The EY Open Science Data Challenge

Orientation session

Building a better working world

Agenda

Welcome		Beatriz Sanz Saiz EY Global Consulting Data and Analytics Leader
Impact of the challenge		Prof. David Lodge Director of Cornell Atkinson Center for Sustainability
The role of satellite imagery and its potential		Brian Killough, PhD EY Senior Consultant, Ernst & Young LLP Former NASA Engineer and Scientist
How to get started		Saurabh Agarwal EY Manager, Tech Consulting, EY GDS LLP
Q&A		Tom Schoeters EY Open Science Data Challenge Talent Lead

Welcome

Beatriz Sanz Saiz
EY Global Consulting Data and Analytics Leader

Please watch the welcome video via the recording here.

Impact of the challenge

Prof. David Lodge
Director of Cornell Atkinson Center
for Sustainability

2023 topic

Food security

Eight hundred twenty million people regularly go to bed hungry.

The UN Food and Agriculture Organization (FAO) predicts that the agriculture industry will need to produce 70% more food while only being able to use 5% more land in order to nourish those already suffering from hunger, as well as two billion people expected to be added to the global population by 2050.

Source: https://www.ey.com/en_us/digital/digital-agriculture-data-solutions

Rice production in Vietnam

Rice is a staple food for 4 billion people, more than half of the world's population.

It is grown in more than **100 countries**, and **80% of rice producers** are smallholders that farm less than **2 hectares** of rice.

Global rice consumption was **440 million tons in 2010**.
It is estimated to increase to **650 million tons** by 2050, a **48% increase.**

But both the amount of land available for rice and the yield growth rates are in decline, irrigated rice yields in developing countries are forecast to decrease by **15% due to climate change.**

Vietnam is the world's fifth-largest paddy rice producer.

Vietnam is also one of the most vulnerable countries to climate change. Calculations from the World Bank suggest that **Vietnam lost $10 billion in 2020, or 3.2% of GDP, to climate change impacts.**

How we will feed the world

Improve predictions of crop yields

- Crops are susceptible to environmental factors and increasingly to climate change. Science can help anticipate such risks and provide adaptation measures that allow farmers to increase crop yields.
- Accurate yield predictions help governments to estimate food supply. It allows policy makers to take data-driven decisions on importing or exporting crops to assure national food security.
- In addition, NGOs and government agencies can proactively take targeted measures to support rice farming communities, to educate farmers on climate-resiliency, and therefore strengthen the national food supply.

More efficient use of farmland

- For the last 20 years, global land area for agriculture has been declining. Increasing efficiency of land use is therefore a must, despite the fact that rice yields, between 2010 and 2050, have been predicted to decrease by 15% due to climate change.
- Data science can help identify the sustainable intensification measures to be adopted at each specific location, which will translate in higher land use efficiency and crop yields.

By working closely together, governments, NGOs and scientists can use data-driven science to stay ahead of the curve and proactively assist farmers to become more efficient, climate-resilient and productive. Longer-term investments can be designed, with the goal of assuring reliable food supply for a growing population.

The role of satellite imagery and its potential

Brian Killough, PhD
EY Senior Consultant, Ernst & Young LLP
Former NASA Engineer and Scientist

What will participants do?

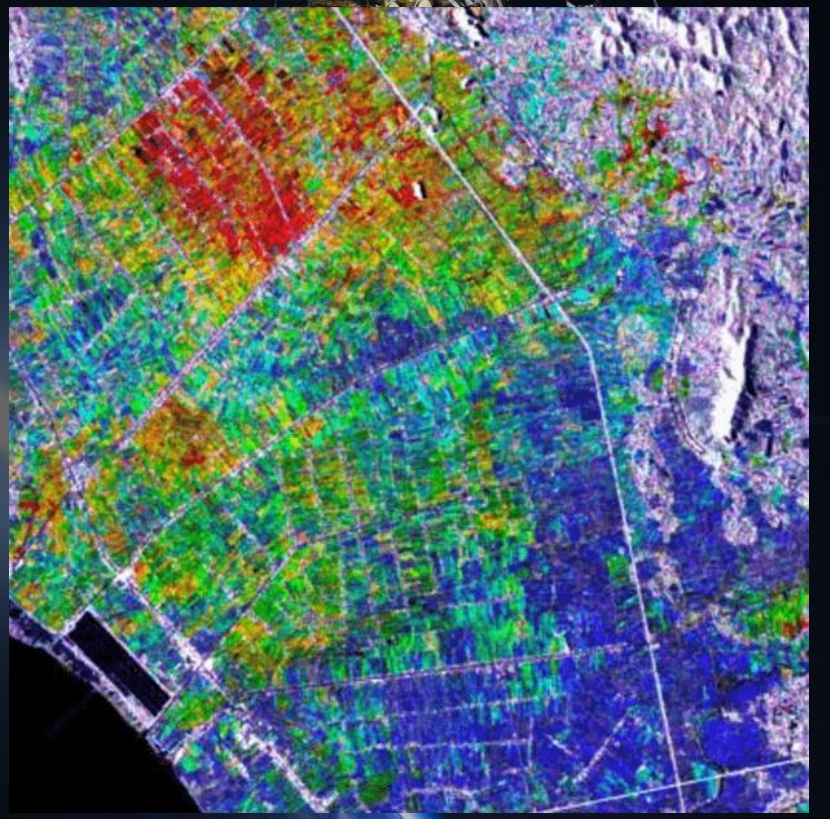
There is a host of data available that can help the people of Vietnam better plan and adapt rice production in a changing climate. **But that data must be properly analyzed and synthesized.**

Participants will use data from Microsoft's Planetary Computer including Sentinel-1 (radar) and Sentinel-2 (optical) satellite data to build identification and prediction models that:

Level 1 Identify the areas where rice is cultivated.

Level 2 For any identified area of rice cultivation, forecast yield.

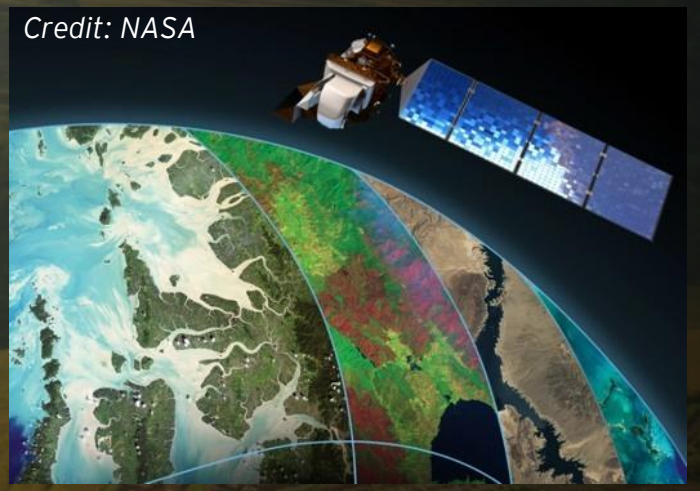
Participants will build skills in understanding and interpreting satellite data, coding in python using Jupyter notebooks and testing machine learning techniques including classification algorithms and prediction algorithms.

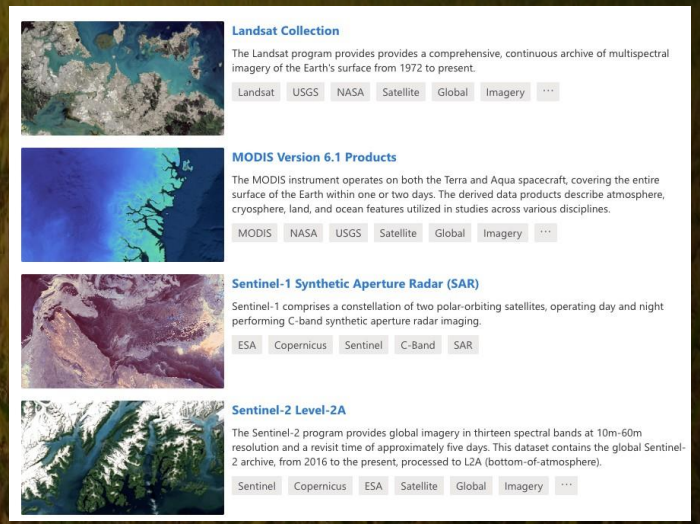


Credit: Brian Killough, EY

What is the value of satellite data?

- Satellites give us a continuous and unique synoptic view of the earth. Such measurements are impossible from the ground at large scales.
- Many environmental conditions can be measured by satellites … water extent, vegetation, land type, temperature, and precipitation.
- For this challenge, we have focused on two types of satellite data … optical (visual bands) and radar (non-visible bands).
- These free and open data come from NASA (Landsat) and the European Copernicus program (Sentinel-1 and Sentinel-2).
- The EY Open Science Data Challenge suggests the use of satellite data from the Microsoft Planetary Computer. This cloud-based data is easily accessed using Python notebook scripts.

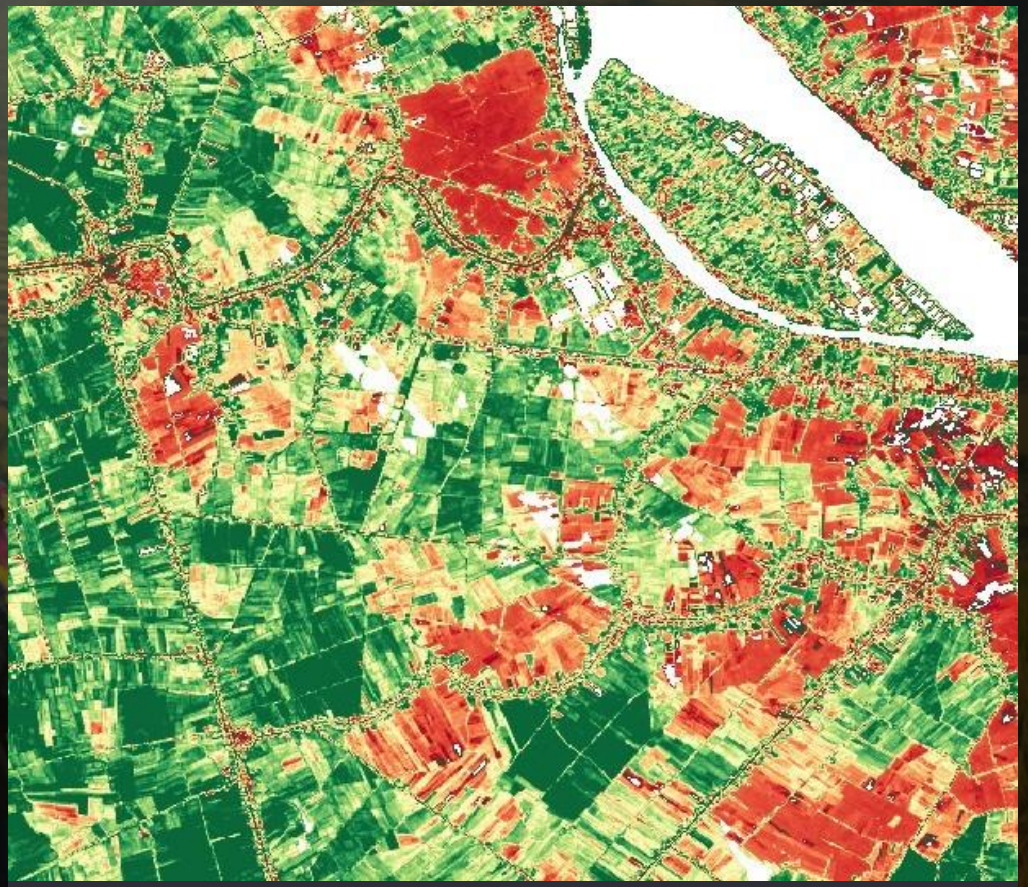




Credit: Microsoft

Optical data over rice crops

- Landsat views a location on the earth every eight days with two missions. Each pixel is 30 meters in resolution.
- Sentinel-2 views a location on the earth every five days with two missions. Each pixel is 10 meters in resolution.
- Optical data is obtained by "taking a picture" of the surface using a multi-spectral instrument.
- This data gives us information about the "greenness" of the crop.
- Optical data cannot penetrate clouds. So, in Vietnam, where it is cloudy two-thirds of the time, we need to filter the clouds out of the images.

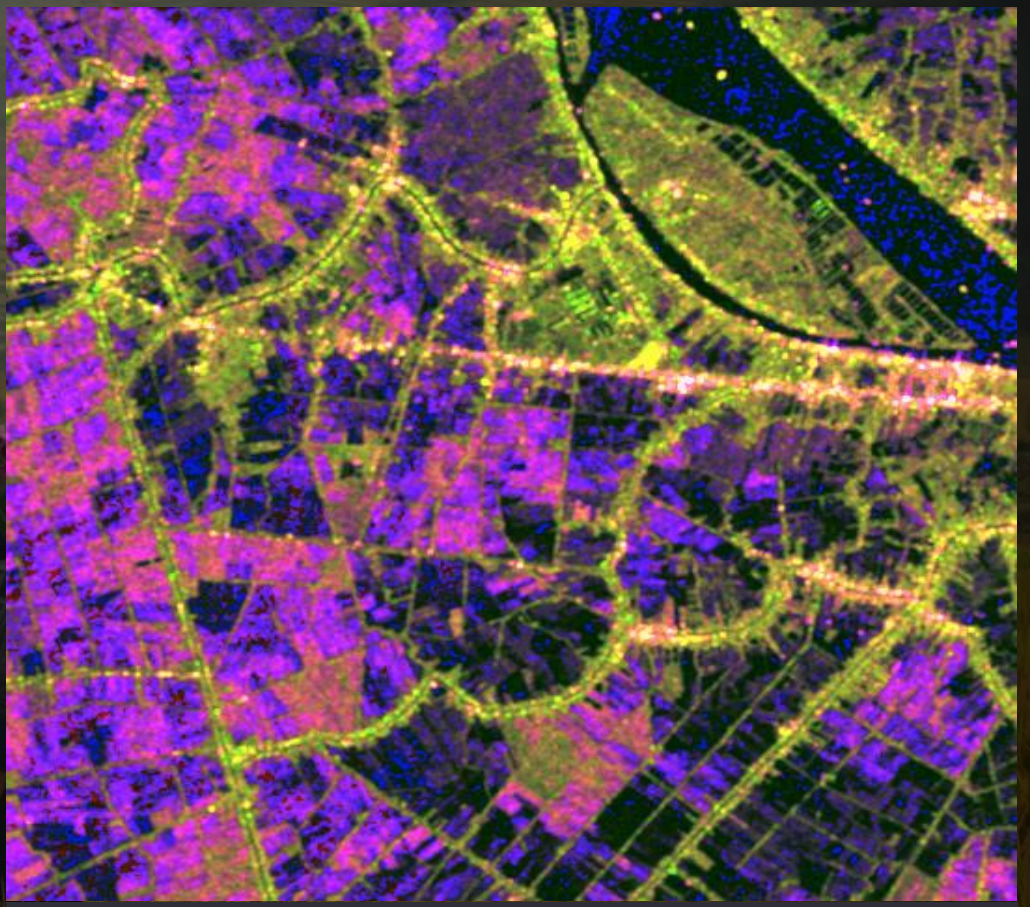


An example Sentinel-2 vegetation index product over rice crops in Vietnam (31-Dec-2021)

Credit: Brian Killough, EY

Radar Data over Rice Crops

- Sentinel-1 views a location on the earth every six days with two missions. Each pixel is 10 meters in resolution.
- One of the Sentinel-1 missions failed in December 2021, since then we only get 12-day revisits.
- Radar data is obtained by sending a pulse of energy to the ground and measuring the returned scattering response at the satellite.
- This data gives us information about the "structure" of a crop which is very different from optical data.
- Radar data has a unique capability to penetrate clouds which yields a consistent view of the ground.

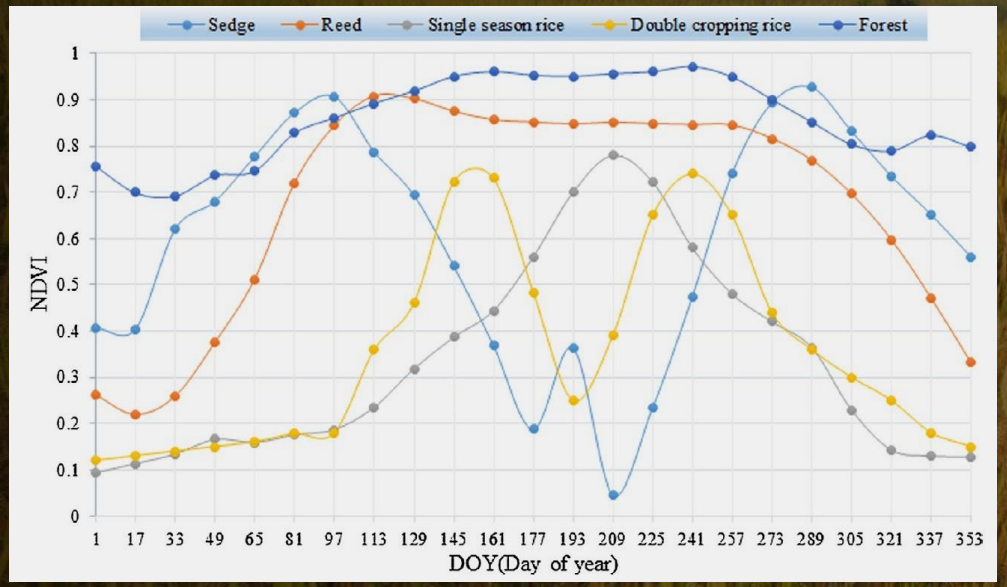


An example Sentinel-1 product (RGB=VV,VH,VV/VH) over rice crops in Vietnam

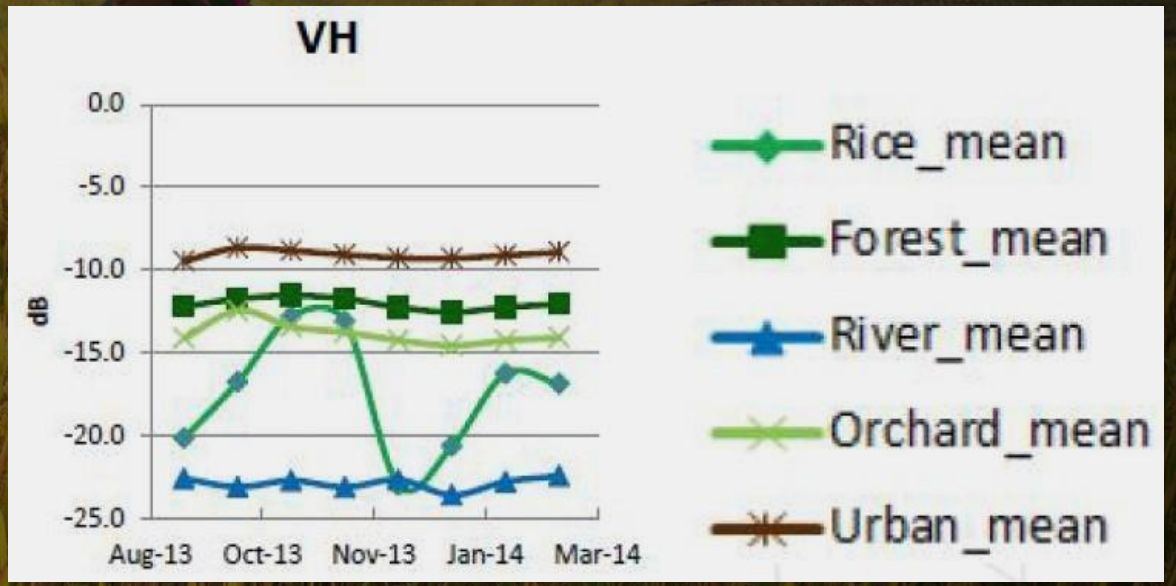
Credit: Brian Killough, EY

How can we use satellite data for rice crops?

- We can use the bands or mathematical combinations of the bands (e.g., indices).
- We can view the phenology or the vegetation growth over time.
- We can compare statistical differences in the phenology response (e.g., min, max, range, period and shape).
- Differences in the band combinations and phenology response can help detect rice versus non-rice. locations (Challenge Level 1) and predict differences in yield (Challenge Level 2).



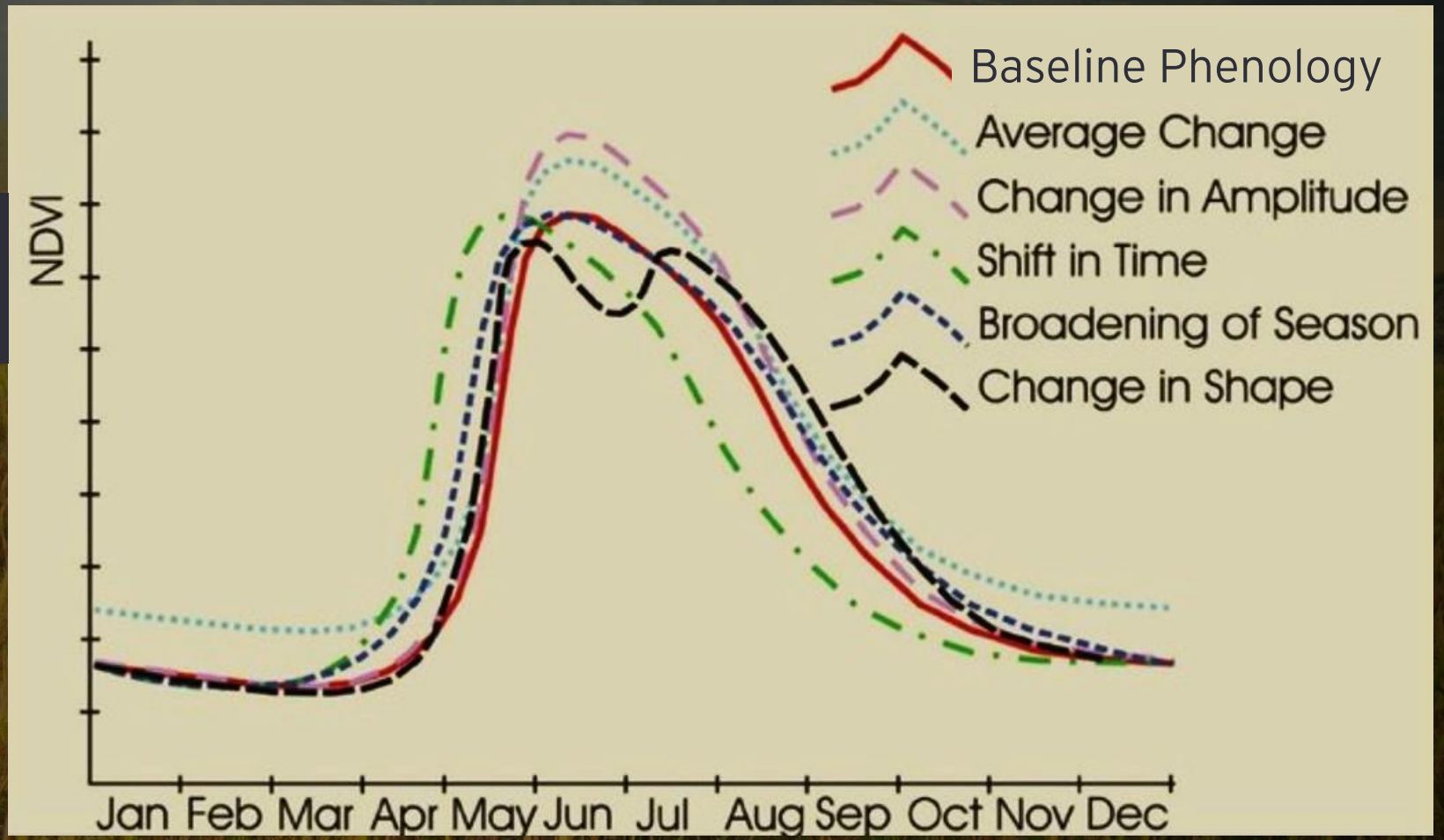
Credit: Yaotong Cai, et.al., 2019.



Credit: Lam Dao Nguyen, VNSC

Phenology (vegetation response) variations can be related to yield to build a machine learning model

Phenology curves can be created for both optical and radar data



Credit: Keith McCloy, Remote Sensing, 2010

How to get started

Saurabh Agarwal
EY Manager, Tech Consulting, EY GDS LLP

Please watch the "how to get started"-video via the recording here.

Please visit challenge.ey.com for more information or contact us via email at datachallenge@ey.com

Q&A

Tom Schoeters
EY Open Science Data Challenge Talent Lead

2023 EY Open Science Data Challenge

Prizes for External Participants

Level 1

Top 1
Submission selected and awarded

One global winner receives
US$ 2,500

Level 2

Top 10
Semi-finalists selected and asked to provide content package

Five global finalists
will be selected by GDS team out of the 10 semi-finalists

Judging panel
selects a global winner and two runners up, who will be invited to the awards ceremony

Global winner receives
US$ 10,000

1st runner up receives
US$ 5,000


EY
Building a better working world
How will you harvest data to help solve world hunger?
Join the EY Open Science Data Challenge
The better the question. The better the answer. The better the world works.
Microsoft

EY

Disclaimer:

- The views of third parties set out in this webcast are not necessarily the views of the EY organization or any of its member firms.
- This material has been prepared for general informational and educational purposes only and is not intended, and should not be relied upon, as accounting, tax, legal or other professional advice. Please refer to your advisors for specific advice.
- Neither the EY organization nor any of its member firm thereof shall bear any responsibility whatsoever for the content, accuracy, or security of any third-party websites that are either linked (by way of hyperlink or otherwise) or referred to in this presentation.

EY | Building a better working world

EY exists to build a better working world, helping to create long-term value for clients, people and society and build trust in the capital markets.

Enabled by data and technology, diverse EY teams in over 150 countries provide trust through assurance and help clients grow, transform and operate.

Working across assurance, consulting, law, strategy, tax and transactions, EY teams ask better questions to find new answers for the complex issues facing our world today.

EYG no. 000755-23Gbl
ED None

ey.com