Population mapping

*Andy Tatem*
Uses of small area population data

• Planning elections
• Calculating GDP
• Local governance
• Traffic planning
• Health systems
• Financial services
• Delivering utilities
• Controlling infectious diseases

• Agricultural subsidies
• Taxation
• Land use management
• Energy strategies
• Supply chain management
• Meeting SDGs
"Traditional" data sources

- Census
- Surveys
- Administrative/registry data
Demographic data challenges

Coarse resolution
Outdated
Incomplete
Inaccuracies, missing populations
Complimentary data sources

“Traditional” data sources

- Census
- Surveys
- Administrative/registry data

“Novel” data sources

A MAP OF EVERY BUILDING IN AMERICA

By TIM WALLACE, DEREK WATKINS and JOHN SCHWARTZ

Oct. 12th, 2018
Geography as a framework for data integration
Gridded data

Grids: flexibility in analysis and data integration

Comprehensive Emergency Obstetric and Neonatal Care (CEmONC) Facilities overlaid on grid of women of childbearing age

Percentage of women of childbearing age per woreda within 50km of a CEmONC population totals
Data integration
A Top down approach

Census population counts

Geospatial covariates

Population disaggregation

Spatial weighting layer created based on covariates, using dasymetric mapping

Gridded population
Census population count data

Red = Urban Built
Orange = Non-Urban Built
Population Density Vietnam Administrative Unit Level 002
> 30 People Per Ha
0

> 30 People Per Ha
0
Settlement/building mapping from satellites
Satellite nightlights
Census population count data
Demographic data challenges

- Coarse resolution
- Outdated
- Incomplete
- Inaccuracies, missing populations
+ missing populations
+ inaccuracies
+ incomplete
+ inter-censal changes
Hybrid Methods to Complete Census Count

- **Complete National Coverage**
- **Partial Enumeration** e.g. Rolling Census
- **Enumerated Population in Limited Sampling Units Only**

**National Census/Projections/Estimates**
A  **Top down approach**

Census population counts

Geospatial covariates

Population disaggregation

Spatial weighting layer created based on covariates, using dasymetric mapping

Gridded population

---

B  **Bottom-up approach**

Microcensus population counts

Geospatial covariates

Population estimation

Prediction of population in unsurveyed areas based on covariates, using statistical modelling

Gridded population

---

Wardrop et al (2018) PNAS
‘Bottom-up’ population modelling approach

Population Data → Modelling Processes → Population Estimates

Settlement

Geospatial Covariates

Wardrop et al (2018) PNAS
Bottom-up population modelling

Population Data

→ Full enumeration of population within a geographically well-defined area

Settlement

Geospatial Covariates

→ Explanatory inputs: full coverage at the resolution of the final, gridded pop map
Settlements /Buildings + mean household size?
Capturing characteristics that determine variations in population density

Densities of schools, roads, market places, conflicts etc

Household sizes, regional groupings, poverty rates

Source: GAO.
Settlement/neighbourhood classifications

-Land use data; OpenStreetMap; Building heights; Mobile network data
Afghanistan example

Population Data

Settlement

Geospatial Covariates

Chamberlain et al (2019) in review
Afghanistan example

Population Data

Settlement

Geospatial Covariates

Number of compounds & area settled

Vegetation index

Night-time lights

Slope

Chamberlain et al (2019) in review
Afghanistan example

Chamberlain et al (2019) in review
Afghanistan example

Population model building and prediction

Population Data → Settlement → Geospatial Covariates

Chamberlain et al (2019) in review
Afghanistan example

Population model building and prediction

Population Data

Settlement

Geospatial Covariates

Uncertainty in EA 2017 Population Estimates

District Aggregation

log(Predicted population total) vs. log(Observed population total)

Chamberlain et al (2019) in review
Hybrid Methods to Complete Census Count

Complete national coverage enumerated population:
- limited sampling units only
- partial enumeration e.g. rolling census

Full Population Model

 enumerated population in limited sampling units only
National Census/Projections/Estimates

Hybrid Methods to Complete Census Count

Full Population Model

- Complete national coverage
- Partial enumeration e.g. rolling census
- Enumerated population in limited sampling units only
Microcensus Surveys

1,143 microcensus locations

14 of 37 States surveyed

Leasure et al (2019) in review
Hierarchical Bayesian model

- Admin boundaries: Region, State, LGA

School density
WorldPop Global
Household size
Settlement Type

- National Average population density
  - low
  - high

Average population density

Region
State
LGA
Settlement Type
Hierarchical Bayesian model

School density

WorldPop Global

Admin boundaries: Region, State, LGA

Household size

Settlement Type

Y

LGA

State

Region

Settlement Type

National

Average population density

low

high
Population estimate: 59 people
95% confidence: 8 - 182
Population estimate: 59 people
95% confidence: < 112 people
Using uncertainty

Requirement: vaccinate 90% of children under five years old

Our estimates indicate that this area in northern Nigeria (shown in blue) most likely contains 159 children under five years old. There is a 90% probability that no more than 212 children in this age group live here.

Leasure et al (2019) in review
Measles Vaccination Post-Campaign Coverage Survey, by EA – Nigeria, 2017-18

Proportion of Surveyed Children Vaccinated
- 0%
- >0% to <90%
- 90% to <95%
- ≥95%

Source: Nigeria MVC Post Campaign Coverage Survey, 2017-18

States basing microplans on modelled population data
National Census/Projections/Estimates

Hybrid Methods to Complete Census Count

Complete national coverage

Nigeria, DRC, Zambia, S.Sudan

Enumerated population in limited sampling units only

Full Population Model
# Capacity Strengthening - Training Pathways

<table>
<thead>
<tr>
<th>GRID3 Themes &gt;&gt;</th>
<th>Foundational Courses</th>
<th>Census / Microcensus</th>
<th>Grid3d Survey</th>
<th>Data Analysis and Visualisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager / Decision Maker</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Analyst / Statistician</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior Statistician</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GIS Analyst</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Census Cartographer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior Cartographer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey Planner</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey Implementer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GIS Trainer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education Researcher / Planner</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Researcher / Planner</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## GRID3 Courses >>

- Understanding the process of implementing mapping and grid systems
- Introduction to grid for development
- Monitoring and evaluation
- Spatial data storage and retrieval
- GIS support for decision making
- Advanced cartography
- Remote sensing
- Grid3d Survey Planning
- Grid3d Survey Analysis and Assessment
- Grid3d Survey and Assessment Support
- Grid3d Survey and Assessment Support for Monitoring and Evaluation
- Grid3d Survey and Assessment Support for Decision Making
- Grid3d Survey and Assessment Support for Spatial Data Storage and Retrieval
- Grid3d Survey and Assessment Support for Remote Sensing
- Grid3d Survey and Assessment Support for Advanced Cartography
- Grid3d Survey and Assessment Support for Monitoring and Evaluation
- Grid3d Survey and Assessment Support for Decision Making
- Grid3d Survey and Assessment Support for Spatial Data Storage and Retrieval
- Grid3d Survey and Assessment Support for Remote Sensing
- Grid3d Survey and Assessment Support for Advanced Cartography
- Grid3d Survey and Assessment Support for Monitoring and Evaluation
- Grid3d Survey and Assessment Support for Decision Making
- Grid3d Survey and Assessment Support for Spatial Data Storage and Retrieval
- Grid3d Survey and Assessment Support for Remote Sensing
- Grid3d Survey and Assessment Support for Advanced Cartography
- Grid3d Survey and Assessment Support for Monitoring and Evaluation
- Grid3d Survey and Assessment Support for Decision Making
- Grid3d Survey and Assessment Support for Spatial Data Storage and Retrieval
- Grid3d Survey and Assessment Support for Remote Sensing
- Grid3d Survey and Assessment Support for Advanced Cartography

## Example Use Cases

- Make informed decisions on national policies and priorities.
- Analyse the socio-economic impact of policies.
- Evaluate the effectiveness of interventions.
- Identify areas for improvement.

## Current population characteristics

- Increase the uptake of open source GIS tools.
- Integrate QGIS training to their respective organisations.
- Mapping school locations and catchment populations.
- Mapping pupil-teacher ratios for assessing the quality of education.
- Mapping health facility locations, catchment populations, travel times by modes of transport.
- Mapping ratios of births to live pregnancies, assessment of ANC provision of skilled birth attendants.

## Supporting Documents

- GRID3 Survey and Assessment Support for Decision Making.
- Grid3d Survey and Assessment Support for Planning and Evaluation.
- Grid3d Survey and Assessment Support for Decision Making.
Population estimation: Supporting the census process

Census data dissemination and analysis
- Filling gaps
- Producing gridded outputs
- Integration with other data

Inter-censal updates and estimates
- Capturing sub-national changes to inform inter-censal estimates

Survey/Census preparation
- Support for census planning
- Updating enumeration areas
- Defining sampling units
Considerations and limitations: population estimation modelling

• Not a replacement for a census!
• Still an area of active research – every setting is different
• Shows promise, but uncertainties can be large and accuracies low at fine spatial scales
• Are the areas being predicted into similar enough to those with the recent data used to build the models? Representative training and validation data is important – plus communicating assumptions
• Incomplete, inaccurate, outdated covariates
• Mobile populations, hidden populations, urban areas
• High level buy-in, close engagement with statistics office, need to be open on methods and limitations
• Communicating uncertainty
Some next steps

• Nigeria, DRC, Zambia, S.Sudan, Burkina Faso, Mali, Niger…..plus more
• Microcensus: Simulated population tests; Geolocated census tests; Household listings from surveys; Repeat surveys
• Models: Simulated population tests; Geolocated census tests; Alternative/hybrid models; Validation approaches
• Covariates: Residential/non-residential modelling; Building heights; Mobile network data; Neighbourhood mapping
• Supporting census/survey processes: Enumeration area delineation/update testing; Gridded population-based sampling
• Capacity strengthening: Population modelling workshops; User interface and tool development
Further information

www.worldpop.org
@WorldPopProject

www.grid3.org
@GRID3Global

E-mail: A.J.Tatem@soton.ac.uk
@AndyTatem