Valerio Avitabile, Maurizio Santoro, Martin Herold, Heiko Balzter, Sytze de Bruin, Joao Carreiras, Nuno Carvalhais, Shaun Quegan, Kevin Tansey, Pedro Rodriguez-Veiga
D4: Ground Data Document
Aim: Cal/Val of the Global product

Outputs (D4):
- Ground Data Document
- Ground Database

Three sources of data:
1. Field plots
2. High-resolution maps
3. Sub-national statistics
Field plots

Existing data from:
- Research networks (Fluxnet, RAINFOR, GEM, AfriTron, etc.)
- Research projects (Brazil, Indonesia, Ghana, Ethiopia, Laos, Guinea Bissau, etc.)
- National Forest Inventory (Europe, Uganda, Mexico, Guyana, Vietnam)
- Forest concessions (DRC, Sierra Leone)

Metadata Quality Criteria:
- Plot coordinates acquired with GPS
- Ground measurements from year 2000
- AGB for all living trees with DBH ≥ 0-10 cm
- Appropriate allometric model
- Allometry from Dbh and wood density

Field plots (Tropics)
# Field plots - Tropics

<table>
<thead>
<tr>
<th>QA/QC</th>
<th>Field plots</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Pre-Screening</td>
<td>Metadata analysis</td>
</tr>
<tr>
<td>2. Upscaling</td>
<td>with Google Earth / Tree Cover (VCF)</td>
</tr>
<tr>
<td>3. Aggregation</td>
<td>Average Biomass @ map resolution</td>
</tr>
</tbody>
</table>
Ground database - v.02

Version 02 (January 2016):
- 28 ground datasets (12,738 reference plots)
- NFI in Europe: acquisition ongoing
Quality Criteria:
- Locally calibrated, res. ≤ 100m, published, etc.

### QA/QC Biomass maps

1. Pre-Screening
   - Metadata analysis
2. Upscaling
   - Aggregation @ map resolution
3. Area selection
   - Areas with higher confidence

<table>
<thead>
<tr>
<th>Continent</th>
<th>Country</th>
<th>Extent</th>
<th>Year (map)</th>
<th>Resolution (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>Uganda</td>
<td>National</td>
<td>1999-2003</td>
<td>30</td>
</tr>
<tr>
<td>Africa</td>
<td>Madagascar</td>
<td>Local</td>
<td>2010</td>
<td>100</td>
</tr>
<tr>
<td>Africa</td>
<td>Mozambique</td>
<td>Local</td>
<td>2007</td>
<td>50</td>
</tr>
<tr>
<td>Africa</td>
<td>Cameroon</td>
<td>Local</td>
<td>2007</td>
<td>100</td>
</tr>
<tr>
<td>Africa</td>
<td>Cameroon</td>
<td>Local</td>
<td>2007-2010</td>
<td>25</td>
</tr>
<tr>
<td>Africa</td>
<td>Guinea Bissau</td>
<td>National</td>
<td>2008</td>
<td>50</td>
</tr>
<tr>
<td>S. America</td>
<td>Peru</td>
<td>National</td>
<td>NA</td>
<td>100</td>
</tr>
<tr>
<td>S. America</td>
<td>Colombia</td>
<td>Sub-nat.</td>
<td>2010</td>
<td>100</td>
</tr>
<tr>
<td>C. America</td>
<td>Mexico</td>
<td>National</td>
<td>2007</td>
<td>30</td>
</tr>
<tr>
<td>C. America</td>
<td>Panama</td>
<td>National</td>
<td>2008 - 2012</td>
<td>100</td>
</tr>
<tr>
<td>Australia</td>
<td>Queensland</td>
<td>Local</td>
<td>2009</td>
<td>50</td>
</tr>
</tbody>
</table>
Sub-national Statistics

Available:
› Canada
  – NFI (2005) plots not georeferenced (AGB, GSV)
› USA
  – FIA plots with approx. coordinates (AGB, GSV)
› Europe
  – Statistics for counties and provinces

To be acquired:
› Russia
› China
› India
› Japan
Version 02 (January 2016):
• 28 ground datasets (12,738 reference plots)
• 14 reference biomass maps
D5: Validation Protocol
Validation approach

Stage 1
- Small validation sample

Stage 2
- Significant validation samples
- Spatial/temporal consistency evaluated
- Results published

Stage 3
- Uncertainty fully quantified
- Rigorous sampling

Stage 4
- Systematically updated

http://lpvs.gsfc.nasa.gov/
Validation approach

**Uncertainty assessment**
- **Performed by:** production teams
- **Inputs:** model and data uncertainties
- **Outputs:** uncertainty metrics and uncertainty maps

**Independent validation**
- **Performed by:** production teams and independent project partners
- **Inputs:** Validation database and GlobBiomass products
- **Outputs:** error statistics, error analysis

**Maps inter-comparison**
- **Performed by:** production teams and independent project partners
- **Inputs:** GlobBiomass products + similar existing products
- **Outputs:** discrepancy maps

**User assessment**
- **Performed by:** user communities
- **Platforms:** Regional teams, Geo-Wiki
- **Outputs:** product assessment and recommendations

---

**DUE GlobBiomass**

**D5**
Validation Protocol

Prepared for European Space Agency (ESA-ESRIN)
In response to ESRIN/Contract No. 4000113100/14/I_NB

Wageningen University and Research Centre, Laboratory of Geoinformation Science and Remote Sensing, The Netherlands

December 2015
Uncertainty:
› Assess error sources
› Assess model precision / repeatability
   • From multiple model realization (95% Prediction interval, Quality flags)

Accuracy Assessment:
› Compare estimates with reference data
› Output: Error statistics
› Consider bias and precision

Approach: identify optimal (statistical, comprehensive) and reachable targets for assessing biomass products, and future research needs
1. Uncertainty assessment
Uncertainty estimation

Sources of uncertainty:
1. **RS data**: Technical limitations of remote sensing instruments
2. **Additional datasets**: Accuracy of input maps, if used
3. **Ground data**: Amount, distribution and quality of calibration data
4. **Model**: Uncertainty of models used to transform RS signals in AGB

Approaches for uncertainty estimation:
› Error propagation theory
  \[ \varepsilon_{\text{AGB}} = (\varepsilon_{\text{measurement}}^2 + \varepsilon_{\text{allometry}}^2 + \varepsilon_{\text{sampling}}^2 + \varepsilon_{\text{prediction}}^2 )^{1/2}, \]
› Monte-Carlo simulations
Uncertainty estimation

Example: Uncertainty of Reference Data

› Error at tree level
  • Measurement error (Dbh, height, species/wd)
  • Allometry error

› Sampling error
  • Representativeness of plots of AGB of the pixel
  • Representativeness of samples of AGB of the area

› Spatial & temporal mismatch

Define unit of error (or Uncertainty):

› Variance, rel. error, etc.

Define spatial resolution of error

› Compute errors at different resolutions
2. Independent Validation
Validation Data

- Screening for quality criteria
  - Min. plot area, same biomass pool, GPS, ± 1 year or stable area, etc.

- Harmonize
  - Convert to same DBH, use same allometry, etc.

- Aggregation to map resolution

Amount of validation data
- Independent dataset: 10-30% of ground data (random, stratified)
- Cross-validation for very small reference datasets
Validation

Biomass maps

- Accuracy metrics:
  - RMSE
  - Rel. RMSE (%)
  - Bias (mean error)
  - 95% C.I. of the mean (error)

- Compute metrics by biomass class
- Compare histograms/PDF with ref. data
- Include uncertainty of validation data
- Assess representativeness of val. data to biomass distribution
Validation

Biomass change maps

- Validation
  - Use reference data (permanent plots), if available

- Assess consistency
  - Visual analysis of change areas with high-res. images (commission errors)
  - Assess changes on ‘stable areas’ (omission errors)

- Assess Uncertainty of change
  - Develop and compare C.I. (Overlapping or separate)
3. Maps inter-comparison
Objective:
- Evaluating relative consistency
- Identify areas with higher disagreements
- Assess strengths and weaknesses of different datasets
- Build confidence in the user communities

Output:
- Scale:
  - At pixel level
  - Aggregated resolution
- Difference metrics:
  - Mean difference, scatterplots, RMSE
- Difference map:
  - Absolute difference
  - Relative difference (%)
4. User assessment
User assessment and feedback

User Assessment is an essential quality control and feedback mechanism

– Objectives:
  • assess the users’ acceptance of the products
  • evaluate the quality and limitations from User’s perspective
  • obtain recommendations to future improvements

– Metric:
  • Questionnaires

– Output:
  • User survey report
User assessment and feedback

Geo-Wiki

A platform to visualize, analyze and improve biomass datasets

The Geo-Wiki Project is a global network of volunteers who wish to help improve the quality of global environmental data. Since large differences occur within existing global environmental datasets (e.g., land cover, biomass, cropland), the Geo-Wiki Project provides a platform to visualize, analyze and improve upon these differences. The Biomass Geo-Wiki has collected a comprehensive set of recent biomass data from around the globe, and makes it freely available for visualization. Users are provided with an instant global overview of available datasets, overlaid on the Google Earth platform with comparable units. This provides an instant gap analysis of global data. Additional data to be uploaded could include geo-tagged pictures, in-situ measurements and more. Finally it would be possible with a critical mass of data to produce a global mosaic of terrestrial biomass.

View publication