

WRITING: Best Practices for Assessing Accuracy of Land-Cover Change and Estimating Area of Change (order of sections can be re-arranged, section leads and contributors can be revised)

15 October 2008

- I. Introduction (Curtis Woodcock and Martin Herold)
  - a. General description of problem addressed – defining the scope
  - b. Objectives of change accuracy (descriptive accuracy of change map, ability to estimate area of change, others? Assessment of net change to augment gross change assessment, or leave it at gross change?)
  - c. Motivating examples, role of GOF-C-GOLD, etc. Specify major ‘cases’ of change assessment we will address, e.g. is it global change products, other?
  - d. Document builds on earlier CEOS document, but key differences include ... specify after document completed. Differences from Strahler (2006): 1) Quality (existence?) of reference data at first time period, 2) Change is rare; 3) Area estimation; 4) others?
  - e. Brief revisit of three major components (response design, sampling design, and analysis), following but not repeating information in Strahler et al. (2006). The write-ups (following major sections) may need to provide some background, but the writing should focus very strongly on how the issues pertain directly to change with minimal redundancy with the information in Strahler et al. (2006) – in other words, we do not need as much background information in the change document.
  
- II. Response design issues (Mike Wulder, Matt Hansen, Frederic Achard, Giles Foody, Phillippe Mayaux)
  - a. Assessment units for change and issues
    - i. Pixels
    - ii. Polygons
    - iii. Blocks
  - b. Reference labeling protocol and issues related to collecting reference data for each assessment unit, for example:
    - i. Regard the assessment unit as homogeneous so reference data must be a single class (a class could be type of change)
    - ii. Unit is homogeneous (single class) according to map but mixture according to reference labeling
    - iii. Single reference label, primary and secondary label, fuzzy reference data
    - iv. Within a block or polygon, may have per pixel (spatially explicit) information on reference class, or possibly(?) just the area of each class in the polygon or block
  - c. Sources of reference data, strengths and weaknesses of each
  - d. Methods to increase consistency of reference labeling (e.g. how to best employ multiple interpreters);

- e. Defining agreement (this should be in Strahler et al. 2006, so perhaps a quick overview). Depending on the assessment unit and labeling protocol, this may be more complex than just a ‘crisp’ label type comparison
- f. Quantifying reference data error (labeling and location error, maybe Giles adds something)
- g. Other topics to include (I’m not sure if I’ve stated these correctly given my poor remote sensing knowledge)
  - i. dense time series of Landsat
  - ii. multi-sensor approach in the future (to include information available for quality control purposes on sensors – Paul Stephens comment)
  - iii. Trade offs or issues related to how much reference data from ground versus how much from other sources
  - iv. There was (I think) discussion of causes of change – I don’t recall what the group decision was on that, but perhaps the issue should be mentioned and the document scope state whether this is or is not part of what is discussed.
- h. Short summary section on recommendations for response design, what are the best practices, and maybe what are practices to avoid.

### III. Analysis issues (Giles Foody)

- a. Basic description via error matrix – usual accuracy measures, importance of standard errors and confidence intervals derived from them (quantifying uncertainty of accuracy estimates)
- b. We may need to include the reminder that the specific accuracy estimation and variance estimation protocols are dependent on the sample (i.e. we can’t do the analyses as if the sample were simple random when strata and clusters are involved).
- c. Collapsed error matrix (2x2 change and no change), mention van Oort (2008) as another reporting format that is in a collapsed form.
- d. Estimating area (Stehman)
  - i. Bias of pixel counts for area estimation
  - ii. Confusion matrix based estimators (Card 1982) and their limitations
  - iii. Regression estimator
  - iv. Synthesis via model-assisted estimation
  - v. Quantifying uncertainty of area estimators
- e. Analysis accommodating reference data error (e.g. Hagen et al. 2003 for a general conceptual framework, but we could also suggest some simpler approaches).
- f. “Trend” in the change (Frederic Achard)
- g. Net change accuracy (Stehman – if topic included)
- h. Short summary section: General recommendations on analysis

### IV. Sampling issues (Stehman, Gallego)

- a. List of desirable design criteria, importance of prioritizing

- b. Major design decisions: strata, clusters, systematic versus simple random protocols
- c. Stratification – rare change
  - i. Motivated by goal of precise class-specific accuracy, particularly for the rare change types
  - ii. Concerns about change omission error as related to stratification
  - iii. ‘Sources’ of stratification
    - 1. Existing change maps
    - 2. Variables correlated with change
  - iv. Allocation of sample size to strata – guidelines for how to quantify
  - v. Situations where stratification may not be warranted
- d. Clusters – motivated by cost
  - i. Cluster size issues (anything special pertaining to change?)
  - ii. Clusters and relation to area estimation objective – stratifying clusters by quantity of change in the block
- e. Possible options for combining stratification and clustering
- f. General recommendations on sampling design

V. Special topics (catch all section)

- a. Cumulative evidence approach, what to do in the absence of adequate reference data (Matt Hansen)
- b. Special needs and concerns of specific major projects, e.g. REDD, Curtis’s work, Mike Wulder’s work, etc.? Include one or two paragraphs per project, enough information to that a reader can see parallels between his/her project and the issues described for the specific projects we include. Not sure this is worth including if we don’t resolve the specific needs and concerns.
- c. Effect of resolution on accuracy of area estimates (and also accuracy of map change?)
- d. We have the list of change accuracy examples that Martin put together. We might include that here, although I think we have to be careful to be clear we are not endorsing any of these methods and we are not even critiquing them. Alternatively, we could evaluate some of them and endorse some as good practice? That may be a lot of work. The question is how can we use the past literature of examples?

VI. Future research and unresolved issues (Contributions from all)