FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

AGRILINKS

Geo-Referenced Data to Inform Earth Observations Modeling for Agriculture: A Discussion among Collectors, Users, and Aggregators

Speakers: Kiersten Johnson, USAID Bureau for Food Security*;* Paul Tanger, USAID Bureau for Food Security; Philomin Juliana, International Maize and Wheat Improvement Center; Jawoo Koo, International Food Policy Research Institute; Estefania Puricelli, NASA Harvest; Narendra Das, NASA Jet Propulsion Lab; Raghavan Srinivasan, Texas A&M University; Rober Hijmans, University of California, Davis

Moderator: Julie MacCartee, USAID Bureau for Food Security

Date: *May 29, 2019*

FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

Kiersten Johnson, USAID Bureau for Food Security

Dr. Kiersten Johnson is an international population and health demographer with 20 years of experience conducting in-depth analyses of Demographic and Health Surveys (DHS) and Service Provision Assessment data for the USAID/Bureau for Global Health's MEASURE DHS project, the USAID/Climate Change Initiative's Climate Change Resilient Development project, and the USAID/Forestry and Biodiversity Office's Measuring Impact project. She has led innovative developments in the integration of NASA's satellite remote-sensing data into the DHS to explore associations among climate, environment, and health and nutrition outcomes. She is currently serving as a monitoring and evaluation technical advisor in USAID's Bureau for Food Security Monitoring, Evaluation and Learning Division for the US Government's Feed the Future Initiative.

Paul Tanger, USAID Bureau for Food Security

Paul is an Agriculture Research Advisor in USAID's Bureau for Food Security where he manages biotechnology and crop improvement projects. Previously as an AAAS S&T Fellow at USDA NIFA, Paul led the launch of a new initiative focused on data science in agriculture, as well as developing open data policies, examining and visualizing impacts of research funding, and coordinating plant breeding investments. Paul's Ph.D. work at Colorado State University in collaboration with the International Rice Research Institute pioneered the use of advanced techniques to measure crop traits in the field for improved agricultural crop breeding. Previously, Paul worked in the technology transfer space, as well as a project manager in the financial services industry.

Narendra Das, NASA Jet Propulsion Laboratory

Narendra Das joined the Jet Propulsion Laboratory (JPL), and has been conducting research in hydrology and microwave remote sensing research on land. He is currently a Research Scientist with the Radar and Algorithm Processing Section at JPL. He is the lead developer of the NASA SMAP mission high resolution soil moisture data. He recently joined the NASA Surface Water and Ocean Topography (SWOT) mission Algorithm team. He is the JPL PI for the NASA SERVIR projects, and co-developer of the Regional Hydrologic Extreme Assessment System (RHEAS) framework. He is instrumental in deployment and operationalizing the RHEAS platform in Lower Mekong Basin countries, Mekong River Commision, and East African countries through the NASA SERVIR project.

Robert Hijmans, University of California, Davis

Robert Hijmans works on data science applications in international development, with an emphasis on modeling spatial variation. His research interests include agricultural biodiversity use, climate change, remote sensing of agriculture, sustainable intensification, and nutrition. He has developed widely used software such as the R package "raster" and spatial databases such as Worldclim and GADM. He directs the Geospatial and Farming Systems Consortium of the Feed the Future Sustainable Intensification Innovation Lab. He has a PhD from Wageningen University and has worked in Africa, Asia, and Latin America.

Philomin Juliana, International Wheat and Maize Improvement Centre

Philomin Juliana is an associate scientist and a bread wheat breeder at the International Wheat and Maize Improvement Centre (CIMMYT), Mexico working with Dr. Ravi Singh, as part of the Delivering Genetic Gain in Wheat (DGGW) and the USAID Feed the Future projects. She holds a Ph.D. and Masters degree in Plant Breeding and Genetics from Cornell University, and a Bachelors in Biotechnology from Tamil Nadu Agricultural University, India. Her current research involves evaluating genomic selection and high-throughput phenotyping based predictive modelling for traits like grain yield, disease resistance and quality in CIMMYT's bread wheat breeding program. She also works on genome-wide association mapping and genomic fingerprinting CIMMYT's wheat germplasm for key traits.

Raghavan Srinivasan, Texas A&M University

Raghavan Srinivasan, Ph.D. is a professor at Texas A&M University and director of the Spatial Sciences Laboratory at Texas A&M. He has become known and respected throughout the world for his developmental work with spatial sciences and computer-based modeling, especially the Soil and Water Assessment Tool or SWAT model. His research and its applications have contributed to long-lasting changes in natural resource assessments and development of management system options, currently being used in more than 90 countries. Over the past nine years, he has conducted more than 60 international workshops for students and professionals in more than 20 countries and the demand is increasing each year.

Estefania Puricelli, University of Maryland, NASA Harvest

Estefania has a Master's in Agricultural Economics from the University of London. She completed her undergraduate studies at the University of Buenos Aires, where she received degrees in Agricultural Economics and Farm Management as well as Economics. She collaborates with GEOGLAM, the G20 initiative that monitors crop conditions globally and was created together with AMIS. Estefania manages the monthly AMIS Market Monitor report, the primary product of the platform. Estefania is part of the Coordination team of the Agriculture Monitoring in the Americas (AMA). The initiative focus on strengthening national systems' monitoring capabilities through Earth observations (EO) in the Americas. Estefania also co-leads the NASA Harvest working group on Markets and Trade.

Jawoo Koo, International Food Policy Research Institute

Jawoo Koo is a Senior Research Fellow at the International Food Policy Research Institute, based in Washington, DC. Jawoo has almost 20 years of experience applying geospatial analyses and large-scale biophysical modeling in agricultural research and development. He received Ph.D. at the Agricultural and Biological Engineering Department at the University of Florida. He co-founded the CGIAR Platform for Big Data in Agriculture program, supporting agricultural scientists to use open, multidisciplinary data and data science techniques for sustainable development impacts. He serves as the Global Coordinator of the CGIAR Consortium for Spatial Information, a geospatial Community of Practice in CGIAR and partner institutes.

FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

Geo-Referenced Data to Inform EO Modeling for Agriculture: A Discussion among Collectors, Users, and Aggregators

Paul Tanger

Agrilinks theme for May: Earth Observations for Food Security and Agriculture

- Showcase the many ways that EO data and applications are being used in ways that directly correspond to Feed the Future's strategic approach
- Creates space for researchers who work with EO data to engage in mutually beneficial dialogs with people whose work in food security and agriculture could benefit from the insights derived from these data
- Encourage to review other content from this month
- https://www.agrilinks.org/post/earth-observations-food-security-and-agriculture-applications-feed-future

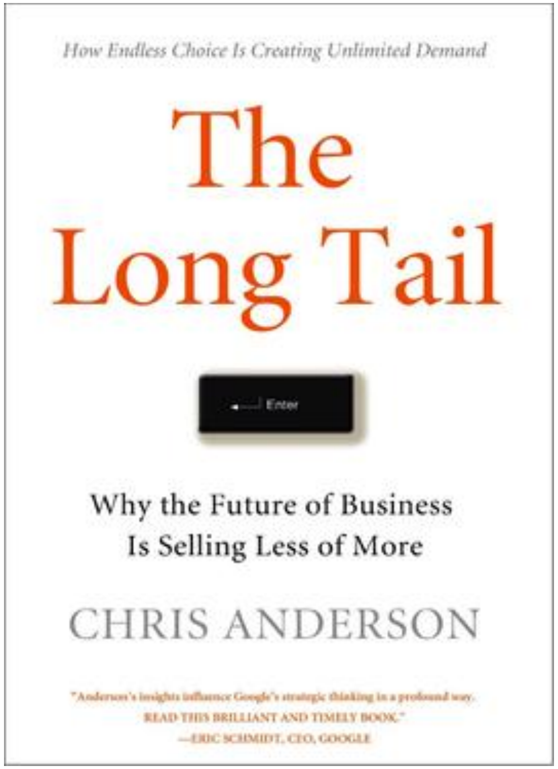
Rationale and Purpose

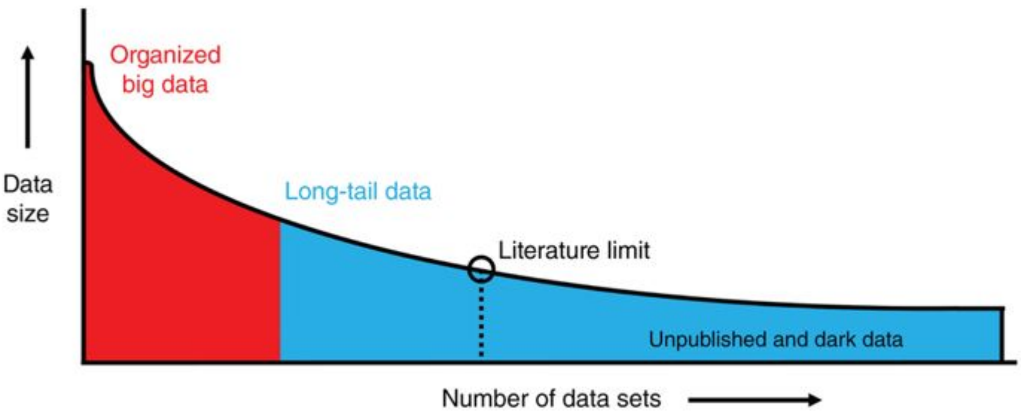
Current capabilities for effective at-scale monitoring of smallholder agricultural systems are limited at best. However, increased availability of high- and medium-resolution Earth observation (EO) data and modern computational capacities mean that the promise of EO data for effective agricultural monitoring has never been greater. Yet, a critical methodological challenge remains: the limited access to and/or availability of reliable, representative ground-referenced data, which are critical for developing and assessing the accuracy of EO-based models.

The purpose of this webinar is to convene practitioners that work on different aspects of ground-referenced data, so that this essential resource can be shared and utilized more efficiently and widely in the agricultural development space. Our panel of speakers will share examples of how they have generated, utilized, curated, and shared ground-referenced data, and will invite participants to share their own experiences and resources.

Rationale and Purpose

The long tail of data and how can we enable more use?





Do we need better organization, or different data?

Scope

- Earth observations and remote sensing
- Ground referenced data
- Use cases:
  - Monitor and respond to disasters
  - Manage resources
  - Detect emerging threats
  - Track trends
- Related topics:
  - PII
  - Sustainability
  - Consumption (decision support tools)

Outline

- Three groups:
  - Generators
  - Aggregators and Platforms
  - Modelers and users
- End users!

Agenda

9:40	Narendra Das, NASA JPL/CalTech
9:45	Robert Hijman, University of California, Davis
9:50	Philomin Juliana, CIMMYT Mexico
9:55	Raghavan Srinivasan, Texas A&M University
10:00	Estefania Ines Puricelli, GEOGLAM, University of Maryland
10:05	Jawoo Koo, IFPRI
10:10	Q&A
10:30	Open discussion
10:45	Kiersten Johnson, USAID

FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

Geo-Referenced Data to Inform EO Modeling for Agriculture: A Discussion among Collectors, Users, and Aggregators

Narendra N. Das,
Jet Propulsion Laboratory (JPL), NASA

Who are we and what are we doing?

- I am a hydrologist and microwave remote sensing person. I also work extensively on hydrologic modeling and crop modeling, and use remote sensing data in the study of hydrology and crop phenology.
- I work as a Research Scientist in the NASA's Jet Propulsion Laboratory (JPL)
- JPL has many Earth based environmental satellites to remotely sense geophysical state variables and parameters?
  - JPL has also developed the Regional Hydrologic Extreme Assessment System (RHEAS) i.e., a coupled hydrologic and crop model. RHEAS platform needs various farm management information for optimal modeling and crop yield data for validation

Geographies and crops

- What geographies are you mostly working in?

Two SERVIR projects: our areas of interest are East Africa countries (Kenya, Tanzania, and Ethiopia)

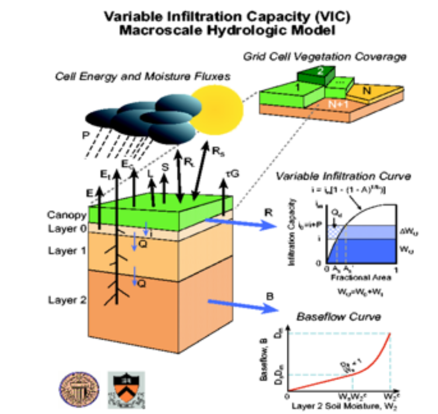
Lower Mekong Basin countries (Cambodia, Vietnam, Laos, Myanmar, and Thailand)

- What crops do you mostly work with?

We are working for cereal crops, mostly rice, wheat and maize.

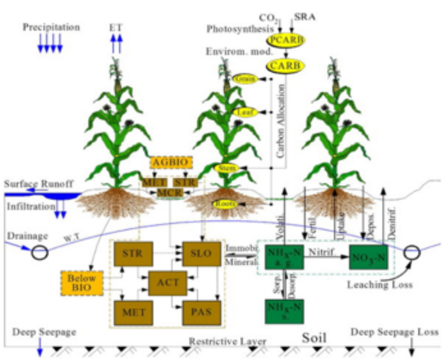
Hydrologic and Crop Modeling

JPL PI for the NASA SERVIR Project



VIC

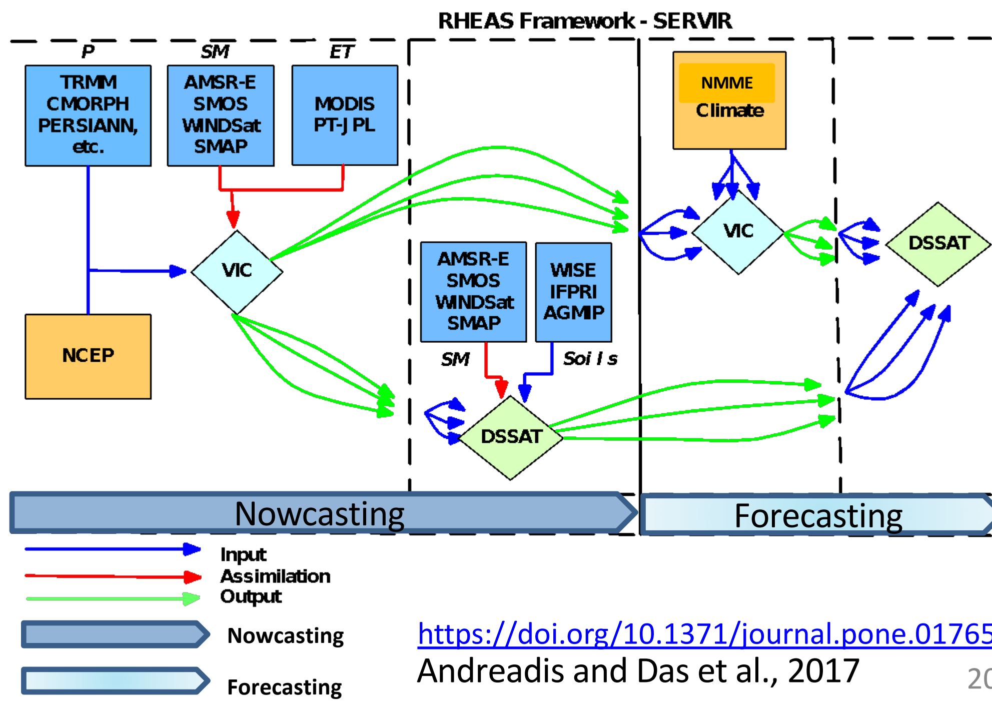
Semi Distributed Process-based model



M-DSSAT

Process-based Crop model, modified to work in ensemble mode. Assimilates profile soil moisture and LAI.

Regional Hydrologic Extreme Assessment System (<u>**RHEAS**</u>), it is a prototype software framework for hydrologic and crop modeling and data assimilation that automates the deployment of water resources and crop yield nowcasting and forecasting.



https://doi.org/10.1371/journal.pone.0176506

Andreadis and Das et al., 2017

FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

Hydrologic and Crop Modeling

A Case Study on Lower Mekong Basin (LMB) Countries

- Mekong: 12th longest river (4,350 km)
- LMB: Lao, Thailand, Cambodia, Vietnam, Myanmar
- Seasons: Wet (May-Oct); Dry (Nov-Apr)
- Annual Rainfall: 1200-2500 mm per year
- Agriculture and Fishing are the main activities
- Primary Crop: Paddy (Rice)

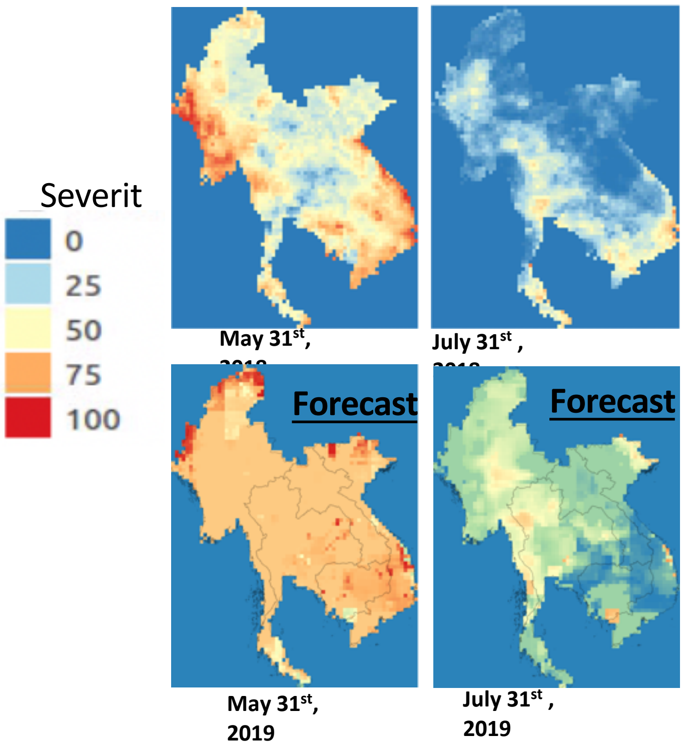
SERVIR Project: *A joint NASA and USAID undertaking that focuses on the developing countries to enhance the end-user capabilities on Hydrological and Agricultural applications for societal benefits.*

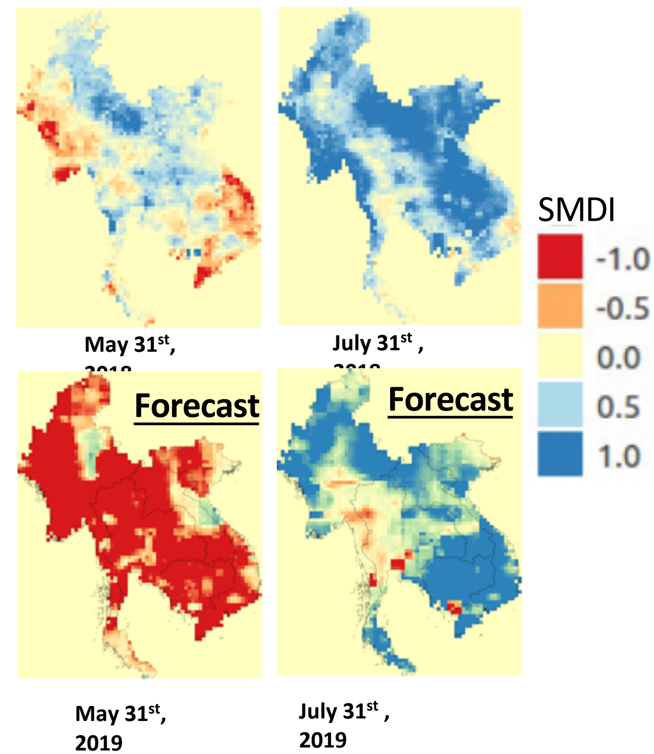
The objective *of this study is to provide drought onset/recovery and rice yield estimate nowcast and seasonal forecast*

Targeted Users: *Agricultural Ministries and*
Government Water Resources Agencies
Mekong River Commission

Hydrologic and Crop Modeling

A Case Study on Lower Mekong Basin Countries, Examples of RHEAS Outputs

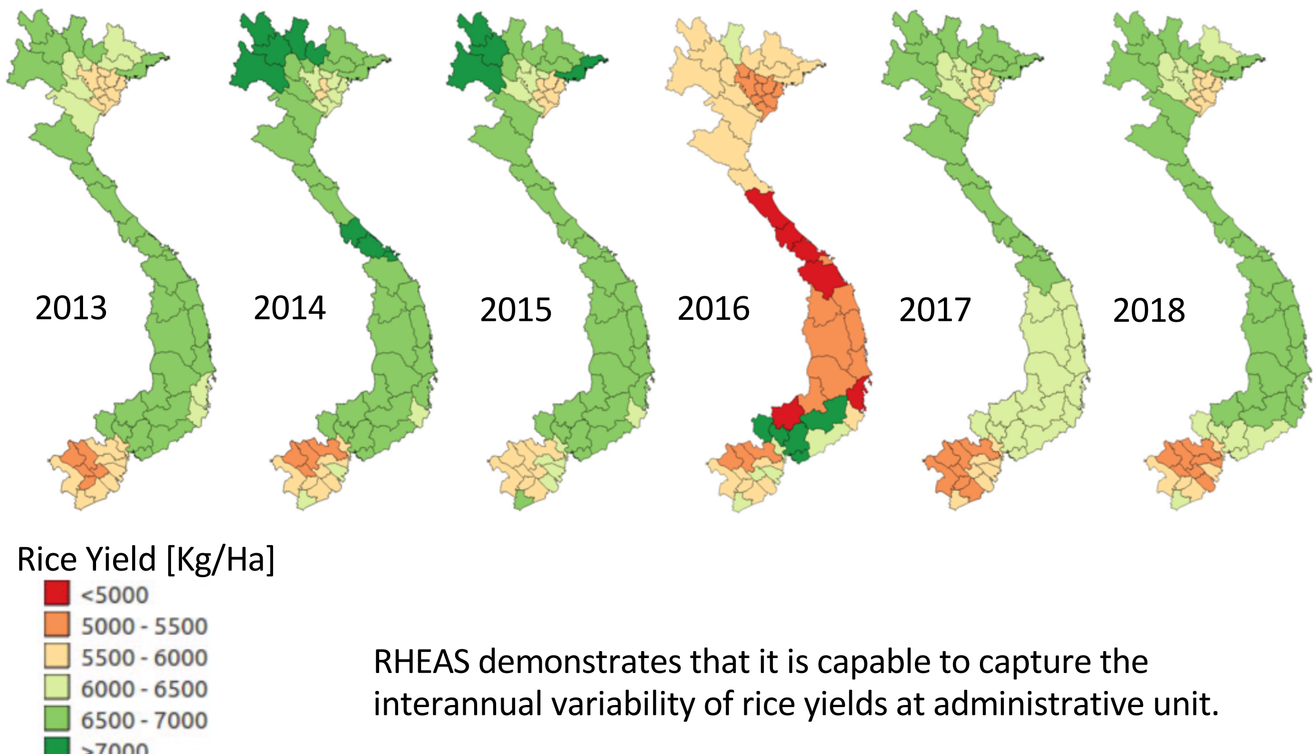


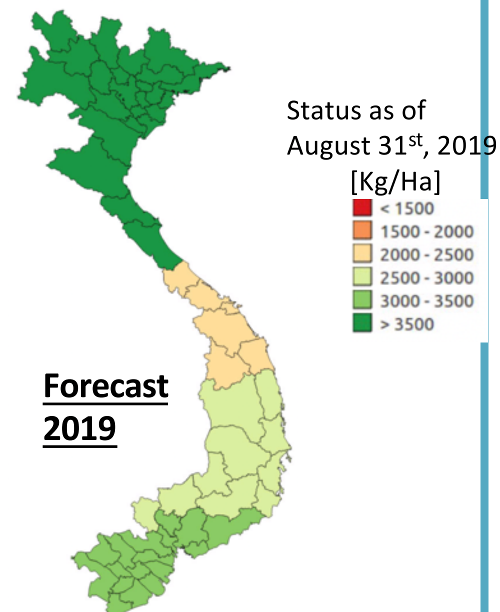


Comparison of Severity and SMDI shows that the forecast for 2019 in LMB countries are drier than in 2018.
The study clearly shows the impact of drought year (2016) on the rice yield in Vietnam.

Hydrologic and Crop Modeling

A Case Study on Lower Mekong Basin Countries, Examples of RHEAS Outputs





Forecast of Rice Yield Crop Status. The crop has not reached maturity. Probable harvest date is in mid October, 2019

RHEAS demonstrates that it is capable to capture the interannual variability of rice yields at administrative unit.

Das et al., 2019, submitted to Agriculture Water Management

Constraints

- What obstacles do you face in getting the data you need for your work?

Credible and quality data of farm management practices. Open access digital data not available or willingness to share.

- What obstacles do you face in maximizing use value of data already collected?

Data are haphazard (spatiotemporal inconsistency) and short time series

Opportunities

- Potential for this area

Active collaborations with participating agencies, and capacity building

- Next steps for you, your area, or other suggestions

Looking forward to work with USAID to maximize the potential and chances to get in situ data and crop related statistics at county/district level

FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

Geo-Referenced Data to Inform EO Modeling for Agriculture: A Discussion among Collectors, Users, and Aggregators

Robert J. Hijmans
University of California, Davis

Who are we and what are we doing?

- Professor, spatial data science in agriculture
- University of California, Davis
  Feed the Future Innovation Lab for Sustainable Intensification
- What is your organization doing in this space?
  - Assembling georeferenced crop and crop yield data
  - EO-based prediction of crop distribution and yield in space and time to support crop insurance
  - Open source software development (R) and capacity building to support this

Geographies and crops

- What geographies are you mostly working in?
  - Sub-Saharan Africa, South Asia, Andes, California
- What crops do you mostly work with?
  - Maize, rice, potato

Constraints

- *What obstacles do you face in getting the data you need for your work?*

An urgent need for more high-quality public crop yield data.

- *What obstacles do you face in maximizing use value of data already collected?*

Data access (even though this is *much better now*)

Opportunities

- Strengthen community data curation
- Strengthen open science --- allowing for entry of newcomers
- Targeted, intensive data collection for evaluation and comparison of modeling methods

FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

Geo-Referenced Data to Inform EO Modeling for Agriculture: A Discussion among Collectors, Users, and Aggregators

Thomas Payne (t.payne@cgiar.org)
Rosemary Shrestha (R.Shrestha2@cgiar.org)
Philomin Juliana (p.juliana@cigar.org)
International Maize and Wheat Improvement Centre (CIMMYT, Mexico)

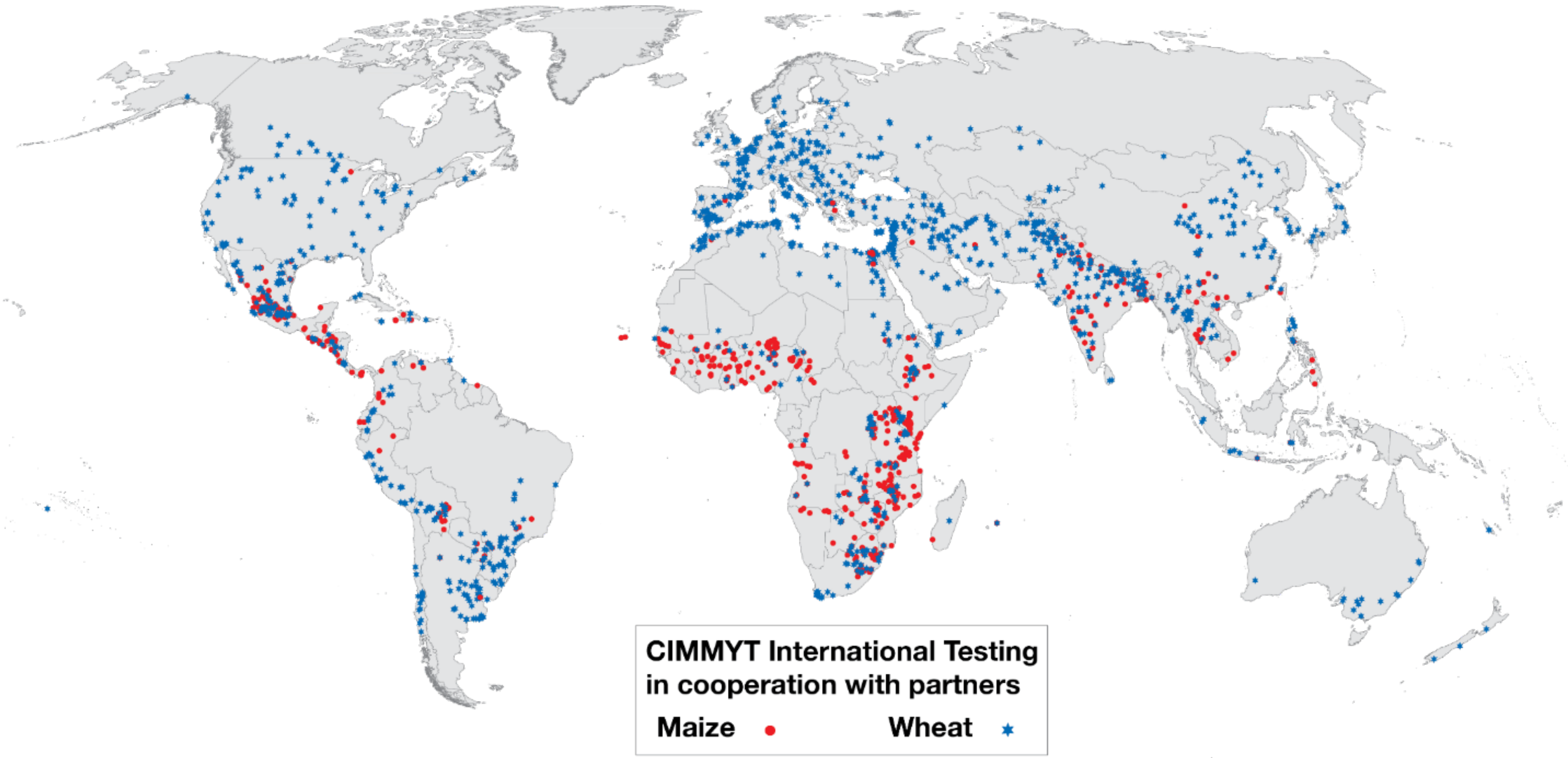
Mission

Maize and wheat science for improved livelihoods.

Vision

A world with healthier and more prosperous people – free from global food crises – and more resilient agri-food systems.

Global seed distribution network

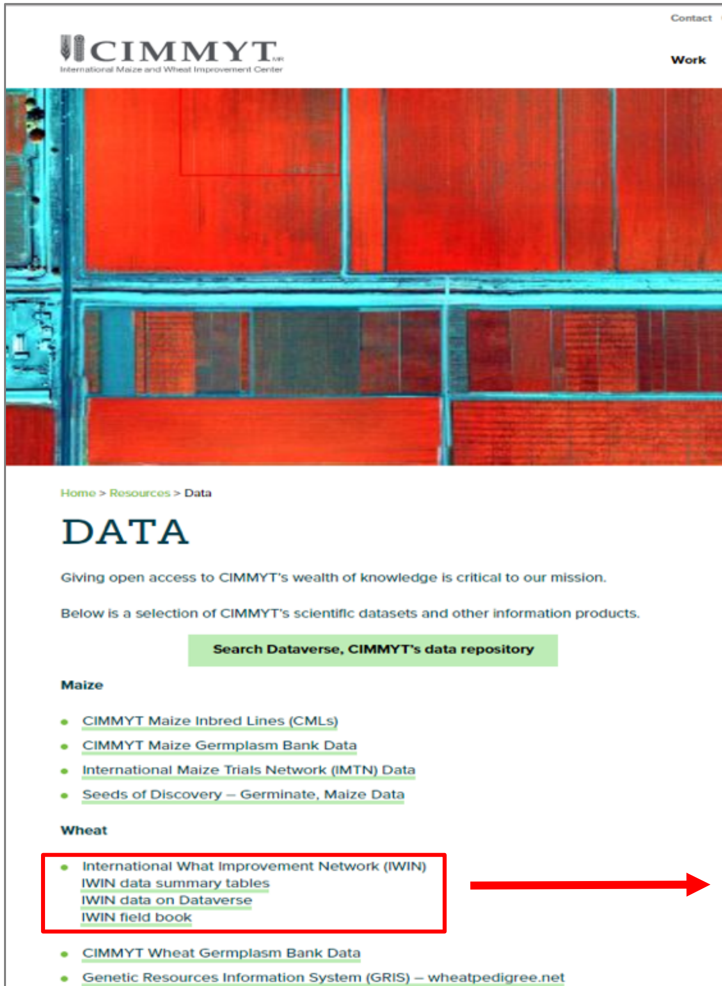


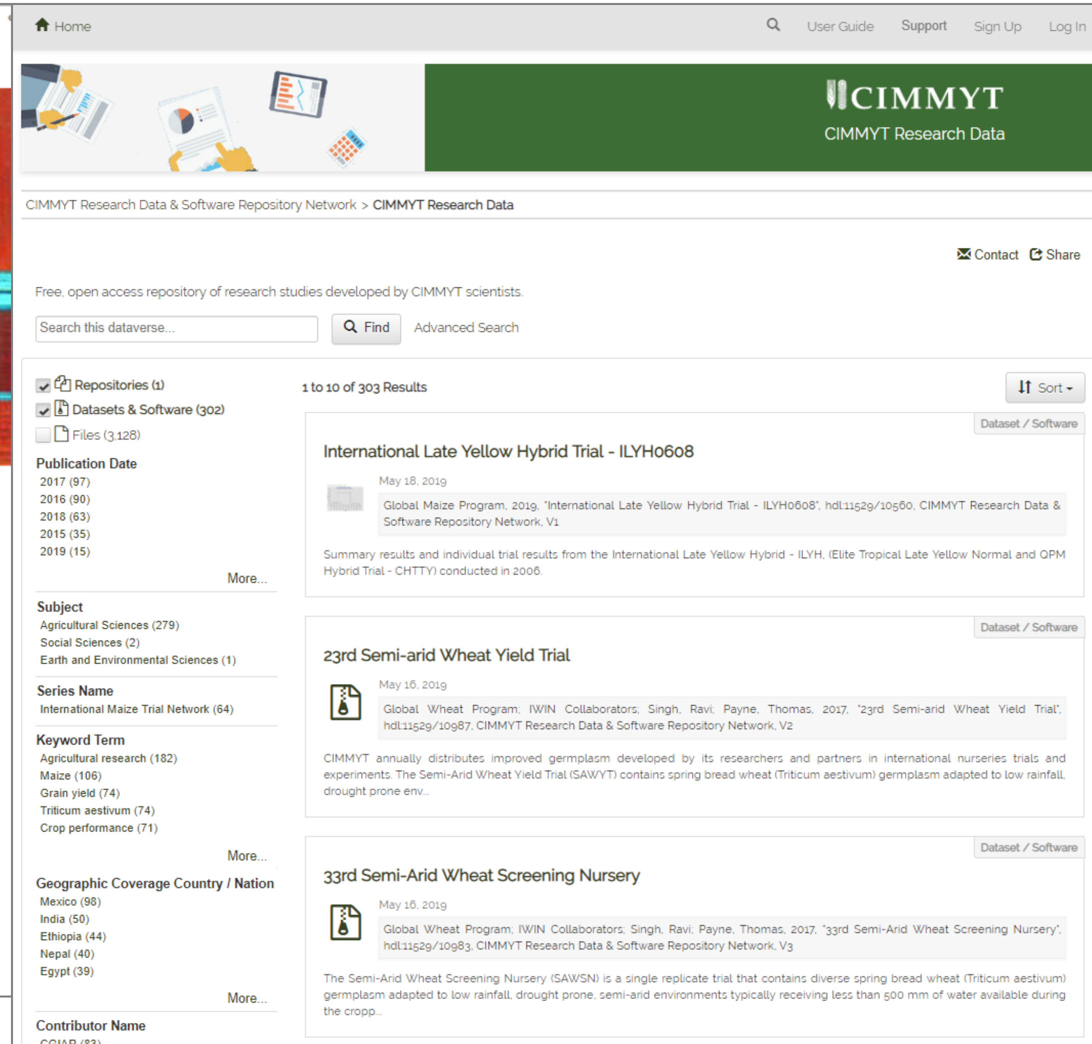
The **CGIAR provides 80% of germplasm** to the developing world.

International Wheat Improvement Network

- Annual worldwide distribution from CIMMYT:
  - Spring bread and durum wheat, and winter bread wheat trials
  - 28 targeted observation nurseries and replicated yield trials
  - 240 collaborators in 80 countries
  - 1300 experiments, >10 tons of seeds, sent by air freight
  - Free-of-charge to researchers. All covered by the SMTA.
- Objectives of IWIN
  - Rapid dissemination of improved wheat germplasm
  - Genetically diverse parental germplasm
  - Direct release after local evaluation
  - Voluntary evaluation data return to CIMMYT, 40-50% rate of return. Precise phenotyping is not an objective for IWIN.

Platform serving ground-referenced data for EO modeling consumption: CIMMYT Dataverse





FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

Home https://data.cimmyt.org/dataverse/cimmytdatadvn User Guide Support Sign Up Log In

Metrics 109 Downloads

Genotypic data from CIMMYT bread wheat breeding lines used in the Feed the Future Innovation Applied Wheat Genomics Version 2.1

Poland, Jesse; Dreisigacker, Susanne; Shrestha, Sandesh; Wu, Shuangye; Singh, Ravi; Mondal, Suchismita; Juliana, Philomin; Crossa, Jose; Rutkoski, Jessica, 2016, "Genotypic data from CIMMYT bread wheat breeding lines used in the Feed the Future Innovation Lab for Applied Wheat Genomics", hdl:11529/10695, CIMMYT Research Data & Software Repository Network, V2

Cite D... Learn about...

Description Genetic profiling of wheat breeding lines from the CIMMYT bread wheat breeding program was carried out... Unimputed genotypic data in the VCF format (CIMMYT-2013-2018.hmp.vcf) for 91,680 markers are availabl... could not upload and publish the unimputed VCF data due to file size limits in Dataverse.

Subject Agricultural Sciences

Keyword Agricultural research, Triticum aestivum, Genotypes, Bread wheat, SNP

Notes Unimputed genotypic data in the VCF format (CIMMYT-2013-2018.hmp.vcf) for 91,680 markers are availabl... could not upload and publish the unimputed VCF data due to file size limits in Dataverse.

Files | Metadata | Terms | Versions

Search this dataset... Find

5 Files

CIMMYT-2013-2018.hmp.txt.zip.gz
application/x-gzip - 822.4 MB - Oct 30, 2018 - 9 Downloads
MD5: 61865f5b1002f5a6e14dc555ba700663
Unimputed genotypic data in Hapmap format for 91,680 SNP markers
Genotypic data | 2018_Data_Release_v.2

CIMMYT-2013-2018_Genotypic_data_readme.txt
Plain Text - 1.2 KB - Oct 30, 2018 - 12 Downloads
MD5: 3eecaab3d20aab5752477dfb7dca5ec8
Details about the information contained in the genotypic data file

GBS pipeline CIMMYT 2013-2018.pdf
Adobe PDF - 149.4 KB - Oct 30, 2018 - 10 Downloads
MD5: 88285f0f421a74775f1f55263cc2e724
Description of the pipeline used to recall the data generated for CIMMYT wheat germplasm from 2013 to 2018.
Protocol | 2018_Data_Release_v.2

imputed_CIMMYT-2013-2018.hmp.txt.zip.gz
application/x-gzip - 232.9 MB - Oct 30, 2018 - 11 Downloads
MD5: 552c2dbf54441cb942f70d2eaf46fc19
Imputed genotypic data in Hapmap format for 52,851 SNP markers
Genotypic data | 2018_Data_Release_v.2

imputed_CIMMYT-2013-2018_vcf.zip.gz
application/x-gzip - 298.5 MB - Oct 30, 2018 - 3 Downloads
MD5: 5f87bfad43a892f69b1f53222786be68
Imputed genotypic data in VCF format for 52,851 SNP markers
Genotypic data | 2018_Data_Release_v.2

Metrics 289 Downloads Contact Share

Phenotypic data from trials conducted by the CIMMYT Bread Wheat Breeding Program Version 6.0

Singh, Ravi; Mondal, Suchismita; Crespo, Leonardo; Kummar, Uttam; Imtiaz, Muhammad; Lan, Caixia; Randhawa, Mandeep; Bhavani, Sridhar; Singh, Pawan K.; Huerta, Julio; He, Xinyao; Rahman, Mokhles; Pinto, Francisco; Perez Gonzalez, Lorena; Juliana, Philomin; Singh, Daljit; Lucas, Mark; Poland, Jesse, 2016, "Phenotypic data from trials conducted by the CIMMYT Bread Wheat Breeding Program", hdl:11529/10696, CIMMYT Research Data & Software Repository Network, V6

Cite Dataset / Software
Learn about Data Citation Standards

Description Phenotypic data were collected in on-station field trials for advanced breeding lines from the CIMMYT Bread Wheat breeding program over several years.

Subject Agricultural Sciences

Keyword Above ground biomass, Agronomic score, Anthesis time, Aphid damage, Barley yellow dwarf incidence, Booting time, Canopy green normalized difference vegetation index, Canopy normalized difference vegetation index, Canopy normalized pheophytinization index, Canopy photochemical reflectance index, Canopy ratio analysis of reflectance spectra carotenoid, Canopy ratio analysis of reflectance spectra chlorophyll a, Canopy ratio analysis of reflectance spectra chlorophyll b, Canopy red normalized difference vegetation index, Canopy simple ratio, Canopy structural independent pigment index, Canopy Temperature, Crop ground cover, Flag leaf lamina length, Flag leaf lamina width, Frost damage, Germination percentage, Grain number per spike, Grain test weight, Grain yield, Harvest Index, Heading time, Helminthosporium species severity, Leaf chlorophyll content, Leaf rust plant response, Leaf rust seedling response, Leaf rust severity, Lodging incidence, Maturity time, Peduncle length, Plant height, Spike length, Spike number, Spikelet number, Stem rust plant response, Stem rust seedling response, Stem rust severity, Stripe rust plant response, Stripe rust severity, Agricultural research, Triticum aestivum

Files | Metadata | Terms | Versions

Search this dataset... Find

11 Files

Download

CIMMYT_KSU_DB_2018_10_12.sql.zip.gz
application/x-gzip - 161.0 MB - Oct 24, 2018 - 9 Downloads
MD5: 47b8d02b425f5be6f7ede13f954b1267
SQL database containing phenotypic data values for advanced CIMMYT bread wheat lines. The data are raw plot-level phenotypic data collected through June of 2018.
Phenotypic database file
Download

conditions_data_2018_10_12.txt
Plain Text - 535 bytes - Oct 24, 2018 - 9 Downloads
MD5: 213ffc1ea270234ed23eaa32b0d15627
Descriptions of the conditions imposed within the different trial environments.
Phenotypic database file
Download

data_dictionary_2019_05_14.txt
Plain Text - 6.6 KB - May 16, 2019 - 0 Downloads
MD5: 716bb1e2243b9aa9a755f242af949c16
A description of the fields and columns of data contained in the database and associated data files.
Data dictionary
Download

germplasm_data_2018_10_12.txt
Plain Text - 17.9 MB - Oct 24, 2018 - 10 Downloads
MD5: d5243b95db712fd27972b4a92dae011f
Basic information about the germplasm that were evaluated in the trials.
Phenotypic database file
Download

locations_data_2018_10_12.txt
Plain Text - 826 bytes - Oct 24, 2018 - 12 Downloads
MD5: fad09c3855ce66574a8cdf472476cfb5
Basic information about the locations where the trials occurred.
Phenotypic database file
Download

Predictive modelling of wheat yield using spectral indices from high-throughput phenotyping

USAID Feed the Future project

Jesse Poland, Suchismita Mondal, Jose Crossa, Matthew Reynolds, Francisco Pinto, Lorena González Pérez, Atena Haghighattalab and Daljit Singh

CIMMYT's yield trials in Ciudad Obregon, Mexico

Matrice 100

Images courtesy: Francisco Pinto

MicaSense rededge camera

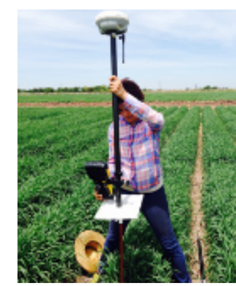
FEED THE FUTURE

The U.S. Government's Global Hunger & Food Security Initiative

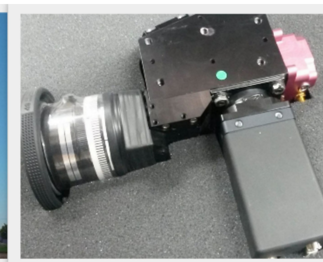
- **Image acquisition:** HTP data is collected using airplanes fitted with a Hyperspectral camera and drones fitted with a multispectral camera.
- **Image processing:** Aligning photos and building sparse point clouds, importing GCPs and geo-referencing the images, stitching images into dense point clouds, classifying points into plants and ground, creating a digital elevation model, ortho-mosaic generation and calculation of vegetation indices for the plots.

Ground control point

Geo-referencing



Piper PA-16 Clipper aircraft with a hyperspectral camera



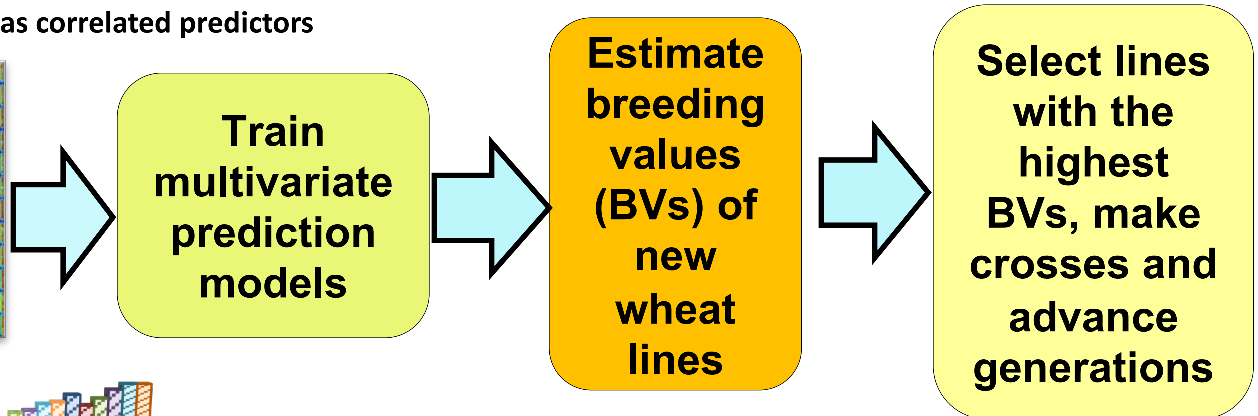
Predictive modelling of wheat yield using spectral indices from high-throughput phenotyping

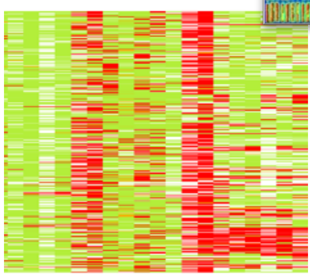
USAID Feed the Future project

Jesse Poland, Suchismita Mondal, Jose Crossa, Matthew Reynolds, Francisco Pinto, Lorena González Pérez, Atena Haghighattalab and Daljit Singh

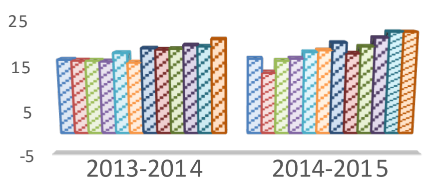
Phenotyping data

Vegetation indices that are related to grain yield are used as correlated predictors





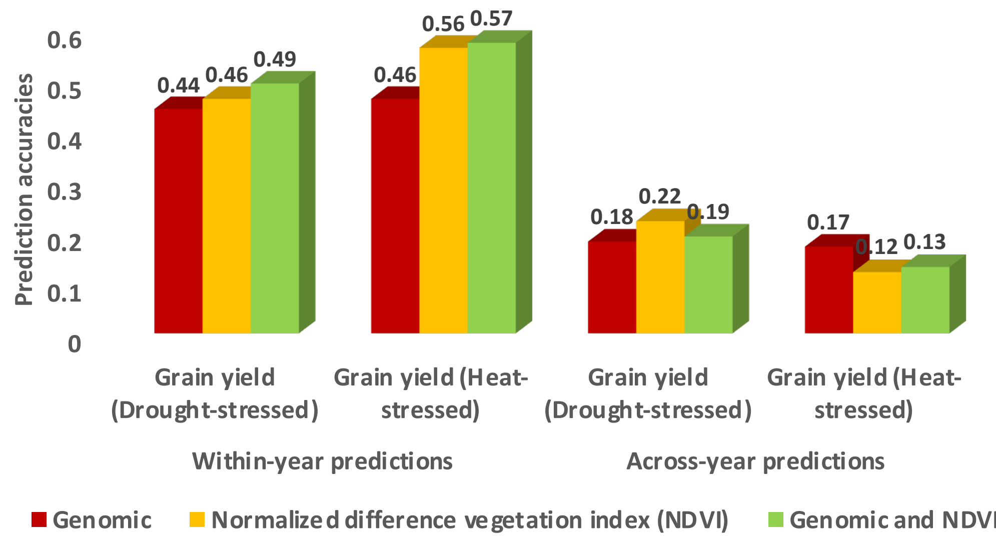
Genome-wide molecular markers



Environmental covariates like temperature, precipitation, humidity etc.

- ❑ Prediction of grain yield across years/environments is challenging.
- ❑ Can crop growth models using meteorological variables, accounting for soil variability and micro-environmental field conditions etc. improve prediction accuracies?

Grain yield prediction accuracies within and across years



FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

USAID
FROM THE AMERICAN PEOPLE

BILL & MELINDA GATES foundation

SAGARPA
SECRETARÍA DE AGRICULTURA, GANADERÍA, DESARROLLO RURAL, PESCA Y ALIMENTACIÓN

UKaid
from the British people

Australian Government
Australian Centre for International Agricultural Research

CROP TRUST

Global Affairs Canada

CAAS

भाकृअनुप
ICAR

syngenta foundation for sustainable agriculture

HENAN AGRICULTURAL UNIVERSITY

USDA

AATF
AFRICAN AGRICULTURAL TECHNOLOGY FOUNDATION
FONDATION AFRICAINE POUR LES TECHNOLOGIES AGRICOLES

BMZ
Federal Ministry for Economic Cooperation and Development

IFAD
Investing in rural people

CSIRO

THE WORLD BANK

GRDC Grains Research & Development Corporation
Your GRDC working with you

European Commission

JIRCAS
JAPAN

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Agency for Development and Cooperation SDC

Ministry of Foreign Affairs

REPUBLIKA NG PILIPINAS

cirad

Kellogg's

BBSRC
bioscience for the future

inifap
Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias

Sida

Republic of South Africa

FAO
FIAT PANIS

DUPONT
PIONEER.

citibanamex

Norad

KAZAGRO
NATIONAL Management Holding

gto
orgullo y compromiso de todos

Fundación Haciendas del Mundo Maya
Conocer y comprender para transformar
NAAT-HA

NEW ZEALAND
FOREIGN AFFAIRS & TRADE
MANATŪ AORERE

Rural Development Administration

AFRICAN DEVELOPMENT BANK

T.C.
GIDA TARIM VE HAYVANCILIK BAKANLIĞI

FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

Geo-Referenced Data to Inform Earth System Modeling for Agriculture: A Discussion among Collectors, Users, and Aggregators

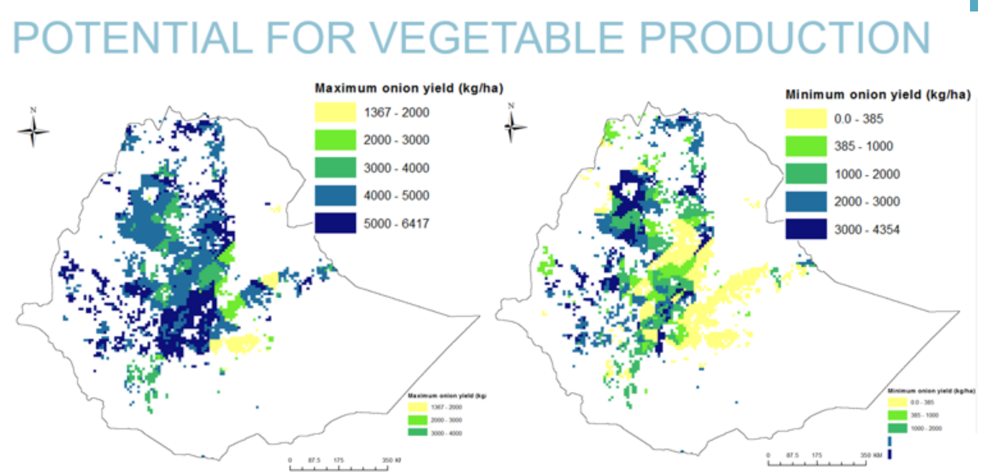
Raghavan Srinivasan, Texas A&M University
R-Srinivasan@tamu.edu

Who are we and what are we doing?

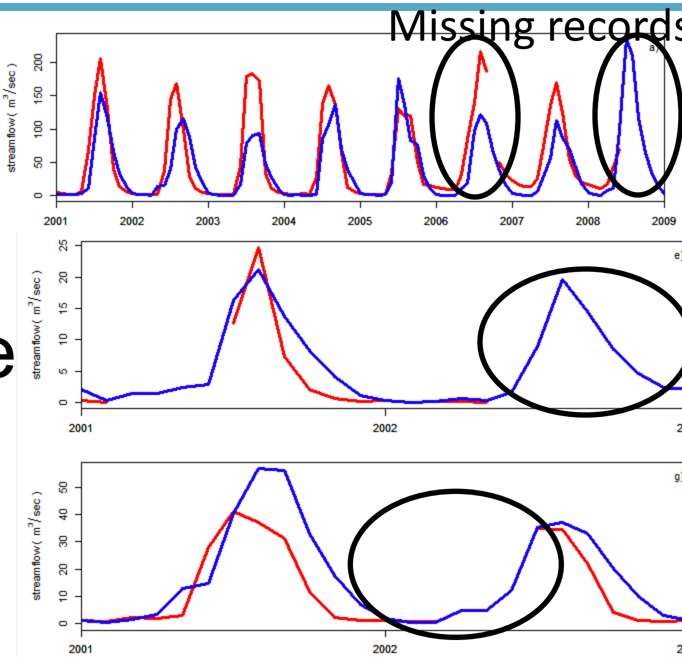
- A professor of spatial sciences and watershed modeling
- TAMU – Developer of the SWAT model for natural resource assessments and development of land and climate management systems
- Our Spatial Sciences Lab
  - Apply both geo-referenced ground based and EO derived data in hydrologic modeling
  - Generate ground-referenced data for use in water and land management
  - assessment of EO-based products (ET, AWC, Biomass…)

Geographies and crops

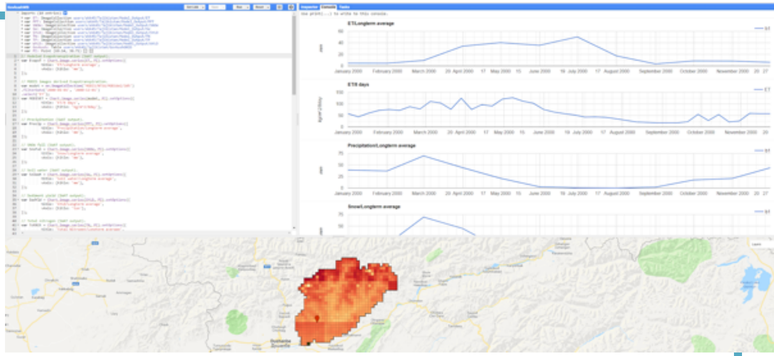
- From ASAR to monsoonal climates
- Small watersheds to river basins including transboundary river systems, and regions (e.g. Africa, Europe and the Continental US.
- Cereal, vegetable, fiber crops, fruit and nuts, beverage and spice, legume crops
- Biomass for fodder and grasses for grazing, bio-energy crops, forest.
- Water balance estimation
  - water availability
  - soil moisture
  - actual evapotranspiration
  - irrigation water requirement
  - Water pollution
- Reservoir water levels
- Wetland status



Constraints



- Some of the obstacles include
  - Quality data scarcity
    - Shot time-series temporal and
    - Low spatial resolution
  - Access to existing data, especially from ministries and agencies in developing countries for ground truthing
  - Emergence of big data, especially from remote sensing, but
    - Data not well validated
    - Large storage and computational requirement for analysis
    - Diverse formats (eg. *.tif, *.NetCDF, *.hd5 etc.)



Opportunities

- Continuous improvement in
  - Spatial and temporal resolution
  - Cloud based computational and storage
- Support EO-based data generation through
  - Validation
  - Flagging anomalies
  - Value added products
  - Capacity building
- Open data policy by funders improved access to data
- Bio-geochemical models becoming source of data
  - Use limited information and generate more data for interpretation
  - Physically based and provide reliable data including uncertainty
  - Help to reasonably predict the future, providing new data and insight

FEED THE FUTURE

The U.S. Government's Global Hunger & Food Security Initiative

Geo-Referenced Data to Inform EO Modeling for Agriculture: A Discussion among Collectors, Users, and Aggregators

Estefania Puricelli

Alyssa Whitcraft, Antonio Sanchez Galves, Mike Humber, Jon Keniston

Agriculture Monitoring in the Americas | A GEOGLAM Regional Network

NASA Harvest | University of Maryland

AMA

AGRICULTURE MONITORING IN THE AMERICAS

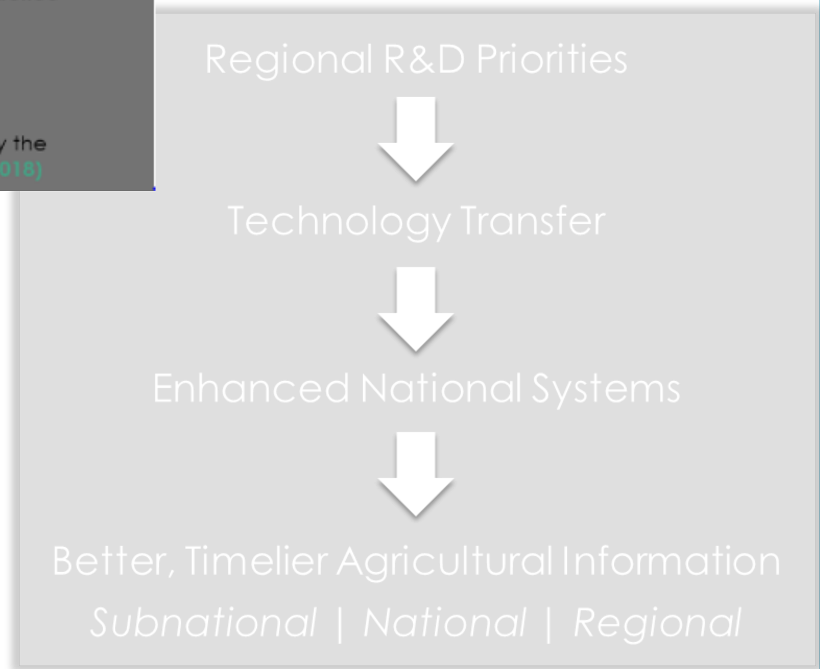
AGRICULTURA MONITOREADA EN LAS AMÉRICAS

AGRICULTURA MONITORADA NAS AMÉRICAS

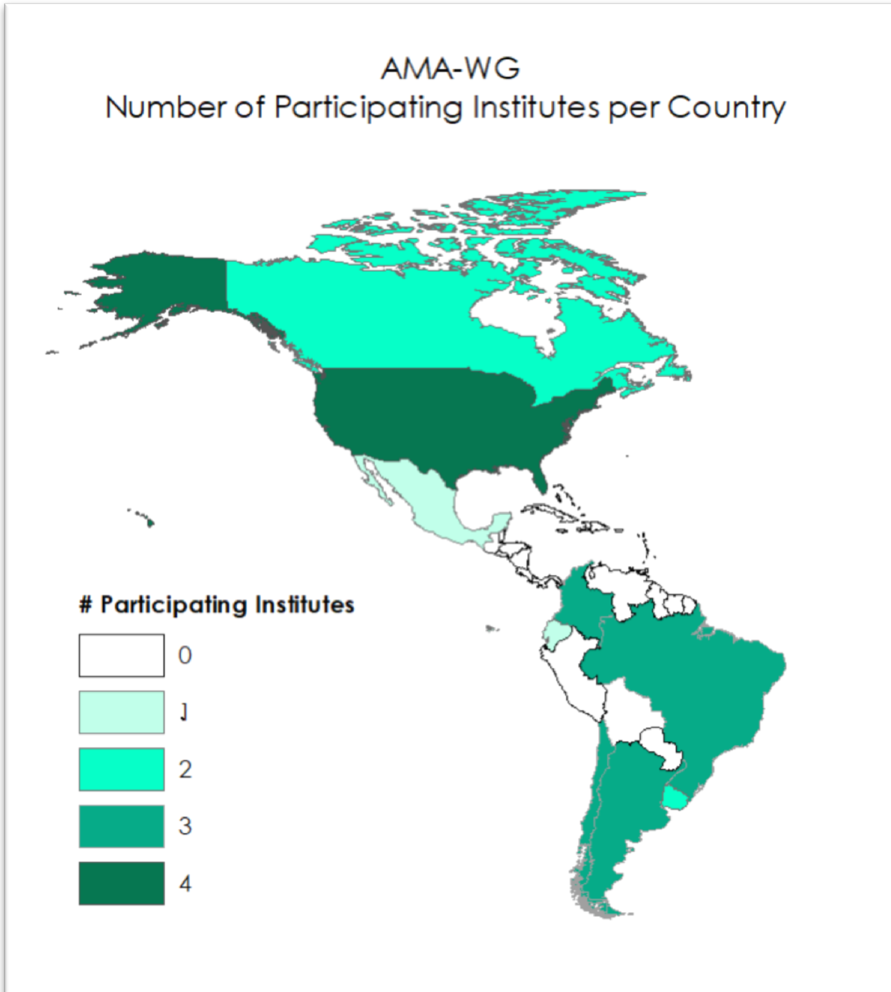
Who are we and what are we doing?

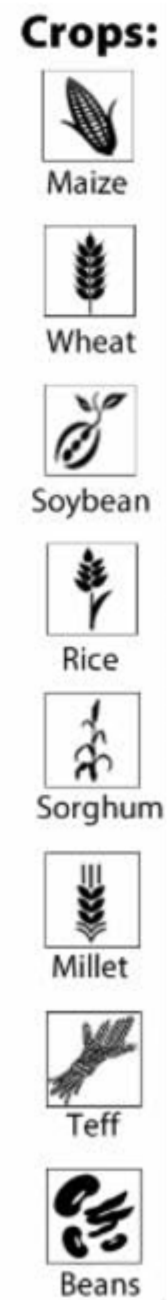
Agriculture Monitoring in the Americas (AMA):

A regional network to enhance national monitoring systems and their participation in GEOGLAM



Geographies and crops



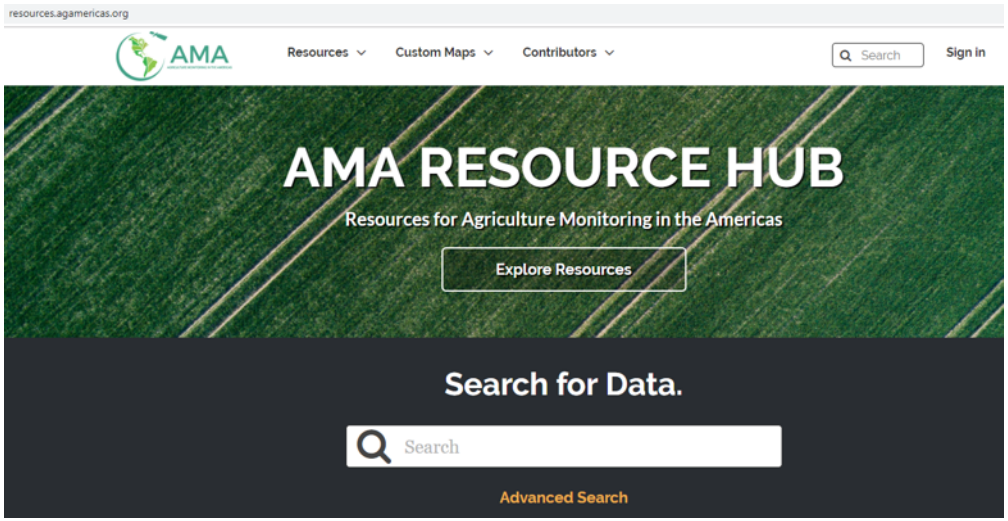


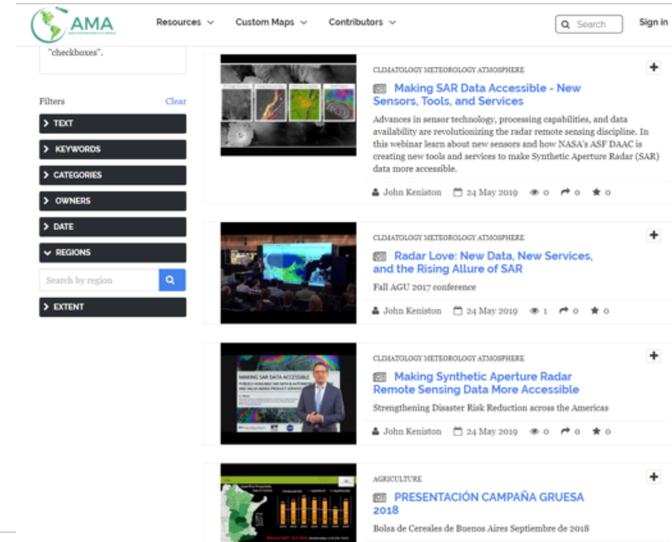
Plus others of national interest

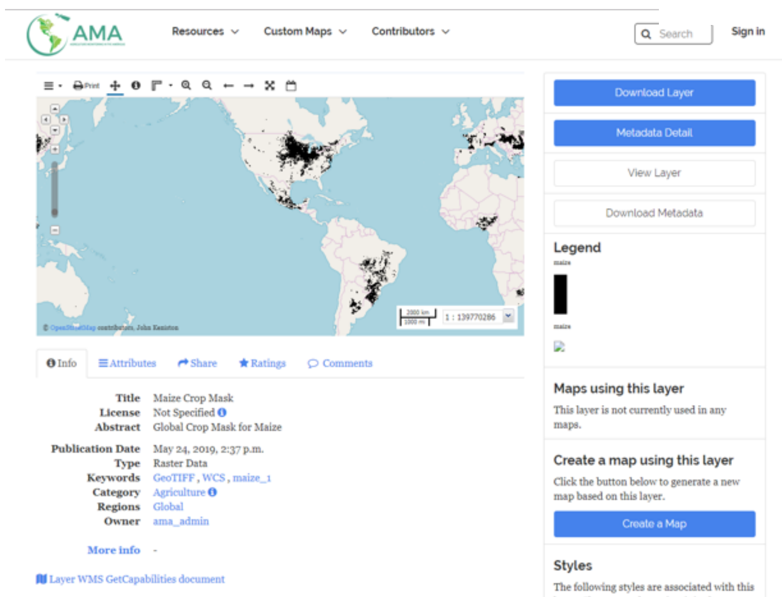
Constraints

- There are many public resources for Remote Sensing (RS) data and use
  - EO-based products (e.g. crop mask, calendar, yield)
  - EO-based tools (e.g. GLAM, Sen2-Agri…)
  - Tabular data (e.g. *in situ* points, production/trade info)
  - CapDev Tools (e.g. webinars, MOOCs, training manuals)
- In the Americas, especially in South America and the Caribbean, not everyone is aware of them
- There are also many really good national resources for RS, but they are hard to find

Opportunities: resources.agamericas.org



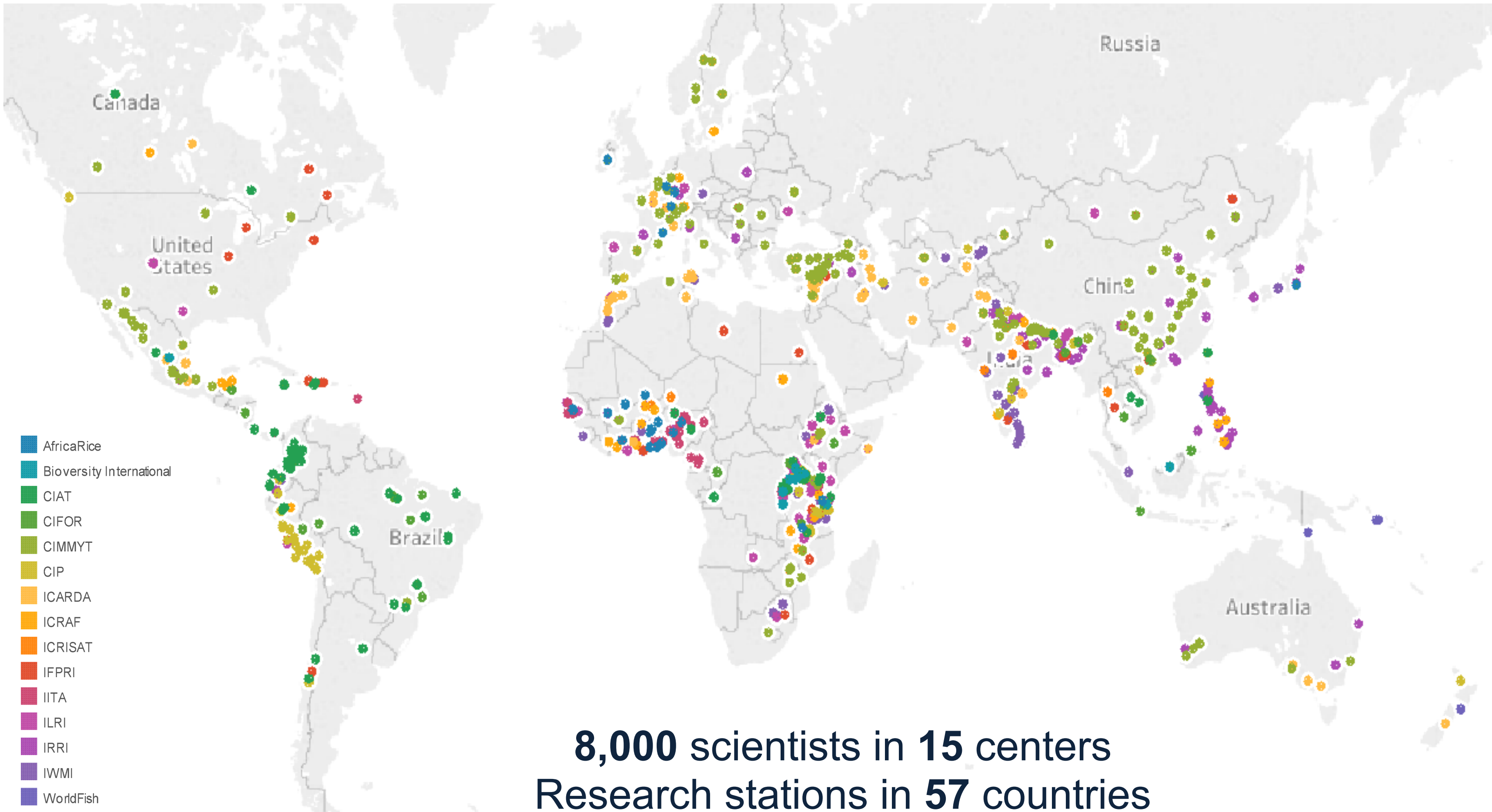




Geo-Referenced Data to Inform EO Modeling for Agriculture: A Discussion among Collectors, Users, and Aggregators

Jawoo Koo
International Food Policy Research Institute
CGIAR Consortium for Spatial Information | CGIAR Platform for Big Data in Agriculture

Where's CGIAR?



8,000 scientists in **15** centers
Research stations in **57** countries
1,000 research locations in **95** countries

FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

Introducing…

CGIAR Platform for Big Data in Agriculture

- One of thee **Research Support Platforms** of CGIAR
- We support agricultural scientists to harness the power of big data to accelerate and enhance the impact of international agricultural research.
- Our activities in three modules:

Organize

Support data generation, curation, management

Convene

Collaborate and convene to incorporate data science

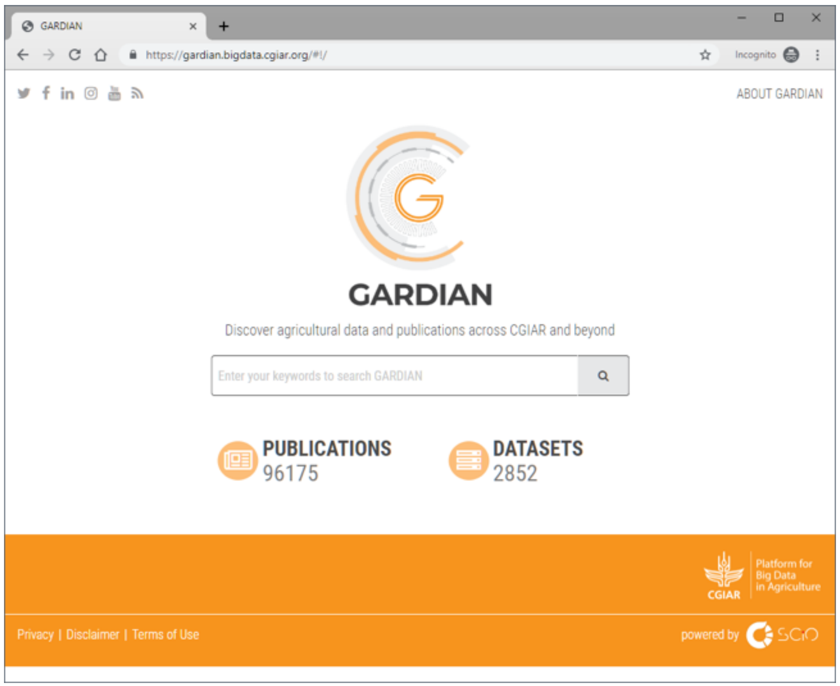
Inspire

Support data science pilot projects in partnership

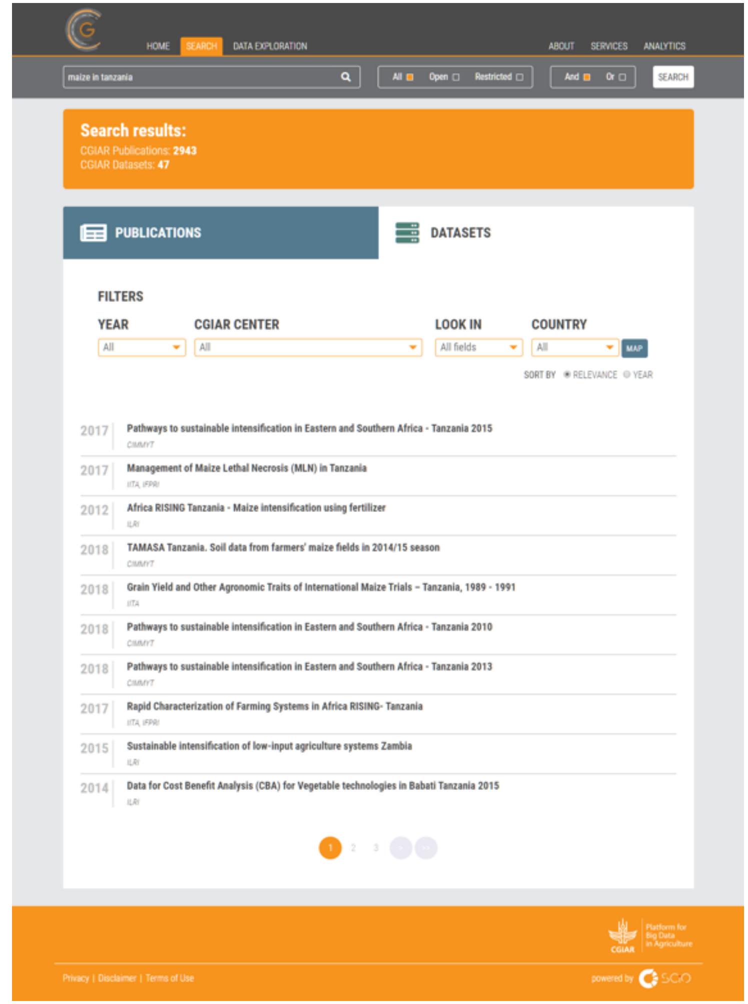
We Make CGIAR Research Data F.A.I.R.

Findable, Accessible, Interoperable, Reusable

Open access is not enough!

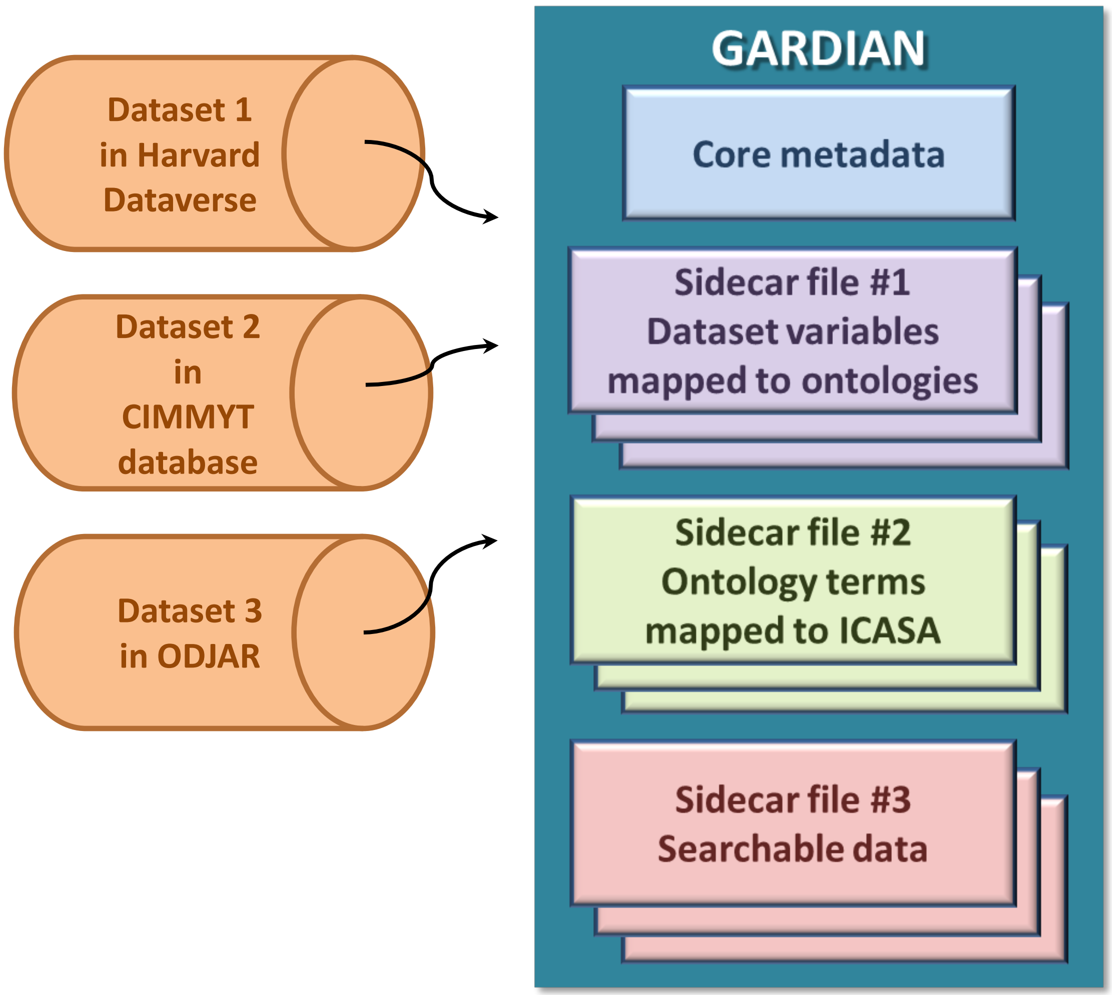


https://gardian.bigdata.cgiar.org



FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

We Make Data More Useful



Allows data to be discovered and accessed.

Allows data to be searched in an automated way and connected with other databases through the semantic web.

Allows data to be transformed using existing AgMIP data translators.

Allows fast, complex searching without accessing and interpreting the data directly.

Credit | Cheryl Porter (U. Florida)

Our Geospatial Modeling Use Cases

Center		Tool		Data		Objective
AfricaRice	uses	Random Forest	to analyze	satellite remote sensing data	for	assessing inland valley landscapes for rice production.
Bioversity Int	uses	Maximum Entrophy	to analyze	geospatial data	for	modeling species distribution.
CIAT	uses	Neural Network	to analyze	satellite remote sensing data	for	detecting human-induced land cover changes in Latin America.
CIAT	uses	Maximum Entrophy	to analyze	geospatial data	for	predicting the suitability of coffee.
CIAT	uses	Support Vector Machine	to analyze	satellite remote sensing data	for	maping croplands.
CIAT	uses	Classification And Regression Trees	to analyze	crop trials data	for	developing data-driven agronomic insights.
CIAT	uses	Machine Learning, Data Mining	to analyze	geospatial data	for	predicting pest and disease epidemics on cassava.
CIAT	uses	Machine Learning Algorithms	to analyze	geospatial data	for	simulating climate change impacts on crop distributions.
CIAT	uses	Dynamic Time Wraping	to analyze	satellite remote sensing data	for	maping cropping patterns.
CIFOR	uses	Classification And Regression Trees	to analyze	satellite remote sensing data	for	detecting burned forest areas in Indonesia and Ethiopia.
CIMMYT	uses	Random Forest	to analyze	satellite remote sensing data	for	predicting plot-level soil properties and land-use.
CIMMYT	uses	Neural Network	to analyze	satellite remote sensing data	for	classifying land-uses and detect features (houses, roads, fields).
CIMMYT	uses	Neural Network	to analyze	photos from crop fields	for	detecting pest, disease, and nutrient deficiencies for crops.
CIMMYT	uses	Random Forest	to analyze	soil spectral measurement data	for	predicting 'wet chemistry' of soil samples (pH, CEC, N, P, SOC).
CIMMYT	uses	Decision Tree, Deep Learning, Gaussian Process	to analyze	nutrient omission on-farm plot-level trial data	for	predicting crop responses to fertilizers.
CIMMYT	uses	Decision Tree	to analyze	panel crop production survey data	for	understanding the drivers of yields.
CIMMYT	uses	Genetic Algorithm	to analyze	soil properties and market price data	for	optimizing the profitability of fertilizer applications.
CIP	uses	Maximum Entrophy	to analyze	weather data	for	predicting the distributions of potato insects and virus.
ICARDA	uses	Random Forest, Neural Network, Support Vector Machine, Lasso	to analyze	genebank accessions data	for	strategizing crop breeding for target traits.
ICARDA	uses	Random Forest	to analyze	satellite remote sensing data	for	mapping rice and fallow cropping systems.
ICARDA	uses	Random Forest, Support Vector Machine, Gradient Boosting	to analyze	satellite remote sensing data	for	mapping crop extent and produciton seasonality.
ICARDA	uses	Random Forest, Monte-Carlo Simulation	to analyze	satellite remote sensing data	for	predicting crop yield variations spatially.
ICARDA	uses	Neural Network	to analyze	mobile/drone-captured photos	for	detecting crop diseases and weeds.
ICARDA	uses	Harmonic Analysis	to analyze	satellite remote sensing data	for	monitoring crop health and produce quality.
ICARDA	uses	Random Forest	to analyze	satellite remote sensing data	for	detecting crop species.
ICARDA	uses	Random Forest, Classification And Regression Trees	to analyze	field-measured data from IOT sensors	for	optimizing agronomic practices.
ICARDA	uses	Neural Network	to analyze	field-measured data from IOT sensors	for	automating irrigation systems.
ICARDA	uses	Iterative Self-Organizing Data Analysis Technique, Random Forest	to analyze	satellite remote sensing data	for	mapping land degradation and desertification.
ICARDA	uses	Random Forest, Neural Network	to analyze	satellite remote sensing data	for	mapping water bodies and monitoring changes.
ICARDA	uses	Decision Tree, Classification And Regression Trees	to analyze	geospatial data	for	mapping herb species distribution.
ICRAF	uses	Maximum Entrophy, Maximum Likelihood, Boosted Regression Tree, Ran	to analyze	biodiversity data	for	predicting species suitability.
ICRAF	uses	Random Forest	to analyze	satellite remote sensing data	for	monitoring land uses and degradation at the landscape level.
ICRISAT	uses	Machine Learning Algorithms	to analyze	genomic data	for	selecting the most suitable crop varieties.
ICRISAT	uses	Machine Learning Algorithms	to analyze	satellite remote sensing data	for	mapping crop types.
IFPRI	uses	Edge Detection Algorithm	to analyze	drone-captured imagery data	for	drawing field boundaries.
IFPRI	uses	Maximum Likelihood, Neutral Network, Support Vector Machine	to analyze	satellite remote sensing data	for	classifying land cover types.
IFPRI	uses	Edge Detection Algorithm	to analyze	satellite remote sensing data	for	mapping crop planting methods.
IITA	uses	Empirical Bayesian Regression	to analyze	satellite remote sensing data	for	predicting soil properties.
IITA	uses	Random Forest	to analyze	satellite remote sensing data	for	classifying land cover types.
ILRI	uses	Random Forest	to analyze	geospatial data	for	mapping livestock production systems.
IRRI	uses	Random Forest, Support Vector Machine	to analyze	geospatial data	for	mapping rice ecosystems and characterizing rice areas.
IRRI	uses	Computer Vision, Neutral Network	to analyze	drone-captured imagery data	for	detecting rice straw burnings automatically.
IWMI	uses	Random Forest, Classification And Regression Trees	to analyze	satellite remote sensing data	for	mapping land cover types.
WorldFish	uses	Machine Learning Algorithms	to analyze	geospatial fishing boat tracking data from small-scale f	for	monitoring temporal and spatial changes in fish production.

Credit | CGIAR-CSI

Data types

- Global (96 countries)
- Maize, rice, wheat, cassava, potato, beans, and trees/coffee (plus other CGIAR mandate crops)
- Land use/land cover, agronomic/breeding trials, farm and natural resource managements, household surveys (700K HHs/year).

Applications

- Supervised classifications (e.g., crop geography, field boundary)
- Mapping outcomes using proxies (e.g., nighttime lights for nutrition)
- Modeling for scenario analyses (e.g., climate change impacts)

Constraints

- **Obstacles in getting the data we need**
  - Not reliable national/subnational statistics data (representativeness, calibration/validation)
  - Inherent challenges in mapping low-input smallholders' agriculture (heterogeneity, small plots, weeds, clouds)
- **Obstacles in maximizing value of data collected**
  - Incentivizing researchers for further use of data by others
  - Implementing interoperability
  - Biases in published data (more successes than failures)
  - Fear of exposing PIIs and any unintended consequences of data used by others

Opportunities

- **Potentials**
  - Avoid unnecessary duplicate collection of data
  - Gaining insights by linking data across domains
  - Increase research visibility and validity
  - Facilitates open scientific enquiry and innovation
  - Provides important resources for education and training
- **Next Steps**
  - Standardizing data from the collection (AgroFIMS)
  - Strengthening EO community in CGIAR (incentivize to share georeferenced data, provide training, mini-grants for pilot projects and key datasets, and shared services)
  - Providing CGIAR data in AI/ML-friendly formats (Sidecars)

FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

JOIN THE DISCUSSION

agrilinks.org

Contact: jmaccartee@usaid.gov

Comment on today's topic: https://www.agrilinks.org/event/geo-referenced-data-inform-eo-modeling-agriculture-discussion-among-collectors-users-and

Tweet tips! twitter.com/agrilinks

Post resources! facebook.com/agrilinks