Questions & Answers Part 1

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 Amita Mehta (amita.v.mehta@nasa.gov) or Sean McCartney (sean.mccartney@nasa.gov).

Question 1: How can we detect floods in urban areas using the InSAR coherence for Sentinel-1? I know that the question is a little far from the webinar objectives but I hope I can find an answer here.

Answer 1:
1. https://www.mdpi.com/2072-4292/11/2/107
2. https://ieeexplore.ieee.org/document/7730976
3. https://www.researchgate.net/publication/318740277_SENTINEL-1_DATA_TO_MAP_FLOODED_AREAS_THE_ROLE_OF_INSAR_COHERENCE_AND_POLARIMETRIC_INFORMATION
4. ARSET training on Urban Flooding:

Live Answer: Urban flooding will be addressed in the next part of the training. If you use SAR, there are limitations. In urban areas, you can get a double bounce that limit flood detection. Because of resolution and cloud cover, optical data may not be applicable for urban flooding. Urban flooding is an issue and we can see how to use that.

Question 2: (Slide 28) How were the number of disaster events determined? What qualifies as a "disaster"?

Answer 2: https://www.unisdr.org/files/61119_credeconomiclosses.pdf

Based on the Centre for Research on the Epidemiology of Disasters (CRED) and the UN Office for Disaster Risk Reduction, compiled from data provided by the United Nations, National governments, US governments, International Federation of Red Cross and Red Crescent Societies, World Bank, Reinsurance companies (e.g. SwissRe, MunichRe, AON Benfield), and the Press.

Live: CRED is compiled by the UN based on various governments. And also based on press reports for the disaster reports that come up in local news media.

Question 3: What are the extents of these data? Globally? National? Continental?

Answer 3: Most data are global in extent, though SLOSH is only for the US and TRMM-TMPA has latitudinal extent from 50S to 50N.

Live Answer: Regarding the remote sensing and modeling data that we discussed - most are global. The exception is SLOSH, which is for storm surge, and is only for U.S. coastal regions. TRMM data (TMPA) has limited coverage 50S to 50N since it's for tropical regions of the world. Other than that, all data mentioned are global.

Question 4: What does the meaning of fidelity to input data mean here (SEDAC slide 7)?

Answer 4: Gridded Population of the World (GPW) does not use ancillary data and/or modeling to allocate the population within a given administrative unit. The assumption is that the population is uniformly distributed within it: proportional allocation. Detailed explanation is available at https://sedac.ciesin.columbia.edu/data/collection/gpw-v4/documentation

Susana: the related input data mean that data is not modeled. Whatever we acquired from the sensor's data - that's the data we used. The only modeling we provide is to assume that the population is uniform and distributed over an area for what that population is the count. If you look later in the presentation we mention other pop grids, and others used ancillary data, model data, for roads, etc. to allocate population within area. That isn't the case with GPW. It's only done for proportional allocation. That's why it's so important to know resolution. The data tends to have more and more higher resolution.

Question 5: Who are the end-users for disaster data products? And what kind of specific data products need to developed or improved for them?

Answer 5: End-users comprise local, regional, and national governments, humanitarian organizations, international institutions, and private companies. The specific data product depends on the analysis needed by the agency or company. It could be long-term forecasting, near real-time forecasting, or a specific application assessing historic risk/resilience to a disaster.

Question 6: Can we follow the same methodology for any region in the world?

Answer 6: Regarding SEDAC datasets, all the ones included in the presentation provide global coverage.

Question 7: How often does NASA update the datasets?

Answer 7: Each data is updated when either a new algorithm or improved products are available -- all have different time tables.

Live Answer: (Amita) It varies with the dataset - there is no specific timetable for all the data. When the mission is launched, during the lifetime of a mission, algorithms are constantly being revised after going through calibration and validation. Newer versions then come out. It depends on the development of algorithms throughout mission lifetime. Sometimes you have multiple versions available. Usually as algorithms improve, newer versions are created. It's the same with models. If there's improvement in model parameterization, then prospective model runs are done - it depends on what dataset you're talking about.

Question 8: Do you conduct a verification or validation of model data toward observers (real) data in applying on your research? How does one conduct it?

Answer 8: (Amita) If you look at all the model data NASA puts out or all the remote sensing data that comes out - for all of them there are validation programs in place. But they aren't global, they're at selected places with detailed in situ data available. That's how these models and remotely sensed data are validated. We always recommend in your own region - for your own application, you do your own validation. In that case, you have to have some in situ data or info that can help you validate the data in the sense that in your local region you may not have a lot of in situ data, but you see flooding occurring on certain days/in certain areas. Go back and look at NASA data and see if it showed flooding. That's indirect validation in your area. But this question is very, very important because it isn't possible to validate globally because there is not enough in situ data to do so. And if there is, they all have different qualities. It's very difficult to come up with one answer for validation. It has to be regionally done.

Question 9: Which is the source of census data and how is it validated?

Answer 9: The most common source of census data are national statistical offices. The UN World Population Prospects is used for adjustment.

Live Answer: (Susana) Those are the primary sources, and we try to be as close to regional sources as possible. For countries without, we use estimates of population, in some cases in Europe, we use those, but we try and use as much of the official country data as possible. Because they're so diverse, we also use the UN world pop prospects for the corresponding year. To know if there are a lot of differences, and provide adjusted and non-adjusted version of the dataset.

Question 10: The gridded population estimates from different sources are quite different. How can we interpret the critical values on the basis of these data?

Answer 10: The population estimates are the results of different methods for converting census and other population data to grids. The raw data is still mostly census data. What's different is the methodology to allocate the population to the grid. The algorithm they use for that allocation is different. That's why we consider it an important purpose. What is the data going to be used for? You may need better resolution or include population data in your model. In that case, how modeled or less modeled the population data can be very important. That's part of what users need to decide, and that's why we have a lot of emphasis on documentation. To identify how the data is processed and where it's coming from.

Question 11: I am interested in mobile applications for data management. Is there any app you recommend to build geographic databases?

Answer 11: (Susana) We have the hazard mapper now as a mobile application for both apple and android. That's the one we have so far. I can't remember an app for that at this point. Not sure what "build geographic database" exactly means. Most of our data is available to coordinate in other map services. Other than that, I might need to put you in contact with our IT dept. They're the ones that developed the app and web services.

Question 12: Can we predict the human population by using SEDAC hazard mapper tool?

Answer 12: (Susana) Well, in order to predict population you need to have 2 points in time. That's the classic approach. The hazard mapper provides the population at 1 point in time, 2015, however, you can go to the regional data and look at 2015-2020 and estimate the difference and extrapolate into the future. It's a very crude way of doing population projections, but our data is not built in order to work with long-range projection into the future (i.e. 2030 or 2040). There are other tools for that. Another approach is using UN world population coverage which is extrapolated to 2100. Use that with the SEDAC data as a baseline and use that as a crude way of forecasting into the future.

Question 13: What's the population coverage of SEDAC application?

Answer 13: (Susana) Global. In terms of hazard mapper, some of the data is only available in the U.S. But the others are global.

Question 14: How should SEDAC vulnerability mapper work?

Answer 14: (Susana) If you're using the app or web-based, you zoom in on an area and can retrieve whatever data is available for that area. You may decide on your delay depending on your scale - you may have too much info. You can hover and see different info for that point. Some days have a lower rate of change. Population data is updated when we have new data. We're now in version 5.

Question 15: The DPSIR (drivers, pressures, state, impact, and response model of intervention) framework is a new method to assess risk and vulnerability, can be combined with disaster analysis and resilience assessment?

Answer 15: (Amita) Not aware of this system.

(Susana) Have heard of the system but haven't worked with it.

Question 16: Can we use the flood data for tsunami inundation?

Answer 16: (Amita) If you use optical data from Landsat or MODIS, they can provide inundation information related to tsunamis.

Question 17: I am from India working in GIS & Risk Assessment. What is the minimum resolution which can be used for flood assessment?

Answer 17: (Amita) Good question - it depends on the watershed you're working with or urban area you're working with. Most remote sensing data - we'll talk about next week, say - are relatively low resolution, 1/10 of a degree to ¼ of a degree, which is quite coarse. But if you use optical data like MODIS & Landsat it can be 30 m (Landsat) - 250 m (MODIS). Combining remote sensing data with hydrologic modeling, you can get good flood assessment that way. Minimum resolution depends on the area you're looking at. An example we'll see next session is for urban flooding where our resolution for rainfall isn't that good, but we used rainfall information averaged over the urban area and then look at terrain, impervious surface, slope within the urban area/city to come up with some areas where there may be more flooding than others. Commercial satellites can provide a few meters resolution, and that works well for flood assessment.

Question 18: How can I get data for another country like Morocco?

Answer 18: Susana already showed how to pick a region through polygons. For remote sensing data, there are tools you can select - spatial and temporal-specific information from different remote sensing datasets. We'll see that in the next session.

Question 19: Are there any efforts underway to integrate real-time user reports of flood conditions, etc. during and following disaster event from apps such as Waze? Could this be an additional means of verification?

Answer 19: (Amita) There is a Global Disaster Alert and Coordination System (GDACS). They have a way when a disaster starts - users can provide information from news media or local media. There are also citizen scientist efforts where people report info back. GDACS is the one center I know that systematically collects info while the disaster is happening in real time. I believe that will happen more and more through apps. People can provide info when disasters are going on. And that happens all the time, and it has to happen somewhere. I think GDACS is the only system systematically doing that. Their data mostly comes through official channels, not citizen scientists.

Question 20: Do you have or is there somewhere an example of drought or extreme heat?

Answer 20: Amita: Next week we'll have a guest speaker to talk about extreme heat cases. For drought info, some of the same datasets can be used. For flooding, you can use the same datasets for detecting drought, also.

Question 21: How much accuracy can show in resilience measurement of an area by the remote sensing system actually?

Answer 21: I'm trying to understand - you're asking with what accuracy remotely sensed data can be used for disaster risk assessment that can lead to resilience. That accuracy varies with dataset, resolution, different regions. There's no one answer. That's why we recommend before taking any actions there has to be direct or inferred validation. Look at the hazard you're concerned with and build resilience using RS data. Try some of this data and see how it works. That's why past data is useful. You can go back and look at past hazards and see damage occurred or didn't - what the relationship between in situ and satellite data in the past.

Question 22: Do I understand correctly that most of the population data sets are basically estimations/averages? For DRA would we combine those with available data in country? (e.g. the latest DHS)

Answer 22:

Live Answer: (Susana) No, they are not. They are estimates, yes, because the data is matched to a target year - so it could be from 1 year before or after. They are not averages. What they are is a transformation of data that comes in a certain system - a vector system, usually - into a grid. It's a transformation of the data, not an average. They're considered good datasets and the data is released at a very high resolution (Brazil, U.S.). Some places have very high resolution down to the census block in a city. Depending on the country and the data source, you can have higher or lower resolution. It could be useful in some countries where you don't have a lot of in-state resources. Factor in what is the framework, what is the active coverage. The sensors, with all the problems, has the advantages of having national coverage. **It is important to read the documentation to see how it was picked**. And it lists the sources and how it was transformed.

Question 23: Is it possible to download the grid data for basic demographic characteristics?

Answer 23: (Susana) Yes, you can download it on the SEDAC website

Question 24: The SEDAC datasets for exposure and vulnerability assessment are updated annually, monthly, or at a specific time?

Answer 24: (SUSANA) Updated when a new run of sensors is coming into place. Version 5 is going to close the data from 2020 run of data. Other datasets are updated when new data is available.

Question 25: Could you please explain how we relate the surface stress condition data i.e lineament change behavior with the multi-factors data for landslide risk analysis in advance?

Answer 25: (Amita) There are models that use RS data to have advanced notice about landslide risk. Next year ARSET is planning to have a second part of this webinar in which a focus will be on landslides, earthquakes, and volcanoes. At that time we'll present on datasets, models, and tools that can help.

Question 26: Is there data on loss to agriculture due to disasters?

Answer 26: (Amita) Landsat and MODIS data have been used to look at post-disaster land surface conditions. If vegetation is destroyed, you can see that in images. It may not be able to tell you about agriculture specifically, but it can tell you how much the vegetation changed. A very nice example is from the vegetation de-greening from Hurricane Maria over Puerto Rico. You can see that in Landsat images from before and after that show clear difference in land cover.

An additional resources would be an upcoming ARSET webinar on SAR for Landcover, including agriculture and flooding. Webinar starts on August 28, 2019.
https://arset.gsfc.nasa.gov/disasters/webinars/2019-SAR

Question 27: How can the Earth observations be used to understand demographic changes or can they not be used?
Answer 27: (Susana) That's one of the questions in the field - how you actually use EO to track demographic changes. IT depends on what demographic changes you're talking about. A number of the grid - basically can track the population distribution. That's the first thing you can see. You can look at how distribution can change over time. In some cases it's very high resolution, in some, not so much. Population characteristics can be more difficult and you need to make a lot of assumptions for those. Another one being used now is the night lights (VIIRS) to track density. Through VIIRS data one can track changes in urban areas, per-se, or track settlements over time. Those can be derived more or less from observations. There's some work linking these characteristics through other changes, like socioeconomic - that's a great question.
(Amita) there's also a dataset based on Landsat that provides info on built-up area, human built-up area, can also help in indirectly telling where demographic changes are occurring.
(Susana) There are at least 2 papers using nighttime lights and Landsat for residential and ambient day.

Question 28: Is there any possible solution that can be provided for assessing earthquake risk site for an area immediately by remote sensing techniques?
Answer 28: (Amita) Now that SAR data is available, earthquake monitoring has been done. I'm not an expert in SAR or earthquakes, but if you attend the SAR webinar coming up in the fall (Oct) or look at past ARSET webinars, you might find the answer.

Question 29: How do you estimate the economic value of the assets under risk?
Answer 29: (Amita) Know there are economic models and they use info about assets and how old they are, how it was constructed, etc. Different parameters are used to assess strength and economic value associated with an asset.
(Susana) There's some global dataset, but they aren't up to date. There's one that is at least 10+ years old. You have the data coming from, for e.g., some of the insurance companies that have prepared this data. There's good data for the U.S. It depends on how you value the data - if you're looking at infrastructure data or other impacts (labor market, work days, etc.)
https://www.rff.org/valuables/

Question 30: Are there any other datasets related to population in SEDAC other than global population, population density and population distribution by age and sex?
Answer 30: There are several other dataset, you can explore SEDAC website under "topics." Or in collection you have the population collection and you also have a collection under population dynamics and there are several datasets to explore.

Question 31: My office has ground satellite receiver of NOAA 18 & 19, how to use that data for DRA?
Answer 31: (Amita) I think NOAA 18 & 19 are operational satellites with sensors like AVHRR, AMSU, and [] there are products derived from these sensors. I'm not familiar with all the details, but they can be used for monitoring land cover, flooding, extreme heat, and precipitation data. AVHRR has a long-term coverage with lower spatial resolution than MODIS but with similar optical bands. And you can look at flooding inundation using that. These are my approximate answers since I don't work with these. Any sensors or datasets that have long-term past coverage and is available in near real-time (NRT) can be very useful for DRA because you can look at past hazards and apply to NRT.

Question 32: What kind of GIS software one should use to download SEDAC gridded population data products? Is ArcGIS 10.6 sufficient? Or ArcGIS Pro better? QGIS?
Answer 32: The data is usually available as grids, other formats could be available for specific datasets. ArcGIS or QGIS can be used. Data is available as a .TIFF file. There's other products - depending on the coverage - depends on how big the dataset is it could be other file formats. You can use ArcGIS or QGIS.

Question 33: Is there any database that includes results of national agricultural censuses? For example, land tenure, extension of different crops, etc.

Answer 33: For the United States (US), USDA's National Agricultural Statistics Service (NASS) conducts hundreds of surveys every year and prepares reports covering virtually every aspect of U.S. agriculture. Production and supplies of food and fiber, prices paid and received by farmers, farm labor and wages, farm finances, chemical use, and changes in the demographics of U.S. producers are only a few examples. https://www.nass.usda.gov/

Outside the US, the USDA Foreign Agricultural Service has databases for agricultural censuses. https://www.fas.usda.gov/

Question 34: Why is the Congo Basin the most vulnerable to activity of lightning and storms of the world?

Answer 34: https://www.nat-hazards-earth-syst-sci.net/18/479/2018/nhess-18-479-2018.pdf

The most intense storms, at planetary scale, are found in the Congo Basin (Zipser et al., 2006). It is probably the most active region in the world in terms of thunderstorms since the number of days of the year with thunderstorm activity is found to be maximum there. Comparison of lightning activity in the Congo Basin lightning is large over this extended area due to the activity of thunderstorms in the basin (Soula et al., 2016).

Live Answer: Based on the paper linked above, it has the most lightning activity - if you look at South America and Africa - interior tropical parts - that's where you have very tall convection occurring and that's where convective clouds form and lightning occurs. If you look at overall cloud distribution, that's where you have very tall cumulonimbus clouds occurring. I think it's because of the heating of the land in the tropics due to direct sunlight and the availability of moisture. When they reach a certain level, they become icy and lightning occurs.

Question 35: What is the best method for detecting storm events between WWLLN: World Lightning Location network and LIS: Lightning Image Sensor?

Answer 35: The World Wide Lightning Location Network (WWLLN) is a ground network of sensors which may be several thousand km distant from the lightning stroke. Each lightning stroke location requires the time of group arrival (TOGA) from a least 5 WWLLN sensors. Typically only about 15 to 30% of strokes detected by one sensor are detected by 5 or more. The accrual rate is not the highest.

The Lightning Imaging Sensor (LIS), is a space-based instrument used to detect the distribution and variability of total lightning (cloud-to-cloud, intra-cloud, and cloud-to-ground lightning). It measures the amount, rate, and the radiant energy of lightning during both day and night. LIS is a calibrated lightning sensor which uses a wide field-of-view expanded optics lens with a narrow-band filter (at 777nm). A Real Time Event Processor (RTEP) inside the electronics will remove the background signal, thus enabling the system to detect a weak lightning signal and achieve up to a 90% detection efficiency. The ISS LIS provides total lightning measurements between +/- 48 degrees latitude which covers a geographic range that includes nearly all global lightning. ISS LIS will also supply near real-time lightning data over data-sparse regions, such as oceans, to support operational weather forecasting and warning. Ideally, both data sets should be used in tandem for detecting storm events as they complement

Question 36: How can we assess risks in the field of civil aviation?
Answer 36: Are you talking about risks because of natural hazards we will look at some of the datasets. The civil aviation follows weather forecasts very closely in deciding their operations. That part is taken care of by civil aviation.

Question 37: How can we integrate all databases for disaster risk mapping? Which software is the best representative and give us a high accuracy result?
Answer 37: That's commonly been used - we'll have examples of online tools that use GIS-based system that provide disaster risk information. So currently, GIS is not the best platform to use where you can integrate different datasets.

Question 38: Cómo influye la escala de representación de la Amenaza, Vulnerabilidad y Riesgos?
How does scale influence the representation of threat, vulnerability and risk?
Answer 38: I assume you're talking about spatial scale. Yes, threat, vulnerability and risk at a city-level or country-level is different than county-level. Process is scale-dependent and impact is scale-dependent, so I guess each case is different. And priorities may be different too, for different scales. How to respond, what kind of risks we're talking about - that may change with communities and scales.

Question 39: Qué metodología recomiendan para la verificación de la cartografía de Amenazas, Vulnerabilidad y riesgos?

What methodology do you recommend for the verification of the mapping of threats, vulnerability and risks?

Answer 39: (Susana) I suppose it refers to [] if there were verification, it's related to the answer to question 6 - it depends on the scale and event you are trying to map. If you're trying to map with different sources - say RS and ground-level observations. Comparing the sources and triangulate the sources. So, you have two different information inputs is one of the standards. So you need - that's clear for RS - it's always good to have ground truth to know what exactly is happening, not only what the RS image is telling you. But using more than one source - especially if you aren't sure of the provenance is a good standard for the field.

Question 40: How I can determine flood risk for one area, or the flood level?

Answer 40: (Amita) this is where past data are useful. If you have in situ data - stream gauges, inundation information, based on in situ data. Looking at past data helps you determine likelihood of floods/inundation. If you have in situ data, you can assess that using RS data then you have to use a hydrologic or flood model to come up with that information. To determine risk, usually looking at past cases or data helps.

Question 41: The datasets that you present are the ones necessary to make a risk map? If I had access to more information what information would it include?

Answer 41: (Amita) If you have in situ data of economic damages that occur or assets that were destroyed or damaged - all that information if you can put in along with RS and socioeconomic data that we're talking about, it will provide better information in your area. If you have specific information about local area that always helps.

Example of vulnerability mapping:

https://sedac.ciesin.columbia.edu/data/collection/wacvm

Example of hazards and impacts: Natural disaster hotspots

https://sedac.ciesin.columbia.edu/data/collection/ndh

Question 42: How to combine detailed information with these data sets? For smaller analysis areas? Do you have any experience about it?

Answer 42: (Amita) So, when we present our datasets, the resolution is somewhat coarse. So how small the area you're talking about isn't clear - if you're talking about RS data, the highest resolution we work with is a few tenths of km. A few tens to

hundreds of meters, say, for optical data and SAR data are also high resolution. So these are the data we can work with for smaller areas. There is also commercial imagery available that has a couple of meters to a few km resolution, but it costs a lot. The data we talk about is all open source.

Question 43: Is it possible to have a future course on hydrologic or flood models?
Answer 43: ARSET did an introductory webinar on hydrological modeling and the variable infiltration capacity model (VIC). When we did that, we just focused on hydrologic components and didn't do routing, so flood modeling wasn't done then. But we've gotten that request, and it might come up in the future. Link to VIC training: https://arset.gsfc.nasa.gov/water/webinars/VIC18

Question 44: How can we use artificial intelligence in SEDAC?
Answer 44: Could you clarify what type of AI?
(Susana) There's a panel of different population datasets available. Some of them use machine learning, particularly random forest. Because SEDAC is almost raw data, it's used as an input for some of those algorithms. Particularly in WorldPop, they work with an algorithm and machine learning and some of the SEDAC data is part of the inputs. If that's what they're asking. There are papers from WorldPop that might provide more detailed answers.