



Questions & Answers Session 3

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 Amber McCullum (amberjean.mccullum@nasa.gov), Alex Zvoleff (azvoleff@conservation.org), and/or Monica Noon (mnoon@conservation.org)

Question 1: Could you explain how the relationship materializes between land consumption rate and population growth rate?

Answer 1: The relationship (LCR/PGR) presents a value which reflects an annual rate at which a given city/urban area is growing spatially versus how its population is changing. A value less than one means that the annual land consumption rate is lower than the population growth rate and vice versa. A value of 1 means that both urban expansion and population growth are happening at the same rate. Cities which experience a loss in urban land or population record negative values for the indicator

Question 2: How would the balance be achieved on the two fronts which are mentioned: 1) Increasing forestation and; 2) Increased Urbanization?

Answer 2: This is very relative to the area of interest. Some cities/urban areas require densification, while others require a bit of reduction in densification. The indicator should thus be looked against other indicators as opposed to stand alone value; and the recommended secondary indicators also help to understand the required actions
Live Answer: (Dennis)

Question 3: How do you determine in Trends.Earth the approximate value of the population living within a 1km² box? Which are the required inputs?

Answer 3: The population grids are created by other organizations, such as CIESN, JRC at the European Commission, WorldPop/Flowminder foundation, etc.

Trends.Earth just uses the existing databases. For the dataset itself however, the required inputs include population data at the lowest unit (e.g shp. Of the enumeration areas), and ancillary data such as built up areas and land uses



Remote Sensing for Monitoring Land Degradation and Sustainable Cities

SDGs

July 9-23

Live Answer: (Alex) GPWV4 is used in Trends.Earth currently, and is produced by season and is a default. In a summary table, you can input your own pop data if you have it. (And you can listen back to that in the webinar, since there are some issues assessing population growth rates from gridded datasets. So we recommend using locally available data that meets units of analysis the best.)

Question 4: Does NASA have a database of geospatial population/demographics distribution? Can it be used for disaster management (e.g. evacuation, disaster severity, etc.)?

Answer 4: Refer to the Socioeconomic Data and Applications Center (SEDAC) [<https://sedac.ciesin.columbia.edu/>] for demographics information relating to remote sensing. We will have an upcoming webinar featuring a guest speaker from SEDAC talking about remote sensing applications relating to disaster risk management using SEDAC data (<https://arset.gsfc.nasa.gov/disasters/webinars/19-DRA>). There is also a Disasters program at NASA, and they have a portal for data here:

<https://maps.disasters.nasa.gov/>

Question 5: Which NASA/ESA satellite platforms and instruments are used for measuring NDVI?

Answer 5: Multiple sensors can be used. MODIS has standard NDVI and EVI (Enhanced Vegetation Index) products available. You can obtain MODIS NDVI products via NASA Worldview here: <https://worldview.earthdata.nasa.gov/>. The standard products from MODIS available on NASA Worldview are monthly and 16-day (with an 8-day rolling NDVI product also available). [Landsat](#) can also be used for NDVI, as it has a higher spatial resolution (30 m), it is useful for things like monitoring large agricultural fields. The ESA sensor [Sentinel-2](#) can also be used for NDVI.

ARSET had an Advanced Webinar “[Creating and Using Normalized Difference Vegetation Index \(NDVI\) from Satellite Imagery](#)” from 2016. You can see all the materials, exercises, and recordings from the webpage.

Live (Alex): Using 30 m Landsat data for SDG 11.3.1, as well.

(Mariano) For these indicators in particular, we’re using Landsat. We’re doing a time series of NDVI for different years. There’s a stack of 20-something bands we use for each year, but it’s from Landsat.



Question 6: What does impervious surface index mean and how do you calculate it?

Answer 6: (Mariano) It's an index that ranges between 0-100 and lower values are more correlated with no areas with no impervious surfaces (vegetation, no human build-up). Higher values mean buildings, roads, things constructed by humans, rocks. That's why we need to use the different thresholds like night time lights to determine what is human vs. natural. Check the Trends.Earth website for a section we added yesterday on SDG 11 on how we computed what we did is we used the global man-made impervious surface dataset produced by NASA. It was a product of impervious surfaces for 2010. And ESA and CCI land cover dataset to train a set of random forest models using Landsat data. We train it for 2010 and apply it to a series of years to get a time series. A step by step explanation is on the website.

Question 7: Are there efforts to take sustainability methods such as this, and expand to a watershed level as a management unit? Therefore, conserving (creating a sustainable area) urban, suburban, agriculture and wild areas?

Answer 7: (Mariano) Probably Dennis may know more about this -

(Dennis) Not from our side, because we are more focusing on look of things since our mandate is on urban aspects of settlements. But there are opportunities emerging from the SDG framework mapping urban growth and land uses, like watershed management, it will be interesting to see within the framework this kind of application and what's required.

Question 8: Are all the world countries considered in the calculation of urban area changes?

Answer 8: (Mariano) So, yes and no. We did train the model globally. So we used, like Alex mentioned, regions to train different sets of models for different regions. We tried to run it globally and it covers most cities. But we didn't find training data for some regions. There were limitations on the datasets. You'll see that on the website (http://trends.earth/docs/en/background/understanding_indicators11.html) - we posted some maps and statistics on cities we weren't able to generate data. Most of them being small island-states. It represents a pretty small number. We did an assessment on 224 cities. <2-3% were ones that didn't have data) There's some without data - some of the data isn't working well. That's important why you assess the results for your area. It's working well for ~85% of cities globally.



Question 9: How do we export these layers to a tif? A trial on the previous exercises leads to a 35 band raster and is difficult to identify which band represents which layer.

Answer 9: (Alex) Yes, that's a good question. Right now there's a little trick to see what those layers are if you go to the load data tool in trends.earth. It will show you the names of each band. I'll try and get this up later today. If you go to the trends.earth website, I'll add under the documentation section a description of each of the outputs with each band and how they're coded. It's currently available in the JSON file with the download, but that isn't very accessible. So stay tuned with more info in the coding.

UPDATE: A new page is available on our website that explains the coding for all outputs from Trends.Earth: <http://trends.earth/docs/en/about/layers.html>

Question 10: I would like to know about the validation process of urban map? Uncertainty?

Answer 10: (Mariano) Assessing a dataset like this one is challenging, like you can imagine. It's why we decided to make a tool that the user can use to customize. We did several things. Firstly, we assessed the error we were finding between our impervious surface dataset and the one we used for training. Comparing those two we found the Root mean square area. For an index ranging between 0-100, an error of 10% is relatively low. It pointed us in the right direction. The other assessment we did is now posted on the website. Where we took every capital city in the world - 200 cities - and used analysis on the largest cities in the U.S. that added another 24 cities. We compared the results of the tool to high res images to see how the build up index matched and see how well it was representing build up. For that we did a subjective evaluation from 0-5, 0 meaning no data, 5 meaning it worked perfectly well. And you can see on the website what are the values, and see which cities had good data, bad data, and what are the parameters. The third verification we're planning to do, relies on you. There's so much assessment we can do - we really need the input of users running the tool. When you go to the tutorial/step-by-step guide, there's a little box where we request your feedback. We put together a google survey, there's a 30 second survey. We're asking what are the parameters you used, and what's your assessment? Was it useless and didn't provide any info? (0 being bad, 5 being perfect) That is going to really inform us on the quality/usefulness of this data. We are trying to produce a verified product, but it's challenging with a global dataset.



Question 11: How can we import our own population data?

Answer 11: (Alex) It depends on what you mean. If you've noticed in the final summary table output for 11.3.1, if you scroll down in the excel sheet, there's a table on population. You can replace those numbers and the excel spreadsheet will recalculate. If you're talking about importing a gridded population dataset you have, that isn't directly supported in Trends.Earth

Question 12: How well does Landsat data pick up vertical build-up on the interior of cities (e.g. increased skyscrapers or high-rise buildings)? How to distinguish built-up areas from bare lands in arid and semi-arid regions using landsat data?

Answer 12: (Mariano) Not well at all. With Landsat, you're picking up the footprint of the built-up area. You can't differentiate between a house and a skyscraper. That's why we're measuring build up, not the volume (in terms of how high). 2nd question: there's 2 things here we're doing. First, built-up areas have a constant signature throughout the year. It doesn't really change. Throughout the seasons it's always the same. In areas where there's bare lands, it's bare but there is some vegetation, changes in humidity when it rains, so there is variability. So we're trying to capture those characteristics by using statistics of time series by reflectance of time series over time. Using median of each bands, and including variability over time. If build up is constant, there will be variability when it rains on bare ground. That is when we're building and training the model. At the end, the last control is what you're doing in the urban mapper - using the nighttime lights. That's the final control to differentiate between the two areas. If you only use impervious surface, the tops of mountains will come up as build up, but these areas have no nighttime lights because no people live there. So we're using that to filter out and only look at cities. You won't find one set of parameters that work for globe or country. You need to look at city-scale. Define the parameters that work well in that area. If you have a mountain 100 m away, it may not work there, but you should be running the analysis on the city boundary.

(Alex) Some of you may have noticed there's a new version of the QGIS tool (v0.66). You might get a notification if you try to run a large analysis - since we recommend you run this at a city scale - if you specify an area greater than 10,000 km² - you'll get a message that pops up. With the older versions, you won't get the message, but your task will fail if you run a large area. So if you get a failure, try a smaller area. And upgrade to the new plugin, which will automatically detect that for you.



Question 13: When defining urban pixels, do you define a 500m buffer or 564m, to correspond to 1 square KM area?

Answer 13: (Mariano) we use 500 m as the buffer radius around each pixel

Question 14: How are the yearly Landsat-stacks created? Are you using all available scenes (maybe depending on cloud coverage) or are there certain months preferred/chosen for specific regions?

Answer 14: (Mariano) What we're doing - the time series we're creating starts in 1998. We're producing a stack for 1998, 2000, 2010, 2018. For each of those years - yes, we filter for cloud coverage and quality. We don't filter the images, we remove the clouds and use all the useful information the scene has. In many areas, esp for earlier years (1998, 2000), availability is a limiting factor. For each year, for 2000, it's combining 1999, 2000, and 2001. So we're using info from year prior and after from the "year." We're filtering, masking, and using images throughout the year and all images available. We have different ways of summarizing them for some of them we use medium - the measure of centrality within the year of viability. We include median and standard dev to determine viability and pixels constant over time and changing in terms of reflectance - which are associated with more natural

Question 15: In the last Q&A session, you mentioned that you have operated in Liberia just as in Eastern African countries. Please under which ministry or office/institution can I connect with your activities? I am currently resident in Liberia, and would love to assist there in any capacity, having acquired this knowledge and many others previously here.

Answer 15: (Monica) First of all, we have a Conservation International office in Liberia. We have a GIS Analyst based out of there, as well. In March 2018, when we held the Capacity training workshops with UNCCD we had a member from Ministry of Environment as the GIS expert in that workshop. We've worked closely with him in the last 1.5 year. He's been using the tool for different purposes. Email us at trends.earth@conservation.org and we can connect you to our colleagues.

Question 16: Why is Random Forest the selected algorithm instead of another IA algorithm?



Answer 16: (Mariano) We use Random Forest - there are others which could have been used, but RF has proven to be successful for this type of analysis. We're running this analysis in Google Earth Engine, and it's one of the models available. If someone has a suggestion that's more appropriate, please reach out. We're constantly trying to improve the dataset.

Question 17: When selecting the population definition we have two options. Which one is higher in general? Which one is optimal for analysis?

Answer 17: (Alex) This is referring to urban indicator option to choose a population dataset - you can choose from UN corrected population statistics or those consistent with national estimates]. Neither is higher in general, which one to choose is dependent on what you're using the data for. If you're working with a government ministry, you may want to use one that matches with that.

Question 18: Is there a way to use a watershed boundary for the buffer instead of just a circular buffer?

Answer 18: (Mariano) Yes. We do not have watersheds pre-loaded in the toolbox, but you can use your watershed shapefile or a polygon in the tool. In the window where you define the study area, you can use your own watershed file to run your analysis on that. We do have the 100x100 km area limitation. So if your watershed is too big, you won't be able to run it.

Question 19: Is the model compatible with tropical countries? We also have a very heterogeneous land cover + bad zoning practice, hence my question.

Answer 19: (Mariano) The model, in terms of impervious surface model, is, I imagine, what the q is referring to. You're going to notice when you check the verification, that it works pretty well in tropical countries - for most of them. There are some regions where cloud coverage + small archive of landsat images, the model isn't performing well. For most tropical areas it works great. Again, you'll notice this when you test the google mapper for your city. You'll see that for a few cities. If the image availability is low, you'll see it isn't capturing well where build up is.



Question 20: Is there a general rule or method for determining the threshold for the different indices (ISI and NLI) because for example we can be two person working about the same region but choosing different values, it will impact the calculation? and is the gridded population data exists for Morocco and if we can exporte urban layer to use it for other analysis?

Answer 20: (Mariano) 1st question: we do not have a set of recommended values. We provide what we consider the best values for different regions. It's true, if two people with different criteria choose different parameters, the results will vary. Not greatly, but it would mean your build up area would increase more or less, but the trends in SDG11 will not change significantly. If you're trying to get built-up area, the numbers will vary. But if you're looking at the trends, it won't be affected very much. The 2nd part of the question: The gridded pop data - yes, exists for Morocco, it's global datasets. How good they are will depend on data provided to build the model, so it has its limitations. All the layers we produce are GeoTIFFS. All the results you download can be used in any other GIS software.

Question 21: What is the basis for defining the percentage of urban and suburban?

Answer 21: (Dennis) We are using the approach called city-urban extent for city urban definition. Within every 1 sq km of each pixel, we look up how many other pixels are built up. If 50% of surrounding pixels are built up, it's an urban area. If it's within 25%-50%, that area is a suburban area.

(Mariano) In the tool, these parameters can be changed.

Question 22: Is there any evidence to use only 1 year before and after the aimed year to deal with the data? :

Answer 22: (Mariano) Our approach - what we're trying to do here - we'd ideally like to only use images from the target year. We started that way, and realized there weren't enough data for many regions. Especially in the U.S. it's a huge archive, but it's a compromise. We started with the target year, and slowly increasing the time period. Our goal is to make that window as narrow as possible, but have data for most places. It was a compromise between ideal and data availability.

Question 23: How long should the processing take? I submitted a job and it is still running after 20 minutes.



Remote Sensing for Monitoring Land Degradation and Sustainable Cities

SDGs

July 9-23

Answer 23: (Mariano) Great question - it depends on the size of the area that you're running it on. Usually the tests we've been running take about 30-40 minutes to run. That being said, we're running these analyses on GEE. There is - when a lot of tasks are submitted, a lot of the tasks take longer. Since a lot of us may be submitting tasks, it may take longer. Have a bit of patience, and as long as you don't get a failed message, you'll be fine. Even if you wait for an hour and it seems like a long time - we discussed this with UN Habitat and NASA - it's an hour, but if you had to do analysis yourself, it would be weeks of work. So even if it takes a little time, don't get frustrated. Have a coffee break.

(Monica) I want to remind you that if you have that download window open, it isn't going to auto refresh with the results are done. So make sure you click refresh list in the downloads window - so it may be done every time so you get the latest results.

Question 24: The urban area mainly mix with sand, bare land, rock outcrop..so how to improve the classification for urban extraction? While working at regional level, built-up area mix with these land features?

Answer 24: (Mariano) Are you - if you could send us an email where you're running it and parameters, we can provide a more useful response. There are some areas where these mixups happen. But for most of the areas it's not. If you find an area where this is happening, but you think the overall dataset is providing insights, you can download the results and do some cleaning up in the GIS software. (Like QGIS) It would be useful to know the area you're testing, so we can give you a more informed response.

For more questions, go to our List Serve for questions and answers on Trends.Earth:

<https://groups.google.com/forum/#!forum/trendsearth>