Map accuracy assessment

- How to assess the accuracy of a map
- Issues we face when evaluating accuracy
- Field data collection – sample methods
- Different measurements of accuracy
MAP

ACCURACY ASSESSMENT

THIS MAP IS "GOOD ENOUGH"!

MAY BE GOOD ENOUGH FOR YOU!

WHAT IS "GOOD ENOUGH"?
Field Sampling

The (field) sampling is very important.

The sample (to reduce data and costs) has to be:

- Representative of the population
- Randomly chosen?
- Complete (we have to minimize drop-out)
- Possible / practical to collect
Data sampling reasons

A: Support for object interpretation in image analysis

B: Support for evaluation and statistical analysis
A: Support for object interpretation in image interpretation

Study the material to be interpreted

What objects do you need to understand?

Think of images/field, field/field and image/image discrepancies:

Size differences?
Seasonal differences?
Color differences?
Surface/reflectance characteristics
A: Support for object interpretation in image interpretation

1. Study the material to be interpreted

   What objects do you need to understand?

   Think of images/field, field/field and image/image discrepancies:

   Size differences?
   Seasonal differences?
   Color differences?
   Surface/reflectance characteristics

2. Make field trip and visit typical areas for the objects that are going to be recognized

   Bring RS data into field for comparison?
   Sample with camera, notes, sketches, samples to lab?
B: Support for evaluation and statistical analysis

Compare field with a finished product

Collect a sample in field to compare with the finished product (e.g. a map) to quantify the quality.
B: Support for evaluation and statistical analysis

The real population is an infinite number of data points

A "sample" of reality:
finite number of the population

Sample requirements:
Representative (miniature of reality)
Random
Falling off (loss) should be minimized
(In geography a demand of being feasible)
B: Support for evaluation and statistical analysis

The sampling point.

How large should a "point" be?
Large enough to cover the errors in other parts of the mapping process.

A: Positioning in field.
   Map: 1:50000, 1 mm = 50 m. Hence, Sampling 100x100 meter "points"

B: Geometric errors in aerial imagery (up to 10m) or in satellite imagery (50 m)

C: Map errors (10 m in a 1:10000) map
B: Support for evaluation and statistical analysis

A: Positioning in field.
  Map: 1:50000, 1 mm = 50 m. Hence, Sampling 100x100 meter “points”
B: Geometric errors in in aerial imagery (up to 10m) or in satellite imagery (50 m)
C: Map errors (10 m in a 1:10000) map

More than one error possible = error propagation

Conclusion:
Sampling ”points” (areas) must be chosen with respect to the possible errors.
Areas for aerial imagery ~10 m
How many points do we need?

Statistical significance.

The more the better but aim for at least 50 points in each class to avoid biased samples. (Overrepresentation of correct or incorrect points)

If a sample is too small there is a risk of:
  Type 1 error: Rejecting a correct map
  Type 2 error: Accepting an incorrect map
Different sample designs

**THE SAMPLE HAS TO BE REPRESENTATIVE!**

**DIFFERENT SAMPLE DESIGNS:**

**SIMPLE RANDOM SAMPLING:** Each element has an equal chance of being selected
- Many points are needed to get 50 in each class
- "Very" representative

**SYSTEMATIC SAMPLING:** Elements are selected at some equal interval over space
- No equal chance of being selected
- Uniform spread of points

**STRATIFIED RANDOM SAMPLING:** Allocation into sub-populations (strata), and then random sampling in each stratum
- How to stratify? No equal chance of being selected
- Less points are needed to get 50 in each class
Different sample designs

**Road Sampling:** Systematic sampling in space or time along a number of (randomly selected) roads
- Not at all representative
- Fast

**Transect Sampling:** Random selection of starting point and direction then systematic sampling in space or time
- Not representative
- Relatively fast (depending on terrain)

We recommend simple random sampling or stratified random sampling!
Assessment of Classification Accuracy

- Most common form of expressing classification accuracy is the error matrix (confusion matrix or contingency table)

- Error matrices compare, on a class-by-class basis, the relationship between known reference data (ground truth) and the corresponding results of the classification procedure.
Assessment of Classification Accuracy

Simple error matrix

Realistic error matrix
Overall and Individual Class Accuracy

- **Overall / Total Accuracy**
  - Computed by dividing the total number of correctly classified pixels (i.e., the sum of the elements along the major diagonal) by the total number of reference pixels.

- **Individual Class Accuracy**
  - Calculated by dividing the number of correctly classified pixels in each category by either the total number of pixels in the corresponding column; *Producer’s accuracy*, or row; *User’s accuracy*.
Overall / Total accuracy

\[
(3 + 5) / 12 = 67\%
\]
Producer’s Accuracy

- Producers Accuracy (Omission Errors)
  - Results from dividing the number of correctly classified pixels in each category (on the major diagonal) by the number of reference pixels “known” to be of that category (the column total)

  - This value represents how well reference pixels of the ground cover type are classified
To estimate accuracy we need an error matrix (also called confusion matrix).

The diagonal represents correctly mapped sampling points.

We can let:

- **A** denote the number of correctly mapped points
- **B** denote the number of "ground truth points"
- **C** denote the number of "map data points"
- **N** denote the total number of points

For each class:

**Class I (Land):**
- \( A = 3 \)
- \( B = 4 \)
- \( C = 6 \)

\( N = 12 \)

**Class II (Sea):**
- \( A = 5 \)
- \( B = 8 \)
- \( C = 6 \)
To estimate accuracy we need an error matrix (also called confusion matrix).

The diagonal represents correctly mapped sampling points.

We can let:

- $A$ denote the number of correctly mapped points
- $B$ denote the number of "ground truth points"
- $C$ denote the number of "map data points"
- $N$ denote the total number of points

For each class:

Class I (Land): $A = 3$, $B = 4$, $C = 6$, $N = 12$

Class II (Sea): $A = 5$, $B = 8$, $C = 6$

Producer/Classification accuracy: for each class, the probability that a randomly chosen point in field has the same class value on the map.

For Class I (Land): $\frac{3}{4} = 75\%$

For Class II (Sea): $\frac{5}{8} = 62\%$
## Producer’s Accuracy

### Table 7.3: Error Matrix Resulting from Classifying Training Set Pixels

<table>
<thead>
<tr>
<th></th>
<th>W</th>
<th>S</th>
<th>F</th>
<th>U</th>
<th>C</th>
<th>H</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Set Data (Known Cover Types)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>480</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>485</td>
</tr>
<tr>
<td>S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>72</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>313</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>353</td>
</tr>
<tr>
<td>U</td>
<td>0</td>
<td>16</td>
<td>126</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>142</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>38</td>
<td>342</td>
<td>79</td>
<td>0</td>
<td>459</td>
</tr>
<tr>
<td>H</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>359</td>
<td>0</td>
<td>481</td>
</tr>
<tr>
<td>Column Total</td>
<td>480</td>
<td>356</td>
<td>248</td>
<td>402</td>
<td>438</td>
<td>1992</td>
<td></td>
</tr>
</tbody>
</table>

**Producer’s Accuracy**
- W = 480/480 = 100%
- S = 052/068 = 76%
- F = 313/356 = 88%
- U = 126/248 = 51%
- C = 342/402 = 85%
- H = 359/438 = 82%

**Overall accuracy** = (480 + 52 + 313 + 126 + 342 + 359)/1992 = 84%

---

aW, water; S, sand; F, forest; U, urban; C, corn; H, hay.

24 % **Omission** error
User’s Accuracy

- **Users Accuracy (Commission Error)**
  - computed by dividing the number of correctly classified pixels in each category by the total number of pixels that were classified in that category (the row total)
  - Represents the probability that a pixel classified into a given category actually represents that category on the ground.
**Back to the Error Matrix and Measurements of Map Accuracy**

**User/Object Accuracy:** For each class, the probability that a randomly chosen point on the map has the same class value in field.

- For Class I (Land): \( \frac{3}{6} = 50\% \)
- For Class II (Sea): \( \frac{5}{6} = 83\% \)

To estimate accuracy we need an error matrix (also called confusion matrix).

The diagonal represents correctly mapped sampling points. We can let:

- \( A \) denote the number of correctly mapped points
- \( B \) denote the number of "ground truth points"
- \( C \) denote the number of "map data points"
- \( N \) denote the total number of points

**Class I (Land):**
- \( A = 3 \)
- \( B = 4 \)
- \( C = 6 \)
- \( N = 12 \)

**Class II (Sea):**
- \( A = 5 \)
- \( B = 8 \)
- \( C = 6 \)
User’s accuracy

**TABLE 7.3 Error Matrix Resulting from Classifying Training Set Pixels**

<table>
<thead>
<tr>
<th>Training Set Data (Known Cover Types)*</th>
<th>W</th>
<th>S</th>
<th>F</th>
<th>U</th>
<th>C</th>
<th>H</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>480</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>485</td>
</tr>
<tr>
<td>S</td>
<td>0</td>
<td>52</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>72</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>0</td>
<td>313</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>353</td>
</tr>
<tr>
<td>U</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>126</td>
<td>0</td>
<td>0</td>
<td>142</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>38</td>
<td>342</td>
<td>79</td>
<td>459</td>
</tr>
<tr>
<td>H</td>
<td>0</td>
<td>0</td>
<td>38</td>
<td>24</td>
<td>60</td>
<td>359</td>
<td>481</td>
</tr>
<tr>
<td><strong>Column total</strong></td>
<td>480</td>
<td>68</td>
<td>356</td>
<td>248</td>
<td>402</td>
<td>438</td>
<td>1992</td>
</tr>
</tbody>
</table>

**Producer’s Accuracy**
- W = 480/480 = 100%
- S = 052/068 = 76%
- F = 313/356 = 88%
- U = 126/248 = 51%
- C = 342/402 = 85%
- H = 359/438 = 82%

**User’s Accuracy**
- W = 480/485 = 99%
- S = 052/072 = 72%
- F = 313/353 = 87%
- U = 126/142 = 89%
- C = 342/459 = 74%
- H = 359/481 = 75%

Overall accuracy = (480 + 52 + 313 + 126 + 342 + 359)/1992 = 84%

*W, water; S, sand; F, forest; U, urban; C, corn; H, hay.

28% Commission error
NOTE

Depending on selected measurement we get different accuracy estimations!

Good or bad? Which measurement to use?
Also Note!

If we map without knowledge

<table>
<thead>
<tr>
<th>Ground Truth</th>
<th>Map</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Ground Truth</th>
<th>Class</th>
<th>I</th>
<th>II</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td></td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>II</td>
<td></td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

Chance results in a certain accuracy:

Overall accuracy = 33%

User accuracy: I = 16%  II = 50%

Producer accuracy: I = 25%  II = 38%

We want to get rid of the influence of chance!
This can be done by calculation of the coefficient of agreement, Kappa ($\kappa$)

-1 $\leq \kappa \leq 1$

-1 = No agreement
0 = Random agreement
1 = Perfect agreement

Kappa can be estimated according to

$$\hat{\kappa} = \frac{N_d - q}{N^2 - q}$$

Where

N = Total number of points

$\delta$ = Sum of correctly mapped points

$q$ = Sum of the products between B and C for each class
KAPPA ESTIMATION

EXAMPLE 1: "OUR FIRST EXAMPLE"

\[ N = 12, \; d = 3 + 5 = 8, \; q = (6 \cdot 4) + (6 \cdot 8) = 72 \]

\[ \chi = \frac{(12 \cdot 8) - 72}{12^2 - 72} = \frac{24}{72} = 0.33 \]

MEANING THAT THE MAP IS 33% BETTER THAN CHANCE

<table>
<thead>
<tr>
<th>CLASS</th>
<th>I</th>
<th>II</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAP DATA</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>GROUND TRUTH</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>TOTAL</td>
<td>4</td>
<td>8</td>
<td>12</td>
</tr>
</tbody>
</table>

THE DIAGONAL REPRESENTS CORRECTLY MAPPED SAMPLING POINTS

WE CAN LET

A DENOTE THE NUMBER OF CORRECTLY MAPPED POINTS
B DENOTE THE NUMBER OF "GROUND TRUTH POINTS"
C DENOTE THE NUMBER OF "MAP DATA POINTS"

N DENOTE THE TOTAL NUMBER OF POINTS

CLASS I (LAND): A = 3 B = 4 C = 6 N = 12
CLASS II (SEA): A = 5 B = 8 C = 6

THIS CAN BE DONE BY CALCULATION OF THE COEFFICIENT OF AGREEMENT, KAPPA (\( \chi \))

\[ -1 \leq \chi \leq 1 \]

-1 = NO AGREEMENT
0 = RANDOM AGREEMENT
1 = PERFECT AGREEMENT

KAPPA CAN BE ESTIMATED ACCORDING TO

\[ \chi = \frac{Nd - q}{N^2 - q} \]

WHERE

N = TOTAL NUMBER OF POINTS
q = SUM OF CORRECTLY MAPPED POINTS
D = SUM OF THE PRODUCTS BETWEEN B AND C FOR EACH CLASS
KAPPA ESTIMATION

This can be done by calculation of the coefficient of agreement, Kappa (κ)

-1 ≤ κ ≤ 1

-1 = No agreement
0 = Random agreement
1 = Perfect agreement

Kappa can be estimated according to

$$κ = \frac{N\hat{d} - q}{N^2 - q}$$

Where
N = Total number of points
\(\hat{d}\) = Sum of correctly mapped points
q = Sum of the products between b and c for each class

Chance results in a certain accuracy:

Overall accuracy = 33%
User accuracy: I = 16% II = 50%
Producer accuracy: I = 25% II = 38%

We want to get rid of the influence of chance!

Example 2: "Classification without knowledge"

N = 12, \(\hat{d} = 1 + 3 = 4\), q = (6.4) + (6.8) = 72

$$κ = \frac{(12.4) - 72}{12^2 - 72} = \frac{-24}{72} = -0.33$$

Meaning that the map is 33% worse than chance
KAPPA ESTIMATION FOR AN INDIVIDUAL CLASS (i)

\[ \hat{\chi}_i = \frac{N di - qi}{NB_i - qi} \]

* FOR EXAMPLE II:

CLASS I: \( N = 12, \ d_i = 1, \ q_i = 6.4 = 24, \ B_i = 6 \)

\[ \hat{\chi}_i = \frac{(12 \cdot 1) - 24}{(12 \cdot 6) - 24} = -\frac{12}{48} = -0.25 \]

\( d_i = 1.33 \text{ EXPECTED BY CHANCE} \)

CLASS II: \( N = 12, \ d_i = 3, \ q_i = 6.8 = 48, \ B_i = 6 \)

\[ \hat{\chi}_i = \frac{(12 \cdot 3) - 48}{(12 \cdot 6) - 48} = -\frac{12}{24} = -0.5 \]

\( d_i = 6 \text{ EXPECTED BY CHANCE} \)
Practice in the Verkeå-project 1

- You will interpret and make a map manually from aerial imagery (small area) and make a classification.
- You will collect evaluation data in field by foot (you design the sampling).
- You will calculate the map accuracies.
The total station

\[ \text{DISTANCE} = \text{TIME} \times \text{SPEED} \]

E.G. \[ \text{METRES} = \text{SEK} \times \frac{\text{METRES}}{\text{SEK}} \]
GNSS – Global Navigation Satellite System

• **NAVSTAR GPS** – NAVigation Satellite Time and Ranging system Global Positioning System US

• **GLONASS** – GLObal NAvigation Satellite System Russian system

• **Galileo** – EU system

• **BeiDou** – Chinese system aka COMPASS

• **IRNSS** - Indian Regional Navigation Satellite System
THE GLOBAL POSITIONING SYSTEM (GPS)

- Space Segment
- Control Segment
- User Segment
SPACE SEGMENT

21 + 3 = 24 SATELLITES
CONTROL SEGMENT

HAWAII
COLORADO SPRINGS
ASCENCION
DIEGO GARCIA
KWAJALEIN
USER SEGMENT = RECEIVER
USER SEGMENT = RECEIVER
KNOWN POSITION

\[ x_{S1} \quad y_{S1} \quad z_{S1} \]

\[ t_{S1} \quad \text{TIME SEND} \]

\[ \text{DIST}_1 \]

\[ t_{R1} \quad \text{TIME RECEIVE} \]

\[ x_u \quad y_u \quad z_u \]

UNKNOWN POSITION
TIME FOR SIGNAL SAT. $\rightarrow$ RECEIVER:

$$dt = t_{RI} - t_{SI} + b$$

CLOCK ERROR (RECEIVER)

SPEED OF SIGNAL:

$$c$$ (SPEED OF LIGHT, APPROX. $3 \times 10^8 \text{ m/s}$)

DISTANCE SAT. $\rightarrow$ RECEIVER:

$$\text{DIST}_1 = c \times (t_{RI} - t_{SI} + b)$$
ALTERNATIVE SOLUTION, DISTANCE SAT. → RECEIVER

\[ \text{DIST}_t = \sqrt{(x_{s1} - x_u)^2 + (y_{s1} - y_u)^2 + (z_{s1} - z_u)^2} \]

(THEOREM OF PYTHAGORAS IN THREE DIMENSIONS)

COMBINATION OF THE TWO WAYS:

\[ c \times (t_{RI} - t_{SI} + b) = \sqrt{(x_{s1} - x_u)^2 + (y_{s1} - y_u)^2 + (z_{s1} - z_u)^2} \]

\[ \text{KNOWN} \quad \text{KNOWN} \quad \text{UNKNOWN} \quad \text{UNKNOWN} \]

4 UNKNOWNS
To estimate 4 unknown factors, we need 4 equations in an equation system.

This means we need 4 satellites to estimate our position in x, y, and z.
ONE UNKNOWN

E.G.  \[ x = 17 + 5 \]
     \[ x = 22 \]

TWO Unknowns

E.G.  \[
\begin{align*}
  x &= 15 + y \\
  x &= 3y
\end{align*}
\]

\[ 3y = 15 + y \]
\[ 2y = 15 \]
\[ y = 7.5 \]
\[ x = 15 + 7.5 \]
\[ x = 22.5 \]

WE NEED ONE EQUATION

WE NEED TWO EQUATIONS

AND SO ON...
To locate itself, a GPS receiver must find the distance to three satellites of known positions.
If the receiver finds that it is X miles from one satellite, it knows that it must be somewhere on an imaginary sphere, with the satellite as the center and a radius of X.
If the receiver can generate these spheres for two satellites, it knows it can only be located where the surfaces of the two spheres intersect.
The two spheres overlap in a ring of possible receiver positions.
By generating a sphere for a third satellite, the receiver narrows its possible positions down to two points.
The receiver dismisses the point located in space, leaving only one possible position.
Three satellites
Accuracy ± 10 Meter
Enhanced by better clocks
Advanced calculations using more satellites
Differential GPS

 Accuracy ± 0.75 Meter

Known Point

Unknown Point