

Potential for Crowdsourced Data for Population Distribution Models and Databases

Presented at
**UN-SPIDER Expert Meeting on
Crowdsourcing for Disaster Management
and Emergency Response**

Budhendra Bhaduri
Corporate Research Fellow

**December 05, 2012
Vienna, Austria**

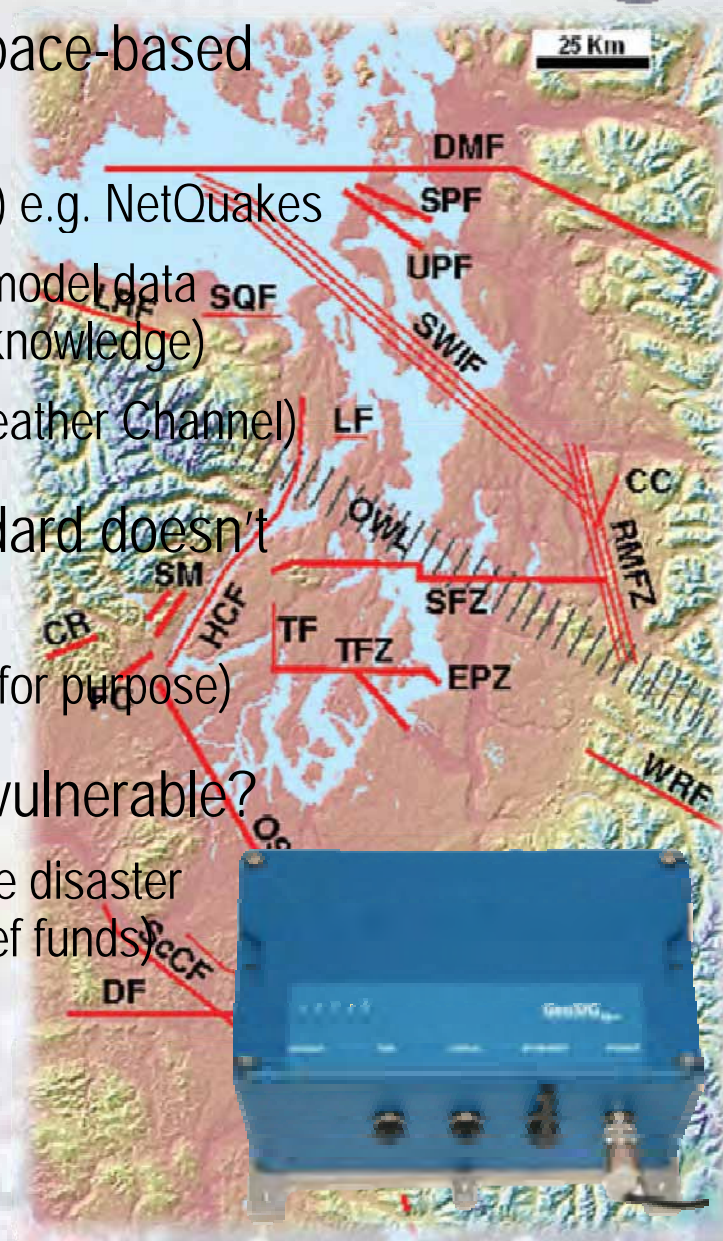


Crowdsourcing: Points to ponder

Geographic Information Science and Technology

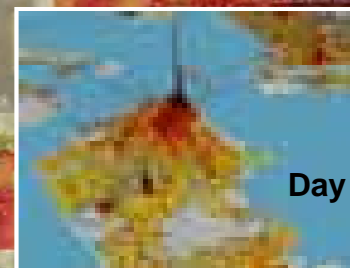
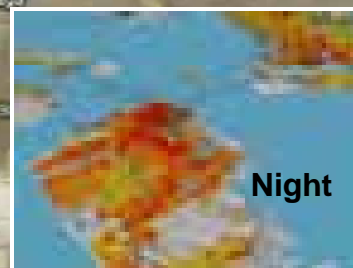
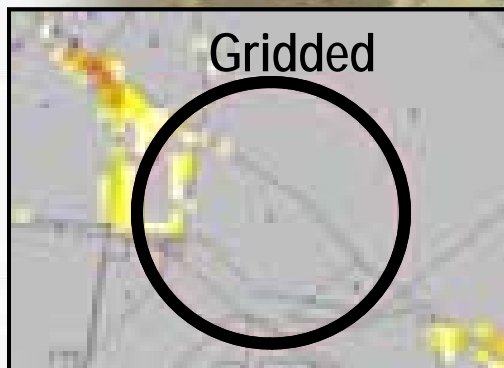
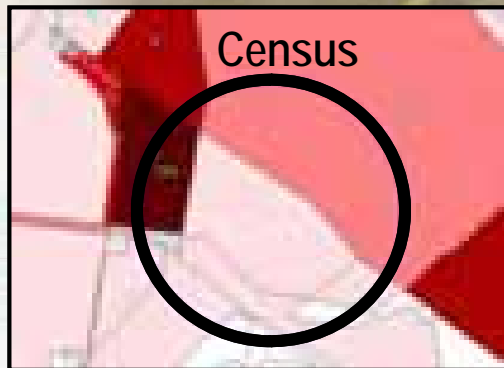


- Crowdsourced information clearly augment space-based data
 - Increase density and resolution of data (Gap filling) e.g. NetQuakes
 - Enhance currency and quality of observation and model data (incidence report, damage qualification, and local knowledge)
 - The media makes great use of it (CNN iReport, Weather Channel)
- Traditional top-down spatial data quality standard doesn't work
 - When's good is good enough (user defined and fit for purpose)
- When does crowdsourcing make the system vulnerable?
 - Reliability of the crowd and crowd fatigue (are there disaster magnitude and frequency thresholds similar to relief funds)
 - Digital divide, victim crowd, and system overuse
 - Social, legal, and ethical concerns

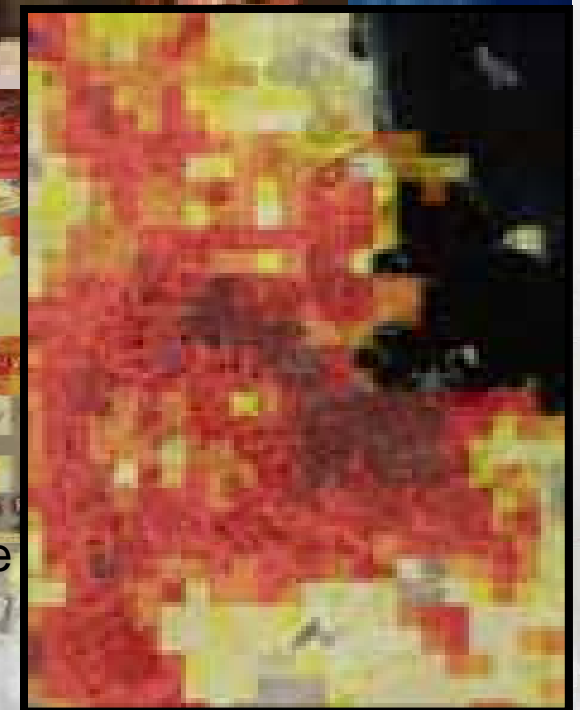


LandScan Population Distribution and Dynamics Model and Database

LandScan Global



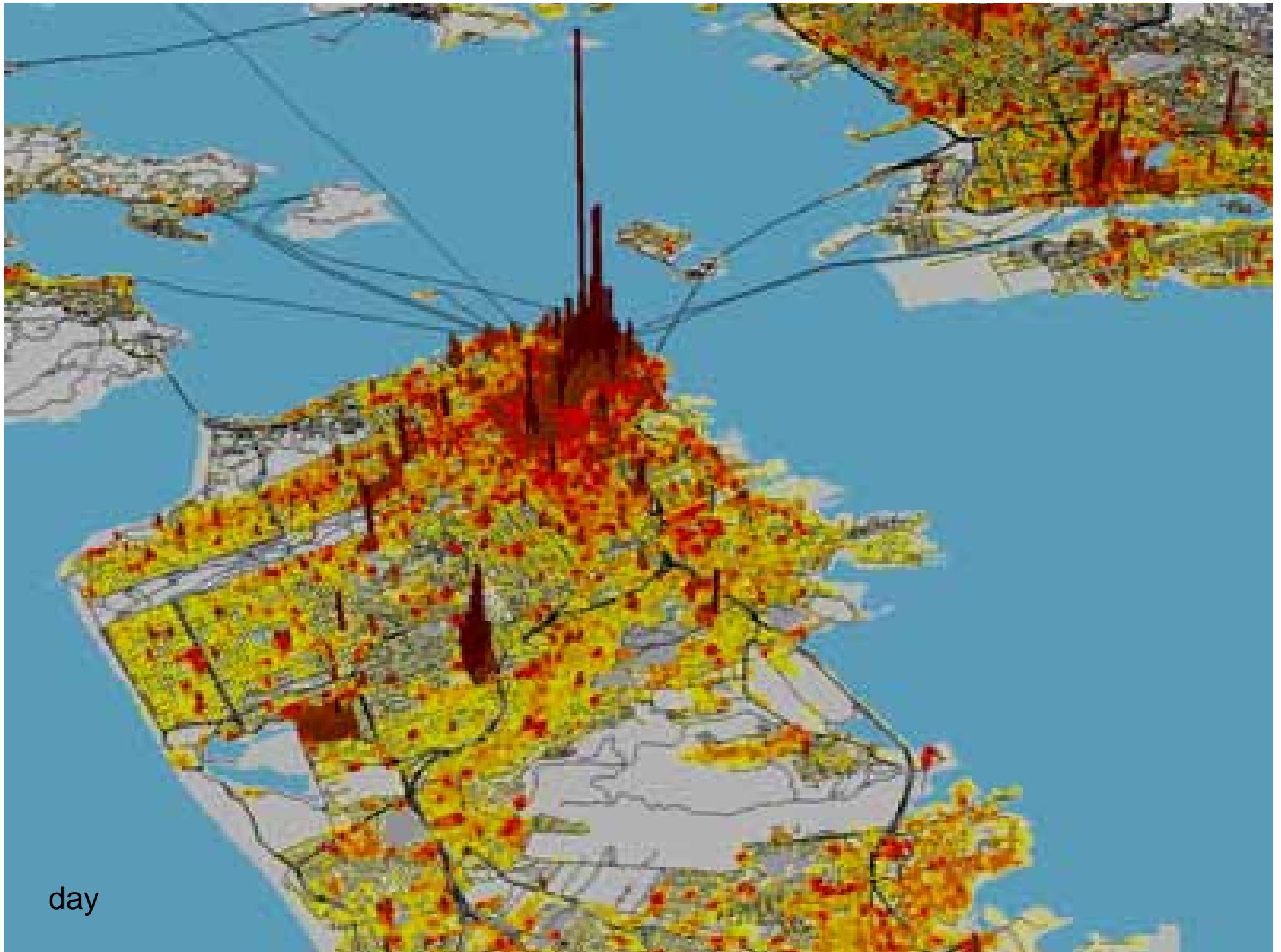
LandScan USA



As the finest population distribution data available for the world and the US, LandScan Global and LandScan USA are the community standard for estimating population at risk

Managed by UT-Battelle
for the Department of Energy



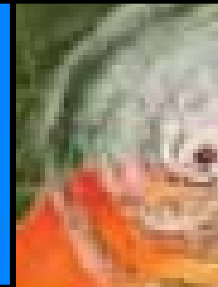
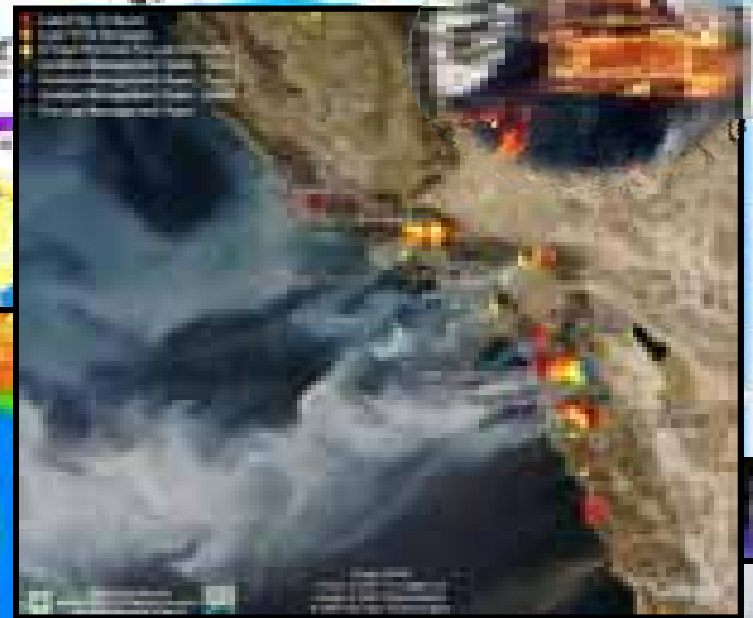


Disasters make population data obsolete

Geographic Information Science and Technology

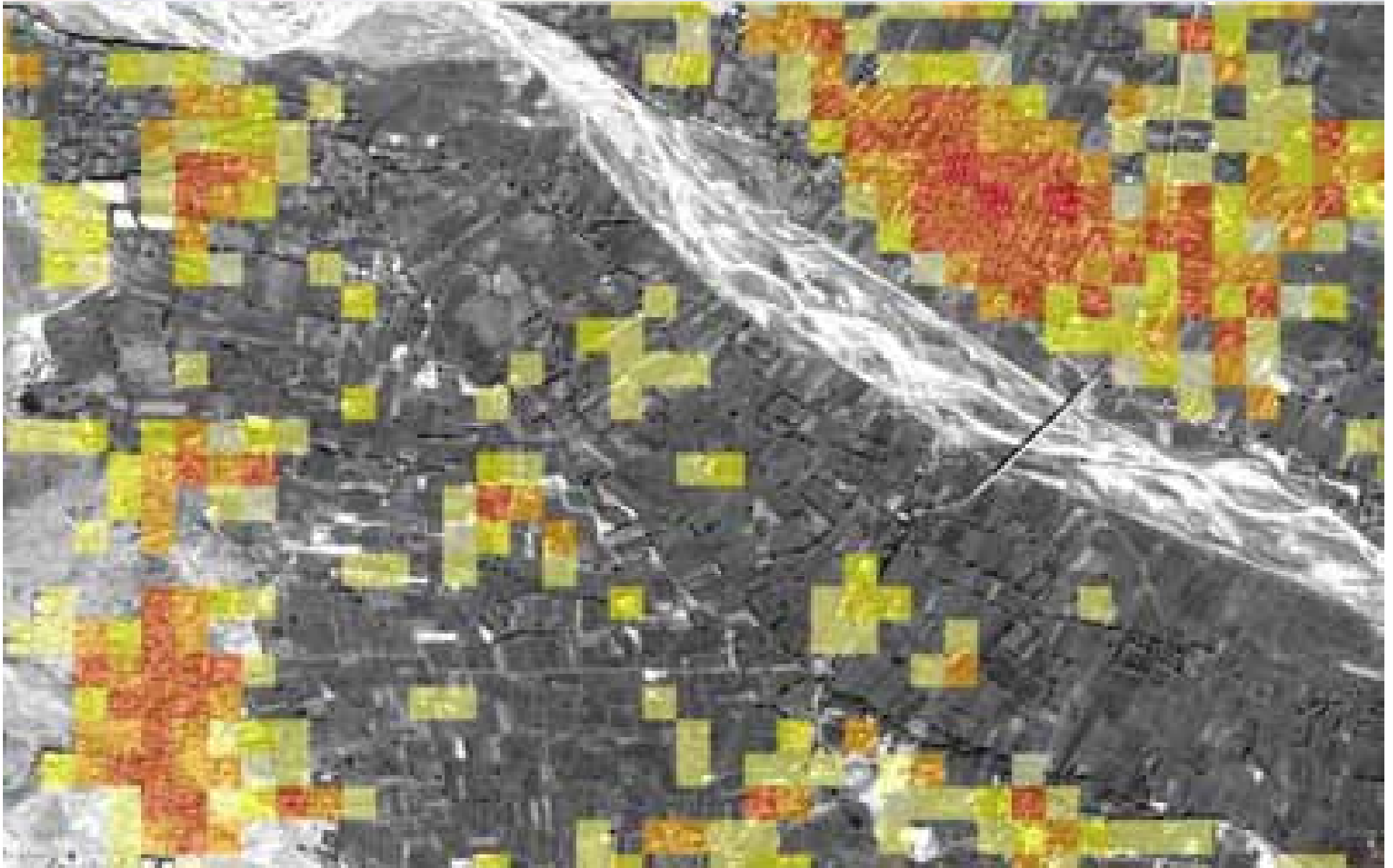
- Loss and dispersion of population
- Capturing population redistribution is critical at many time scales
 - Earthquake aftershocks (minutes-days)
 - Hurricanes (weeks to months)
 - Sea level rise (years to decades)
- Space based observation only interprets land cover
 - Flood damage is often difficult to detect for structures
- Is crowdsourcing a strategy?
 - Active (including self disclosure)
 - Passive (social media)

Managed by UT-Battelle
for the Department of Energy



Spatial refinement of LandScan Global

Geographic Information Science and Technology



Addis Ababa, Ethiopia

- 2 Xeon Quad core 2.4GHz CPUs + 4 Tesla GPUs + 48GB
- Image analyzed (0.3m)
 - 800 sq. Km
 - RGB bands
- Overall accuracy 93%
 - Settlement class 89%
 - Non-settlement class 94%
- Total processing time
 - 27 seconds

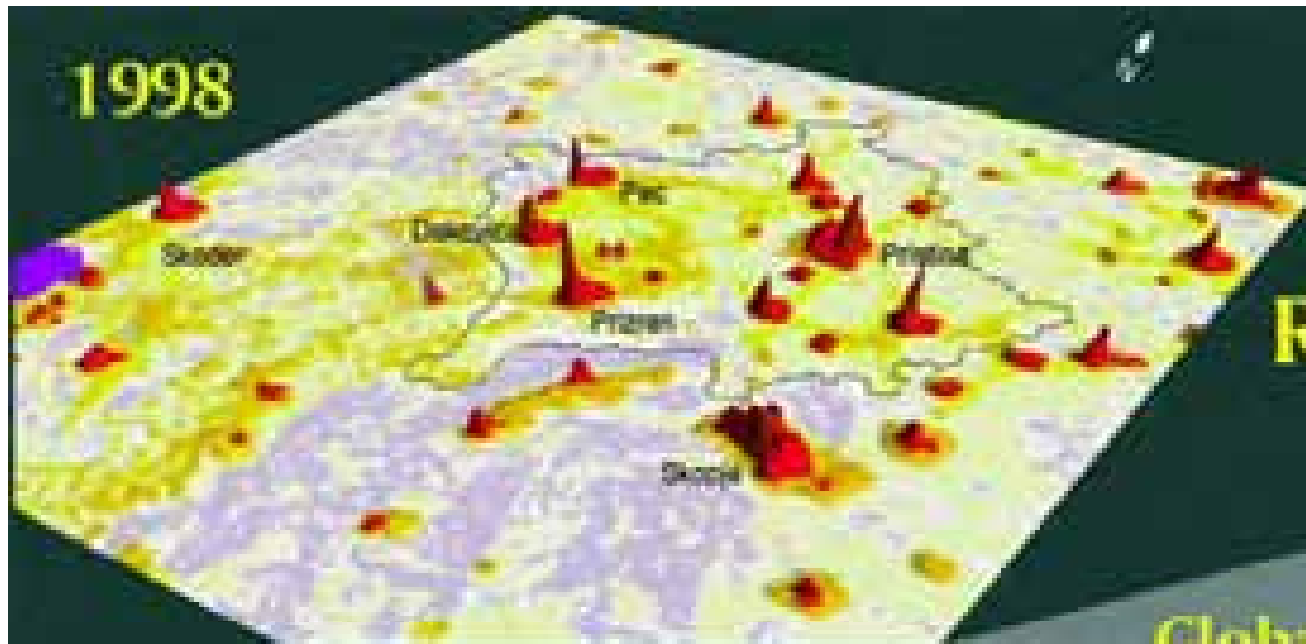


1000

100



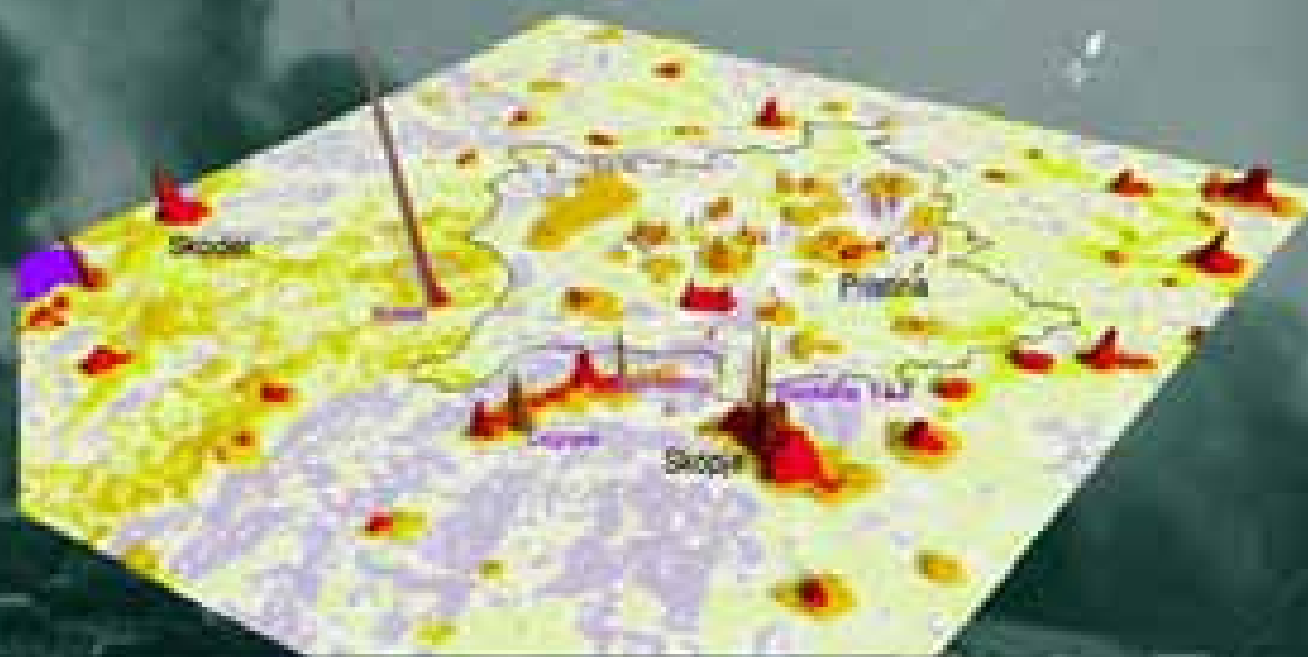
1998



Kosovo Refugee Crisis

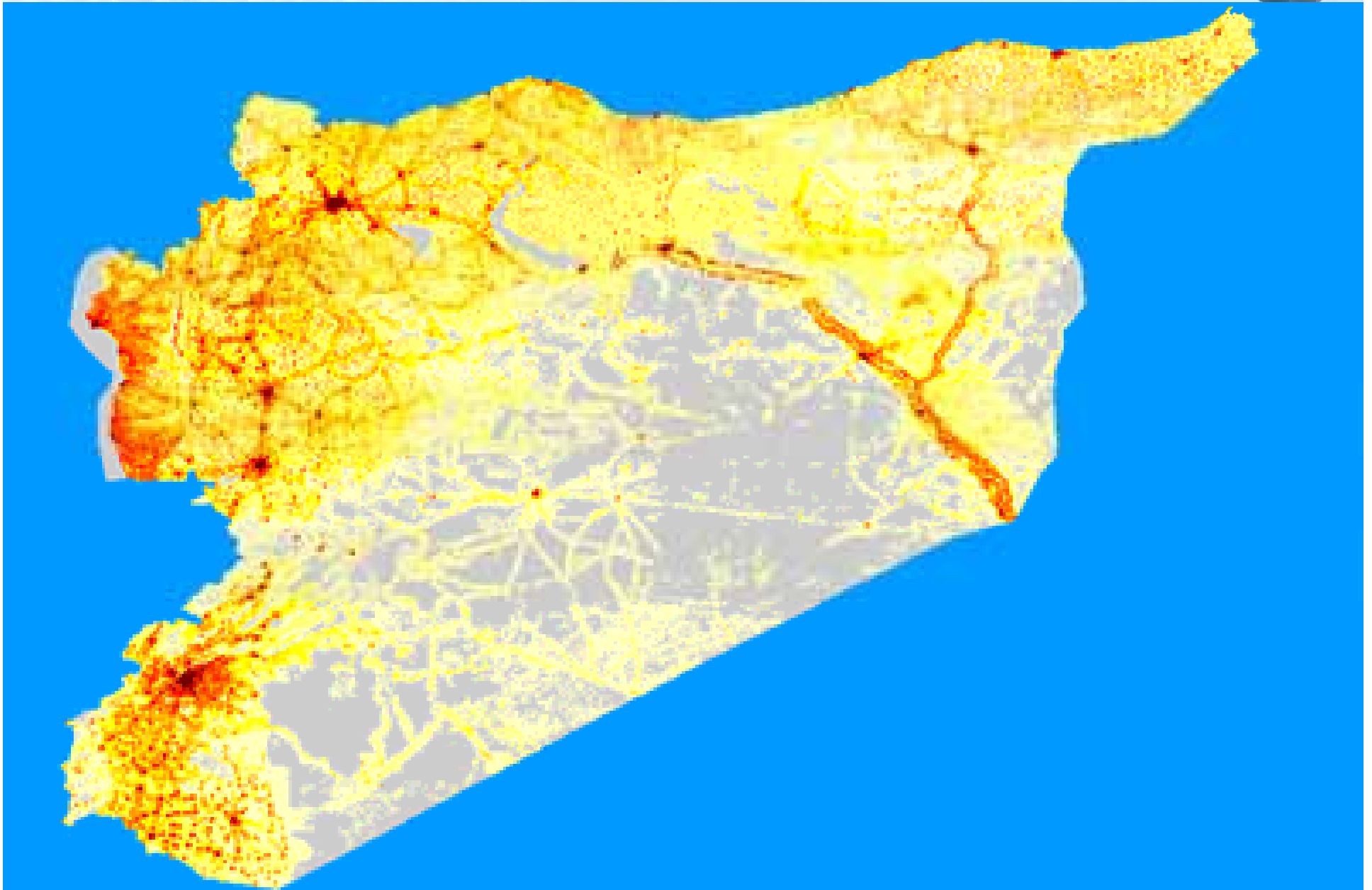
LandScan
Global Population Project
Oak Ridge National Laboratory

May 25, 1999



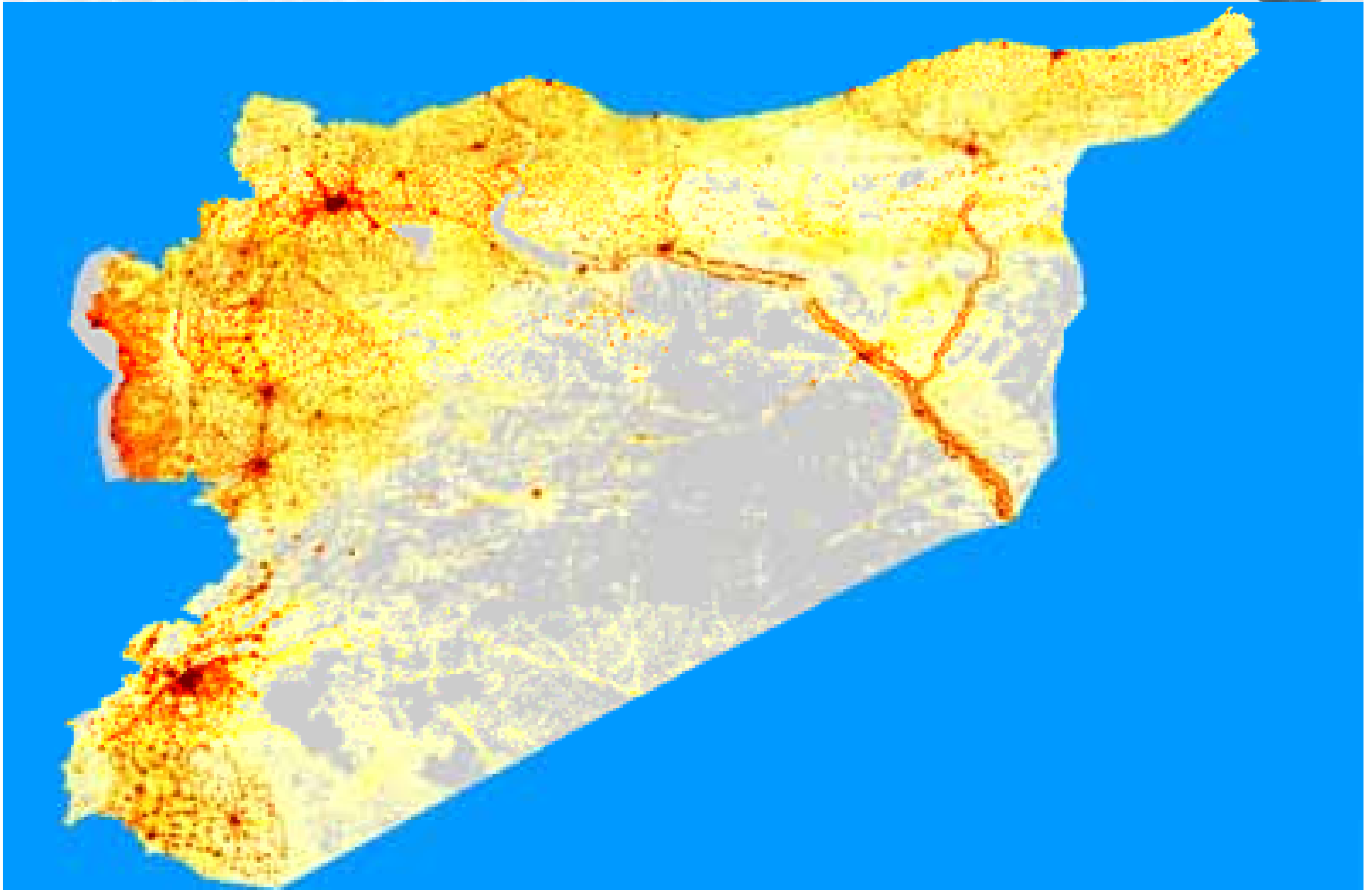
Syria: LandScan 2011

Geographic Information Science and Technology

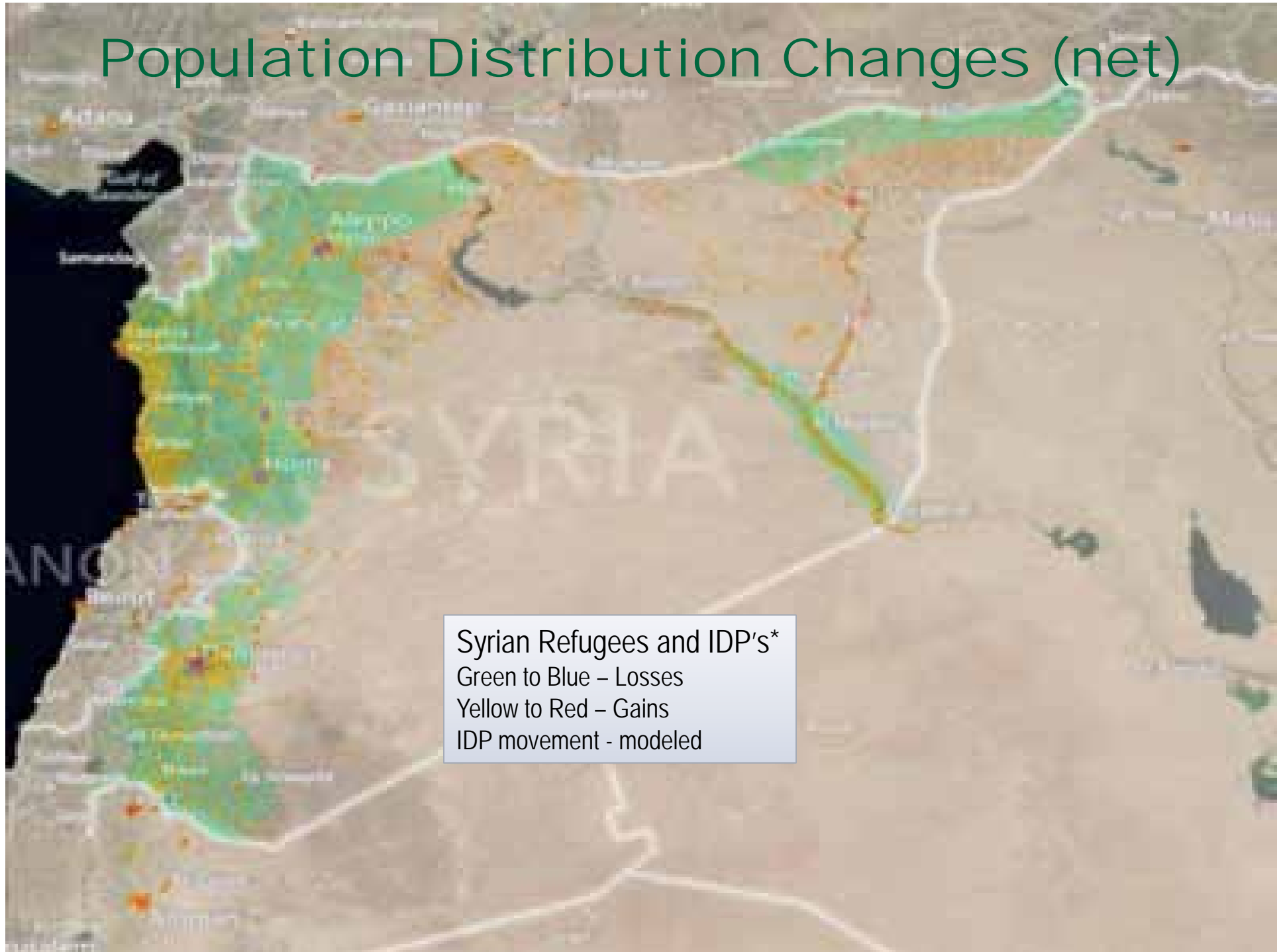


Syria: LandScan Aug.31, 2012

Geographic Information Science and Technology



Population Distribution Changes (net)



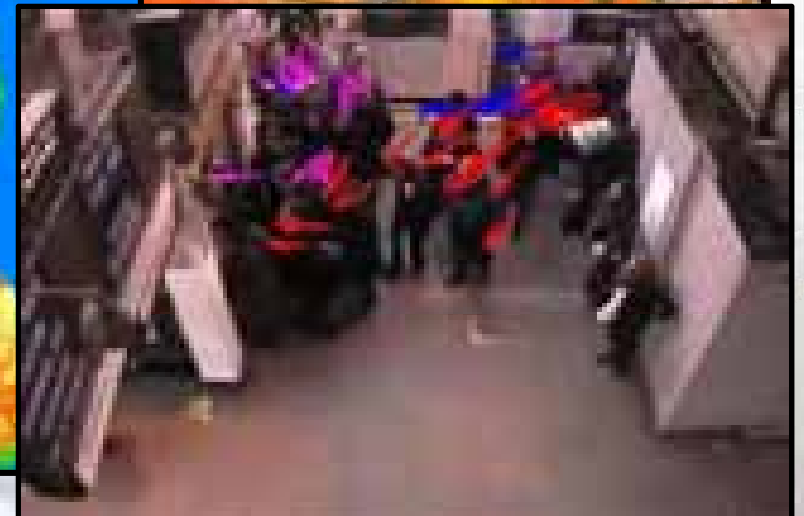
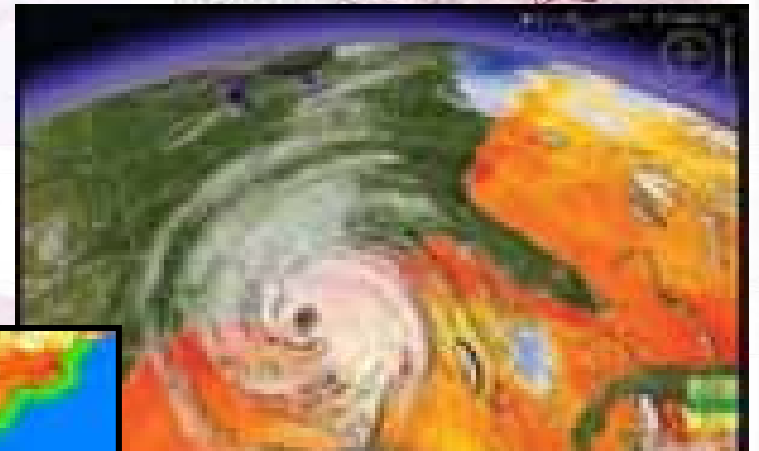
Assessing Population Dynamics

Geographic Information Science and Technology

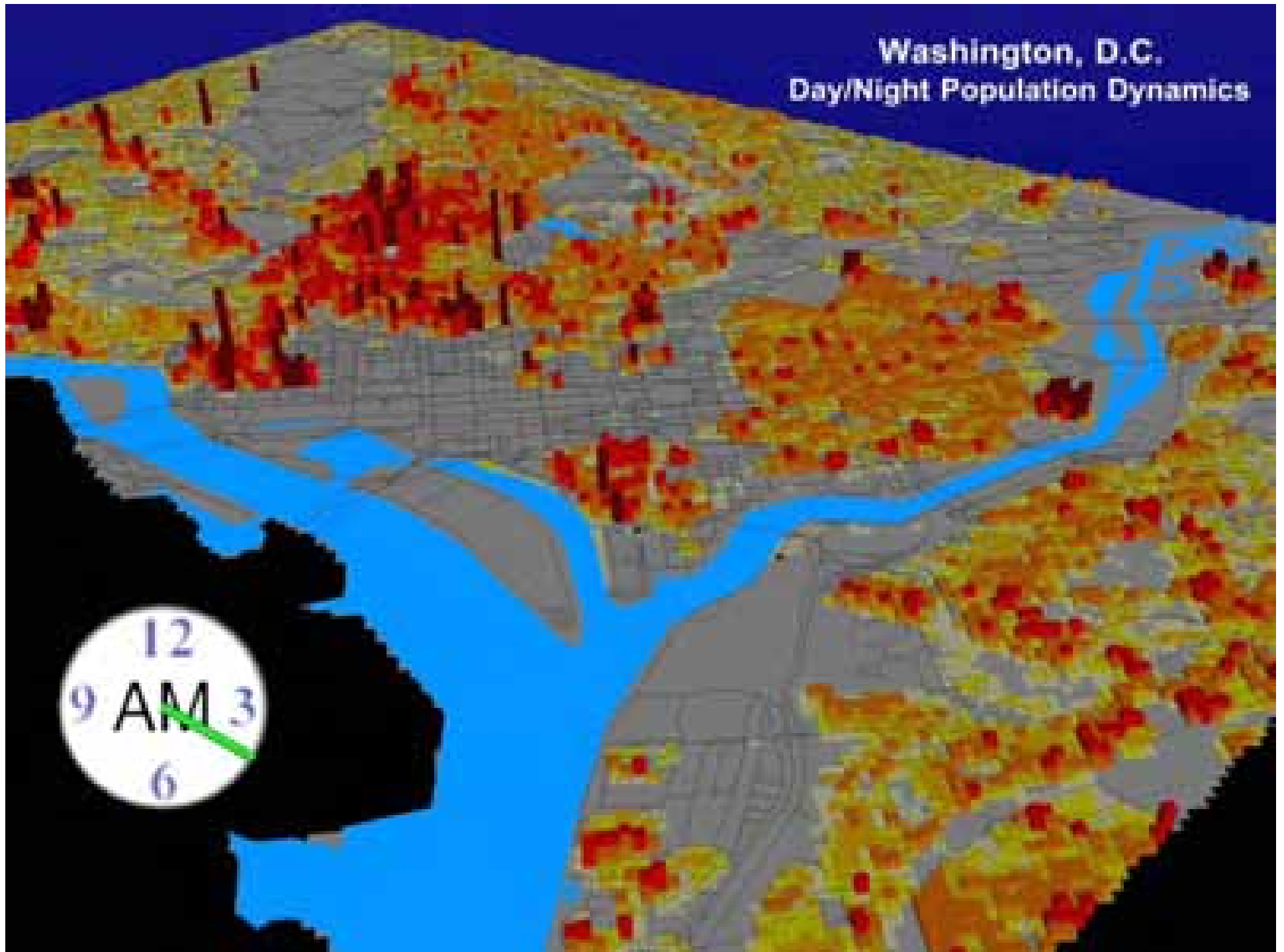


- Dynamic tracking of people and vehicle fleet movement from streaming multisensor data
 - Video, cell phones, social media
- Sociocultural input
 - Accounting for refugees through remote sensing is often challenging
- Migration can be captured but the challenge is circular migration (resettlement)
 - Information flow and media coverage significantly drops with time
 - Where do they come back from?

Managed by UT-Battelle
for the Department of Energy



Washington, D.C.
Day/Night Population Dynamics



Real Time Rome: MIT Senseable City Lab

Geographic Information Science and Technology



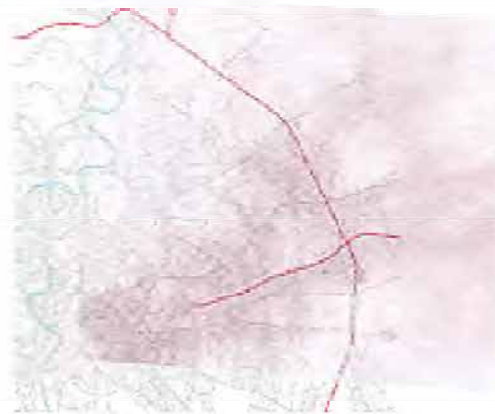
00 - 00 AM



04 - 00 AM



08 - 00 AM



12 - 00 PM

“.....Ratti's team obtains its data anonymously from cell phones, GPS devices on buses and taxis, and other wireless mobile devices, using advanced algorithms developed by Telecom Italia, the principal sponsor of the project. These algorithms are able to discern the difference between, say, a mobile phone signal from a user who is stuck in traffic and one that is sitting in the pocket of a pedestrian wandering down the street. Data are made anonymous and aggregated from the beginning, so there are no implications for individual privacy.”

Social networking and self disclosure

Geographic Information Science and Technology



- "Latitude" is being marketed as a tool that could help parents keep tabs on their children's locations, but it can be used for anyone to find anyone else, assuming permission is given."
- "...allow you to share that location with friends and family members, and likewise be able to see friends and family members' locations"
- "To protect privacy, Google specifically requires people to sign up for the service. People can share their precise location, the city they're in, or nothing at all."



Social networking and self disclosure

Geographic Information Science and Technology



- “Foursquare just snagged its six millionth member...”
- “Foursquare, the social network that allows members to communicate with acquaintances by “checking in” to locations they patronize, is breaking with its own traditions by allowing users to “check in” to the Super Bowl even if they’re not attending the game in person.”



Challenges and Opportunities

Geographic Information Science and Technology



Social

- Recruiting the crowd could benefit from high-profile volunteer catalysts
- The crowd may not be aware of engagement opportunities
- Success may be locally variable because of cultural differences

Legal

- Expectation of privacy is a variable standard
- Legal standards are not clearly defined and understood
- Self disclosure could be an effective way to address privacy

Ethical

- Does this involve deceptive principles (instrumenting national parks, GPS and battery life)?
- Should we promote the crowd as only volunteers?
- Self disclosure may come with expectations of service guarantee

We must have a strategy for data curation



Geographic Information Science and Technology



Man
for u

Courtesy: Eric Young

Additional Resources

Geographic Information Science and Technology



- **Role of Volunteer Geographic Information in Advancing Science: Quality and Credibility**
 - GIScience 2010 and 2012 workshops
 - <http://www.ornl.gov/sci/gist/workshops/2012/index.shtml>
 - Springer eBook forthcoming in 2013
- Hurricane Sandy image courtesy: Eric Young (Penn State)
- Contact: bhaduribl@ornl.gov