THE NEED OF DATA FOR INSURANCE: LESSONS FROM BANGLADESH

Vijayasekar Kalavakonda
Senior Insurance Specialist, Finance & Markets Global Practice
The World Bank Group

G4AWeek

21 - 24 September 2015

The Hague

NEED FOR DATA

YIELD INDEX INSURANCE

* Historical Yield/ Production Data
* Minimum 7-10 years of data

PARAMETRIC INDEX INSURANCE

* Historical weather data (for weather contracts – temp, precipitation…)
* Minimum 20-30 years of data

WORLD BANK GROUP

Cropping stages and data needs - MAIZE

*Maize yields are particularly sensitive to rainfall during the tasseling stage and the yield formation stage – rainfall during the latter phase determines the size of the maize grain

Sowing and establishment period is also critical crop survival

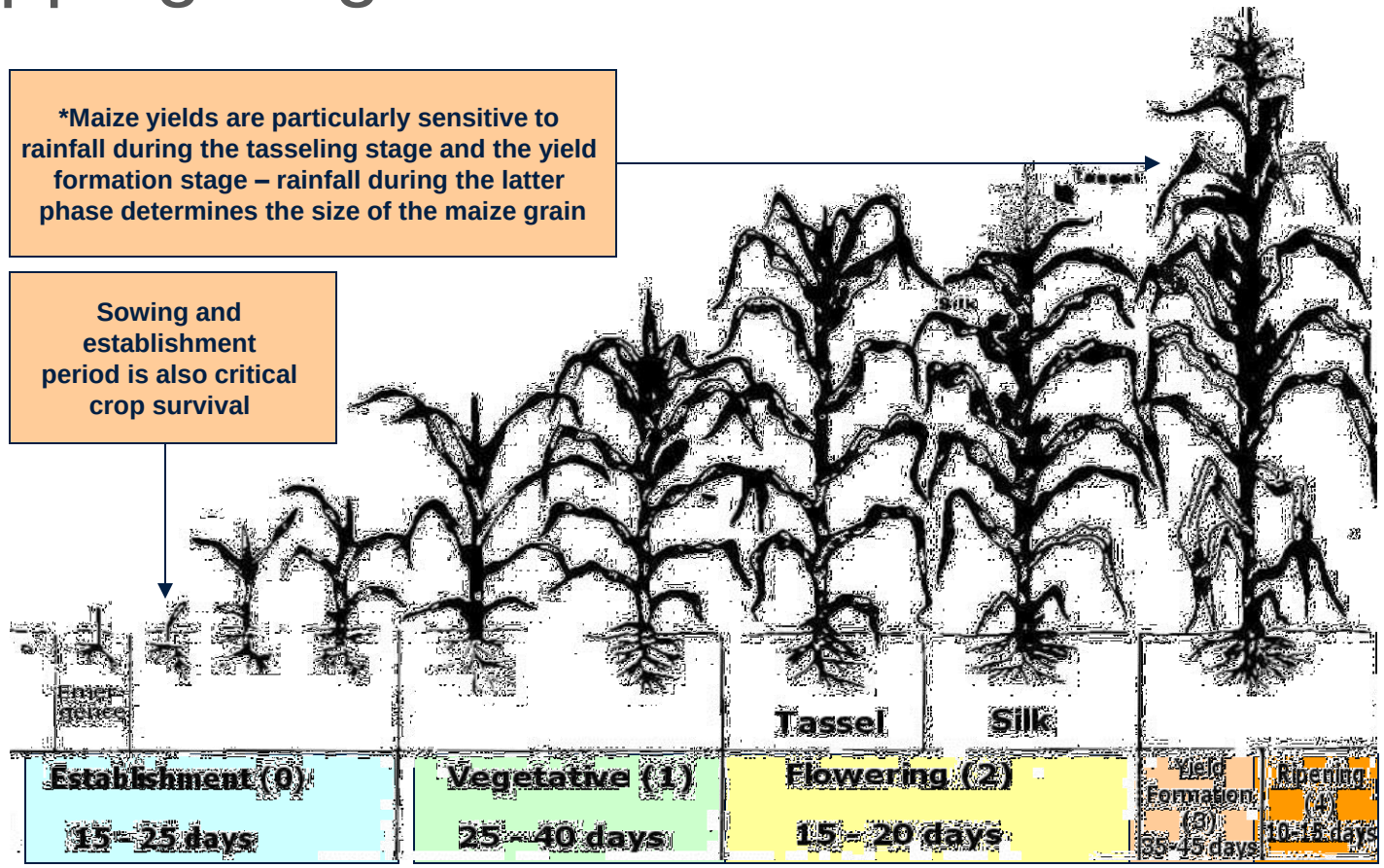


Diagram taken from the FAO's maize water requirement report*

- A rainfall index is normally split into 3 or more crop growth phases
- Objective: maximise the correlation between index and loss of crop yield

Source: William Dick Presentation

WORLD BANK GROUP

Maize Rainfall Index

	Phase 1	Phase 2	Phase 3
	Seedling Emergence to Knee High	Vegetative	Physiological Maturity
Days	30	21	30
Trigger (mm.)	35	50	60
Limit (mm.)	15	20	30
Tick size (Baht/mm./rai)	42	21	21
Sum Insured (Baht/rai)	1,200	1,600	1,700

Source: William Dick Presentation

WORLD BANK GROUP

Data Requirement – varies with need

LEVEL/PURPOSE	DATA GRANULARITY
MACRO Level - planning for food stock & logistics management	No need for high resolution data
MESO Level – protecting the portfolios of FI's or Agri Businesses	Not looking for one-to-one correlation with individual clients. Have a large spatial distribution of clients & risks and hence resolution of data up to 25 x 25 KMs can "possibly" do
MICRO level – correlating losses with individual farm experience	Require HIGH RESOLUTION DATA and very high-degree of confidence level

WORLD BANK GROUP

DATA REQUIREMENTS vs REALITY

	REQUIREMENTS	AVAILABILITY
Yield Data	Typically 7-10 years & at village or sub-district level	Varies from country-to-country BUT insufficient data & delays in data
Weather Data	20 – 30 years of HIGH RESOLUTION data, typically 5x5 or 10x10 KM	Lack of historical data and data from meteorological data from LIMITED Weather Stations
Quality of Data	Timely and fool proof (or trust worthy)	Yield data takes between 2-3 months (at a minimum) after harvest, whereas weather data is FAST Data collection is manual – both weather station & yield hence lack of trust in data

WORLD BANK GROUP

DATA REQUIREMENTS

1. Strong correlation with what is insured
2. Independently verifiable, i.e. based on well-described data sources and processing methods
3. Reliable delivery into future + available in near real-time
4. Available for sufficiently long period to properly represent climatic variability → payout probability and pricing
5. Information gathering at limited cost for insurer

Source: Anton Vrieling, University of Twente (2015)

WORLD BANK GROUP

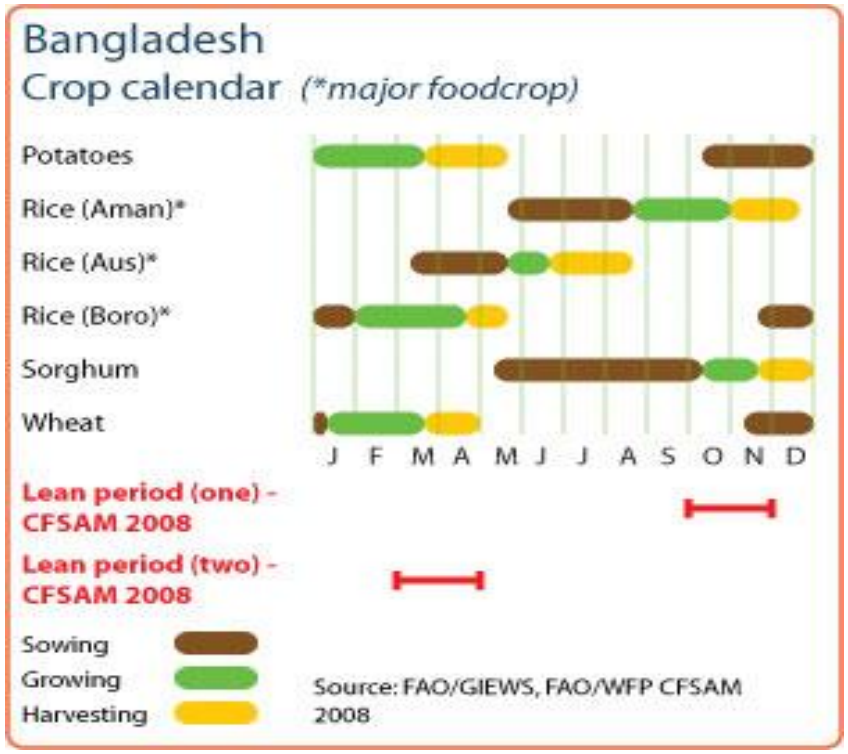
CASE STUDY

BANGLADESH

WORLD BANK GROUP

CROPPING SEASON vs WEATHER SEASONS

CROPPING SEASON



WEATHER SEASON

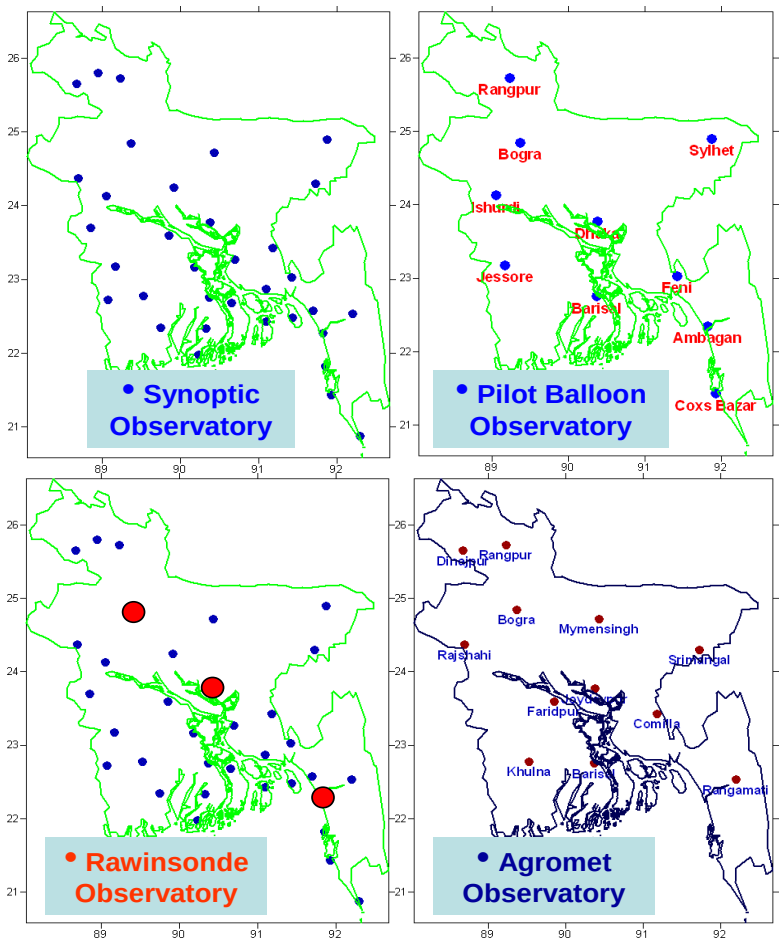
Seasons	Period	Weather Events	Rainfall
Summer (Pre-monsoon)	March to May	Nor'wester, Tornado, Hail, Cyclone, Heat Wave	19%
Rainy Season (Southwest Monsoon)	June – September	Heavy rain, Monsoon Depression, Flood	71%
Autumn (Post-monsoon)	October – November	Cyclone, Tornado	8%
Winter (Northeast Monsoon)	December – February	Abnormal Dryness (Drought), Cold Wave	2%

Footer Information

WORLD BANK GROUP

Observational Facilities

1. Synoptic observatories : 35
2. Pilot Observatories : 10
3. Rawinsonde Observatories : 3
4. Agromet observatories : 12
5. RADAR Stations : 5 (operational - 3 are Doppler Radar)
6. Earthquake Monitoring Stations : 4



WORLD BANK GROUP

Key Challenges in using existing data/data sources

1. Low Resolution:
   - the weather stations density is very low about 35 (old) and 6 (new) – average of 100 – 150 KM coverage area per weather station
   - concentration of weather stations more in the south and compared to the North and North West which are the key agricultural zones of Bangladesh.
   - Similarly, the satellite data resolutions are also pretty low. Satellite data gives a pretty good accuracy for 10-day (or decadal) and monthly rainfall in any given area.
2. Data Quality:
   - data accuracy and gaps in data are major problems – all weather stations are manual reading errors are common.
   - requires deploying data cleaning and data validation tools before the raw data could be used
3. Weak Weather Infrastructure,
   - Based on experience best results are achieved with an average coverage of 10 – 15 KMs per weather station. This would call for deploying more than 500 weather stations in the North and North West region of Bangladesh alone.

PROCESS FLOW – Building historical data with 25x25 KM (Resolution) Gridded Data

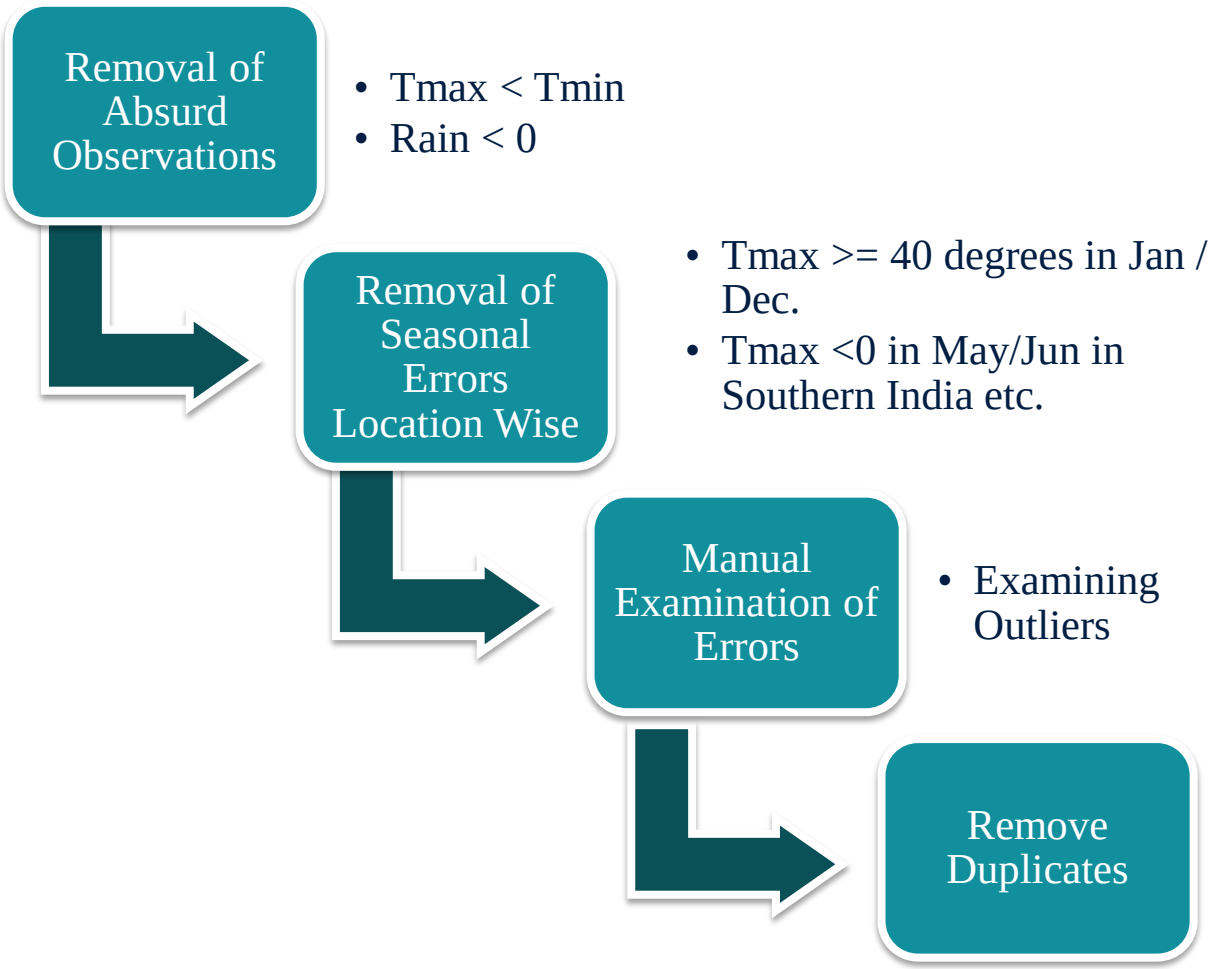
Daily station data for the three parameters (rainfall and temperature (min/max)) for last 30 years from the 35 synoptic meteorological stations

Data corrected for typographical errors. Statistical techniques applied on the data to search and fix these values

Aim of the project is to develop a gridded data product using station data – Using Kriging methodology to interpolated data to create square grid of 0.25^0

WORLD BANK GROUP

QUALITY CONTROL PROCESS



WORLD BANK GROUP

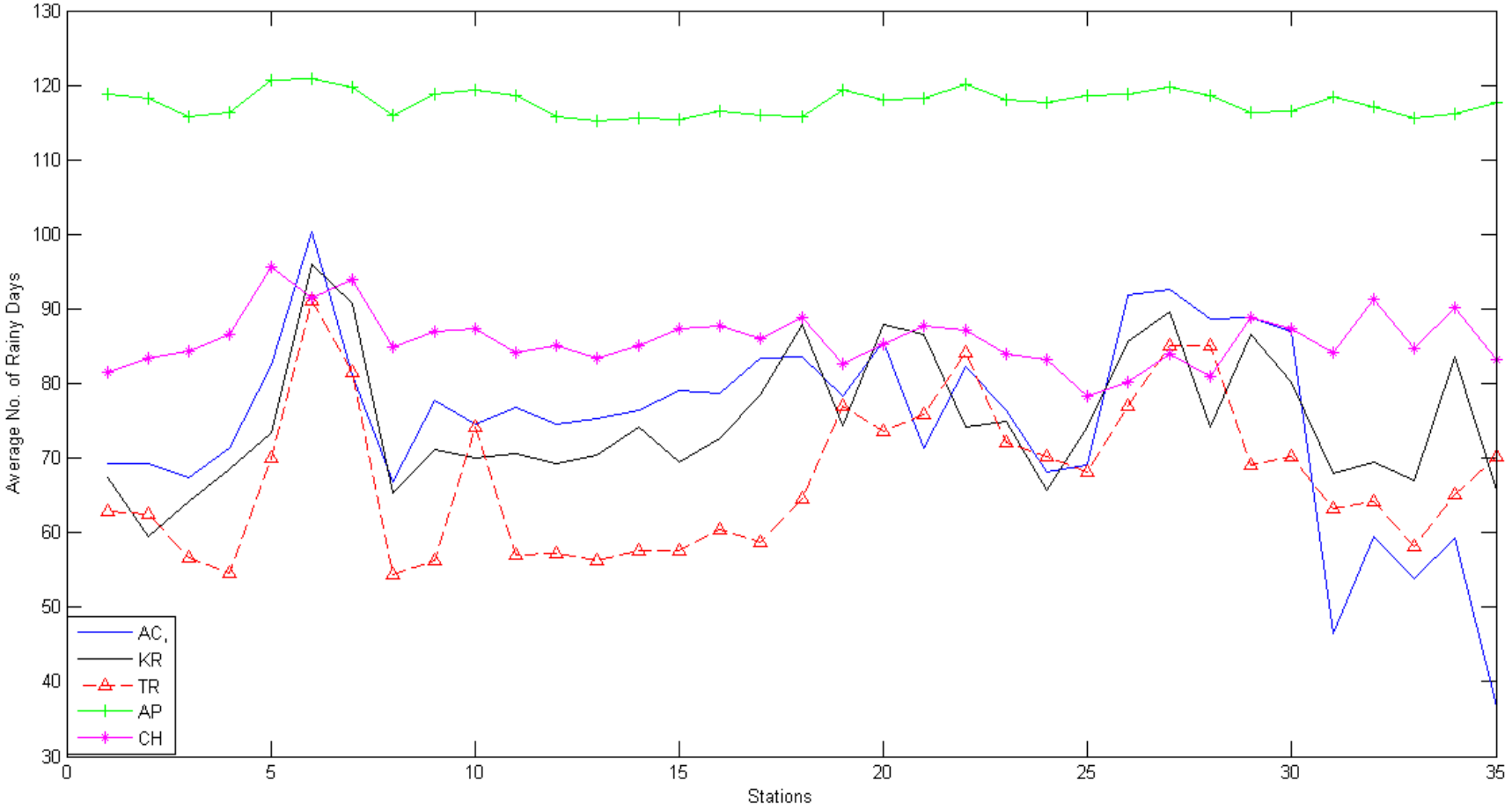
VALIDATING GRIDDED (SYNTHETIC) DATA

Data sets	Parameters	Availability	Spatial Resolution	Comments
TRMM (TR)	Precipitation	2000 Onwards	0.25 degree box	Overestimates heavy rainfall. Good at 10days/monthly scale. So can be used to develop products at longer time scale, but not at daily scale. Cannot be used for claim settlements as the data can be challenged. But can be used for monitoring season.
CHIRPS (CH)	Precipitation	1981 onwards	Two resolutions 0.25 and 0.05 degree grid boxes	It is TRMM data merged with Geostationary IR satellites and then blended with observations. Hence, possess properties of TRMM but slightly closer to the observations. Therefore, cannot be used for claim settlements
Reanalysis datasets	Temperature/ Precipitation/ Wind/ Humidity	1948 onwards	Various resolutions ranging from 5 degree to .5 degrees	These are weather model's initial conditions at various time steps. Very coarse. Can be used to know average temperature/precipitation for a particular month/season and to run weather models.

FOCUS GROUP DISCUSSIONS – data collected from farmers (400 Farmers)

WORLD BANK GROUP

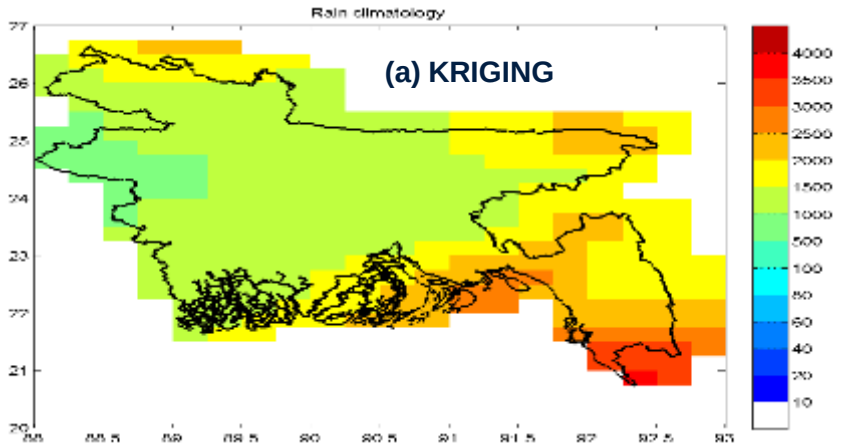
RESULTS: Average number of rainy days over all years

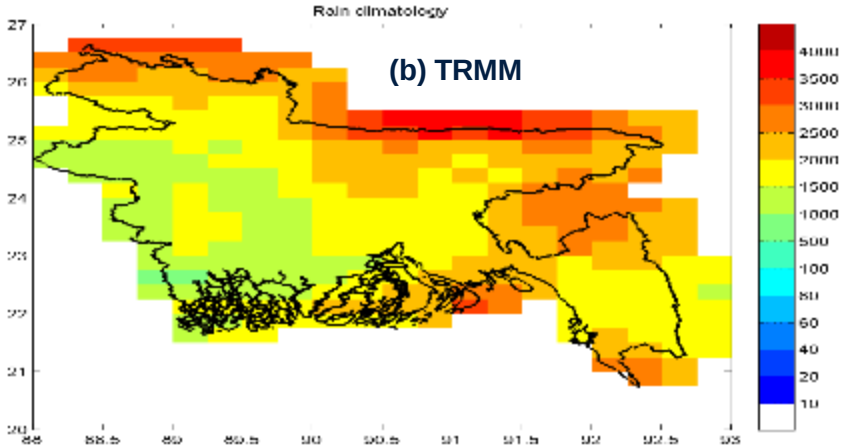


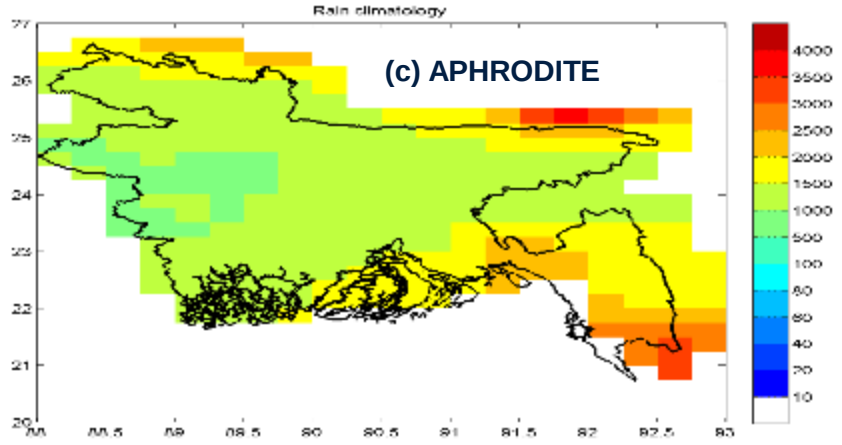
AC: Actual Data; KR: Krigging; TR: TRMM Data; AP: Aphrodite; CH: CHIRPS

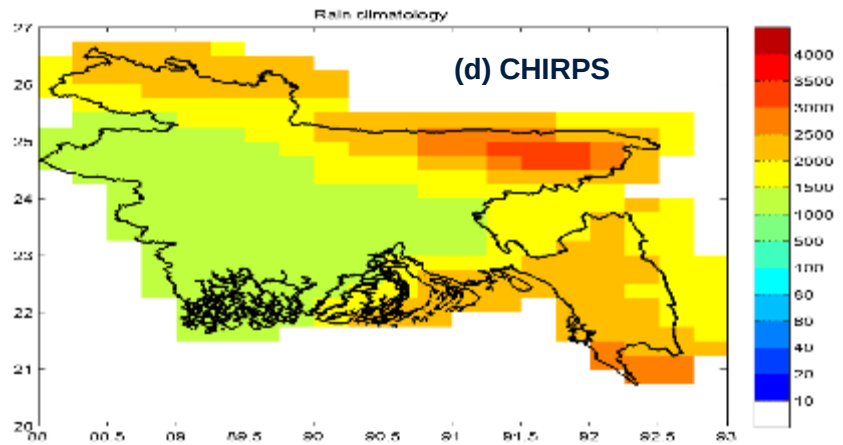
WORLD BANK GROUP

RESULTS: Monsoon season (JJAS) climatology of rainfall over Bangladesh









WORLD BANK GROUP

OBSERVATION – KRIGGING DATA

1. All data sets show similar climatological features for precipitation during the monsoon season.
2. The number of rainy days in APHRODITE data are noticeable more than other data sets.
3. **Kriging data is more close to observations than other data sets.**
4. TRMM data is satellite derived product and might miss some events due to constraints on number of satellite passes per day over a particular region.
5. CHIRPS data is prepared using TRMM data and NOAA's infrared satellite and then blended using station observations. The number of stations over Bangladesh used for blending during the period (1981 to 2014) are changing.

CHALLENGES

1. Data continuity & compatibility
2. Data quality Rainfall: performance vis-à-vis *in situ*?
3. Spatial detail (resolution): what do we monitor?
4. Low basis risk (or high insurance quality) – index to correlate with losses and meet demand
   a) Index construction: vegetation/rainfall + options!
   b) Calibration / validation options
      - Lack of good crop/livestock statistics
      - Insurance is often single peril, but yield reflects multiple perils…
5. COST – High quality data costs money which imposes significant burden on farmers in terms of affordability

Source: Anton Vrieling, University of Twente (2015)

CONCLUSION

1. Need to address the data challenge – building ground level infrastructure will take time & money – RS Data can be an alternative?
2. Challenge – getting high-quality (i.e., high resolution) data at regular intervals i.e., DAILY vs. DECADEL vs. MONTHLY?
3. Cost of Data – High Resolution Data costs money, which increases the premium cost significantly
4. DATA in useable form for NON-METEOROLOGISTS/RS Experts