Advancing Data for Street-Level Flood Vulnerability: Extraction of Variables from Google Street View in Quito, Ecuador to Enhance Observation Driven Flash Flood Risk Assessment

Raychell Velez1, Lauren Carey1, Christopher Aime1, Diana Calderon1, Carolyne Hultquist2, Andrew Kruczkwiewicz3,4, Greg Yetman2, Yuri Gorokhovich1, and Robert S. Chen2.
1 Department of Earth, Environmental, and Geospatial Sciences, Lehman College 2 Center for International Earth Science Information Network, Earth Institute, Columbia University 3 International Research Institute for Climate and Society, Earth Institute, Columbia University 4 Red Cross Red Crescent Climate Centre

Novel geospatial and crowdsourcing technology provide ways to collect vulnerability data. We develop a framework to collect data with a human-machine interface technique using satellite imagery and Google Street View (GSV). Outputs of our method are coupled with EO-driven flash flood modelling to develop an anticipatory action program for flash floods. We focus on Quito, Ecuador, an urban area in a valley surrounded by mountainous terrain with high flood exposure. The study advances data for the development of regional training datasets that will aid in automating the mapping of flood vulnerability in urban areas for flood risk assessment and early warning.

Flowchart of Methods

Introduction

We use a crowdsourcing method to build a database of flood vulnerability variables (e.g., drains, sill height) collected from GSV. Turking is a form of crowdsourcing which breaks down tasks into simpler components with easier decision processes and distributes these focused tasks to many individuals. Our variables were selected as critical factors in flood risk assessment, particularly for floods occurring in urban settings. We developed an interface using the local Mechanical Turk (MTurk) environment that required two primary inputs: (1) geographic coordinates associated with remote sensing imagery and GSV, and (2) a survey document that includes our variables of interest for collection. A basemap layer to better assess co-location with building location. Points not coinciding with a building location were assigned to the closest building. We note our approach requires a standardization matrix (table) to document variable categories to ensure consistency in data entry for all participants. Once standardized, variables were classified using image examples from GSV and general descriptors such as the sill height, the roof type, and the visible condition of a building. Finally, we use GIS to check consistency and describe spatial patterns of flood vulnerability data.

Data and Methods

Data and Methods

Survey Design
1. Define variables
2. Standardize through codebook
3. Develop template task script of survey variables

Selection of Sample Sites
1. Create random points in urban area
2. Ensure point near building
3. Ensure point has GSV coverage

MTurk Setup
1. Export remote sensing images for each point location from GSV and load in AWS
2. Load AWS S3 bucket into MTurk
3. Run module with template task script

Data Collection and Analysis
1. Use MTurk to collect data for each location
2. Evaluate data collected for each variable
3. Visualize variables in context

Results

Acknowledgements

We acknowledge funding from NASA Goddard Space Flight Center (GSFC) contract 80GSC18C0111, Continued Development and Operation of Socioeconomic Data and Applications Center as a Distributed Archive Center, led by the Center for International Earth Science Information Network (CIESIN) at Columbia University, and NASA contract 80NSSC18K0342.