Advancing the citizen scientist's contributions to documenting and understanding natural hazards:

a proof of concept for linking crowdsourced and remotely sensed data on landslide hazards in El Salvador

Eric Anderson[1]*, Robert Griffin[2], Kel Markert[1]

[1]The University of Alabama in Huntsville Earth System Science Center, and NASA/SERVIR Science Coordination Office
[2]The University of Alabama in Huntsville Atmospheric Science Department

*eric.anderson@nasa.gov

AGU Fall Meeting 2017

SERVIR

THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

CILSS adpc RCMRD ICIMOD

USAID FROM THE AMERICAN PEOPLE NASA

Acknowledgements

- This talk synthesizes lessons learned through several teams working through NASA Internship program, NASA DEVELOP program, and University of Alabama in Huntsville Research or Creative Experience for Undergradates (RCEU) program and MS Earth System Science research
  - https://intern.nasa.gov/
  - https://develop.larc.nasa.gov/
  - https://www.uah.edu/undergraduate-research/rceu-summer-program
- Also considers lessons learned through development of SERVIR-Mekong Land Cover Monitoring System and Collect Earth Online
  - https://servir.adpc.net/tools/land-cover-monitoring-system
- Ministry of Environment and Natural Resources (MARN) / General Directorate of the Environmental Observatory (DGOA) in El Salvador
- Free visible imagery from DigitalGlobe provided through Google Earth
- The joint NASA- and USAID- SERVIR program
  - https://servirglobal.net/

Outline

- Brief background and rationale for involving citizen scientists for better understanding of hazards
- Challenges of characterizing landslide hazards
- Objectives
- Data collection methods and results of statistical tests
  - Steps to post-process volunteer data
  - Pitfalls - false positives and false negatives
  - Improving our characterization of the landslide hazard
- Discussion – benefits and challenges of incorporating data from citizen scientists
- Next steps

Background and rationale

- Scientists, practitioners, policymakers, and citizen groups, share a role in ensuring “that all sectors have access to, understand and can use scientific information for better informed decision-making” (Sendai Framework 2015-2030).

- When it comes to understanding hazards and exposure, **inventories** on disaster events are often limited.

- There are many opportunities for citizen scientists to engage in improving the collective understanding—and ultimately reduction—of disaster risk.

Characterizing landslide hazards is difficult

- From US Geological Survey, we need to know 4 things about landslides
  1. When will they happen?
  2. Where will they start?
  3. Where will the go?
  4. What could be affected?

- Dai et al (2002) and van Westen et al (2006) provide excellent summaries of landslide risk management and challenges with hazard zonation

- Heuristic hazard mapping methods are very common in regional hazard zonation, but they often lack a temporal component.
  - Rely on expert opinion/knowledge of physical processes and the region; difficult to scale
- Statistical/probabilistic approaches are becoming more common, but this is only possible when rich **inventories** of landslide events exist.
  - Correlations may be misleading if physical processes aren’t understood

Objectives

- Demonstrate a concept on incorporating crowd-sourced data to improve landslide hazard model performance

- Discuss caveats, challenges, possible steps forward

- Study site: El Salvador
- Focus: rainfall-triggered landslides from 1998 to 2011

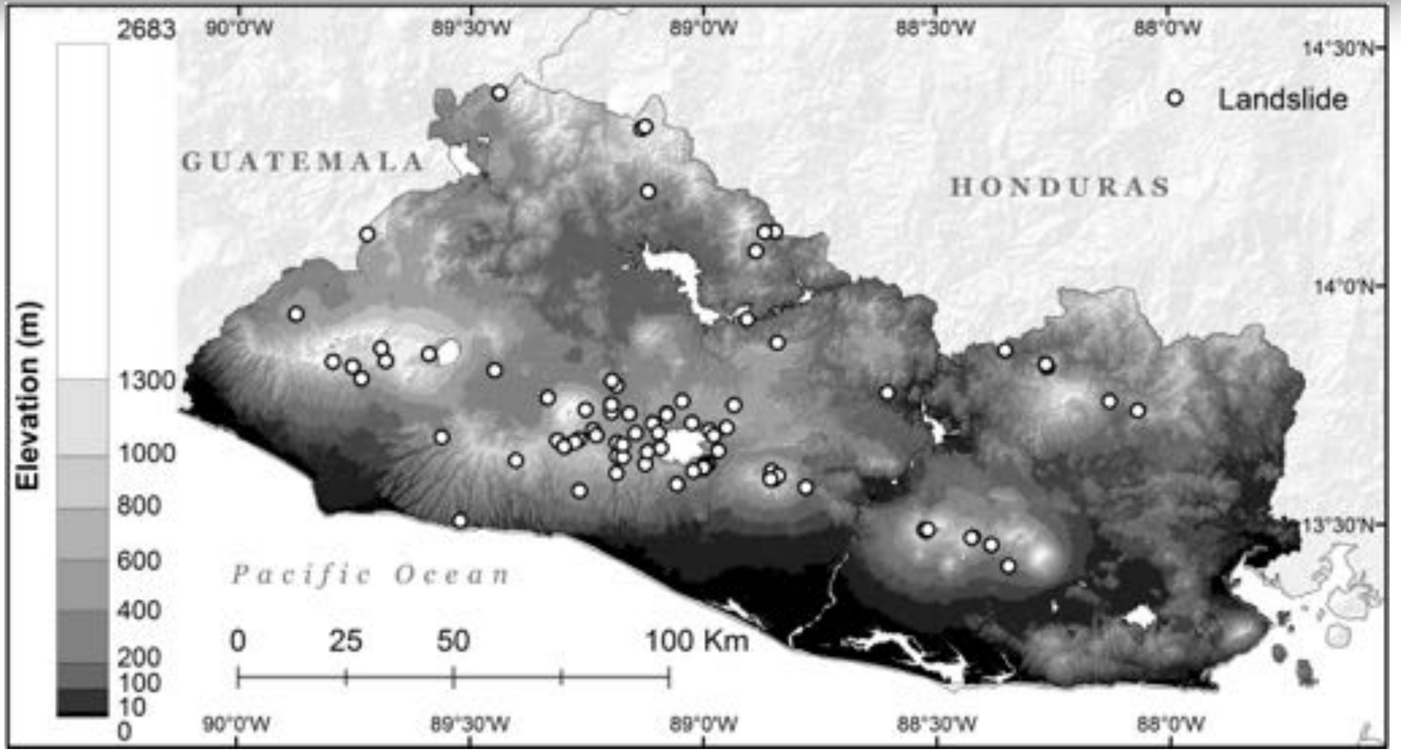
OpenStreetMap

Data collection methods (1)

1. National inventory of 90 rainfall-triggered landslides in El Salvador from 1998 to 2011

2. Identified additional 670 landslide events through human interpretation of high resolution imagery

3. Post-processed volunteer data (QC step)



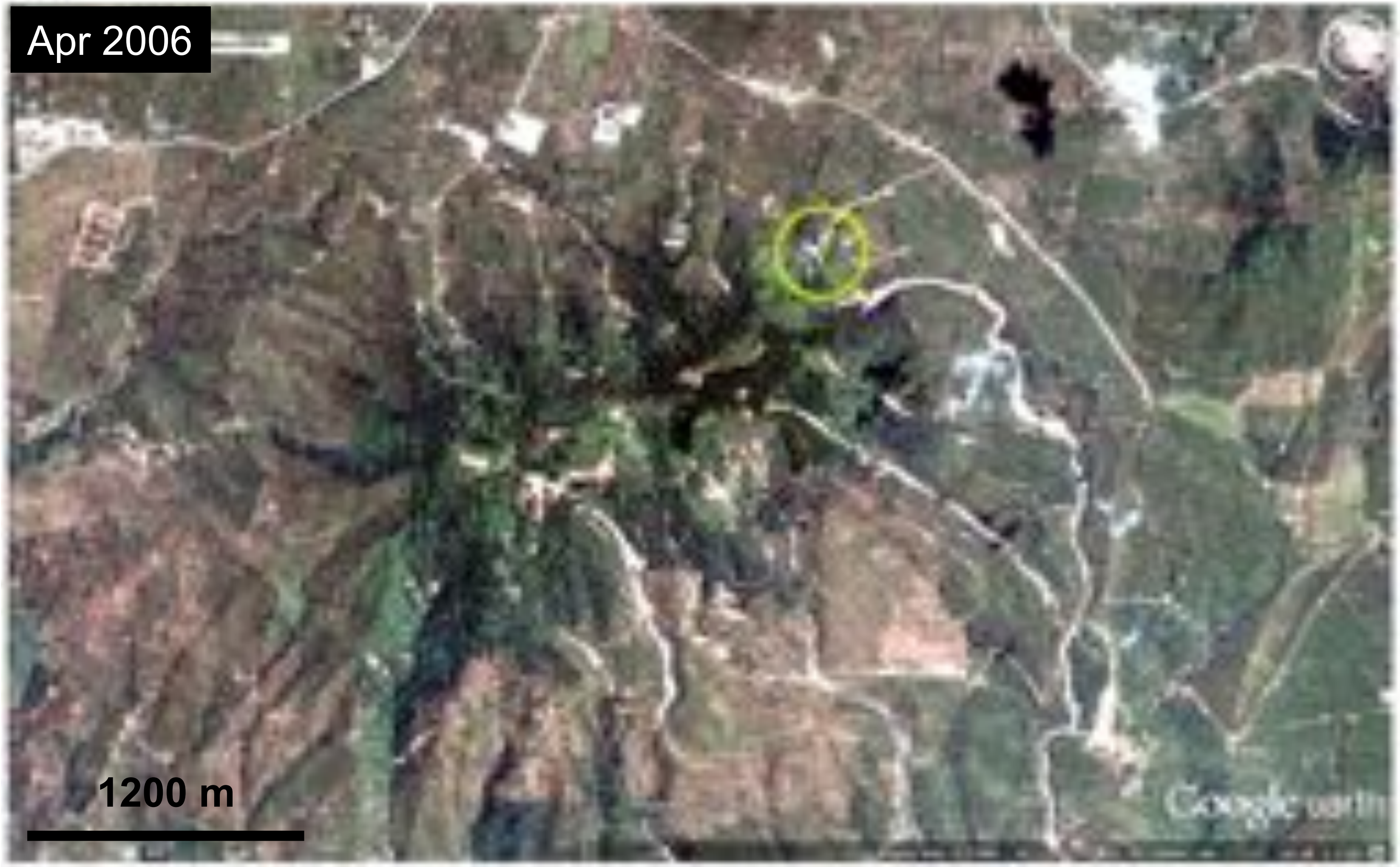
Data collection methods (2)

- Identify ***initiation point*** of individual landslide events through human interpretation of high resolution imagery in the Google Earth time slider

Adapted from Anderson 2013

Data collection methods (2)

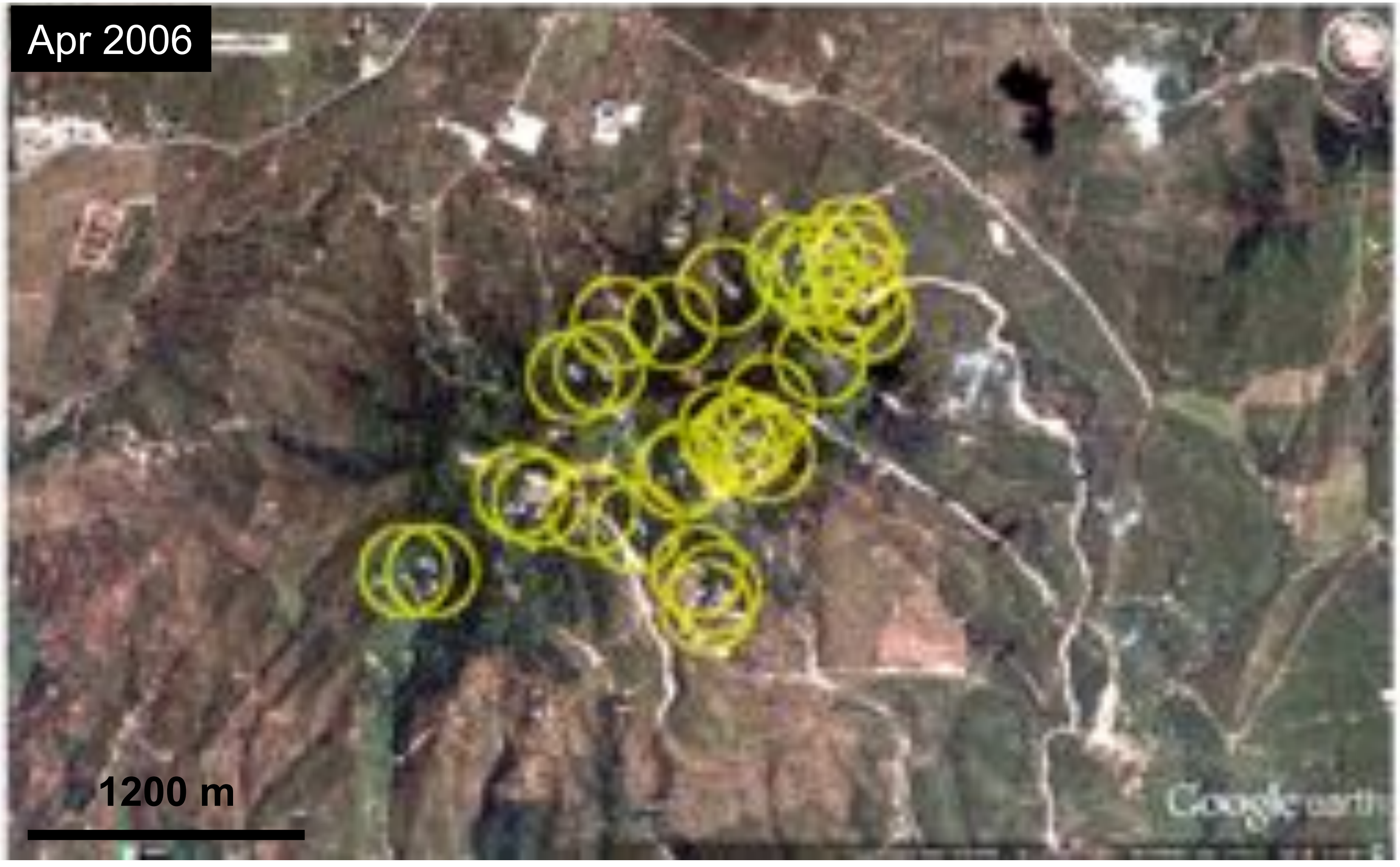
- Identify ***initiation point*** of individual landslide events through human interpretation of high resolution imagery in the Google Earth time slider



Adapted from Anderson 2013

Data collection methods (2)

- Identify ***initiation point*** of individual landslide events through human interpretation of high resolution imagery in the Google Earth time slider



Adapted from Anderson 2013

Data collection methods (2)

- Identify ***initiation point*** of individual landslide events through human interpretation of high resolution imagery in the Google Earth time slider

Adapted from Anderson 2013

Data collection methods (2)

- Identify ***initiation point*** of individual landslide events through human interpretation of high resolution imagery in the Google Earth time slider

Before

Data collection methods (2)

- Identify ***initiation point*** of individual landslide events through human interpretation of high resolution imagery in the Google Earth time slider

After

13 of the original 90 catalogued landslides presented similar conditions

Data collection methods (3)

- Post-processed volunteer data (QC step) provided by expert
  - Visual check to add/remove obvious errors
  - Rough modification of initiation point in GE
  - Detailed modification of initiation point in GIS – ensure initiation point as identified in GE aligns with layers to be potentially included in hazard model
    - e.g., point is situated on correct aspect with respect to flow of landslide
    - Corrects location errors introduced by inconsistent orthorectification of input images

Results

- Identified additional 670 landslide events in El Salvador between 1998-2011, using national inventory as a guide for general location and time
  - New total of 760 cataloged events

- Using logistic regression, tested impact of volunteer data on our ability to characterize the landslide hazard
  - El Salvador case: improvement from 66.7% overall accuracy to 81% accuracy (including rainfall triggering aspect)
  - Rwanda case: 79.6% overall accuracy (Piller 2016)

Discussion - Limitations

- Distinguishing *landslide type* and *trigger factor* is difficult or impossible through visual identification of landslides using optical satellite images
  - Ancillary data or expert knowledge could help in some cases

- Assigning a **date** to landslides identified this way is nearly impossible
  - DRIP-SLIP technique attempting to address with RS https://github.com/NASA-DEVELOP/DRIP-SLIP

- QC Grading or scoring volunteer contributions as a QC measure
  - Self-assigned certainty categories
  - Scoring system per volunteer

Discussion – Opportunities

- ✓ "Local observers" in El Salvador who augment in situ hydro-meteorological measurements
- ✓ **Growing access to Earth observation data to the lay person**
- ✓ Immense interest in connecting citizen scientists to remote sensing data through hackathons such as the NASA Space Apps Challenges
- ✓ Benefits of citizen scientist in combing images for landslides
  - ✓ Once trained, can probably do 90% of the work for a fraction of the cost, leaving 10% of QA/QC to expert (or even a higher-ranked citizen scientist)
- ✓ Even more advanced remote sensing techniques—whether optical image change detection, image segmentation, SAR-based change detection or interferometry—need inventories or samples of actual events, so the citizen scientist could come in here too

Discussion - Opportunities

- Collect Earth Online,
  - http://ceo.sig-gis.com/
  - https://github.com/openforis/collect-earth-online

COLLECT EARTH ONLINE

- Google Earth Engine - potential for more dynamic updating / re-assessment of hazard models, driven by volunteer data.

- Planet Labs sub-weekly hi res imagery

Conclusions

- Inventory data often limit understanding of landslide hazards, and ability to determine causal factors

- Landslides are typically straightforward to identify in optical satellite imagery, and can be done in free platforms

- Significant QC steps are required before immediate ingestion into hazard models, but we have demonstrated improvements

- Involving citizen scientists –
  - Quantitative improvement in understanding of the hazard?
  - Participatory process that helps people understand risk and ultimately take actions to reduce their risk?