Validation and Assimilation of OSM for the Global Roads Open Access Data Set

Alex de Sherbinin & Paola Kim-Blanco

CIESIN, Columbia University
CODATA Global Roads Data Development Working Group

State of the Map US – June 2015

Center for International Earth Science Information Network
EARTH INSTITUTE | COLUMBIA UNIVERSITY

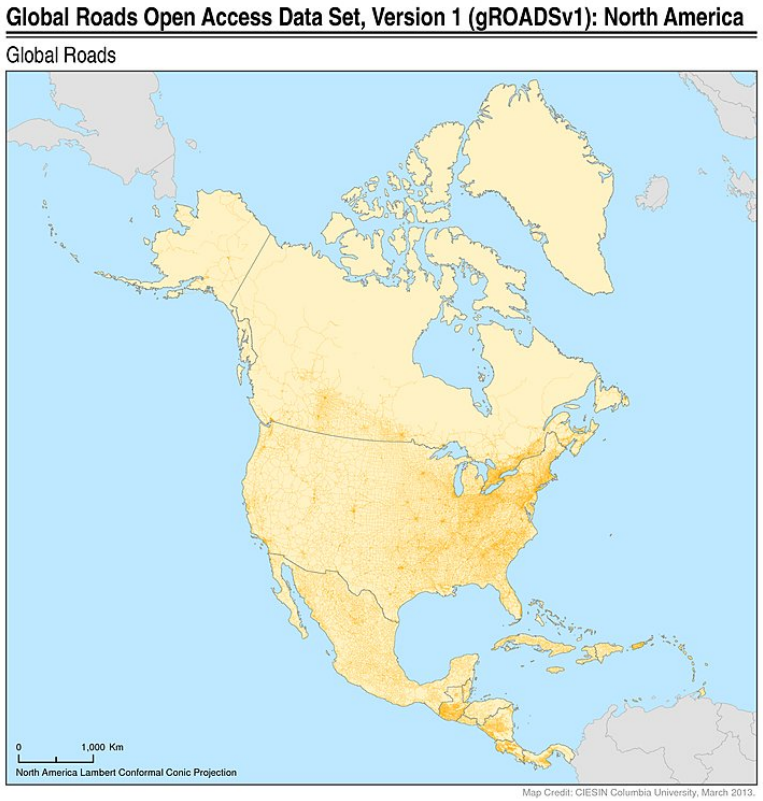
ICSU CODATA

CODATA

Summary

- OSM data are increasingly **viewed as authoritative** by development / humanitarian agencies to carry out relief and development work
- However, a question that often arises is, **how complete are the data in OSM for country X?**
- To help answer this question, we are collaborating with OSM to **assess road length, spatial coverage, attributes, and the geometry of OSM data**
- Goal is to **assimilate** OSM data into Global Roads Open Access Data Set (gROADS), version 2

Global Roads Open Access Data Set (gROADS)



- Goal: To develop a global roads open access data set (gROADS) that is:
  1. fit for modeling / research purposes (not navigation)
  2. Based on a globally consistent model (UNSDI-T)
  3. spatially accurate (~50m positional accuracy)
  4. topologically integrated
  5. focused on roads between settlements (not streets)
  6. well documented
  7. freely distributed (on attribution only basis)

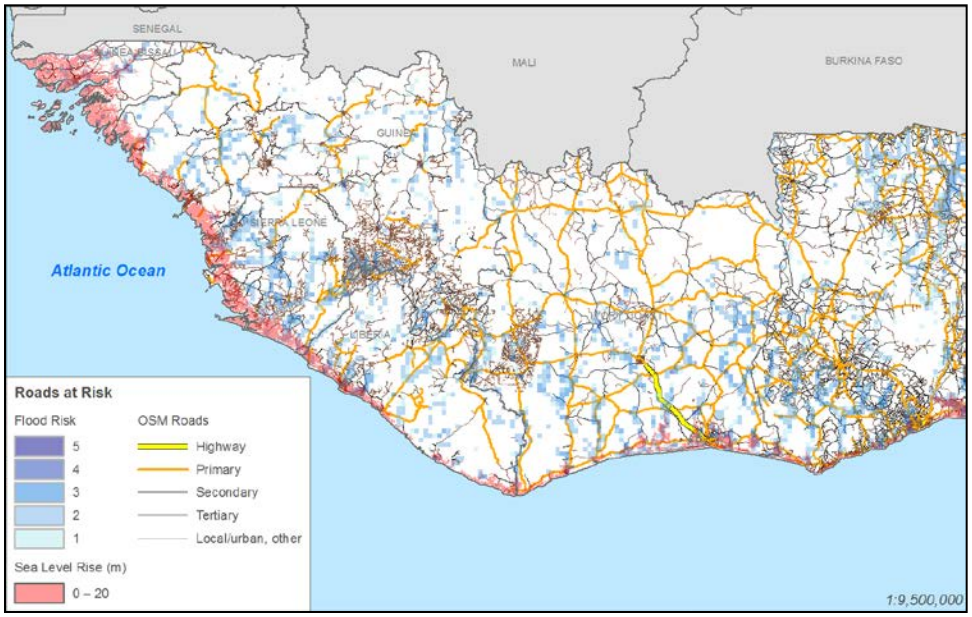
Center for International Earth Science Information Network
Earth Institute | Columbia University

Prior OSM Validation work

- Barron et al. 2014, “A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis”, *Transactions in GIS*.
- Goodchild and Li. 2012, “Assuring the quality of volunteered geographic information”, *Spatial Statistics.*
- Neis et al. 2012. “The Street Network Evolution of Crowdsourced Maps: OpenStreetMap in Germany 2007–2011”, *Future Internet.*
- Mooney and Corcoran. 2012. “Characteristics of Heavily Edited Objects in OpenStreetMap”, *Future Internet.*
- Koukoletsos et al. 2011. “An automated method to assess Data Completeness and Positional Accuracy of OpenStreetMap”, GeoComputation 2011.
- Gires and Touya. 2010, “Quality Assessment of the French OpenStreetMap Dataset”, *Transactions in GIS*.
- Cipeluch et al. 2010, “Comparison of the accuracy of OpenStreetMap for Ireland with Google Maps and Bing Maps”. Conference paper.
- Zielstra and Zipf. 2010, “A Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany”, 13th AGILE International Conference on Geographic Information Science 2010, Guimarães, Portugal.

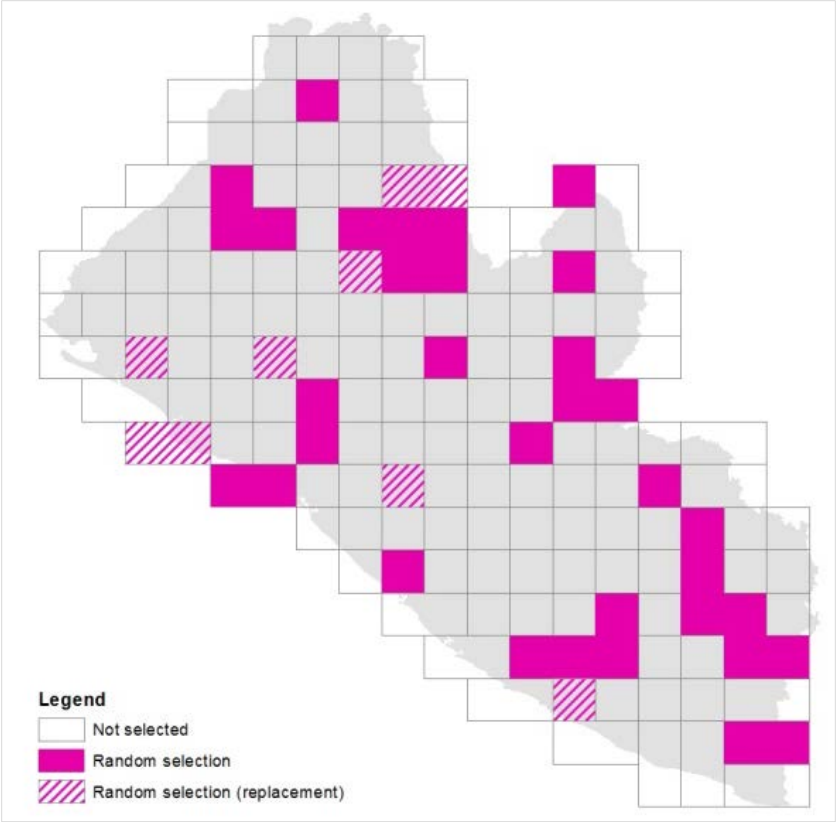
West Africa Analysis

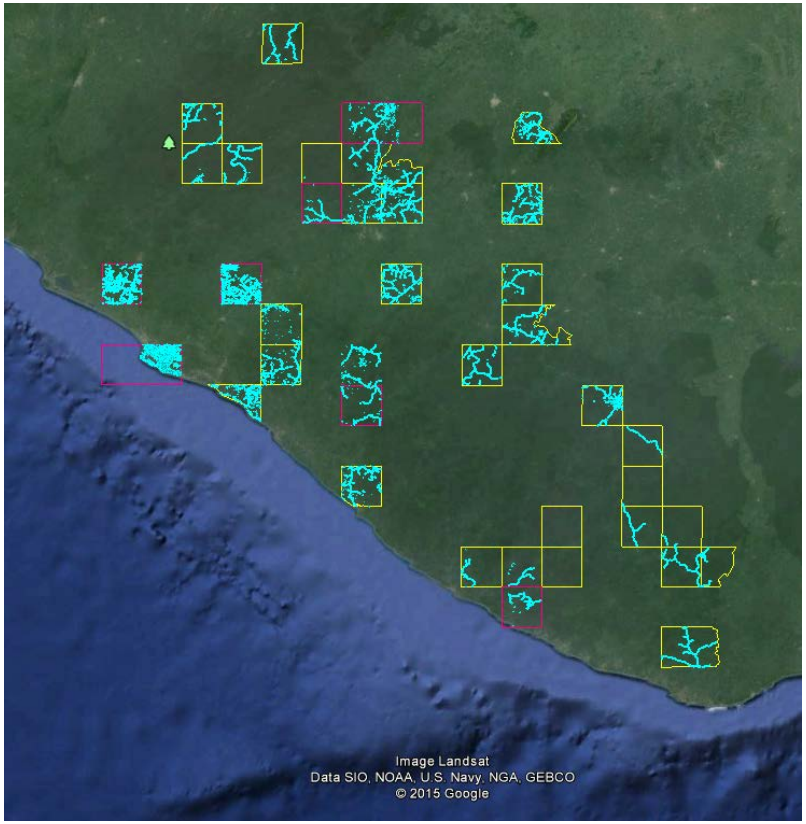
- Downloaded data in **May 2014** for a USAID project
- Downloaded data from Geofabrik in **May 2015**
- A very low income (OSM lagging) region
- Several countries (Guinea, Liberia and Sierra Leone) experienced a major outbreak of Ebola in 2014-15



Liberia: example random sampling procedure

Accuracy testing





Positional accuracy of OSM is generally very good

Results from accuracy testing

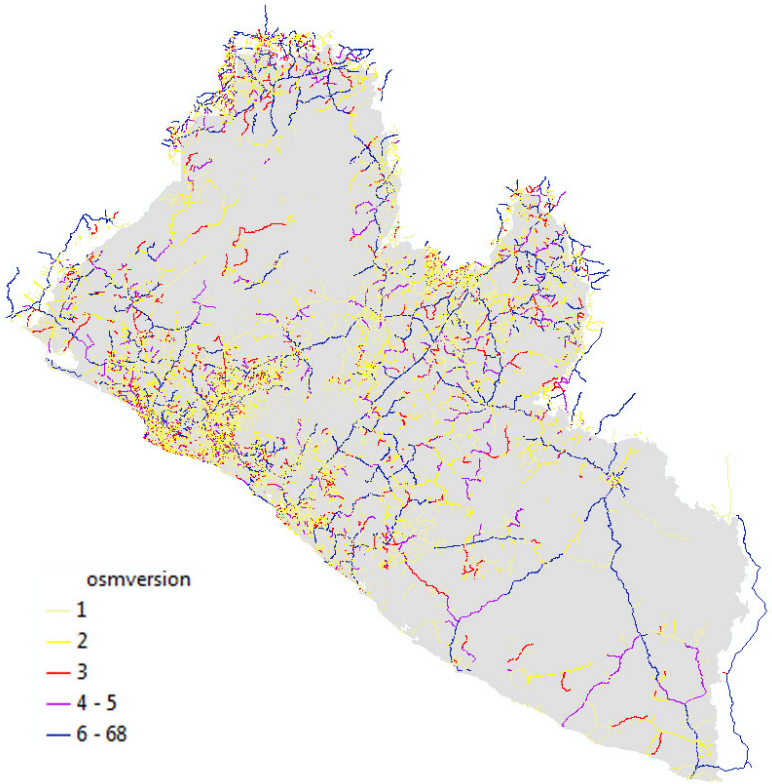
	Country	Ebola affected	RMSE (meters)	NSSDA accuracy (meters)
1	Ivory Coast	No	49	84
2	Cameroon	No	31	54
3	Guinea-Bissau	No	18	30
4	Liberia	Yes	22	38
5	Sierra Leone	Yes	13	23
6	Guinea	Yes	21	36

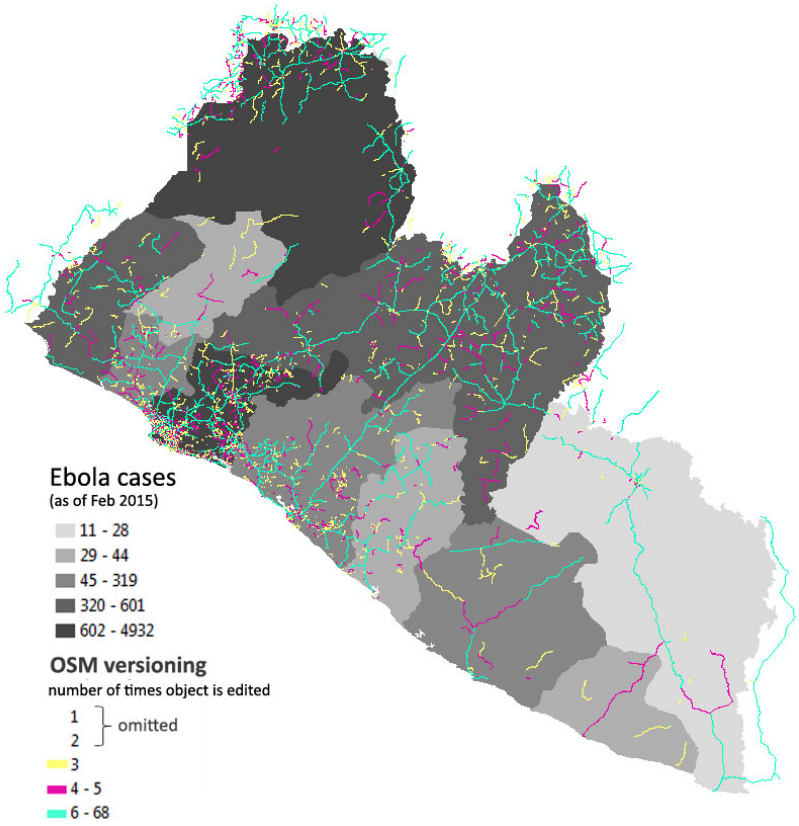
Some seemingly systematic spatial displacement in certain regions

Road length in ebola affected countries more than tripled, compared to a 50% increase for the other countries

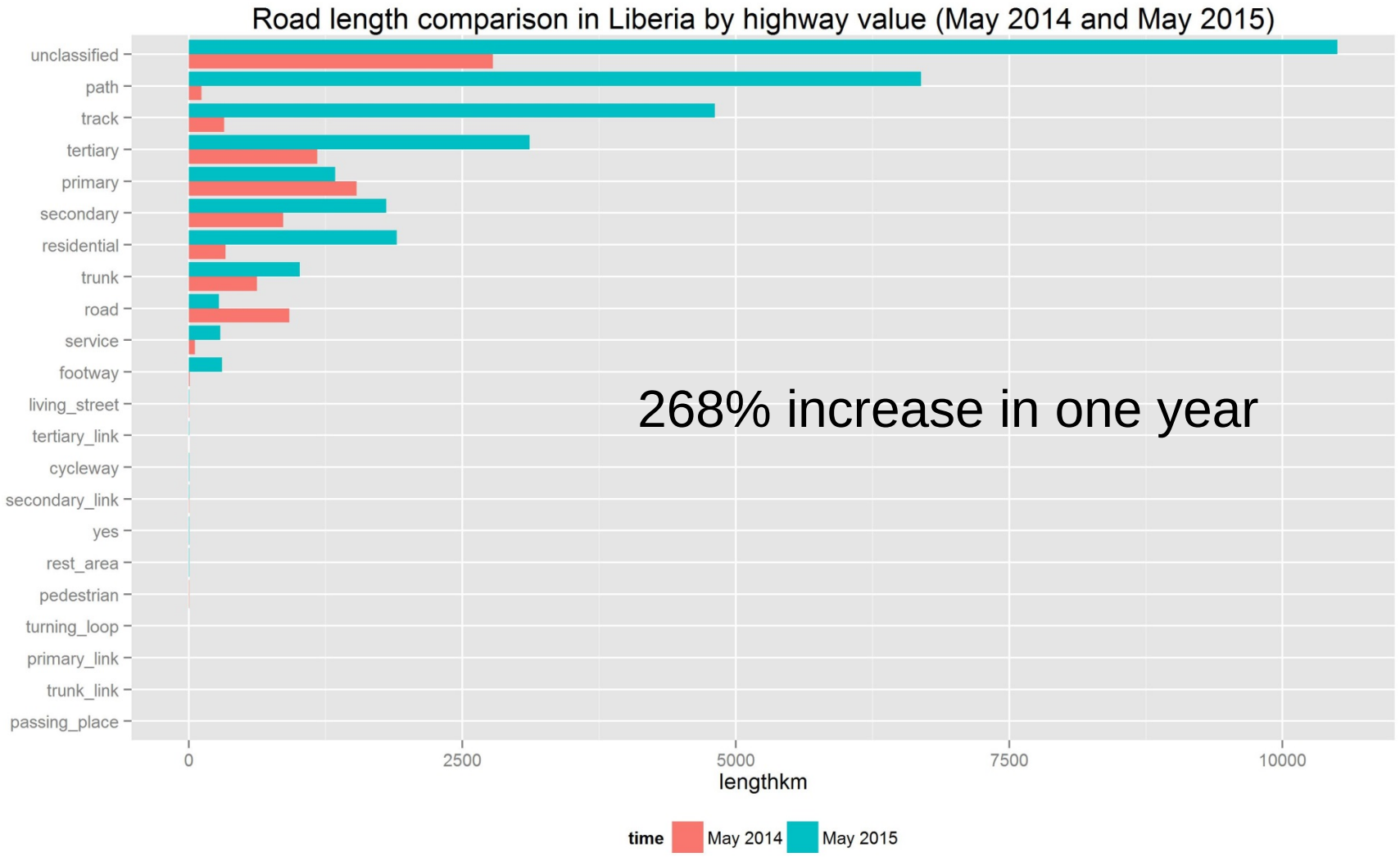
Country	May 2014 length (km)	May 2015 length (km)	Ebola affected	% change
Ivory Coast	32,769	59,081	0	80%
Cameroon	60,124	84,386	0	40%
Guinea Bissau	3,109	4,209	0	35%
Liberia	8,715	32,037	1	268%
Guinea	33,742	99,111	1	194%
Sierra Leone	11,459	50,401	1	340%
Non-Ebola countries	**96,003**	**147,677**	**0**	**54%**
Ebola countries	**53,917**	**181,551**	**1**	**237%**

Liberia: number of edits per road segment





Liberia: length increase by type



Validation and assimilation challenges

- OSM is rapidly evolving → a moving target
  - Validation at any point in time will soon be obsolete
  - Especially in countries that later experience disasters
- Absence of independent “objective” data in low income countries
  - Few authoritative commercial providers
  - High resolution imagery patchy in many countries
- Levels of detail and road types vary by country and may require simplifying and recoding for consistency

Ideas are welcome! www.groads.org