Scaling participatory approaches in index insurance

Special thanks to **ACToday**, CCAFS, CIF, NASA, NOAA NERC DRiSL SatwinAlert, Syngenta Foundation, WFP R4, and World Bank projects for examples

Presented by

Daniel Osgood, Lead Scientist

Financial Instruments Sector Team

deo@iri.columbia.edu

THE APPLICATION: INDEX INSURANCE

- Climate change: more bad years
- Adaptation: increase productivity in normal years to cover bad year loss
- Insurance must be built, evaluated in terms of the problem it is intended to solve
- Insurance is only one part of the solution, must be engineered to be compatible with other tools
- Really hard to know what coverage to reduce that is covered by other risk management for affordability

But strategies that increase productivity in most years face increased risk in bad years

Threat of 1 drought year out of 5 prevents other 4 from being much more productive

Key to adaptation is to relax risk of bad year to unlock productivity options

Insurance: help reduce risk to unlock productivity as part of package.

Designing insurance to meet this need is challenging

WIDE RANGE OF APPLICATIONS, ONE THING IN COMMON: **ALL INFO "NOT PERFECT"**

01 *Remote sensing data often used because there are few other data sources, raingauges sparse, error prone*

Paucity of validation/calibration data, incorrect estimates, lack of knowledge of extent of errors

02 *Sometimes farmer recollection only source of historical information*

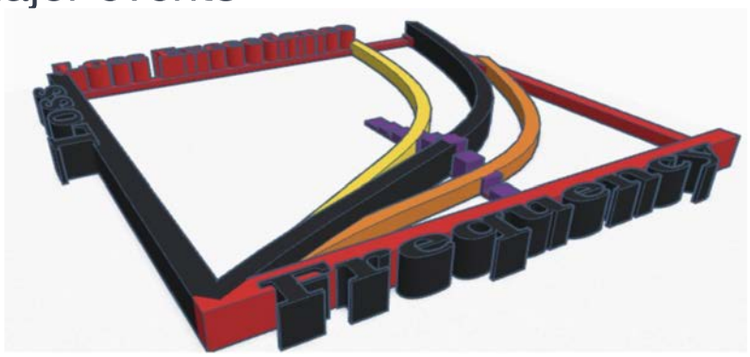
Production datasets may not relate to farmer experiences

- Problem may be costs to maintain production, not production level
- May be cost of failing to repay loan, low production or shifting to food crops without loan better vs medium production of cash crop inadequate to cover costs

03 *Models may miss key local features*

Phenology (crop timing) may be assumed incorrectly by couple of weeks

- Eg due to elevation, temperature, availability of inputs, labor.
- Couple of weeks off can lead to very different water stress results, missing major events



04 *Yield data sparse, short, inaccurate*

05 *Co-Production essential*

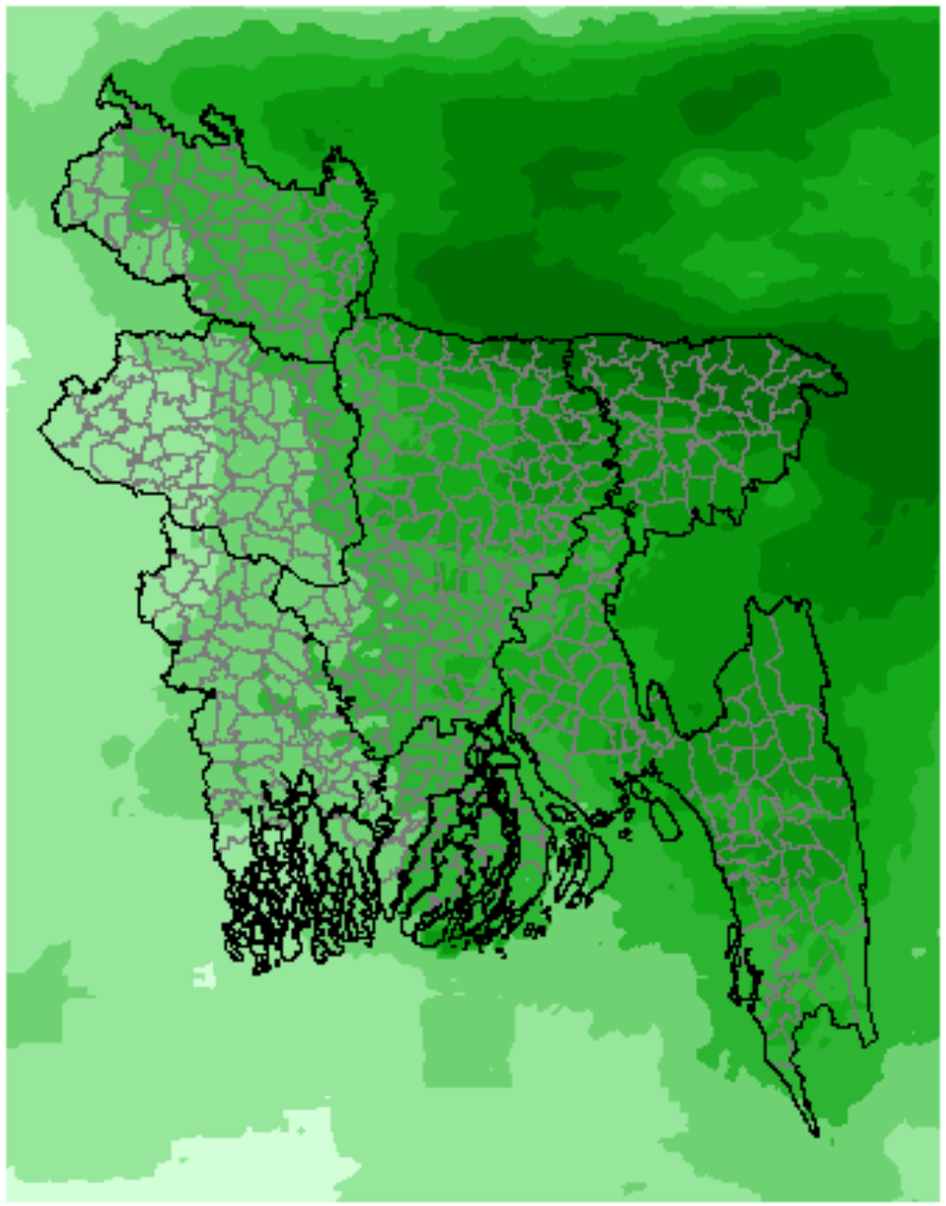
Data validation must be 2 way: farmers must understand quality of remote sensing data used so that they know how well risks addressed, and what gaps they are exposed to

Terabytes of climate/weather big data in index insurance

- **Many FREE sensors/products**
- Rainfall Estimates: CHIPRS, TAMSAT, ARC2, ENACTS…
- Vegetation: EVI, NDVI, NDWI ...
- Soil Moisture (Passive Microwave, Active Microwave, Water Budget)
- Evapotransporative Stress, Flood, High resolution imagery

- **BMD Data Library, ENACTS (gridded merged satellite/raingauge – thx ACToday)**

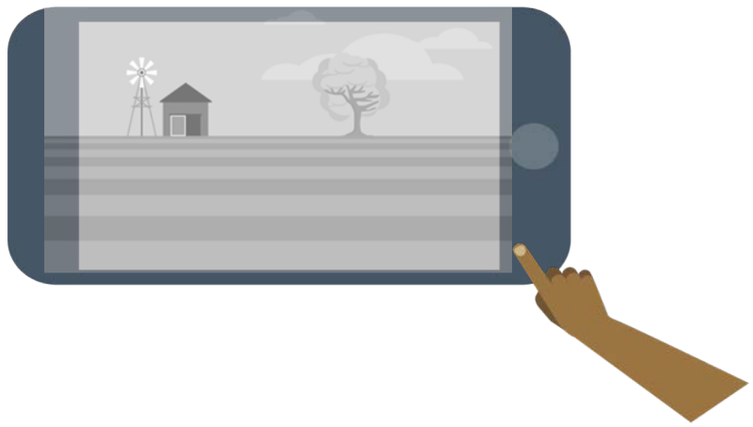
What does this have to do with farmers?



Hundreds and hundreds of village design meetings

Multiple layers of crowdsourced information

- Use existing field technologies as much as possible-– transform to crowdcore of design
- Strategically layer by cost using less accurate, lower cost mechanisms to compliment more intensive sampling
- Expert surveys
- Picture Based Monitoring Apps, Agonomist supervised (IFPRI led)
- Moderated community design structured meetings, tablet/smartphone 2 way links to database
- Individual reporting: Smartphone, SMS, Robo Calls, Games, Contests
- **What does this have to do with satellites/models?**



A PATH FORWARD: FOCUS ON FOREST BEFORE TREES

Compounding evidence design strategy on critical events

- Parametric financing: for worst events
- In bad year, farmer wants meaningful payout
- Focus on techniques that robustly hit worst years
- We must understand statistically, scientifically, and intuitively
- **Compounding evidence design strategy**
- **Crowd at the core** of science based design

- *But don't be too optimistic, this is a climate and development audience, we need to be disappointed about something*

	Payouts in 2 out of 3 Years	
	Loss	Pay
1	1	1
2	0	0
3	0	0
4	0	0
5	0	0
6	1	1
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
12	0	0
13	0	0
14	0	0
15	1	0
	Correlation:	0.784465

Easy to get wrong, important to get right: Accidental overfitting by human or machine
Correlation may be high simply because identified single most obvious thing

Danger even with quality datasets and advanced modeling

Insurance will not work unless designed/evaluated for its purpose for client

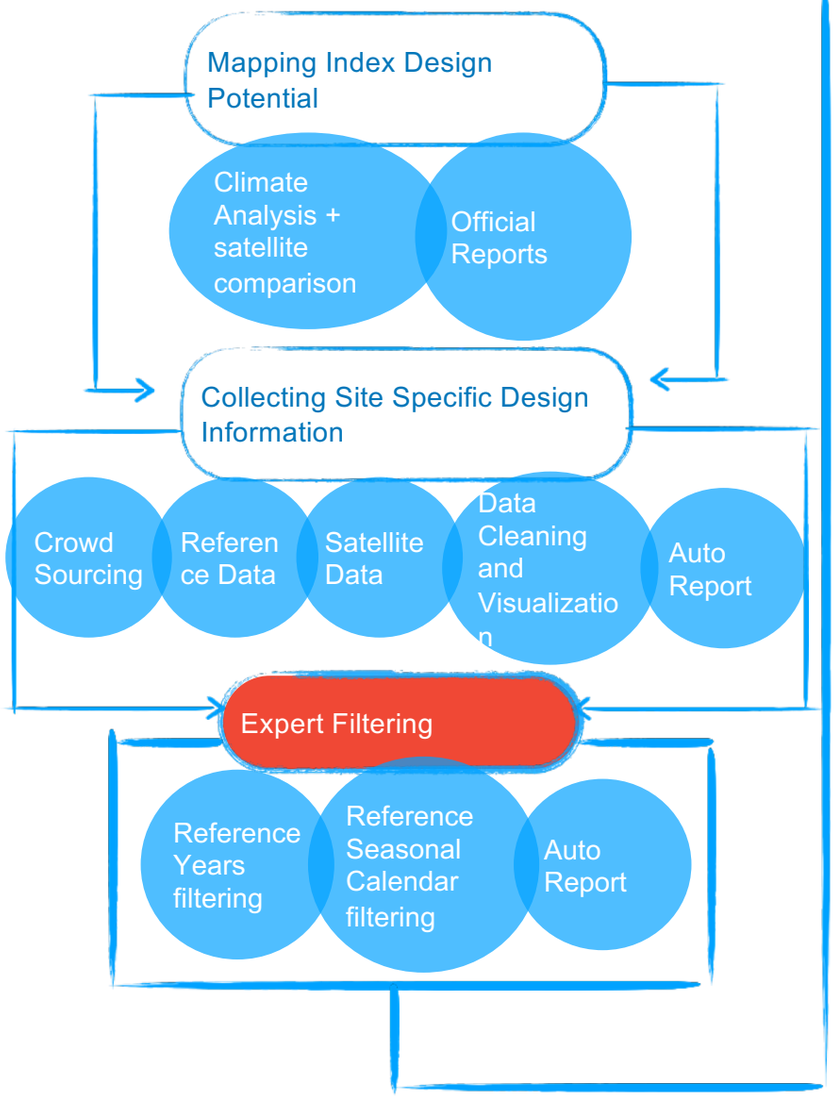
Payout in worst year=very high correlation			
Payout in worst place=very high correlation			
	But is it good insurance?		
	Is it good model?		
	Does model only echo the obvious?		
Year or			
Location	**Loss**	**Pay**	**Pay**
1	0.20	0	0
2	0	0	0.01
3	0	0	0.01
4	0	0	0.01
5	0	0	0.01
6	1.00	0.30	0.30
7	0	0	0.01
8	0	0	0.01
9	0	0	0.01
10	0	0	0.01
11	0	0	0.01
12	0	0	0.01
13	0	0	0.01
14	0	0	0.01
15	0.10	0	0
	Correlation:	**0.98**	**0.97**
		Only worst pay	Wrong pay everywhere except worst

Crowd at core? Research (thx NASA, NOAA, NERC, ACToday…)

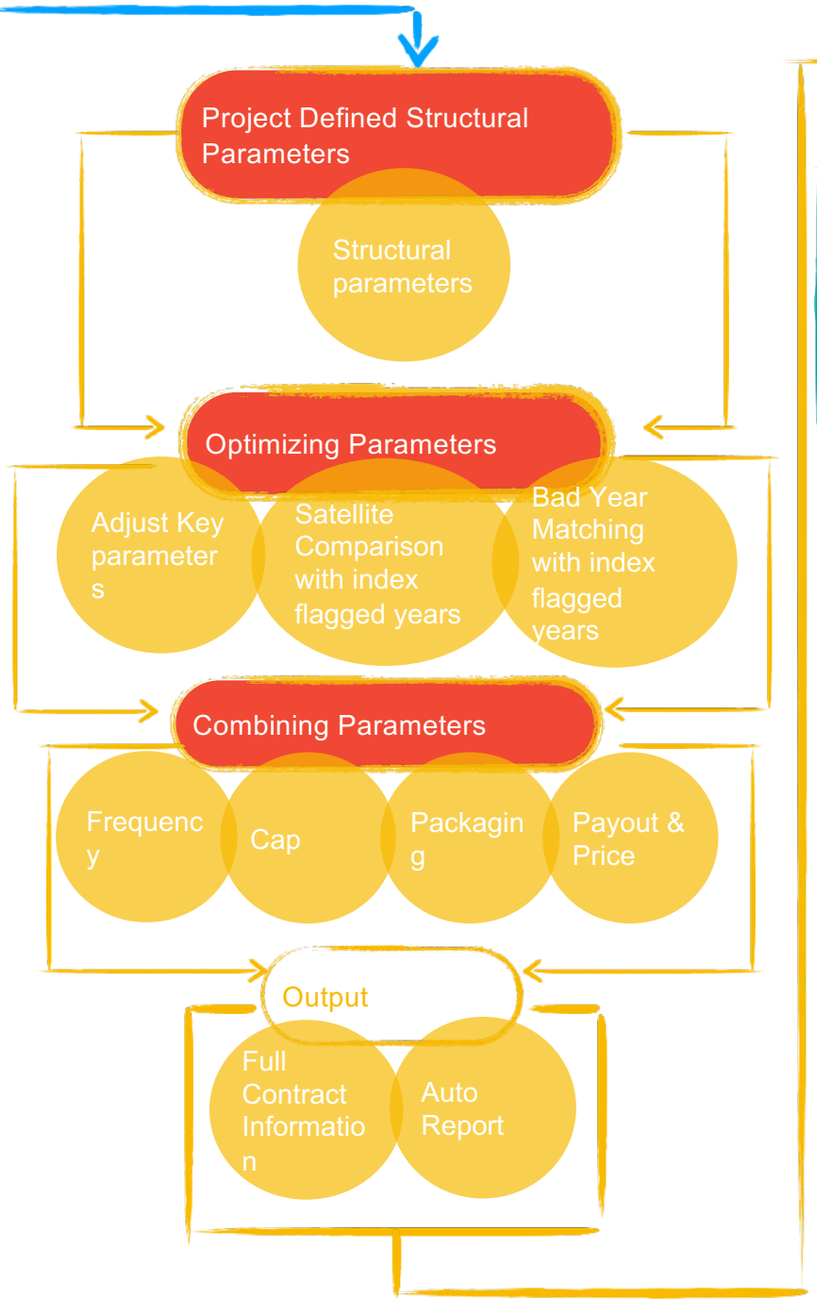
- Is there evidence of information in farmer recollection?
- Are there systematic biases?
- What strategies might address these biases?
- Do we get complimentary perspectives across gender, networks?
- Are there different psychological recollection strategies across gender?
- Do individuals report different information than groups (groups may be cleaner information, but important voices may be missing)
- What strategies might work with robo-calls, is there any useful information?
- Can AI, CI, and satellite information be used to clean farmer recollection?
- Do incentives based on reporting quality improve information?
- Do Non-financial incentives in games, certificates improve reported information quality?

The Goal: Embed Crowd at the core of locally driven science based processes

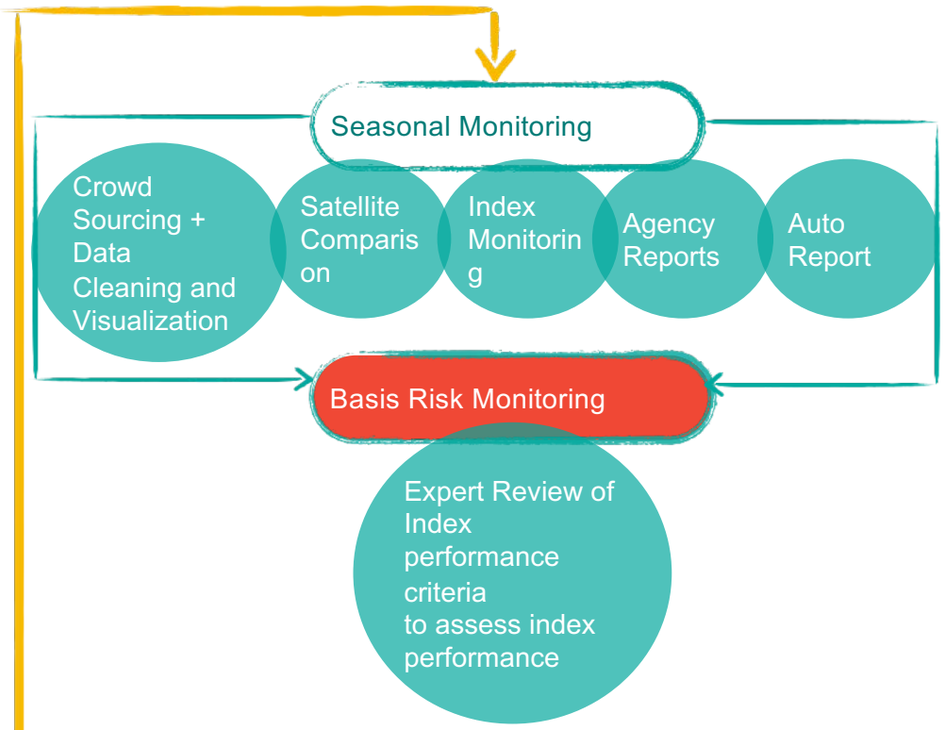
Pre-Index Design



Index Optimization



Agricultural Season



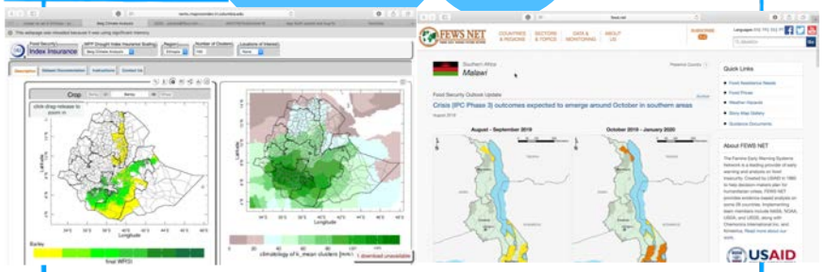
Embed Crowd at core of Annual Index Process: SELF OPTIMIZING INDEXES

Pre-Index Design

Mapping Index Design Potential

Climate Analysis + satellite comparison

Official Reports



Collecting Site Specific Design Information

Crowd Sourcing

Reference Data

Satellite Data

Data Cleaning and Visualization

Auto Report

Expert Filtering

Reference Years filtering

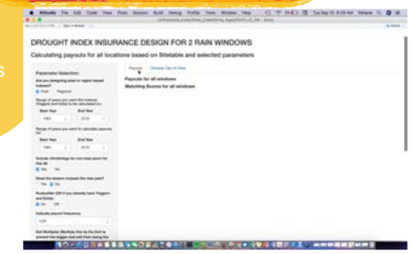
Reference Seasonal Calendar filtering

Auto Report

Index Optimization

Project Defined Structural Parameters

Structural parameters



Optimizing Parameters

Adjust Key parameters

Satellite Comparison with index flagged years

Bad Year Matching with index flagged years

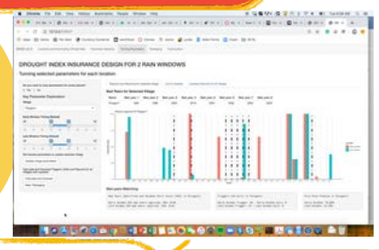
Combining Parameters

Frequency

Cap

Payout & Price

Packaging



Output

Full Contract Information

Auto Report

Agricultural Season

Seasonal Monitoring

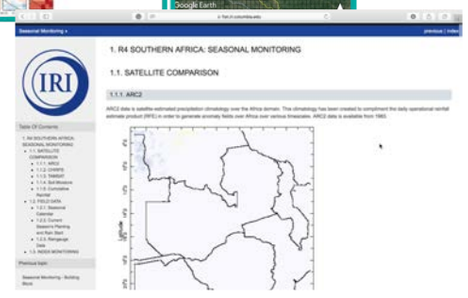
Crowd Sourcing + Data Cleaning and Visualization

Satellite Comparison

Index Monitoring

Agency Reports

Auto Report



Basis Risk Monitoring

Expert Review of Index performance criteria to assess index performance

Roles-determined through participatory processes

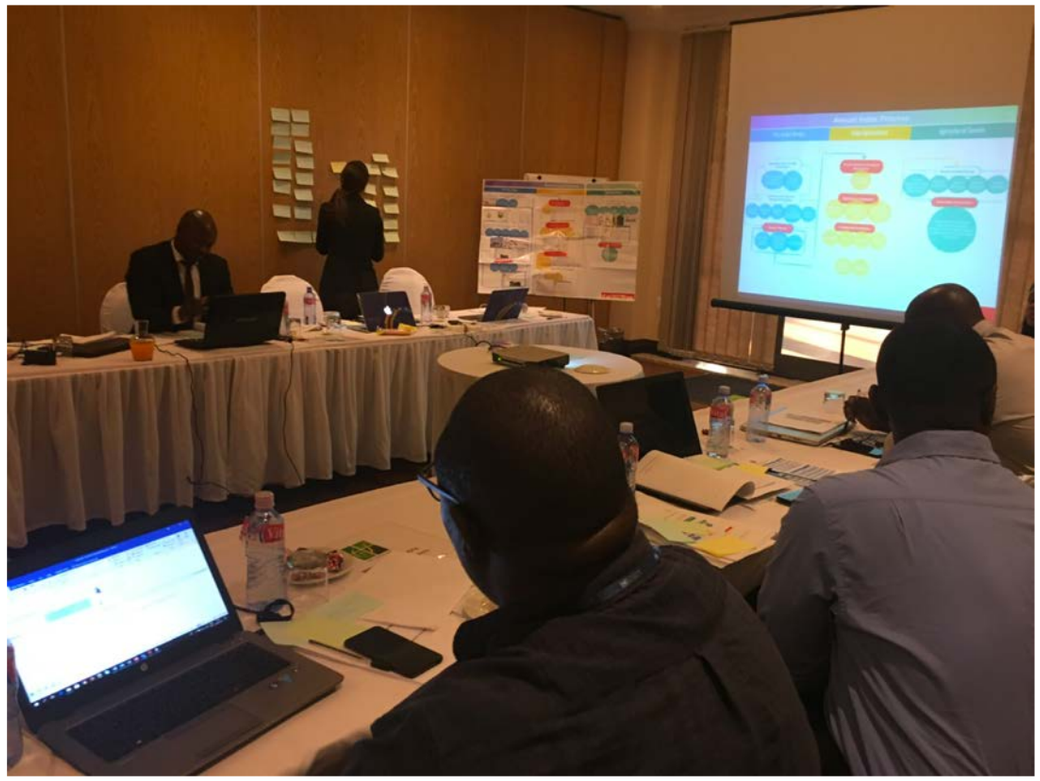
- Eg. Partner R4/ACToday roles workshops

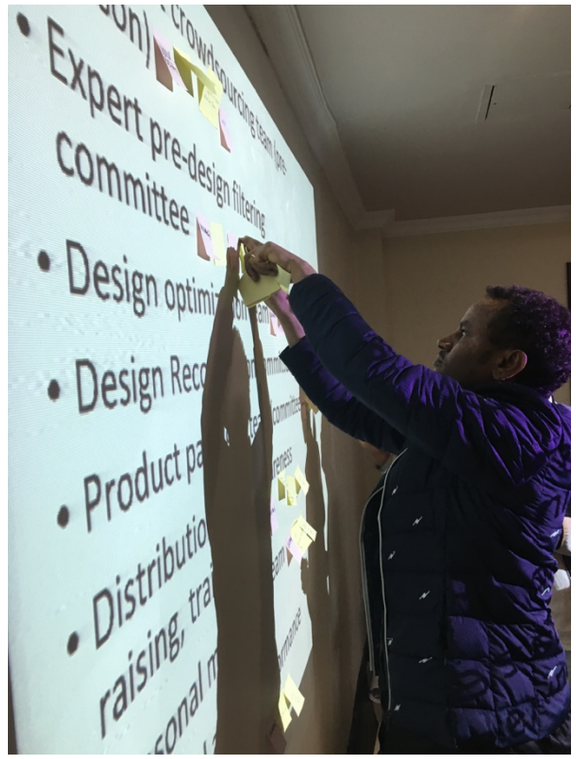
- Gov roles don't need to focus on subsidies:

  - Expertise, leading activities, convening, regulating, data, hosting platforms (eg ENACTS), field staff, field monitoring….

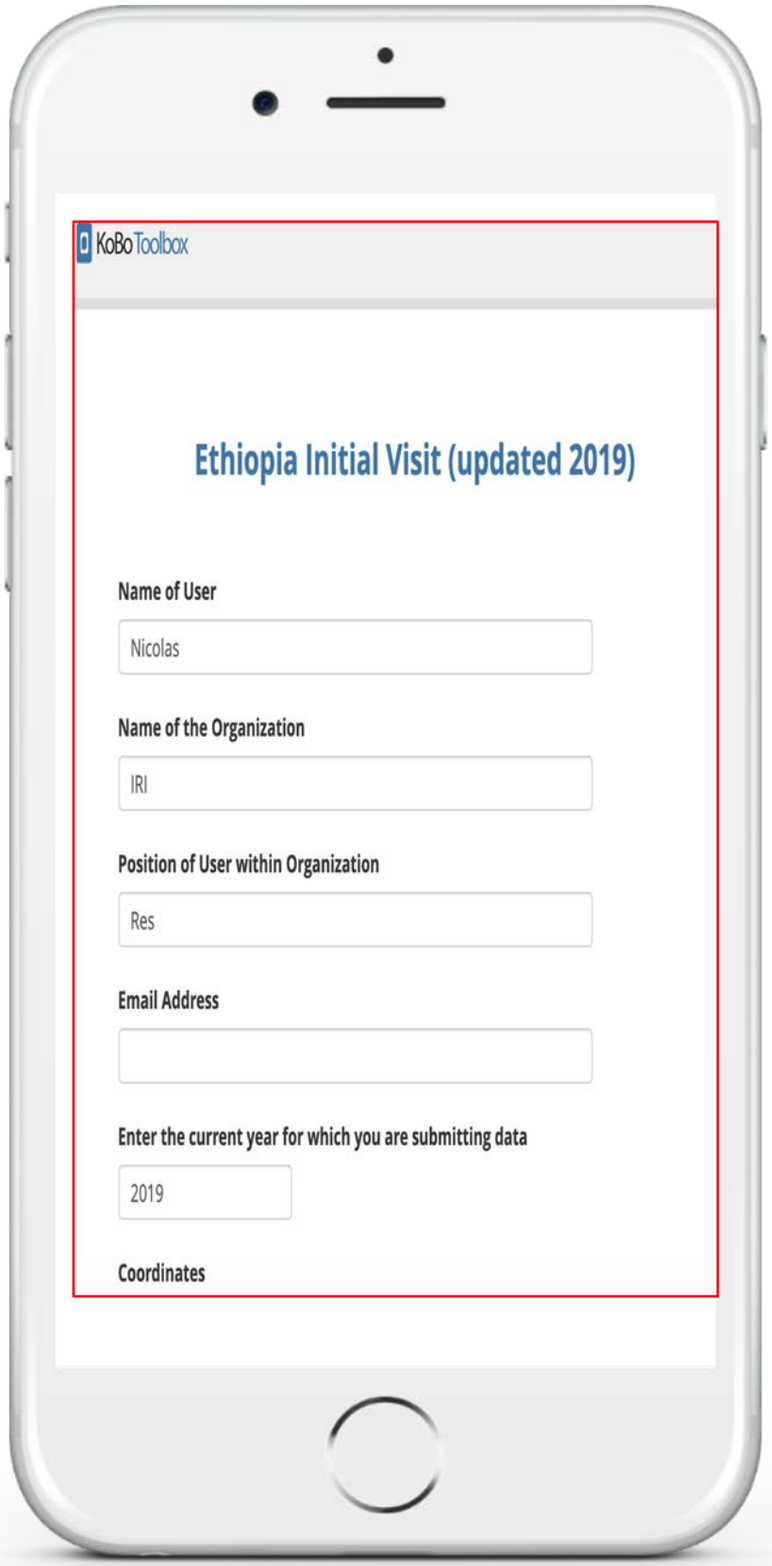
- Subsidies only to where most valuable (data, education, infrastructure)

- Having strong, valued, understood, cogenerated products is more effective than direct premium subsidies





VILLAGE REPORTING and technology



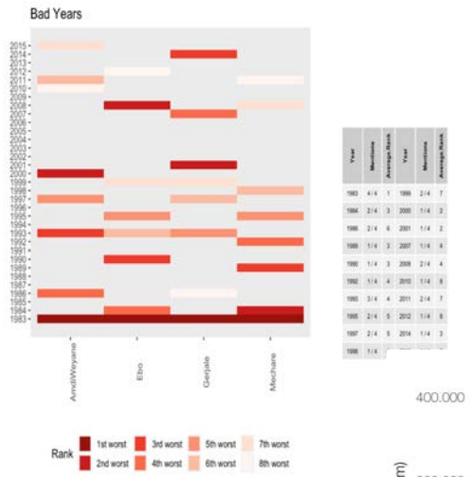
Information is gathered through the interactive exercise games designed specifically to identify the worst years in farmer memory.

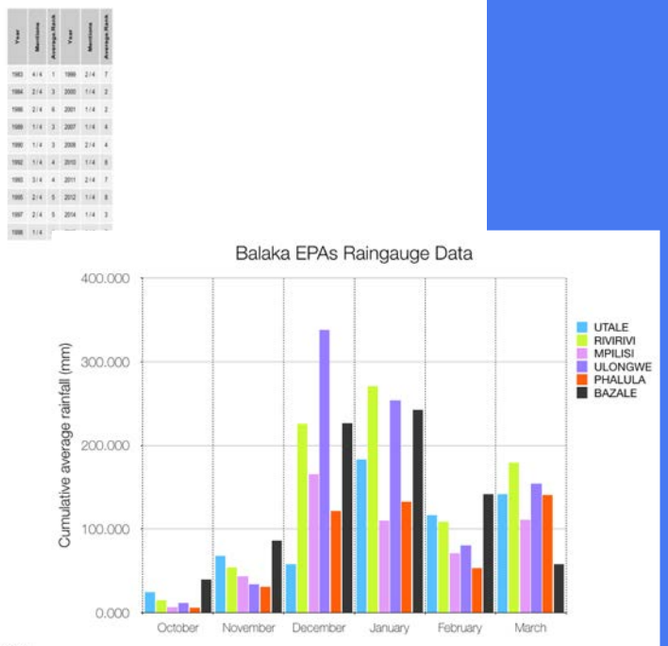
These exercises were normally collected on paper or computer. This process is now testing the uses ODK forms to collect data on tablets or smartphones.

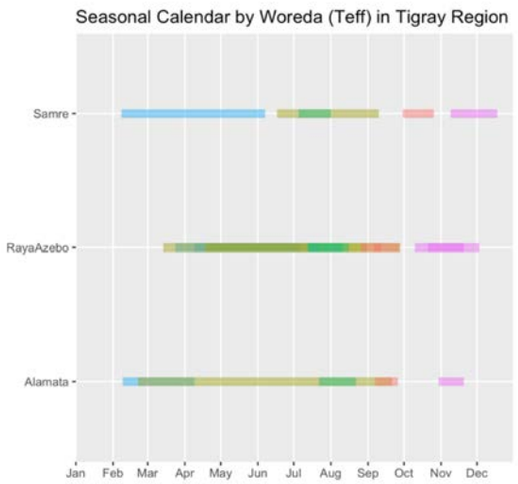
This way, we can get real time data while our partners are in the field. As well as find ways to scale our work from hundreds of villages to thousands and millions.

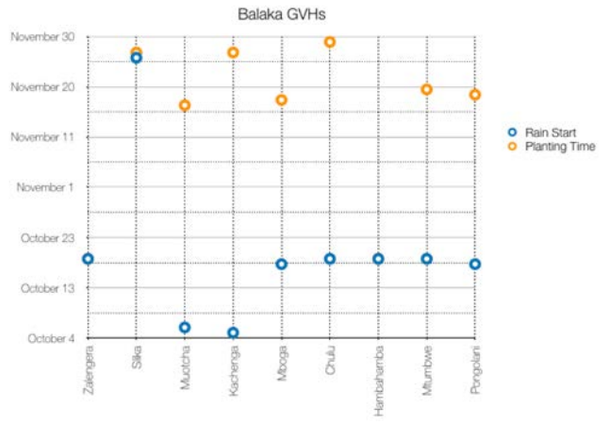
CROWDCORE INFO

There is a lot of information and data available when conducting index insurance design. These tools are for the design and help what affordable mean









Farmer Bad Years Heat Map

The heat maps are a visual representation of the farmer reported bad years

Seasonal Calendar

The Seasonal calendar is mainly obtained from the initial visits, and the information of the seasonal calendar is showcased in a visual representation

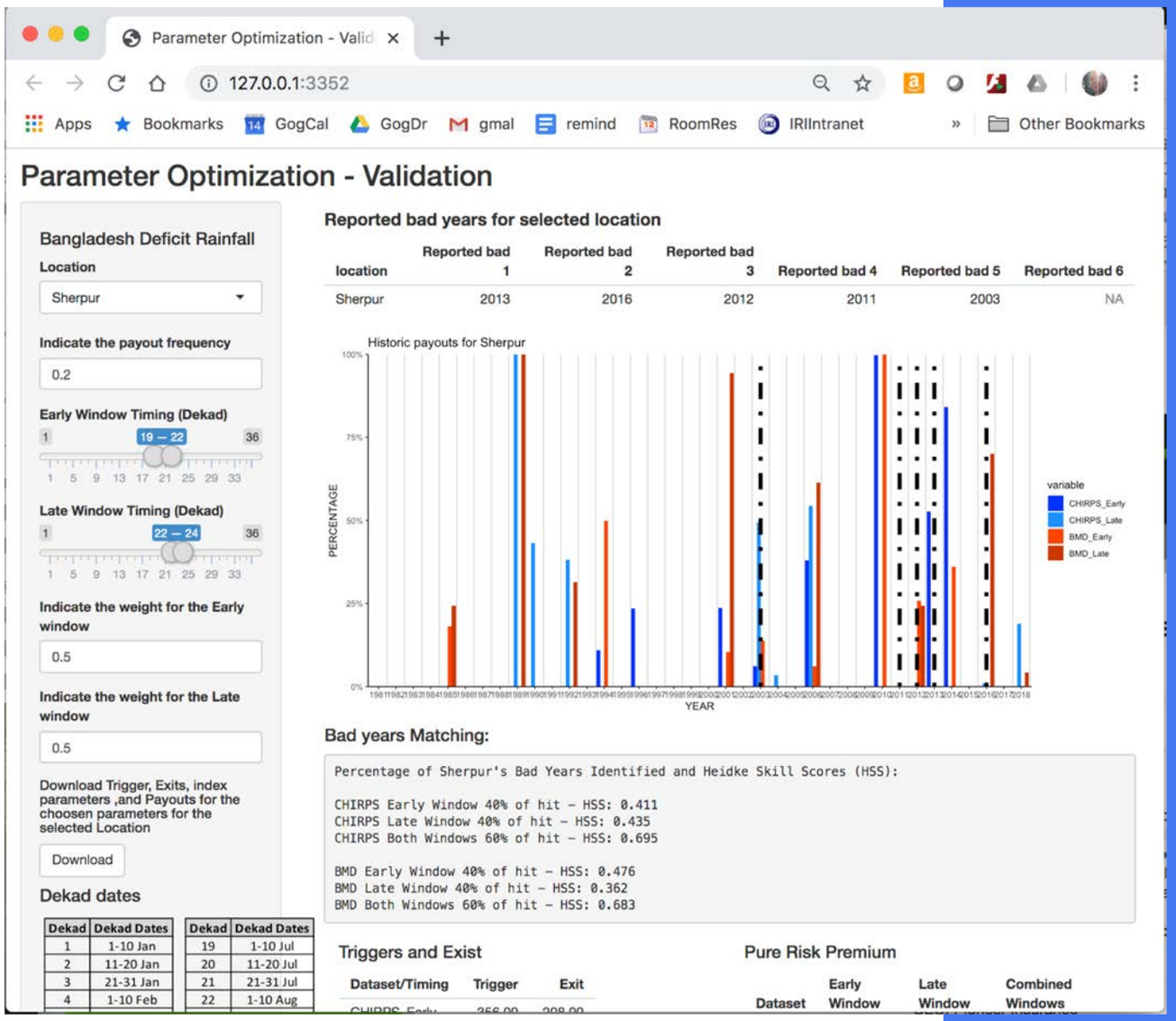
Raingauge Data

A very important piece of information that helps validate the index designed. This rain gauge data collection is in process

Is there evidence from satellites, IPC, Models, etc. to validate crowdsourced years?

		Season			Early										Late									
Year	Raya Mehr+ Belg Disaster Risk Profile	Yield		WRSI	Rainfall Estimates					Vegetation Estimates		Soil Moisture	Evapotransporative		Yield	WRSI	Rainfall Estimates					Vegetation Estimates		Soil Moisture	Evapotransporative	
Roman-Ethiopian	DPPC.gov.ET	A	B	Maize	ENACTS	ARC	CHIRPS	CHIRP	TAMSAT	EVI	AVHRRg	VUT	Hain	HainModis	Teff	Teff	ENACTS	ARC	CHIRPS	CHIRP	TAMSAT	EVI	AVHRRg	VUT	Hain	HainModis
1983-1976				0.11	0.41	0.4	0.51	1	0.24		0.79		0.29			0.17	0.35	0.1	0.24	0.05	0.16		0.54		0.1	
1984-1977				0.14	0.24	0	0.7	0.32	0.03		0.71		0.42			0.03	0.03	0	0.08	0.38	0.11		0.71		0.23	
1985-1978				0.22	0.43	0.6	0.76	0.43	0.35		0.17		0.55			0.11	0.49	0.4	0.24	0.68	0.27		0.33		0.58	
1986-1979				0.67	0.51	0.2	0.73	0.3	0.43		0.38		0.13			0.56	1	0.3	0.24	0.73	1		0.12		0.45	
1987-1980				0.06	0.11	0.1	0.11	0.19	0.16		0.08		0.45			0.22	0.3	0.2	0.27	0.16	0.05		0.21		0.03	
1988-1981				0.81	1	0.9	0.59	1	1		0.92		0.16			0.75	0.54	0.5	0.38	0.32	0.51		0.04		0.35	
1989-1982				0.39	1	0.3	0.16	0.43	1		0.46		0.68			0.14	0.3	0.3	0.16	0.62	0.22		0.5		0.48	
1990-1983				0.08	0.41	0.7	0.43	1	0.41		0.25		0.65			0.08	0.11	0.1	0.03	0.11	0.14		0.08		0.32	
1991-1984				0.61	1	0.8	0.78	1	1		0.33		0.32			0.47	0.05	0.1	0.05	0.03	0.03		0.25		0.19	
1992-1985				0.69	0.16	0.2	0.43	0.08	0.05		0.21		0.19			0.92	0.59	1	0.35	0.24	0.24		0.75		0.26	
1993-1986				0.31	0.24	0.7	0.43	1	0.14		0.04		0.58			0.56	0.14	0.6	0.41	0.51	0.59		0.17		0.06	
1994-1987				0.47	1	0.4	0.68	1	1		0.58		0.1			0.75	0.62	0.5	1	0.27	0.41		0.67		0.61	
1995-1988				0.42	0.32	0.3	0.14	0.3	0.41		0.96		0.71			0.67	1	0.6	1	0.43	0.46		0.42		0.74	
1996-1989				0.89	0.46	0.4	0.43	0.43	0.54		0.83		0.06			1	0.3	1	0.59	1	1		0.58		0.55	
1997-1990				0.25	1	0.8	0.43	0.43	0.49		0.12		1			0.31	0.08	0.2	0.11	0.08	0.19		0.29		0.16	
1998-1991				1	1	1	1	1	1		0.88		0.23			0.83	1	0.8	1	1	1		0.79		0.97	
1999-1992				0.97	1	0.9	1	1	1		0.5		0.26			0.42	1	1	1	1	1		0.88		1	
2000-1993				0.5	1	0.3	0.19	1	1	0.7	0.67		0.87	0.2		0.5	0.41	0.5	0.43	0.41	0.43	1	1		0.29	0.27
2001-1994				0.61	0.16	0.6	0.43	1	0.14	0.8	1		0.74	0.27		0.31	1	0.6	0.51	0.54	0.49	0.89	0.83		0.84	0.2
2002-1995				0.28	0.05	0.2	0.03	1	0.08	0.3	0.29		0.81	0.13		0.44	0.65	0.7	1	1	1	0.32	0.46		0.39	0.07
2003-1996	1			0.56	1	0.5	1	1	0.32	0.6	0.75		0.84	1	0.33	0.67	0.19	0.3	0.3	0.3	0.38	0.58	0.92		0.68	0.47
2004-1997				0.03	0.24	0.1	0.46	1	0.19	0.4	0.54		0.94	0.6	0.17	0.06	0.35	0.2	0.32	0.22	0.08	0.11	0.38		0.71	0.13
2005-1998	1	0.57		0.22	0.57	0.6	0.49	1	1	0.3	0.62		0.97	0.4	0.5	0.42	0.22	0.7	1	1	1	0.26	0.62		0.9	0.67
2006-1999	1			0.86	1	0.8	0.43	0.3	0.22	0.5	0.42		0.35	0.8	0.75	0.86	1	0.4	1	1	1	0.63	0.96		0.42	0.8
2007-2000				0.83	1	1	1	1	1	1		0.82	0.48	0.93		0.69	1	0.7	1	0.7	1	0.37		0.36	0.81	1
2008-2001	1	0.29		0.75	0.11	0.5	0.54	1	0.46	0.1		0.09	0.9	0.87	0.25	0.58	1	0.4	1	1	1	0.21		0.18	0.94	0.87
2009-2002	1	0.14		0.44	1	0.5	1	1	1	0.4		0.27	0.77	0.33	0.42	0.19	0.19	0.2	0.14	0.19	0.35	0.47		0.27	0.13	0.53
2010-2003		0.71		0.67	0.3	0.8	0.43	1	1	1		1	0.39	0.67	0.92	0.89	0.49	1	0.57	0.78	1	0.74		0.73	0.52	0.93
2011-2004		0.43	0.8	0.6	0.4	0.2	0.43	1	1	0.8		0.36	0.52	0.47	1	0.31	0.57	0.5	1	0.46	0.54	0.68		0.45	0.87	0.6
2012-2005		0.86	0.6	0.22	0.57	0.7	1	1	1	0.7		0.64	0.61	0.73	0.58	0.42	1	0.1	0.49	0.14	0.32	0.05		0.09	0.77	0.4
2013-2006		1.00	0.2	0.75	0.3	0.9	0.62	0.11	1	0.2		0.73	0.03	0.53	0.67	0.81	1	0.7	1	0.76	1	0.42		0.91	0.65	0.73
2014-2007			0.8	0.81	0.35	0.2	0.05	0.05	0.27	0.2		0.55		0.07	0.83	0.97	0.38	1	1	0.59	1	0.79		1		0.33
2015-2008			1	0.33	0.03	0.1	0.08	0.03	0.3	0.1		0.18			0.08	0.42	0.43	1	0.54	0.57	1	0.16		0.55		
2016-2009				0.53	1	1	0.84	0.22	1	0.9		0.91				0.78	1	0.4	1	0.65	0.57	0.53		0.64		
2017-2010				0.97	1	0.5	0.57	0.16	0.51	0.5		0.45				0.97	1	1	0.46	0.49	0.62	0.95		0.82		
2018-2011				0.92	1	0.1	0.65	1	1	0.6						0.67	1	1	1	1	1	0.84				
2019-2012					0.49	0.4	0.81	0.16	1								0.51	1	1	0.35	0.3					

BASIS RISK MINIMIZATION TOOLS



Dekad	Dekad Dates	Dekad	Dekad Dates
1	1-10 Jan	19	1-10 Jul
2	11-20 Jan	20	11-20 Jul
3	21-31 Jan	21	21-31 Jul
4	1-10 Feb	22	1-10 Aug

Thx Syngenta Foundation
Bangladesh Insurance Program

Crowdsourcing at core of Satellite Design Machine

Comparing different Station, Satellite, model, and yield/vulnerability data for different locations

… to check how the important parameters relate to farmer “bad years” collected from the field.

Optimizing most impactful parameters (eg seasonal timing) intuititively, and with advanced metrics

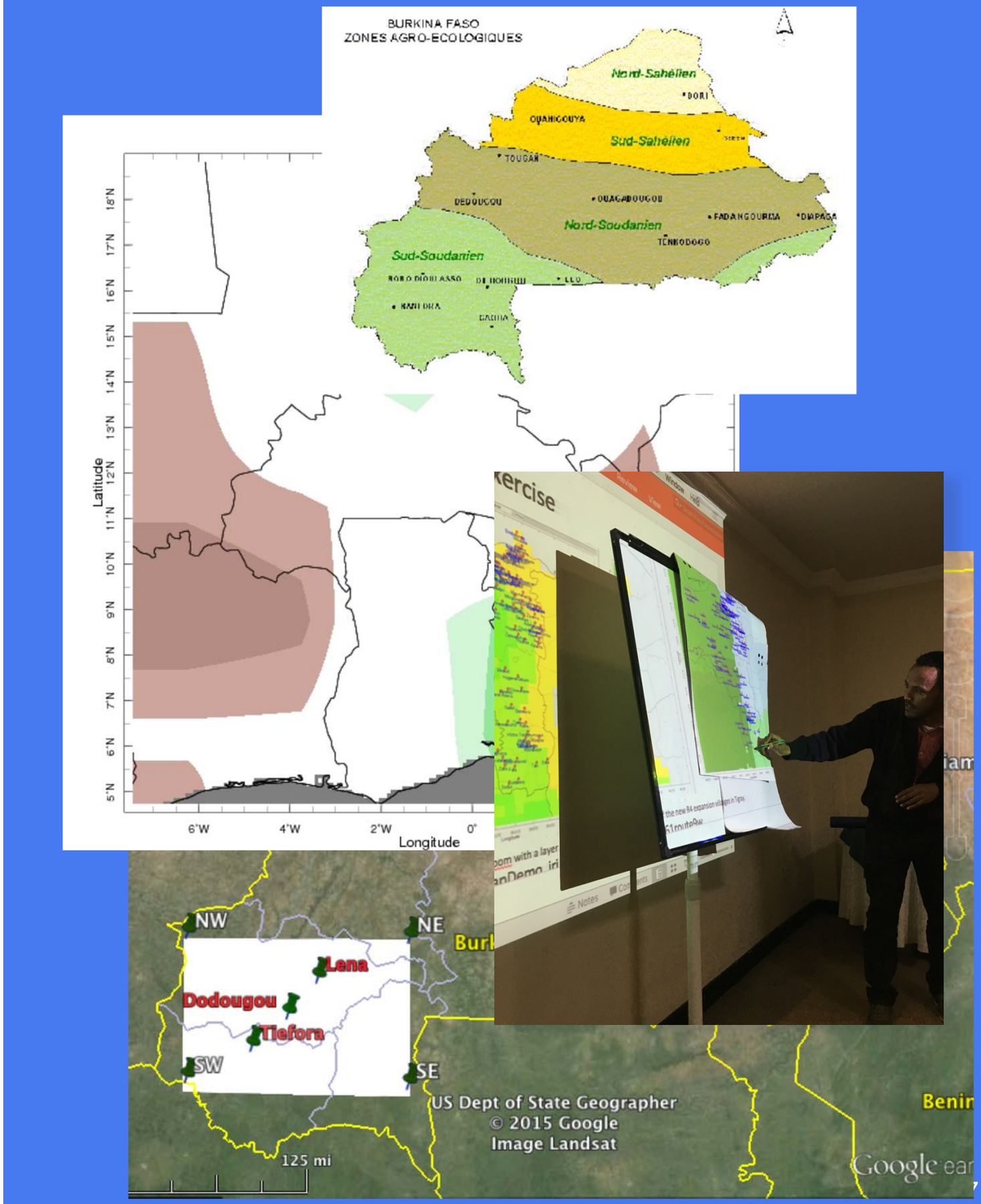
MAPPING AND SAMPLING VILLAGES FOR SCALE

How dense should crowdsourcing be?

For crowdsourcing intercomparision study of national level programs, need to determine how dense sampling should be to effectively inform parametric products

Target data gathering at lowest cost strategy

Set insurance regions, types at appropriate scales



INTER-COMPARISON ACROSS SCALES

Continent level crowdsourcing

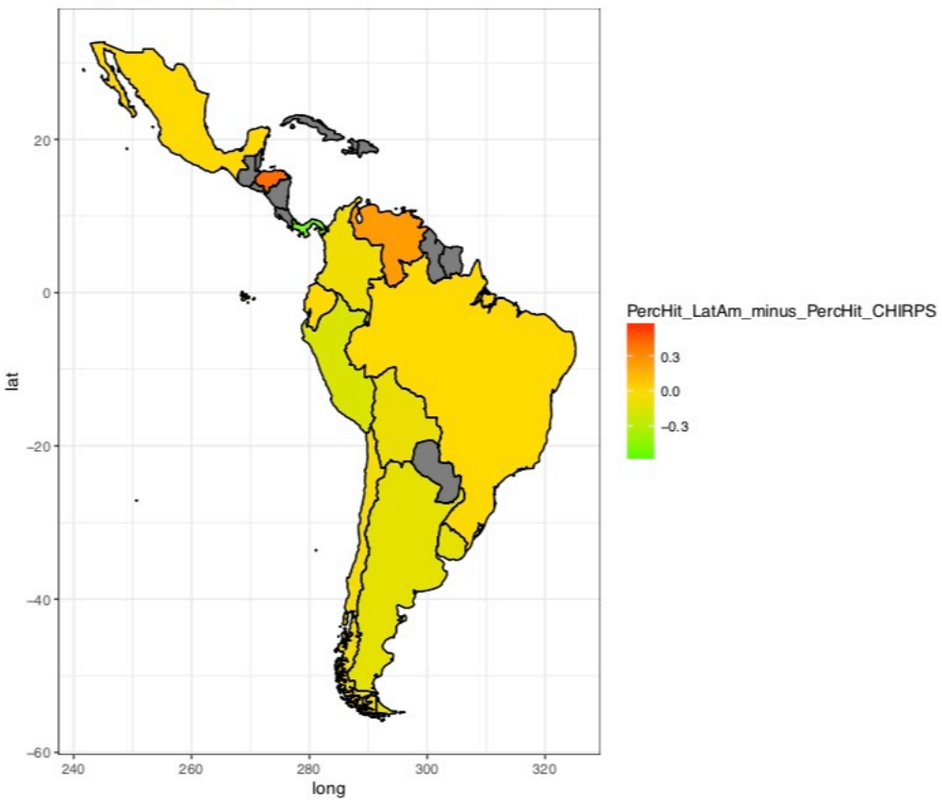
Can disaster events reporting be used to drive remote sensing applications? A Latin America Weather Index Insurance Case Study. (Brahm, 2019)

Combine, farmer crowdsourcing, disaster databases, satellite estimates for drought, excess, flood study, new data product

(Thanks Insuresilience CIF)

Map 13: Best percentage of hit performing dataset by Country for excess rainfall risk

Map 14: Difference in percentage of hit between LatAm and CHIRPS datasets for excess



Community Intelligence driving *"self" optimizing* indexes through *crowdcore* algorithms?

Thanks, using some gratuitous buzzwords