Sources of errors in biomass estimation

Dr. Lutz Fehrmann
Regional Course on REDD+, MRV and Monitoring
11-15 July 2011
Sokoine University of Agriculture, Morogoro, Tanzania

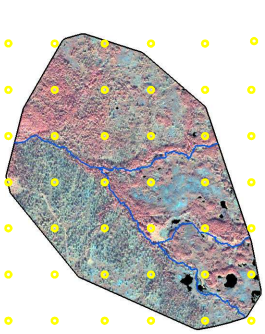
Objectives of this Session

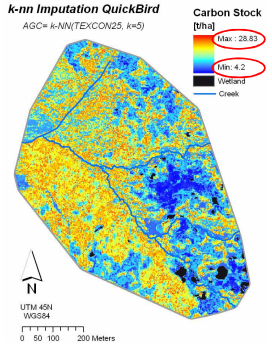
- Summarize the relevant error sources
- Example of classification errors

Which error sources are relevant?

- **Measurement error**
  - Can not be quatified without independent check cuising
- **Sampling error**
  - Variability of biomass per plot
- **Model error**
  - Errors related to the applied biomass model (and model selection!)
- **Co-registration error**
  - Matching of field observations and pixelwise spectral reflectance of satellite imagery
- **Classification error (also in Stratification)**
  - E.g. error in area estimates for distict strata, missclassification of sample points
- **Regionalization error**
  - When producing the map

Wall-to-wall mapping of biomass





Wall-to-wall mapping of biomass

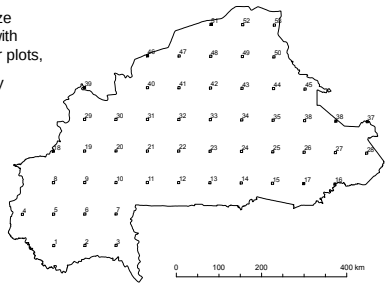
- Imputation techniques (like knn) can only assign / predict values that are in the range of field observations!
- Crucial: However the imputation/classification is done, it always results in a colorful map that seems plausible

Other error sources

- When using RS data, other possible error sources are often ignored:
- Data are only usable after multiple pre-processing steps that might be source of additional errors:
  - Geometric correction
  - Cloud detection
  - Atmospheric correction
  - Topographic Normalization
  - Image balancing

An example of classification errors

- A field based estimation of land use classes (following FAO definitions) for Burkina Faso:
  - Small sample size in combination with rel. Large cluster plots,
  - Sample size only 46! clusters in a systematic grid,
  - No NFI! Was conducted in course of a reserach project



7

An example of classification errors

FRA 2010 categories	FAO 1990	FAO 2000	FAO 2005	FAO 2010	This study	
	Cover [km² and %]					SE%
Forest	68 470	62 480	59 490	56 490	116 847	**9.9**
	24.9%	22.8%	21.7%	21.0%	**42.6%**	
Other wooded land	58 610	54 350	52 220	50 090	4 467	**41.4**
	21.4%	19.8%	19.1%	18.0%	**1.6%**	
Forest and other wooded	127 080	116 830	111 710	106 580	121 315	**9.6**
land	46.4%	42.6%	40.8%	39.0%	**44.2%**	
Other land	146 520	156 770	161 890	167 020	146 729	**8.5**
	53.5%	57.2%	59.1%	60.9%	**53.6%**	
... of which with tree	51 350	55 180	57 100	59 020	13 398	**17.4**
cover	18.7%	20.1%	20.8%	21.5%	**9.1%**	
Total land area	273 600	273 600	273 600	273 600	**270 060**	-
	99.9%	99.9%	99.9%	99.9%	**97.8%**	
Inland water bodies	400	400	400	400	5 957	**99.2**
	0.1%	0.1%	0.1%	0.1%	**2.2%**	
Total area of country	274 000	274 000	274 000	274 000	**274 000**	-
	100%	100%	100%	100%	**100%**	

8

An example of classification errors

- 2010 Assessment was based on a remote sensing study (Landsat)
  - Distinction between shrubs and trees was not possible! Terrestrial sampling shows completely different results!?
  - Forest definition (crown cover %) was assessed on completely different spatial resolution!
    - Terrestrial sampling: crown radii were measured and crown area related to plot area
    - Remote sensing study: ????

9

Some conclusions on sample size

- In order to reduce the SE% for the estimate of forest land from 9.9% to 5%:
  - for a sample size of $n_{9.9\%}$ = 46 observed field plots and assuming simple random sampling with the common error probability of α = 0.05 (setting t = 2), **one would estimate a necessary sample size by the factor (9.9/5)² = 3.9 times larger than in our study**, that is about $n_{5\%}$ **= 3.9*46 = 179**.
- The major question is, whether a government would be ready to cover expenses 3.9 times higher for a 2 times higher precision?

10

Where to invest to reduce errors?

- Interesting research question – with far reaching significance in practice.
  - More field observations.
  - Better measurement devices.
  - Higher resolved remote sensing imagery?
  - Lidar everywhere?
  - Better biomass models (based on more measurements)?
  - Better co-registration between RS and field plots?
  - ….?
- The question is: where in the entire estimation process to allocate more resources to achieve the best improvement of precision per €:

11