2020 Census Program: Reengineering Address Canvassing

Barbara P. Buttenfield, CSAC member

Department of Geography
Meridian Lab
University of Colorado – Boulder
babs@colorado.edu

Questions on 2020 Census Readiness Canvassing Test

- Is proportion of national blocks (12,600) sufficient test?
  - As of 2010, 11,166,336 (US + territories) = 0.113% (not even 1%)
  - Expanded canvassing test argues for automating In-Office Canvassing (but you know that already, I think)
- Does test in single locale cover all potential addressing issues?
  - Decision rules for one part of the country might not apply uniformly
- Multiple tests for metro areas with differing demographic trends
  - Fast growth (e.g., Las Vegas)
  - Fast decline (e.g. Detroit)
  - Relative stability (e.g., Pittsburgh)
  - Demographic restructuring (e.g., influx of immigrants from war conflict, migration of retiring Baby Boomers) might be evident in address growth or decline over a ten year period, even if it is not always trackable refuges or recent immigrants.

Quality Expectations

- In-Field checking
  - 25% goal but testing under 1% - what check rate achieved in past?
  - Consider regressing check rates to qualify base rate (recent automated technologies should improve check rates, so regression gives baseline)

- In-Office Review
  - Imagery checks use what source(s)?
  - Again, I urge use of automated pattern recognition to identify changes, instead of manual image checks
  - Reduce Active and On-Hold category size with ancillary data
  - Data integration issues will arise (differing attribution categories over time, for example) and will take time to address, but still less expensive than additional In-Field checks.

Ancillary Data Sources

- GIS data sources for residential areas / structures
  - Parcel databases (updated locally and often at cycles < 1 year)
  - NLCD impervious surface? (most recent version is 2011, 30 m resolution
  - Global Human Settlement Layer (GHSL) an EU product, 38 m resolution

  Last two won't ID individual addresses but could give current information on changed residential areas

- Fine resolution imagery data sources (1m or finer)
  - IKONOS (launched 1999), QuickBird (2001), OrbView (2003)
  - WorldView (v1=0.44m pan, 2007; v2 = 50cm color 2011; v3 30cm 2015)
  - Airborne LiDAR (1m, 3m) – USGS partial coverage for nation
  - DigitalGlobe, Blackbridge Geomatics (Planet Labs); ApolloMapping

Automated Imagery Review

- Automated pattern recognition of imagery (Shih 2010)
- Statistical pattern recognition (Duda et al 2001)
- Object-Based Image Analysis, or OBIA (Blaschke 2010; Bian 2009)
- Image segmentation (Neubert et al 2008, Hay & Castilla 2008, Lang, 2008)
  - Point-based, edge-based, region-based methods
  - ‘Segments” = regions generated by geometric criteria (texture) or statistical (homogeneity in 1+ bands or feature space dimensions).
  - Segments add spectral info (mean values per band, mean ratios, variance)
  - Segments add spatial info (distances, neighborhood delineation, topologies) that can be compared against expectations (e.g., what is average space between house and detached garage, versus between houses)
  - Spatial info especially informative in suburban and rural areas, where some GIS data sources (e.g., NLCD) tend to underestimate residential areas

Software, Analysis Methods and Quality Control

- OBIA & Image Segmentation Software
  - eCogntion (2000) à Definiens (2007?)
  - Feature Analyst, SAGA, ENVI Feature Extraction, ERDAS Imagine 9.3
  - Open Source: Hierarchical Split Merge Refinement (HSMR) tools
- Analysis Methods
  - CART (classification and regression trees)
  - Machine Learning
  - Support Vector Machines
  - Artificial Neural Networks

These are heuristic not confirmatory methods. Uncertainty assessment involves confusion matrices, random perturbation, sensitivity analysis, simulation, bootstrapping, and similar methods to establish probability distributions and confidence intervals for the given samples.

References for new image checking methods

- Blaschke, T 2010 Object Based Image Analysis for Remote Sensing. *ISPRS Journal Photogrammetry and Remote Sensing* 65(1): 2-16. http://dx.doi.org/10.1016/j.isprsjprs.2009.06.004
- **Blaschke et al (Eds.) 2008. O*bject Based Image Analysis*. Springer**.
- Duda, RO, Hart, PE and Stork, DG 2001 *Pattern Classification.* Wiley.
- Hay GJ & Castilla G 2008 Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. In: Blaschke et al (Eds.): 93–112.
- Shih, FY 2010 *Image Processing and Pattern Recognition: Fundamentals and Techniques.* Wiley.
- Lang S 2008 Object Based Image Analysis for Remote Sensing Applications. Blaschke et al (Eds,): 1-25.
- Zhou W and Troy A (2008) Object Oriented Approach for Analysing and Characterizing Urban Landscapes at the Parcel Level. *Int'l Journal Remote Sensing* 29(11):3119-3135.

References for image accuracy assessment literature

- Albrecht F 2008. Assessing Spatial Accuracy of Object-Based Image Classifications. In: Car, Griesebner, Strobl (Eds.) *Proceedings of Geoinformatics Forum '08.* Heidelberg: 11-20.
- Grenier M, Labrecque S, Benoit M, and Allard M 2008. Accuracy Assessment Method for Wetland Object-Based Classification. *Proceedings GEOBIA 2008: Pixels, Objects, Intelligence: 285-289.*
- Liu Y, and Zhou Q. 2004. Accuracy Analysis of Remote Sensing Change Detection by Rule-based Rationality Evaluation With Post-Classification Comparison. *Int'l JnlRemote Sensing* 25(5): 1037-1050.
- Luscier JD, Thompson WL, Wilson JM, Gorham BE, and Dragut LD. 2006. Using Digital Photographs and OBIA to Estimate Percent Ground Cover in Vegetation Plots. *Frontiers in Ecology and the Environment* 4(8): 408-413.
- Möller M, Lymburner L, and Volk M. 2007. The Comparison Index: A Tool for Assessing the Accuracy of Image Segmentation. *Int'l. Jnl of Applied Earth Observation and Geoinformation* 9(3): 311-321.
- Neubert et al 2008 Assessing Image Segmentation Quality. In Blaschke et al (Eds.):169-184.
- Platt RV and Rapoza, L. 2008. Evaluation of an Object Oriented Paradigm for Land Use / Land Cover Classification. *Professional Geographer* 60(1): 87-100.
- Zhang QF, Molenaar M, Tempfli K, and Shi W. 2005. Quality Assessment for Geospatial Objects Derived from Remotely Sensed Data. *Int'l Jnl. Remote Sensing* 26(14): 2953-2974.