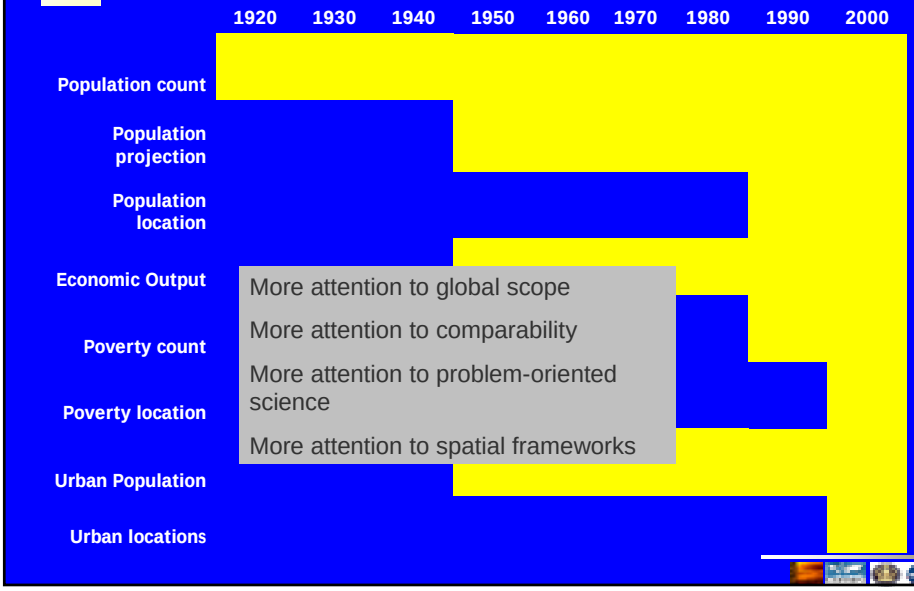
What's new (or not so new) in Population and Poverty Data Initiatives

Deborah Balk

Global Spatial Data and Information User Workshop

21 September 2004

Evolution in global collection of population and poverty data



Part I:

Where are the people? Efforts to spatially render population data

Haven't I seen that slide before?

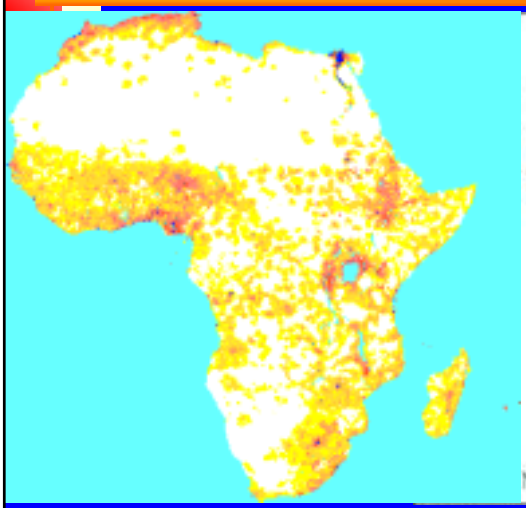
- BUCEN's CIR database
- Africa (UNEP/GRID, 1991)
- Global Demography Project (NCGIA & CIESIN, 1994)
- 1 degree global grid (Environment Canada, 1995)
- Europe (RIVM, 1995)
- Africa update and Asia (NCGIA, UNEP/GRID & WRI, 1996)
- Latin America (CIAT)
- LandScan (ORNL, 1999)
- GPW v2 (CIESIN et al, 2000)

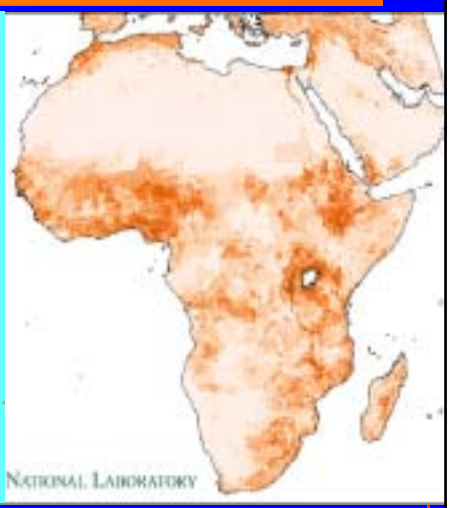
Ten Years of Progress, CIESIN et al.'s Gridded Population of the World (GPW)

Version (pub)	GPW v1 (1995)	GPW v2 (2000)	GPW v3 (2003)
Estimates for	1994	1990, 1995	1990, 1995, 2000
Input units	19,000	127,000	~ 375,000

http://beta.sedac.ciesin.columbia.edu/gpw/

Modeling Efforts to Georeference Population

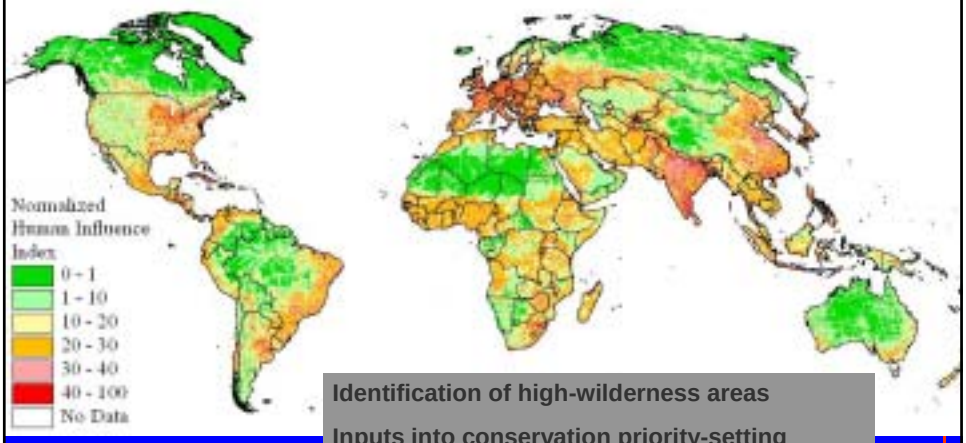




Accessibility Model
UNEP, CIAT, WRI, & NCGIA, 1996+

Landscan
ORNL 1999+

Human Footprint
(Wildlife Conservation Society and CIESIN, 2002)

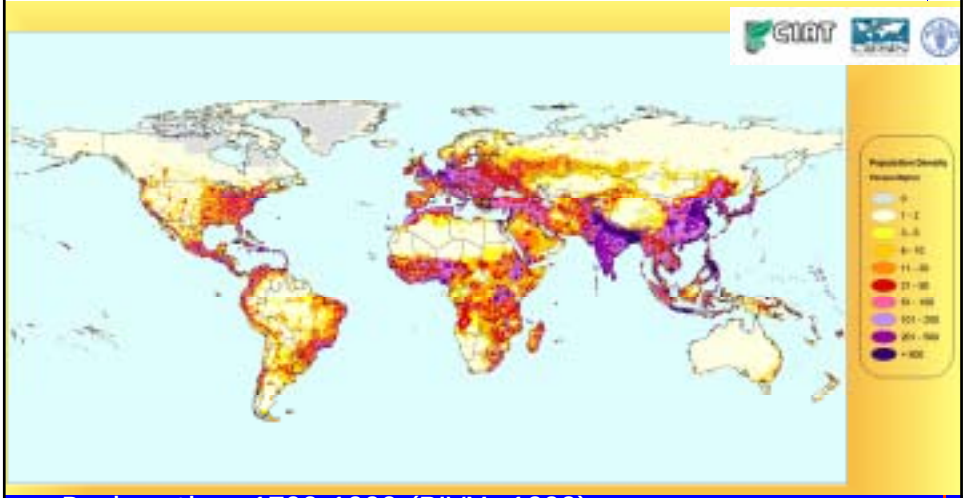


Identification of high-wilderness areas

Inputs into conservation priority-setting

Useful predictor of extinction threats

Population Projections



- Backcasting: 1700-1990 (RIVM, 1999)
- Forecasting: to 2015 (CIESIN, FAO, & CIAT, 2004)

Underlying Data

- Mandatory
  - Population counts
  - Administrative boundaries
- Modeling also requires some but not necessarily all of these:
  - Urban areas (GRUMP, Access)
  - Roads (Access, LandScan, Footprint)
  - Elevation (LandScan)
  - Slope (LandScan)
  - Land cover (LandScan, Footprint)
  - Assumptions about growth (Projections)

Differences reveal need for more mature institutional coordination

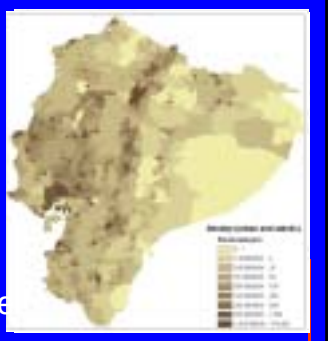
- Current data users must choose between two extremes
  - Highly pristine raw data
  - Highly modeled processed data
- Difficulties
  - getting “under the hood”
  - creating custom aggregations or modeled outputs
  - sharing data inputs such that a range of outputs could be supplied: mix and match

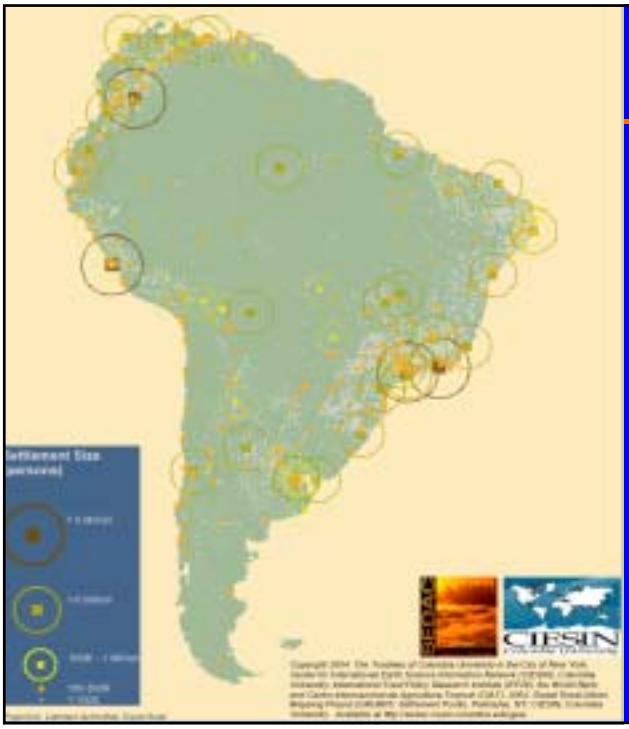
Part II:

Where are the cities?
How are cities connected?

Global Urban-Rural Mapping Project (GRUMP)

- Objective: To met an unmet need to delineate urban and rural extents
  - Collaboration between CIESIN, IFPRI, World Bank, & CIAT
  - Builds on GPW infrastructure, adds urban areas from satellite data
    - With a massive additional data collection and processing effort
- Three databases:
  - Settlement Points (over 70,000)
  - Urban Extents (over 23,500, w/pop of 5k+)
  - Pop Grid reallocated to urban areas at 1 km
- Facilitating new and exciting analysis
- Alpha testing of data now underway
- Data delivered through GPW web service

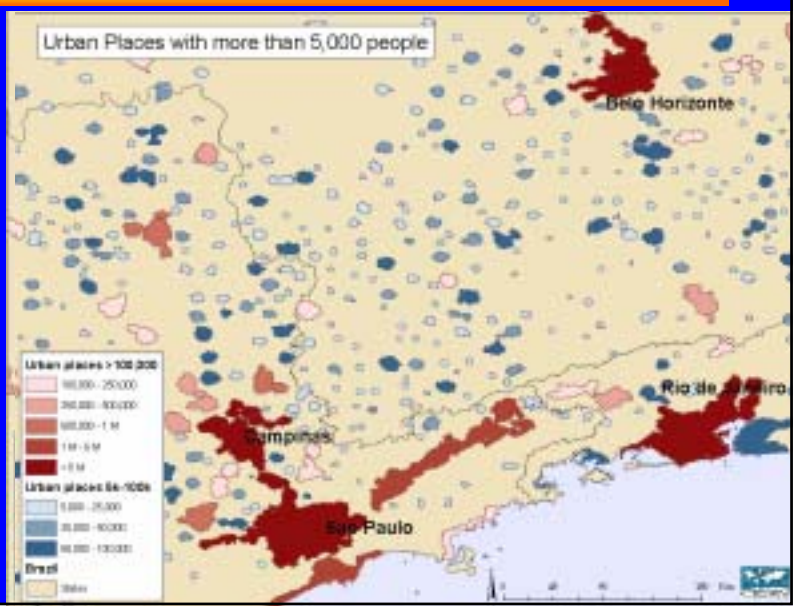




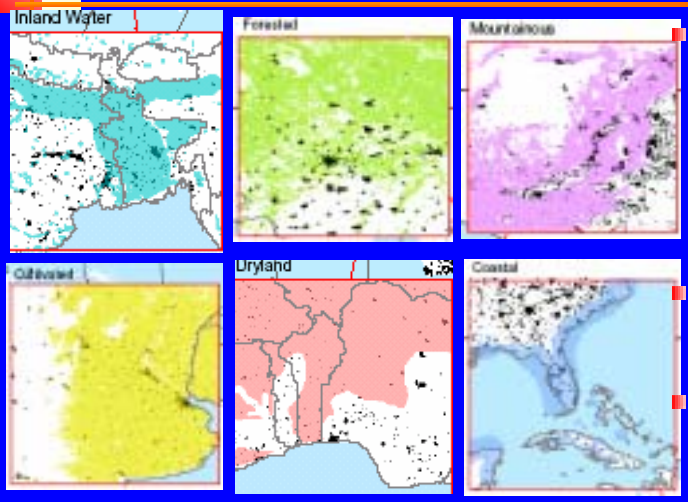
- Settlements
  - No lower limit in data collection
- By increasing the settlement threshold, a vastly greater number of urban areas are revealed

Understanding the spatial dimension

- Close up of Brazil using the 100K person cut off
- Note the variety of shape
  - Much more than points convey
- Threshold at 5k



Urban area overlays with ecosystems

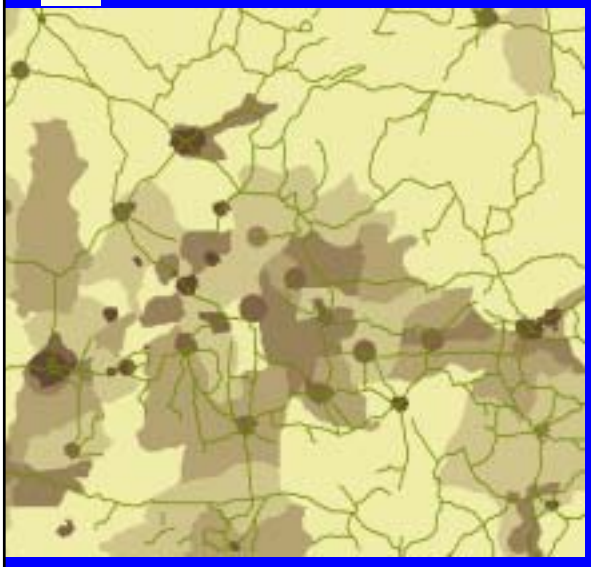


- Coastal systems are disproportionately urban (about 65%)
  - Sustain the highest global pop densities in both urban and rural areas
- Coastal land area is disproportionately urban, around 10%
- Not so urban:
  - Mountain
  - Forest
  - Dryland
    - Except locally

GRUMP's strengths & weakness

- Strengths
  - Combines existing data in new ways
    - Populated place and administrative area
    - Urban extents as derived from night-time lights satellite
  - First-ever globally consistent population data of urban areas
- Weaknesses
  - Data streams less vested
    - Cities database inadequate below 100,000 persons
  - Lights not designed to indicate urban areas
    - So they work less well in places with less light (e.g., Africa)

Roads connect the cities



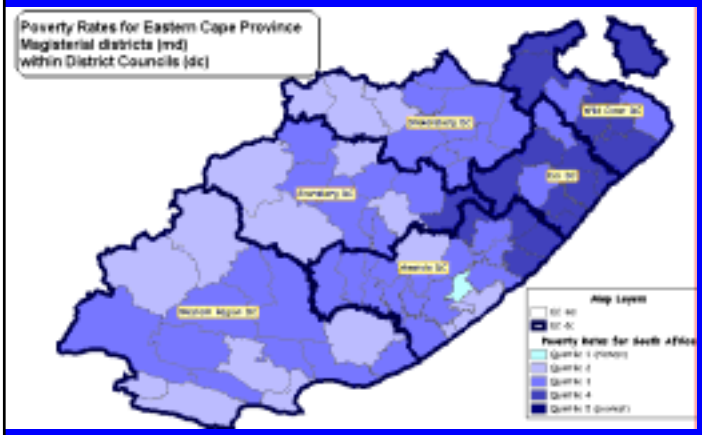
- Applications:
  - Validation of GRUMP urban extents
    - Green lines are VMAP roads; Good match for places found
  - Updated Africa and Latin American roads data for accessibility models
- Unmet need:
  - Everybody wants roads data!
  - Makes coordination on population data look like a cake-walk

Part III:

Where are the poor?

Country-studies of poverty

- Small area estimates for about 25 countries
  - Uses econometric methods to integrate surveys + census data
  - Some time series



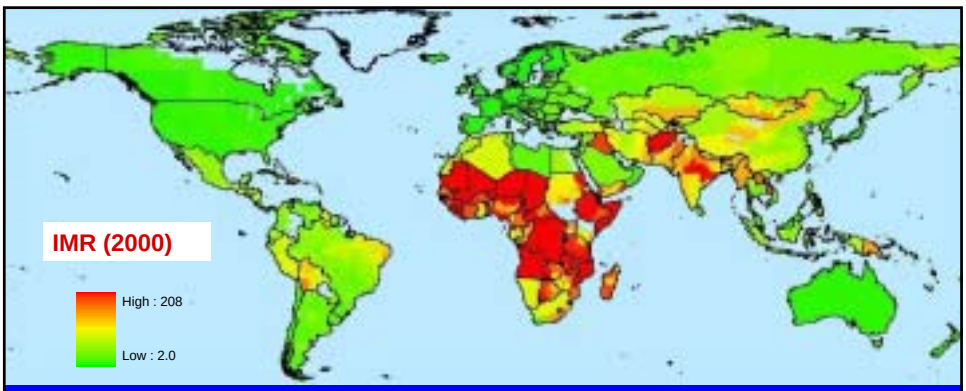
- Poverty Reduction Strategy Programme: country engagement
- Highly disaggregated maps
- Limited distribution

Case studies

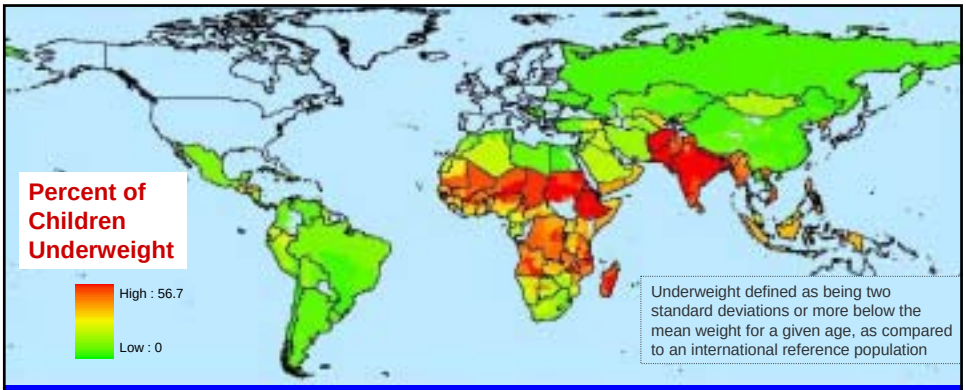
- www.povertymap.net
- **FAO, UNEP and the CGIAR**
- Nine case studies using new data and methods in
  - Mexico, Ecuador, Nigeria, Malawi, Kenya, Sri Lanka, & Bangladesh
- Lots of national-level maps on poverty-related information
  - from survey data, e.g., access to safe water, mortality, etc
  - from national accounts, e.g. GNP
  - some temporal change
- Compendium of efforts (WRI)

Regional & global approached

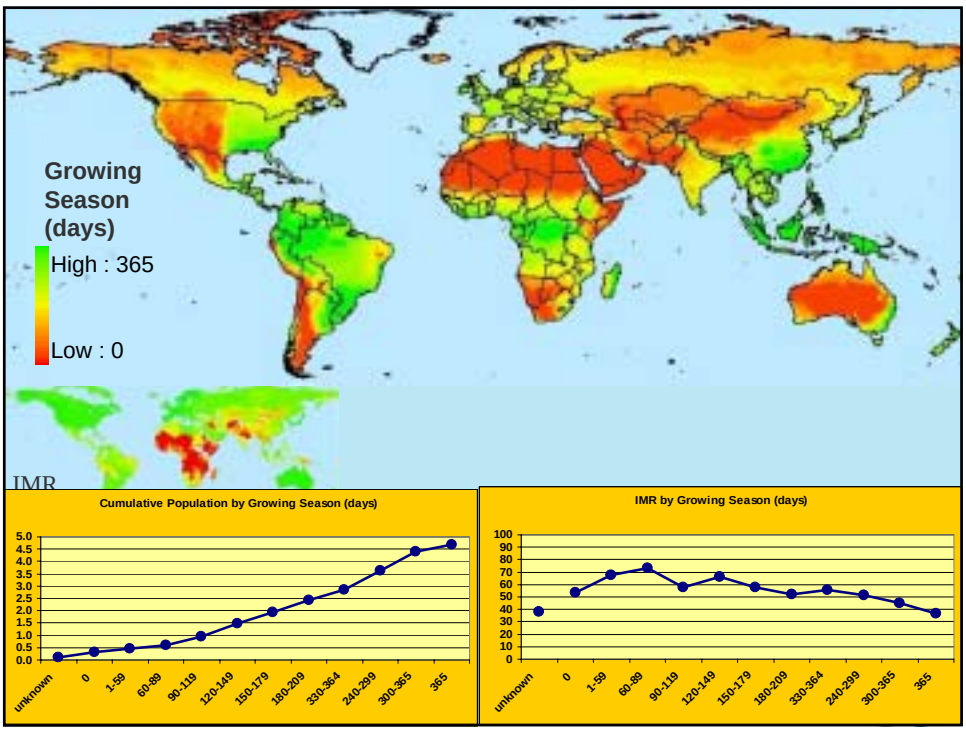
- Many of these case studies and the SAE analyses produce direct estimates of poverty but no global effort does, therefore, the search for reasonable global proxies.
- CIESIN, in support of the UN Millennium Project, began using a variety of data—mostly from standard household surveys—to construct databases of poverty correlates:
  - Outputs: Infant Mortality and Child Nutrition
    - Requires large scale conversation of surveys to subnational spatial units
  - Inputs: Biophysical parameters, infrastructure



- Sources
  - Demographic and Health Surveys (41 countries)
  - Multiple Indicator Cluster Surveys (5 countries)
  - National Human Development Reports (14 countries)
  - National Statistical Offices (16 countries)
  - UNICEF Childinfo – (115 countries)
- Subnational representation
  - 8,029 units (6,886 in Brazil and Mexico alone)
  - 77 countries have subnational data; 115 national only
  - 80% of world population has subnational data
  - Average 14 units per country (outside Brazil and Mexico)
- Converting rates to counts
  - For each subnational unit, estimates of live births, infant deaths calculated based on *gridded* population, *national* fertility data, and *subnational* IMR.
- Calibration
  - Subnational IMR values adjusted to be consistent with national UNICEF 2000 IMR values

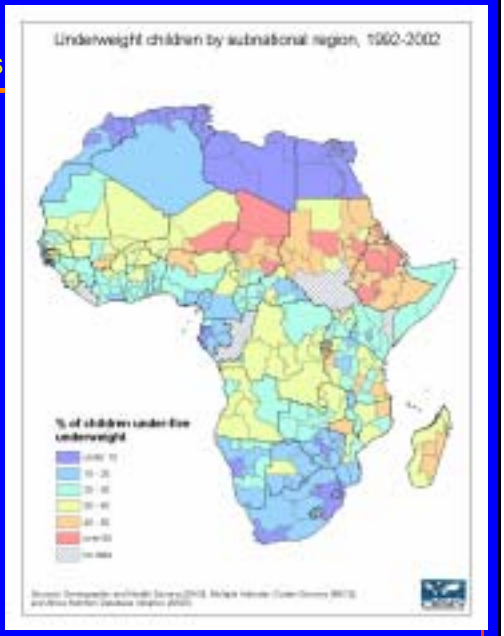


- Sources
  - Demographic and Health Surveys (46 countries)
  - Multiple Indicator Cluster Surveys (27 countries)
  - National Human Development Reports (3 countries)
  - African Nutrition Database Initiative (2 countries)
  - UNICEF ChildInfo – (54 countries)
- Subnational representation
  - 678 units
  - 73 countries have subnational data; 54 national only
  - 65% of world population has subnational data
  - Average 9 units per country
- Converting rates to counts
  - For each subnational unit, estimates of under-5 children, underweight children calculated based on *gridded* population, *national* age structure, and *subnational* underweight percentage.
- Calibration issues
  - Data come from multiple years, and there are no independent series that permit adjustment to a common year

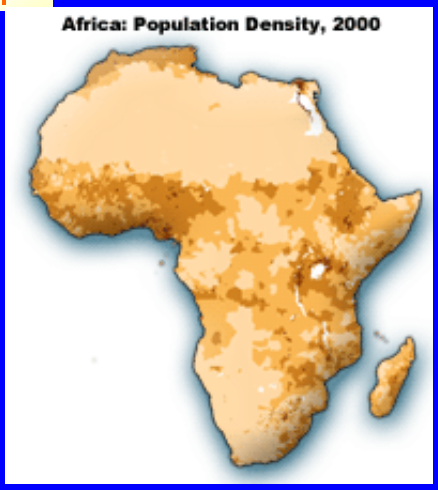


Use of malnutrition data to guide problem-focused diagnosis

- Match survey to boundary data
  - Survey sources: DHS, MICS, ANDI
  - Sub-national units created at finest resolution for which data are statistically robust
- 382 sub-national units (SNUs)

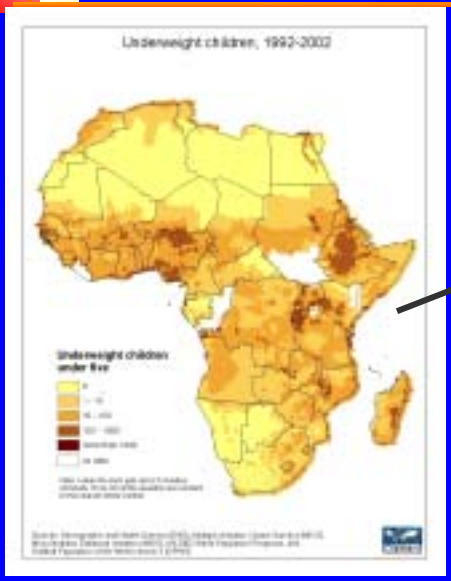


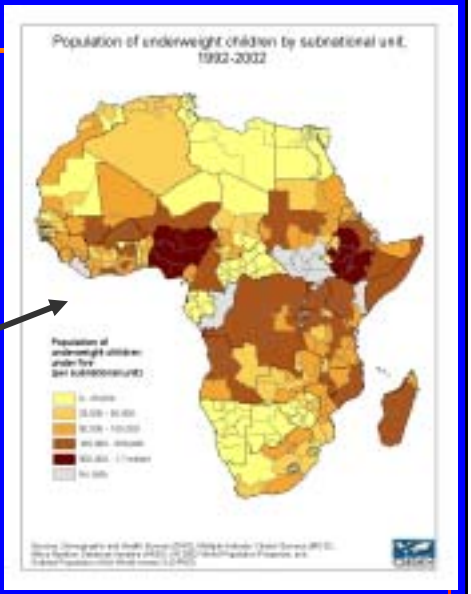
Step 2



X

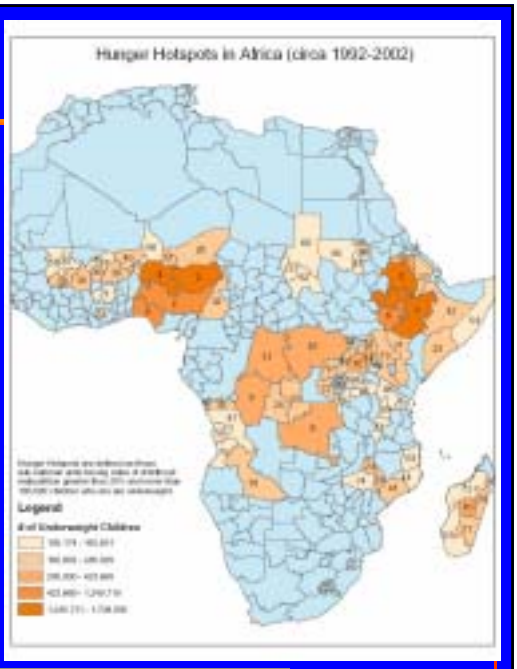
Step 3





Step 4

- Selection of "hotspots": Defined by the Hunger TF as those sub-national units with rates of childhood malnutrition >20% and more than 100,000 children who are underweight
- 75 sub-national units met this criteria



See http://www.unmillenniumproject.org/html/tforce_2.shtm

Weaknesses

- Priority setting is poorly coordinated
  - Extents, theme, proxies, methods?
- Open access is imperfectly realized
  - How to share SAE data, for example?
- Cross-scale coordination is suboptimally implemented
  - Comparability of methods largely unknown
  - WRI/Bank recent construction of poverty mapping compendium is the first attempt to share information
- Resources are inefficiently allocated
  - Duplication of efforts (e.g., multiple child nutrition maps) to be avoided
- Overall, the socioeconomic global data landscape is extremely patchy

Score card: how are we doing?

	Global Extent	Data Quality	Data Availability	Institutional Cooperation
Population	☺☺☺	☺☺☺	☺☺☺	☺☺😐
Boundaries	☺☺	☺☺	☺☹	☺☺
Urban Areas	☺☺☺	☺☺	☺☺	☺☺
Roads	☺☺☺	☺	☹	☹
Poverty	☺	☺☺☺	☹	😐

Two paths forward

- Status quo, reintensified
  - Stove piping by disciplines and bureaucracies
  - Inadequate integration
  - Failed opportunities for cross-fertilization
  - Redundancy and unproductive competition
- Quantum increase in cooperation & institutionalization
  - Coordination between groups
    - **Pool input data sources**
    - **Division of tasks**
    - **Guidelines on common products, coding, formats, projects**
    - **Get endorsement from National Statistical Offices and UN agencies**
    - **Productive interaction on methodological hurdles**
  - More relevant, efficient & timely outputs