
Pontus Olofsson, Giles M. Foody, Martin Herold, Stephen V. Stehman, Curtis E. Woodcock and Michael A. Wulder

a. Department of Earth and Environment, Boston University, 685 Commonwealth Avenue, 02215 Boston, MA, USA
b. School of Geography, University of Nottingham, University Park, Nottingham NG7 2RD, UK
c. Laboratory of Geo-Information Science and Remote Sensing, Wageningen University, Droevendaalsesteeg 3, 6708 Wageningen, The Netherlands
d. Department of Forest and Natural Resources Management, State University of New York, 1 Forestry Drive, Syracuse, NY 13210, USA
e. Canadian Forest Service, Natural Resources Canada, Victoria, BC, 12 V8Z 1M5, Canada

*Corresponding author
Email olofsson@bu.edu
Tel +1-617-353-9734
• Aim to provide practitioners with hands-on recommendations for designing accuracy assessments of change maps, and how to use accuracy information for estimating area and uncertainty

• “Good” (not “best”) practices provided – recommendations satisfy accepted scientific practice

• Practitioners can point to us if assessment designs questioned (i.e., aim to transfer some of the responsibility from practitioners to us)

• Recommendations for all steps of the process are given including (I) sampling design, (II) response design and (III) analysis; (IV) example provided to illustrate recommended workflow
• Draft in final stage of in-house review
• Aim to submit paper to peer-reviewed journal this or the following week

• Additional paper that provides clear examples:

I. Sampling Design: Good Practices
• Implement a probability sampling design to provide a rigorous foundation to sampling inference

• Document and quantify deviations from probability sampling design

• Choose a sampling design on the basis of specified accuracy objectives and desirable design criteria – both simple random and systemic selection protocols are acceptable options

• Stratify by map class to reduce standard errors of class-specific accuracy estimates

• Use cluster sampling only if it provides a substantial cost savings or if the objectives require a cluster unit for the assessment

• Evaluate the total sample size using the standard error formula and a specified target standard error
Allocate sample to strata such that sample size for rare change classes are increased (~100) to achieve an acceptable standard error for estimated user’s accuracies and allocate the remaining sample size roughly proportional to the area occupied by the common classes.

*Stratified random sampling* using the map classification to define strata is a simple, but generally applicable design that will typically satisfy most accuracy and area estimation objectives and desirable design criteria.
II. Response Design: Good Practices
- The overhead cost required for field visits is likely not justified
- The reference data should provide sufficient temporal representation consistent with the change period of the map
- Reference data should be of higher quality than the data used for creating the map, or if using the same source, the process of creating the reference classification should be more accurate than the process of creating the map
- Data from the Landsat open archive in combination with high spatial resolution imagery provide a low-cost and often useful source of reference data (national photo archives, satellite photo archives (e.g., Kompsat), and the collections available through Google Earth™ are possible high resolution imagery sources)
- Specify protocols for accounting for uncertainty in assigning the reference classifications
- Assign each sample unit a primary and secondary label (secondary not required if highly confident)
Include an interpreter specified confidence for each reference label (e.g., high, medium, or low confidence)

Ideally have three interpreters label each unit sampled, and implement protocols to ensure consistency among individual interpreters or teams of interpreters

Specify a protocol for defining agreement between the map and reference classifications that will lead to an error matrix expressed in terms of proportion of area
III. Analysis: Good Practices
- Report the error matrix in terms of estimated area proportions
- Report the area (or proportion of area) of each class as determined from the map
- Report user’s accuracy (or commission error), producer’s accuracy (or omission error), and overall accuracy (Equations 1-3)
- Avoid use of the kappa coefficient of agreement for reporting accuracy of land change maps
- Estimate the area of each class according to the classification determined from the reference data
- Use estimators of accuracy and area that are unbiased or consistent
- For simple random, systematic, and stratified random sampling when the map classes are defined as strata, use stratified estimators of accuracy and a stratified estimator of area
• Quantify sampling variability of the accuracy and area estimates by reporting standard errors or confidence intervals

• Use design-based inference to define estimator properties and to quantify uncertainty

• Assess the impact of reference data uncertainty on the accuracy and area estimates
Where are we and is there a next step?

Basically done at this stage – minor edits and submission, followed by likely requests for revision

Are there other issues that would benefit?

One possibility in the REDD+ context might be some standard change products
Designing a Global Reference Database for Assessing Accuracy of Land Cover Maps

Pontus Olofsson, Curtis E. Woodcock, Christopher Holden, Mark A. Friedl and Damien Sulla-Menashe

Dept of Geography and Environment, Boston University
675 Commonwealth Avenue, Boston, MA 02215, USA

Stephen V. Stehman

Dept of Forest and Natural Resources Management, State University of New York, 1 Forestry Drive, Syracuse, NY 13210, USA

Martin Herold

Laboratory of Geo-Information Science and Remote Sensing, Wageningen University, Droevendaalsesteeg 3, Wageningen, The Netherlands

Chandra Giri and Tom Loveland

USGS

GOFC-GOLD

Global Observation of Forest and Land Cover Dynamics
Goals

- Provide reference data that could be used to help evaluate the accuracy of any LCCS compliant land cover maps (and probably others as well!)
- Provide an accuracy assessment framework that can be augmented regionally or thematically
- Promote the importance of validation activities
- Ultimately, facilitate the process of improving efforts to develop synthetic land cover maps (using multiple land cover maps to produce a new single “improved” land cover map)
Sample Design

- Applicable to assess accuracy for a variety of land-cover maps - independent of existing land-cover maps
- Ability to increase sample size for rare land-cover types
- Ability to enlarge (augment) sample size in targeted regions or land-cover classes
- A random stratified sampling meets these criteria
- Each sample is a 5x5 km block


The Köppen climate system and population data – basis of the stratification:

- Köppen map based on station data – manual editing of climatic borders
- Original Köppen map (Peel et al., 2007) – 32 classes – collapsed to 13 classes
- Intersected by population data: 5 persons/km² => 26 populated and unpopulated classes
- Collapsed to 20 classes and areas of > 1000 persons/km² as a 21st class (built up)
- Areas with water removed by intersection of land/water mask
Table 3: The final strata, its global distribution and allocation. A “p” in front of the strata name denotes populated version.

<table>
<thead>
<tr>
<th>No.</th>
<th>Strata</th>
<th>Distr.</th>
<th>Prop. all.</th>
<th>Final all.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tropical rainforest</td>
<td>2.4%</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Tropical seasonal forest</td>
<td>2.0%</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Savannah</td>
<td>5.0%</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>Desert</td>
<td>14.4%</td>
<td>72</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>Steppe</td>
<td>8.3%</td>
<td>41</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>Mediterranean</td>
<td>1.6%</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>Temperate evergreen forest</td>
<td>1.2%</td>
<td>6</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>Marine west-coast</td>
<td>1.6%</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td>9</td>
<td>Continental forest</td>
<td>4.3%</td>
<td>22</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>Boreal forest</td>
<td>12.7%</td>
<td>63</td>
<td>50</td>
</tr>
<tr>
<td>11</td>
<td>Cold boreal forest</td>
<td>1.2%</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>12</td>
<td>Tundra</td>
<td>3.3%</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>Frost</td>
<td>1.2%</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>pTropical rainforest</td>
<td>2.2%</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td>pTropical seasonal forest</td>
<td>1.9%</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>16</td>
<td>pTropical savannah</td>
<td>11.0%</td>
<td>55</td>
<td>40</td>
</tr>
<tr>
<td>17</td>
<td>pDesert</td>
<td>6.0%</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>18</td>
<td>pSteppe</td>
<td>7.0%</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>19</td>
<td>pTemperate evergreen forest</td>
<td>5.2%</td>
<td>26</td>
<td>40</td>
</tr>
<tr>
<td>20</td>
<td>pContinental forest</td>
<td>6.7%</td>
<td>34</td>
<td>50</td>
</tr>
<tr>
<td>21</td>
<td>Urban</td>
<td>0.6%</td>
<td>3</td>
<td>35</td>
</tr>
</tbody>
</table>
Locations of sample sites

- Tropical Rainforest
- Tropical Seasonal Forest
- Tropical Savannah
- Desert
- Steppe
- pTropical Rainforest
- pTropical Seasonal Forest
- pTropical Savannah
- pDesert
- pSteppe
- Mediterranean
- Temperate Evergreen Forest
- Marine West-coast
- Continental Forest
- Boreal Forest
- pTemperate Evergreen Forest
- pContinental Forest
- Urban
- Tundra
- Snow and Ice
- Cold Boreal Forest
Response Design

- Land cover legend and workflow for reference map production defined
- Object-based analysis of 0.5m pan-sharpened data (mostly from Digital Globe)
- Legend based on the following required classes

<table>
<thead>
<tr>
<th>Tree</th>
<th>Herbaceous</th>
<th>Non-vegetated land</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Broadleaf evergreen</td>
<td></td>
<td>6. Row crop</td>
<td></td>
</tr>
<tr>
<td>2. Broadleaf deciduous herbaceous</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Shrub</td>
<td></td>
<td>10. Water</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>11. Snow/Ice</td>
<td></td>
</tr>
</tbody>
</table>
Site 272 - Sugar Land, TX

Scene acquired: 5-25-2005
Sensor: Quickbird

Legend
- Unclassified
- 2 - Broadleaf deciduous
- 6 - Row crop
- 7 - Other herbaceous
- 8 - Bare
- 9 - Built up
- 10 - Water
- 12 - Unusable

Boston University 6/10/2011
Site 128 - Bellingham, WA

Scene acquired: 10-26-2007
Sensor: Quickbird

Legend
- Unclassified
- 2 - Broadleaf deciduous
- 3 - Needleleaf evergreen
- 5 - Shrub
- 6 - Row crop
- 7 - Other herbaceous
- 8 - Bare
- 9 - Built up
- 10 - Water
- 12 - Unusable
USGS Effort (Chandra Giri, Tom Loveland et al.,)

Simplified response design: fewer classes
(forest, other veg, water, bare)

Classification of the 2.5m multispectral data

225 sites done (all of western hemisphere as well as a 98 site representative sample!)
Current status

- Sampling and response design defined – sampling designed published
- A prioritized subset of 98 images has been identified
- 33 reference maps created as of March 2013
Next Steps

- Find funding!
- Recruit more help for map production
- Write paper on response design
- Continue working on the analysis design
- Upcoming workshop at Wageningen University this summer (Nandika)