Questions & Answers Part 2

Please type your questions in the Question Box. We will try our best to get to all your questions. If we don't, feel free to email Brock Blevins (brock.blevins@nasa.gov) or Greg Yetman (gyetman@ciesin.columbia.edu) or Taylor Hauser (hausertr@ornl.gov)

Question 1: HAZUS integrates into ArcGIS, but ArcGIS is not open (and not everyone will have access); is there a possibility to look at HAZUS when you don't have access to an ArcGIS installation?
Answer 1: The HAZUS software and algorithms are open, they are available to run outside of ArcGIS as Python scripts (https://github.com/nhrap-hazus/hazus). Running the software outside of ArcGIS does require knowledge of Python and programming, and there is no menu-based system to interact with the data and output results outside of ArcGIS. The default data used in HAZUS is also available for download and use in other software packages, like open source GIS or statistical software (https://msc.fema.gov/portal/resources/hazus, scroll down to the data section).

Question 2: How many critical building infrastructure buildings are high risk? What mitigation approaches are moving forward?
Answer 2: Unfortunately, we don't have at hand the total number of critical infrastructure buildings at risk to modeled flood. There is a list of the recommendations for adaptation and mitigation in the technical report appendix (http://fidss.ciesin.columbia.edu/fidss_files/documents/FIDSS_Site-Methods_Report-Draft.pdf), including the recommendations for critical infrastructure. Recommendations are related to changing infrastructure.

Question 3: Does QGIS have damage curve calculation curve methods?
Answer 3: As far as I know it does not have built in damage curve calculations. However, the FEMA flood damage curves can be used in a Python script, and Python scripts can be integrated with QGIS. You will have to adapt the script.

Question 4: Such a great WebGIS application. How do you make a similar flood impact interactive map? Can you please share the repository?
Answer 4: Thanks! We published the flood data and building footprint data using ArcGIS Enterprise (also known as ArcGIS Server). A lot of testing and revision of the services was required to make certain that they performed well and were easy to read on the map when combined.

The web application was built with Esri's experience builder (https://www.esri.com/en-us/arcgis/products/arcgis-experience-builder/overview). We did not create a repository for the code as it's a relatively straightforward implementation of experience builder; if you have access to the Esri software and similar data you could create a similar application for a different area. It is also possible to create a similar application without ArcGIS Enterprise (Server), but it would require using ArcGIS Online to publish services.

To create a similar application using open source software, you could investigate using GeoServer or MapServer to create the data services, and use an open source mapping client to view and interact with the services. There are a number of open and commercial clients, including Mapstraction, MapBuilder, Mapbender, GeoExt and Mapfish).

We can make the data available,

Question 5: In the section about building footprint detection, I thought I saw that the pan-sharpened image is 2.4m ground resolution. Is this detailed enough to complete the building footprint when the building is incomplete by for example tree tops? And how does the model perform when the buildings are closer together than the pixel size?
Answer 5: This is difficult to answer without seeing the output. Higher resolution does not guarantee higher quality detections. One main drawback of using satellite imagery for feature extraction is trees. Our model will not completely delineate a building that is covered by a tree. It will be a partial detection. We did try to implement leaf off imagery when able.

'How did the model perform when buildings are closer together than the pixel size?' We found that the view angle of the imagery had the greatest effect. If the buildings

were closer together and we had off nadir imagery the CNN would draw both buildings together connected by the façade of one of the buildings.

Question 6: How do you select the best machine learning model (as we have many models available)? What I want to do is train the Sentinel images with some drone orthomosaic data I have for building footprint extraction.
Answer 6: Typically one would select a few different models that are suitable for your needs, then train them all and compare their performance to one another on a validation data set.

Question 7: Is the code used to train the CNN to delineate the buildings open access?
Answer 7: No.

Question 8: I have a question regarding the rapid advancements in machine learning algorithms for detecting buildings and extracting polygons from Earth observation images. Some of these changes happen quite fast. How do you stay current with these developments, and what criteria do you use to determine when it's necessary to update your models?
Answer 8: We will run a general model first and evaluate the output. If it meets our needs we will not update. If the output needs improvement, we will update the model via localized training data. Updating the architecture of the models is done every two years or so.

Question 9: Can you provide details on the regularization algorithm? Is it available in GIS software and can be run on a PC or needs a computing server? I also see reference to GAUNTLET - A Tool for Calculating Building Morphologies. Is it an open source tool? Can you provide references to such a tool (e.g., scientific publications, links)? Please also provide details on the ResType model.
Answer 9: The regularization algorithm we used was esri's regularize footprint tool. https://pro.arcgis.com/en/pro-app/latest/tool-reference/3d-analyst/regularize-building-footprint.htm. We are still investigating open source options.

We are in the process of obtaining an open-source license for Gauntlet. Gauntlet and ResType are new developments, and we have a few upcoming papers that are either accepted or in review. I will be happy to announce their publications when they are released.

The Restype model is a binary classification model that analyzes the morphologies generated by GAUNTLET. Its classification is either residential or non-residential. We use this to fill data gaps where we have no parcel information about the structure. We have a paper accepted and the expected publication date is late December.

Question 10: How did you connect the structure in the image to an address in Google Streetview for the Mechanical Turk task?
Answer 10: Based on the latitude and longitude, Streetview lets you use lat and long.

Question 11: Measure sill heights themselves or terrible idea?
Answer 11: For where we were working (Quito), we didn't have access to the sites to do estimates. Unfortunately, we also did not have the resources to survey the residents/owners and ask for sill height. If it were possible to measure the height (or ask owners/residents to do so), the data would likely be higher quality. It could also be useful as validation data to evaluate how accurate analyst estimates are by using the surveyed data as comparison to StreetView analysis.

Question 12: In the use of Hazus for flood impact analysis, what do you think the major limitations will be when trying to deploy it for developing countries?
Answer 12: The damage curves were developed for the U.S. based on floods in the U.S., so the damage estimates are likely to be inaccurate for developing countries. If damage data were available, it may be possible to adjust the curves (or construct new ones) to improve the analysis.

Question 13: How can we automate the process of Google Street View using computer vision like the floors? Additionally, it would be useful where we have a dense, homogeneous urban environment.
Answer 13: I believe that computer vision/machine learning approaches have been used to detect occupancy type (commercial/residential) and building heights/number of

stories, but I'm not very familiar with how successful those efforts have been. Google itself may have resources (Google Code) relevant for this, although I'm not certain.

Question 14: Is the third example's output available to the public? Where can we see more of its visual results?
Answer 14: We plan to publish the results once the paper is published, hopefully over the next few months. There may be revisions to the paper (and data) based on the reviews, so we are waiting for it to be finalized before publishing.

Question 15: Could you tell us more about any other web app optimization problems you solved beside regularization?
Answer 15: Data services in NYS, we had to tune the performance.

Question 16: Is building insurance sill height info available?
Answer 16: Internationally we did not have that information. In the US, companies may perform a similar method.

Question 17: How sensitive are the results to the precision of the measurements? E.g., are sill heights sensitive at the inch, or foot, (or yard)? Measure with a micrometer, cut it with an Axe (reality)?
Answer 17: Within a foot or so. This will vary a lot depending on the building or street view imagery.

Question 18: Regarding characterizing vulnerability using Google Streetview, was the data (Mechanical Turk, etc.) later used to categorize the region into different levels of vulnerability? If so, could you please describe this further or demonstrate this?
Answer 18: This is ongoing work. Random samples from Quito.