end of slides next lecture