What flood event map accuracy is required to enable governments, aid agencies, and insurance companies to protect vulnerable lives and livelihoods?

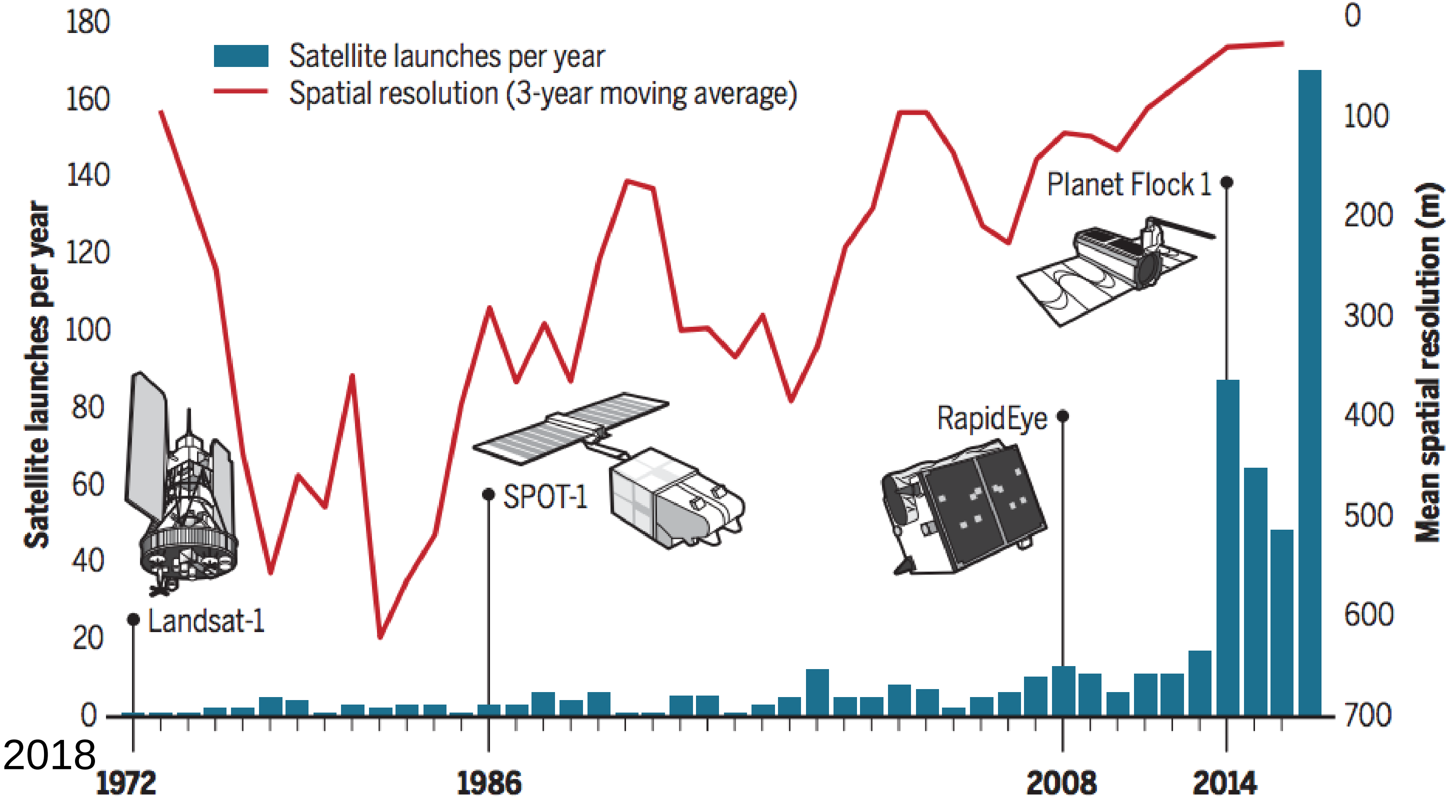
Beth Tellman

@cloud2street

Sam Weber, Jeff Ho, Jon Sullivan, Bessie Schwarz, Colin Doyle

Cloud to Street

Exponential increase in earth observing satellites



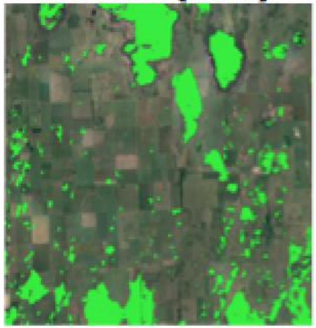
Finer et al 2018

Microsats, drones and the imagery revolution

MODIS (250m)

250m, 2 daily 2000-present

Landsat (30m)

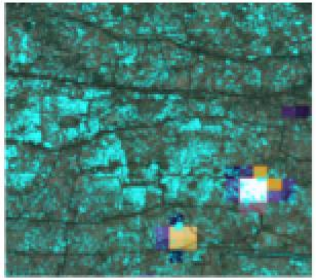


16 days, 1984-present

Sentinel-1 (10m)

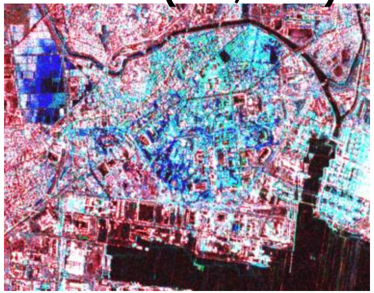
10m, 5-12 days, sees through clouds, 2014-present

PlanetScope (3m)



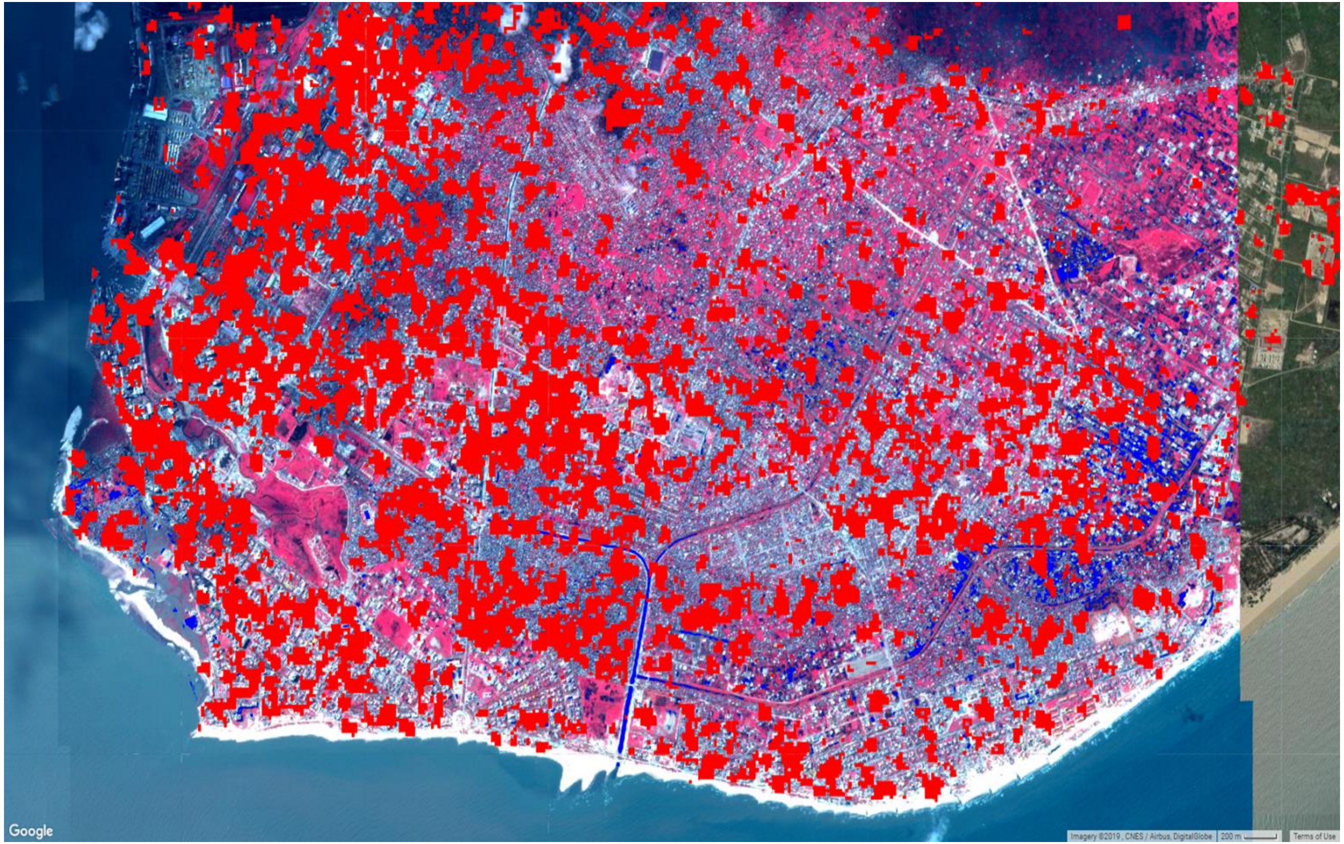
SkySat (80cm)

ICEYE (1m, SAR)



Drone capture: Houston, 2017

Urban Flooding- Sentinel-1 (10m) vs. Skysat (80cm) March 23, Biera



Legend

Damaged Buildings - LIST

Flood - SkySat

Flood map science to decisions

Algorithm published

Code available

Map repository, dashboard or volunteer .pdf and .tiff

Data to decision pipeline

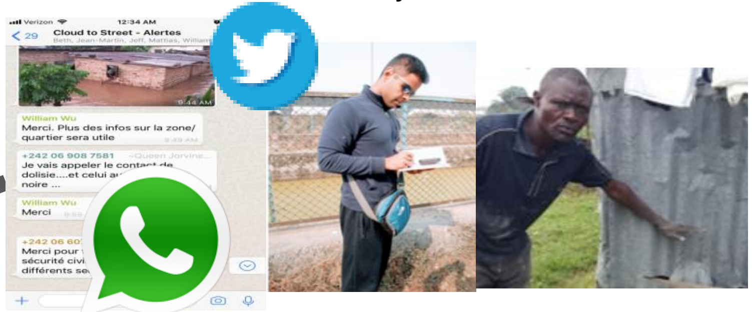
Cloud to Street + other boundary orgs (ICIMOD, CEMADEN, UN-SPIDER, ARC…etc

Flood protection decision from flood map

Data to decision pipeline- Flood Monitoring in the Republic of Congo

https://congo-flood-monitoring.cloudtostreet.info/recent-data

roundtruthing through field agents, the news, the community or social media



NASA

esa
European Space Agency

planet.

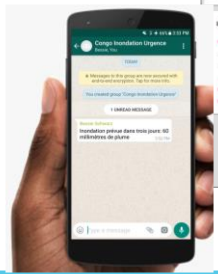
ICEYE

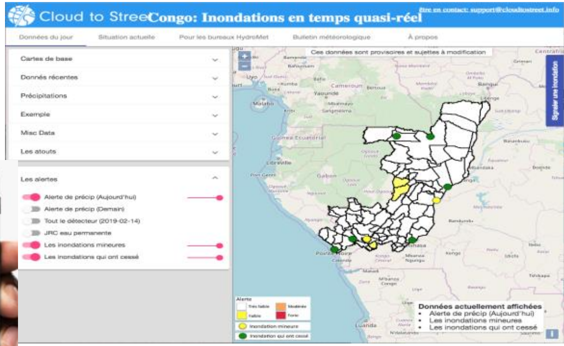
DigitalGlobe

TERRASAR-X SERVICES

Locally-optimized flood detection, with maps fused into one

Automated AI and physics based algorithms in the cloud





Interactive web portal + WhatsApp alerts

Are the existing algorithms to extract surface water good enough to enable flood protection?

For whom?

Well…that depends…

Agenda

1. How remote sensors measure accuracy and why it doesn’t work for making decisions from flood maps

2. For whom are we (or should be!) measuring accuracy?

3. A framework and proposed methods to make science usable for the people who make flood resilience decisions

Typical Remote Sensing Accuracy Assessment

Confusion Matrix

n = 165	Predicted: No	Predicted: Yes	
Actual: No	Tn =50	FP=10	60
Actual: Yes	Fn=5	Tp=100	105
	55	110	

Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment

Robert Gilmore Pontius Jr & Marco Millones

-made for land change maps that don't have clouds

-random stratified sample overestimates accuracy

-Critical Success Index biased towards overestimating flood models (Stephens et al 2015)

-biased towards LARGE slow moving long duration floods

Which satellite can enable affordable insurance products

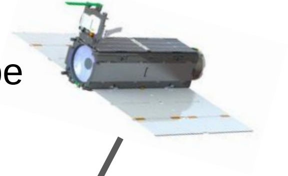
esa

European Space Agency

Sentinel-1
A&B

planet.

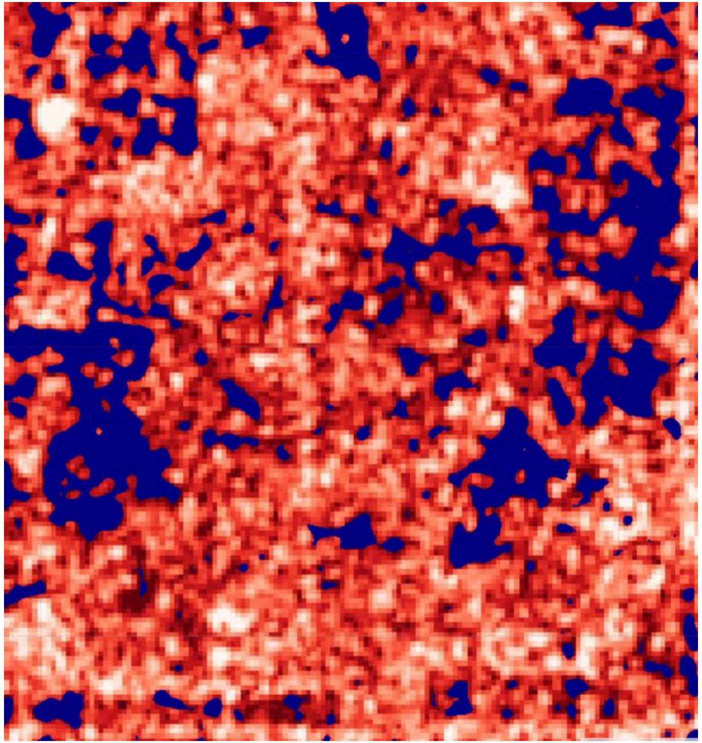
PlanetScope



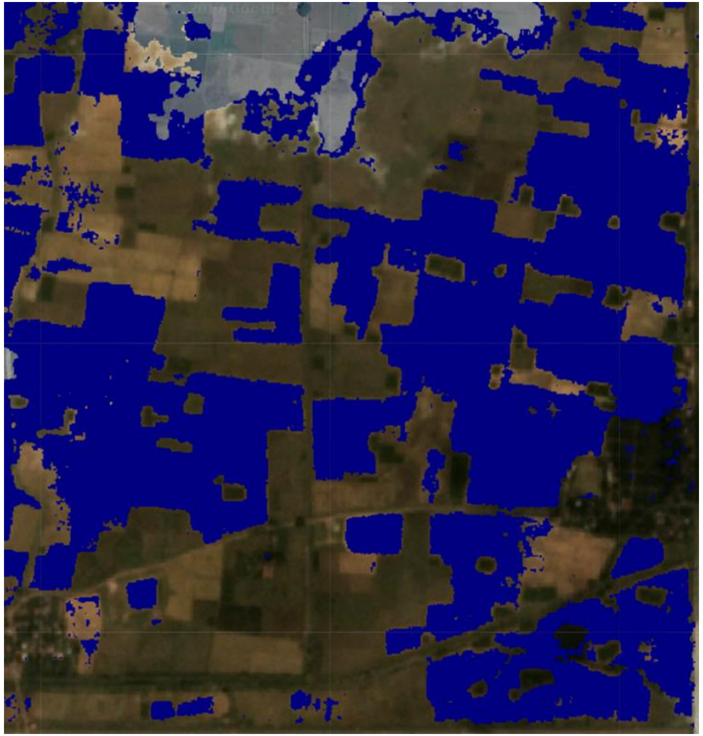
InsuResilience

giz Deutsche Gesellschaft
für Internationale
Zusammenarbeit (GIZ) GmbH

Sentinel-1

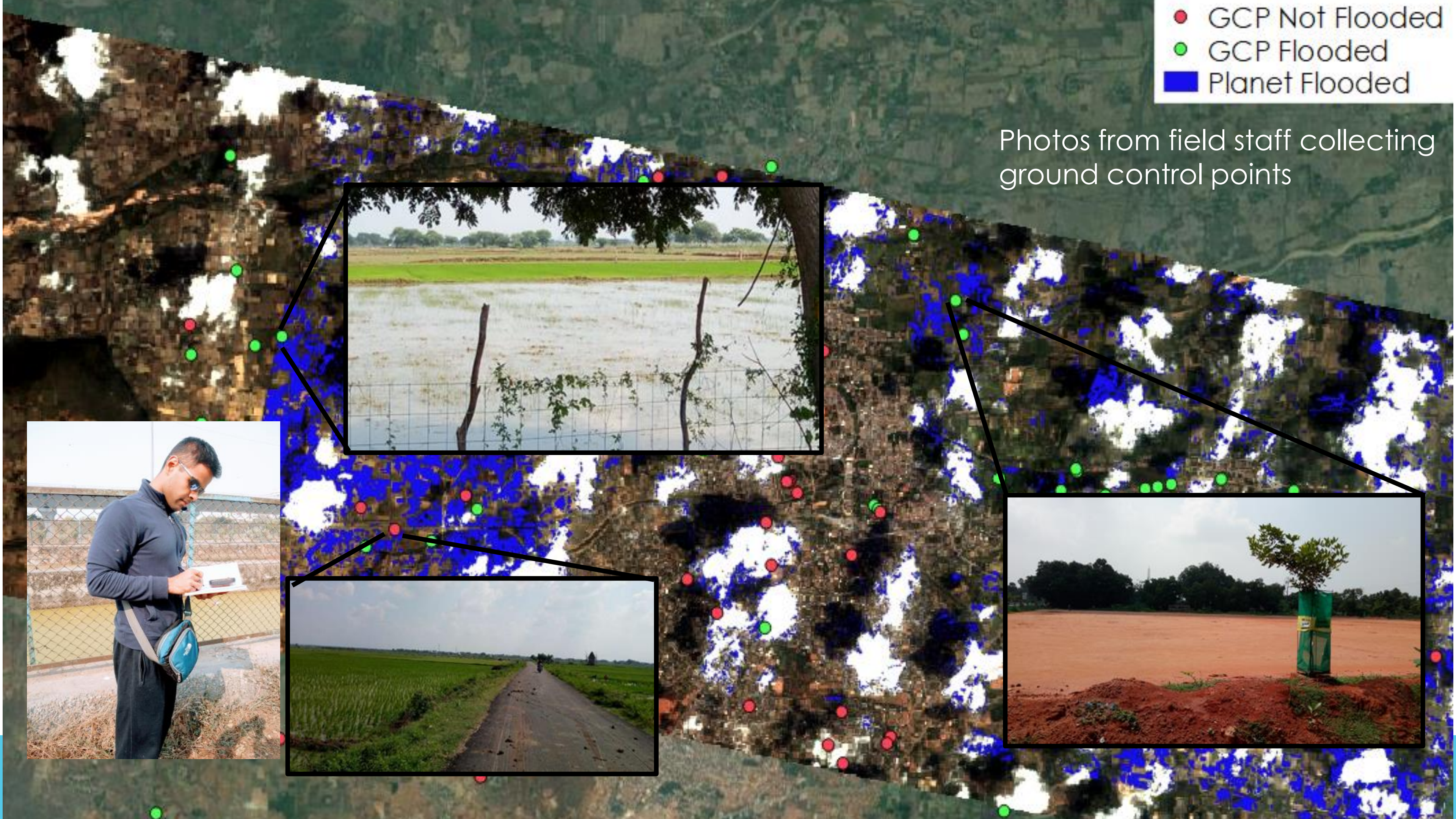


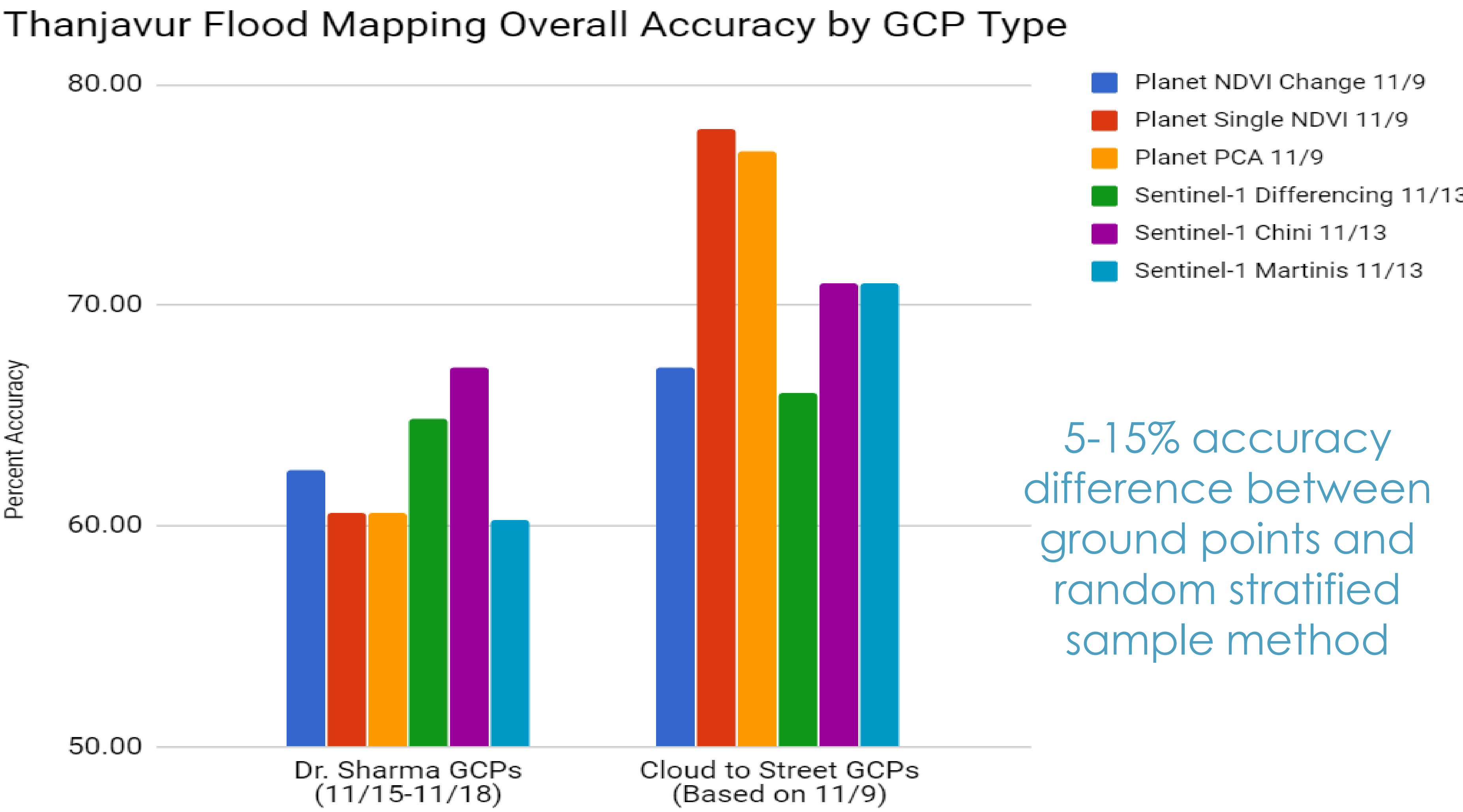
Planetscope



Low High Flood

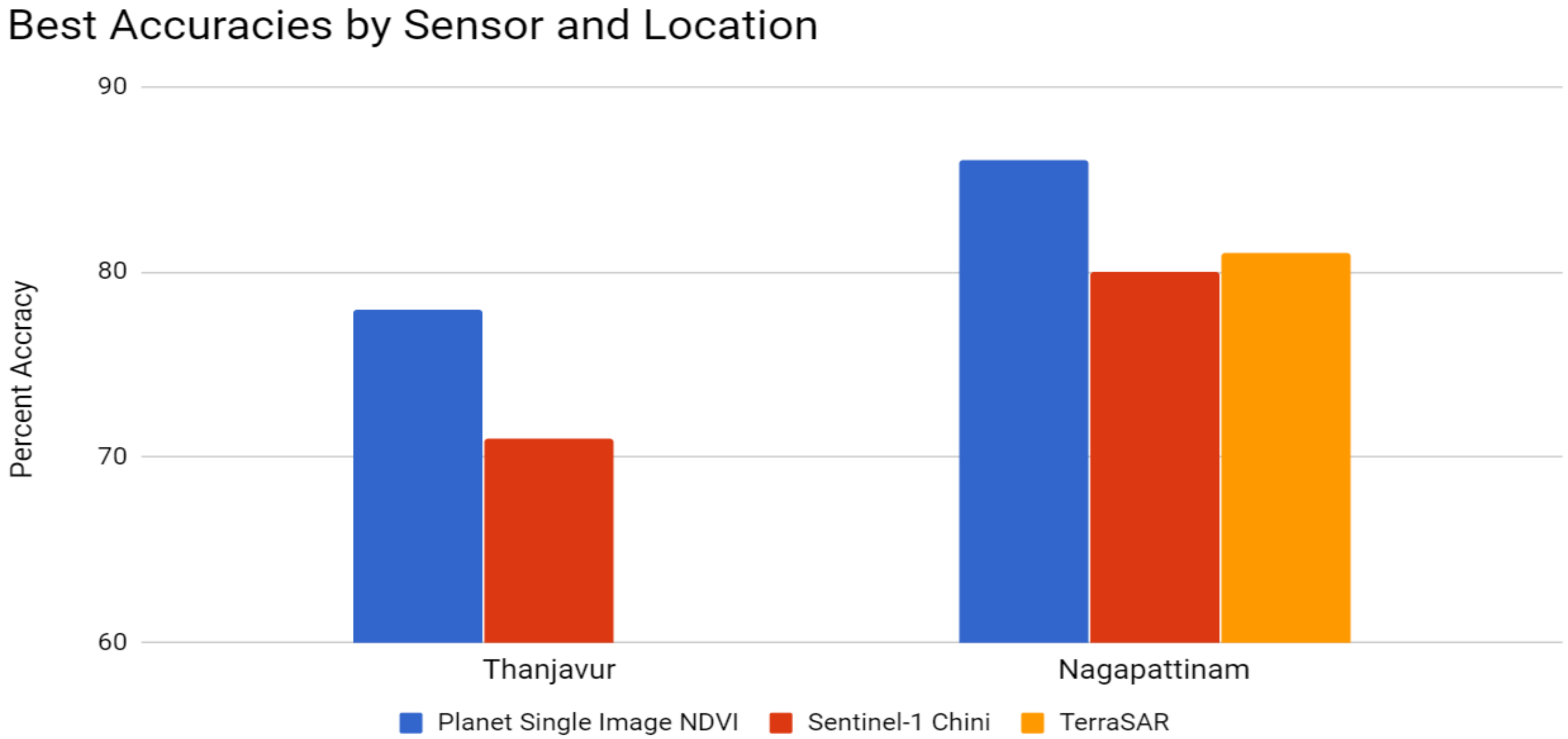
Backscatter


GCP Not Flooded
GCP Flooded
Planet Flooded
Photos from field staff collecting ground control points



5-15% accuracy difference between ground points and random stratified sample method

PlanetScope as high as 86%, Sentinel-1 80%, TerraSAR StripScan 81%
CLOUDS, REVISIT TIME, IGNORED



Why isn't the accuracy of these maps (72% & 80%) as high as it is in the publication (89%- Chini et al 2017)?

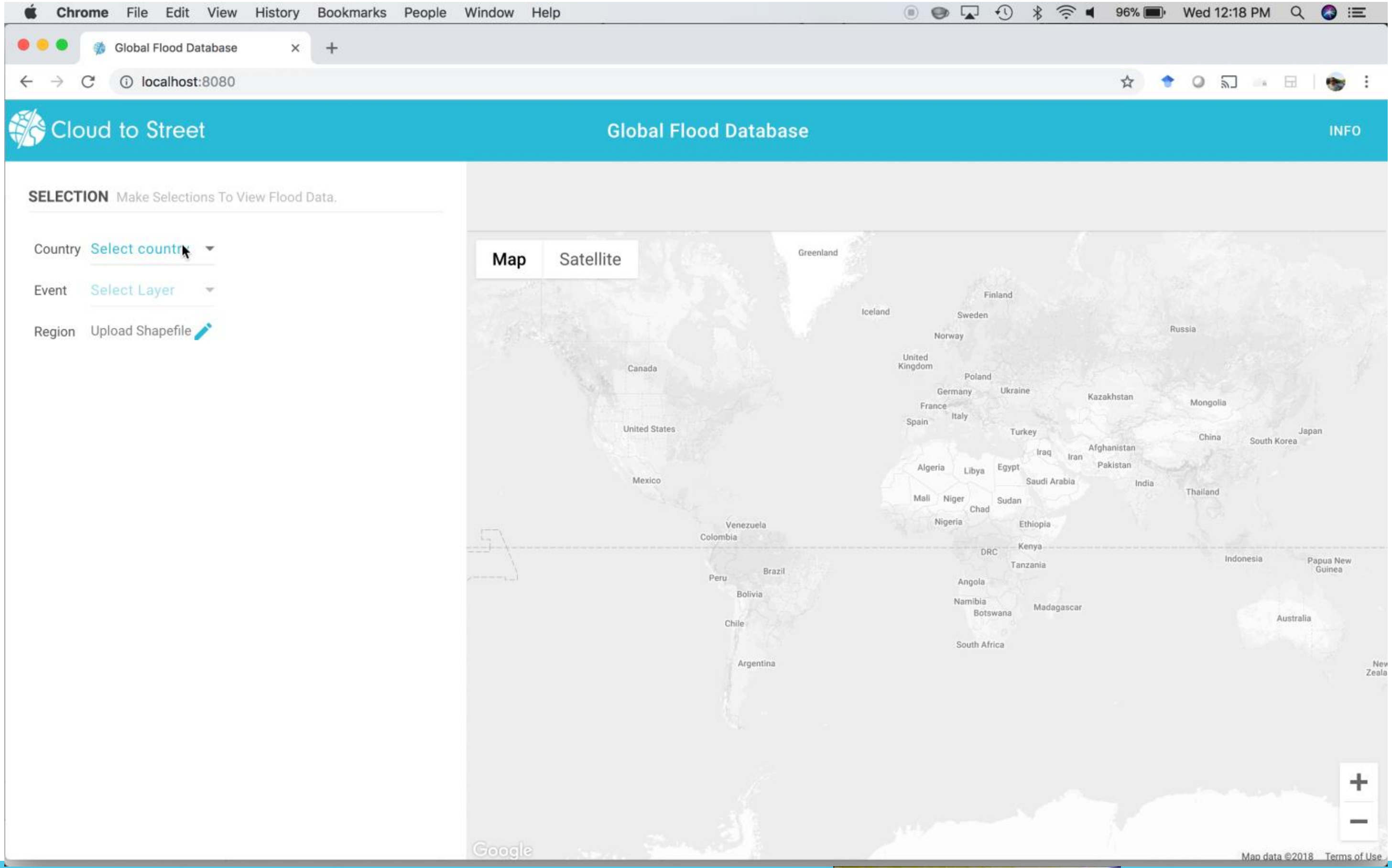
Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

-publication bias towards good maps, low sample sizes

-biased towards the biggest (EASIEST) floods to map

-wide ranging regional variability...rarely tested

Global Flood Database: 896 high quality floods at 250m resolution 2000-2017 (83% accuracy)



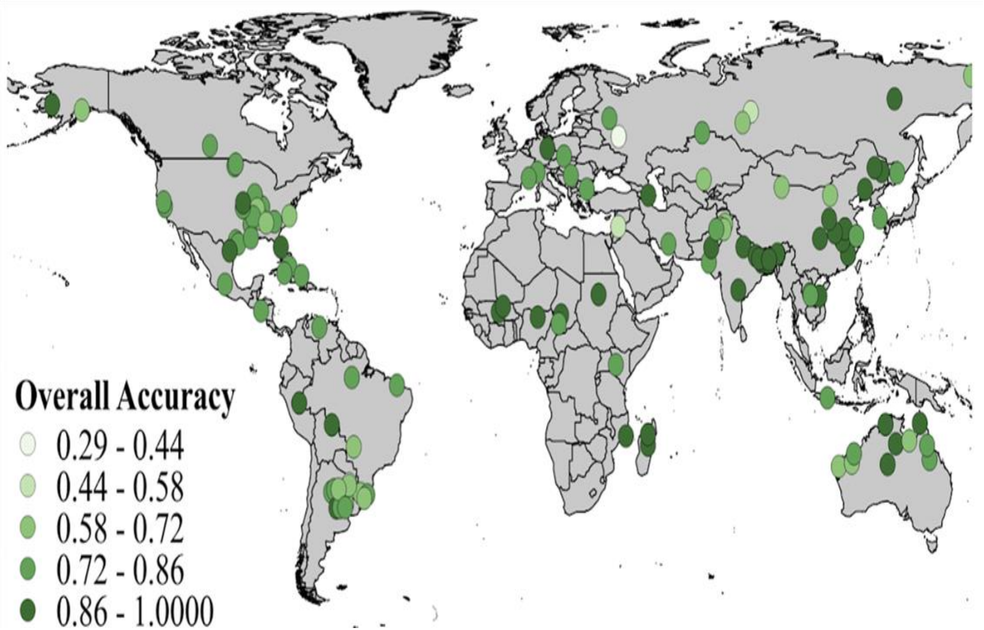
@cloud2street

Google Earth Engine

Dartmouth Flood Observatory

Cloud to Street

Global Flood Database variance in event accuracy and "quality"

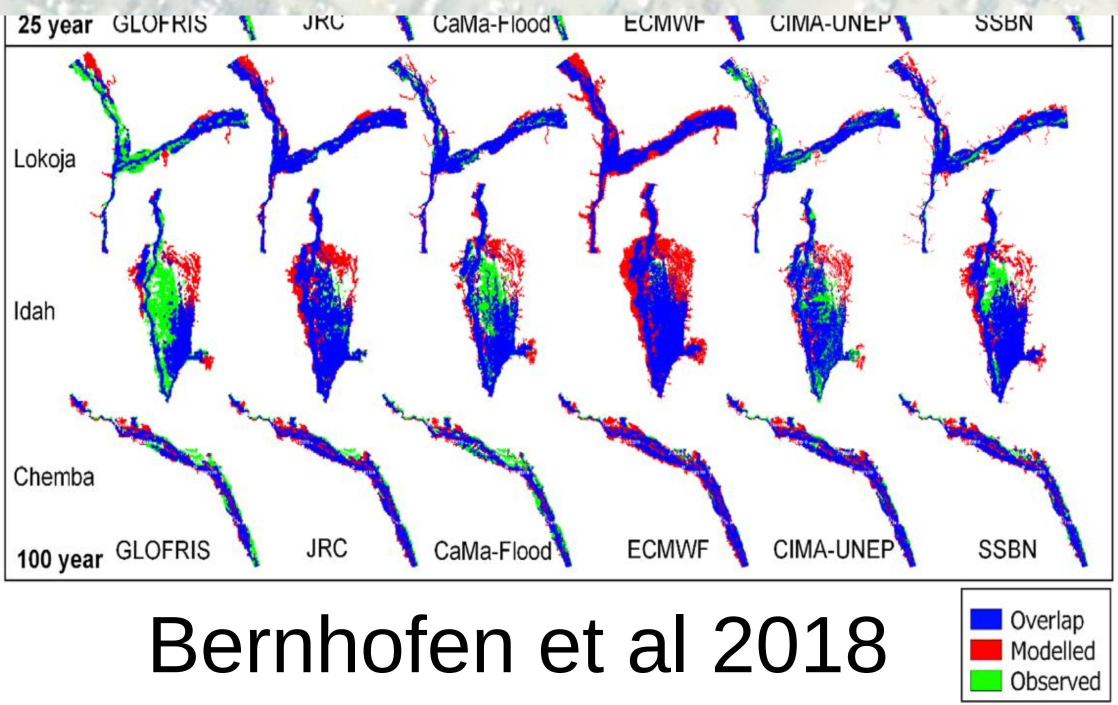


Using MODIS DFO algorithm

(Brakenridge and Anderson 2006)

Remote Sensing to Flood Model Accuracy Assessment

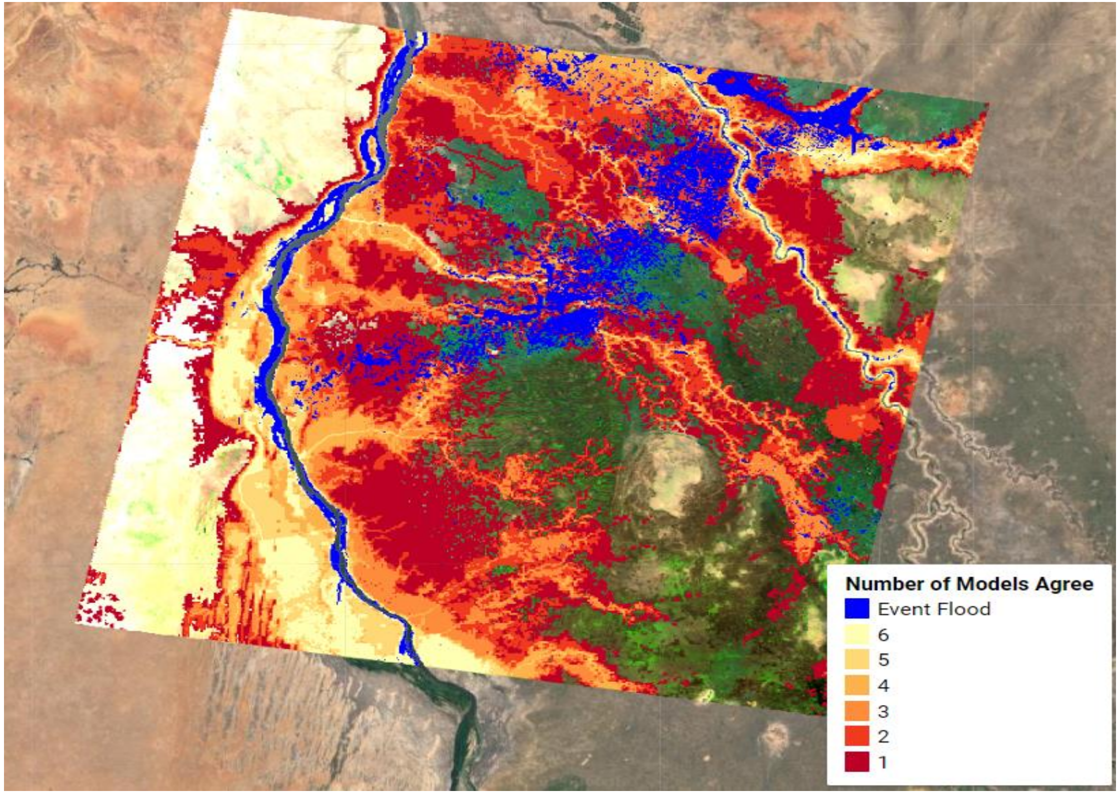
-CSI .4-.7 is that good enough for...?



Bernhofen et al 2018

Index	Mathematical notation
Critical Success Index (CSI)	$CSI = \frac{F_m \cap F_o}{F_m \cup F_o}$
Hit Rate (HR)	$\mathrm{HR} = \frac{F_m \cap F_o}{F_o},$
Bias	$\mathrm{Bias} = \frac{(F_m \cap F_o) + F_m}{(F_m \cap F_o) + F_o} - 1$

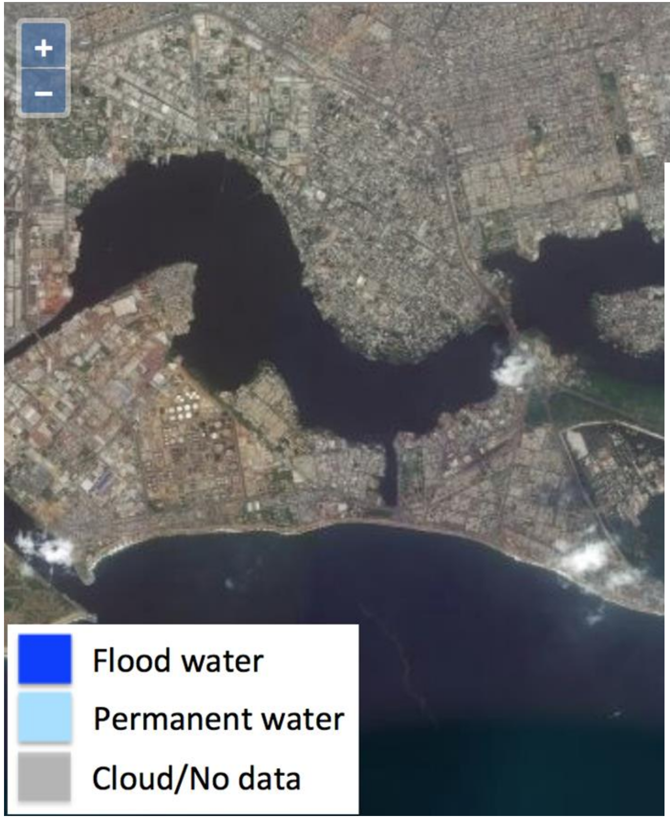
Comparing Events (Nile, 1998 flood) to Global Flood Models

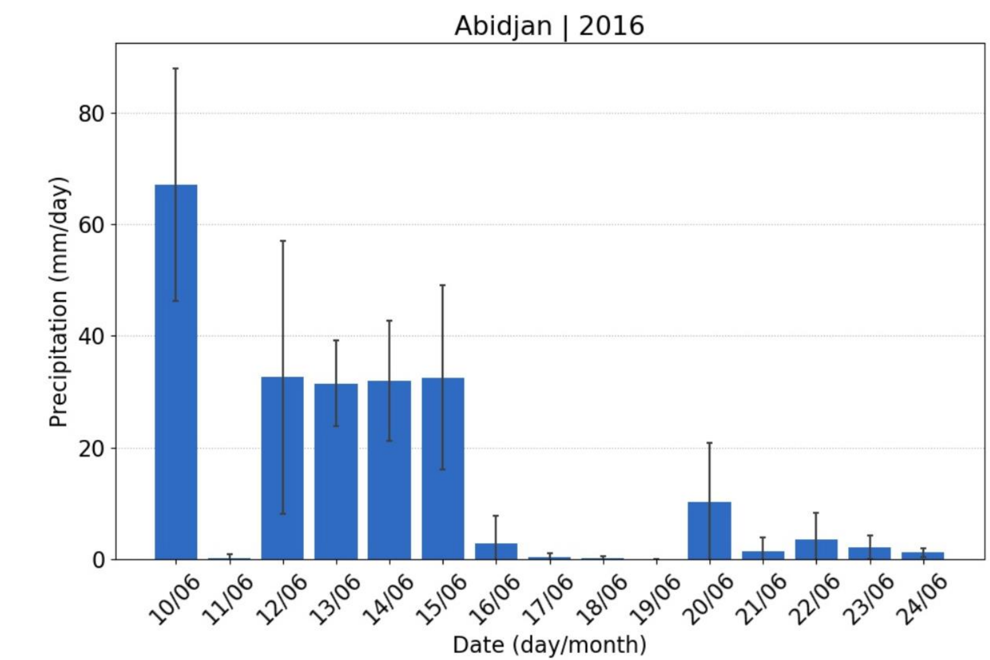


-CSI consistently low (.11) even when ranging flood return times from 25-1000…

-global flood models miss this flooding pattern in the Nile

Abidjan, Ivory Coast, 2016



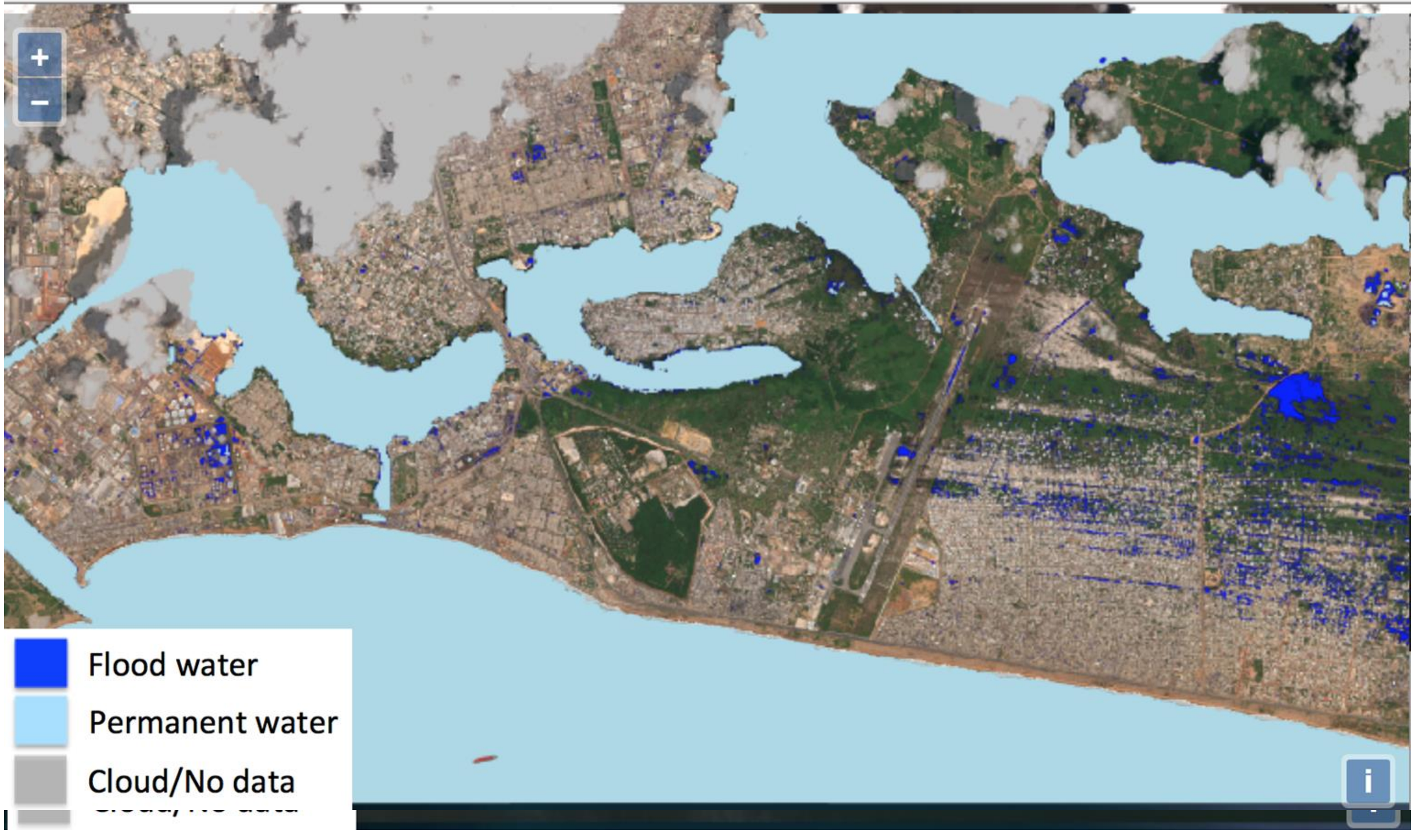


@cloud2street

https://abidjan.cloudtostreet.info

Cloud to Street

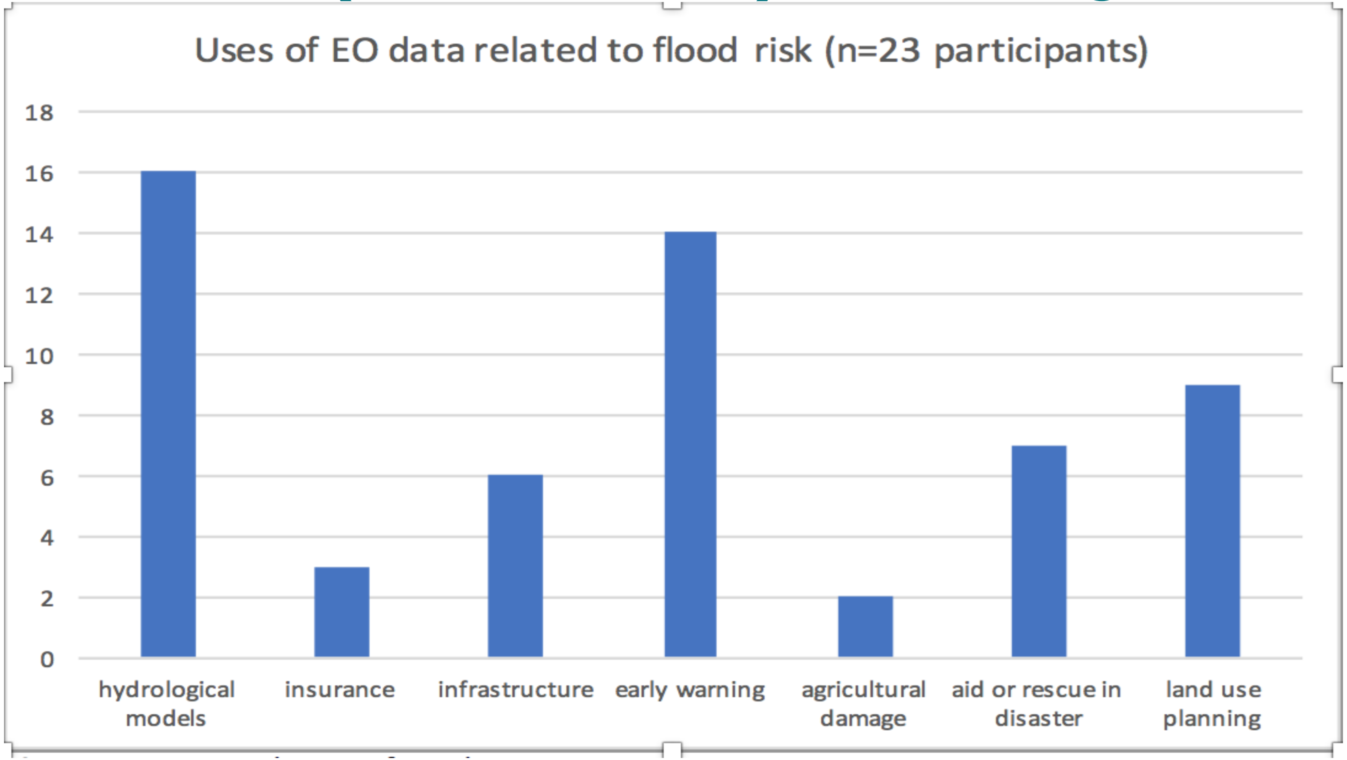
Abidjan, Ivory Coast, 2016



They [Insert Development Agency Here] say the same thing each time...The maps have holes.

Coverage- does the area we can't see matter?
Did we catch the peak flood?

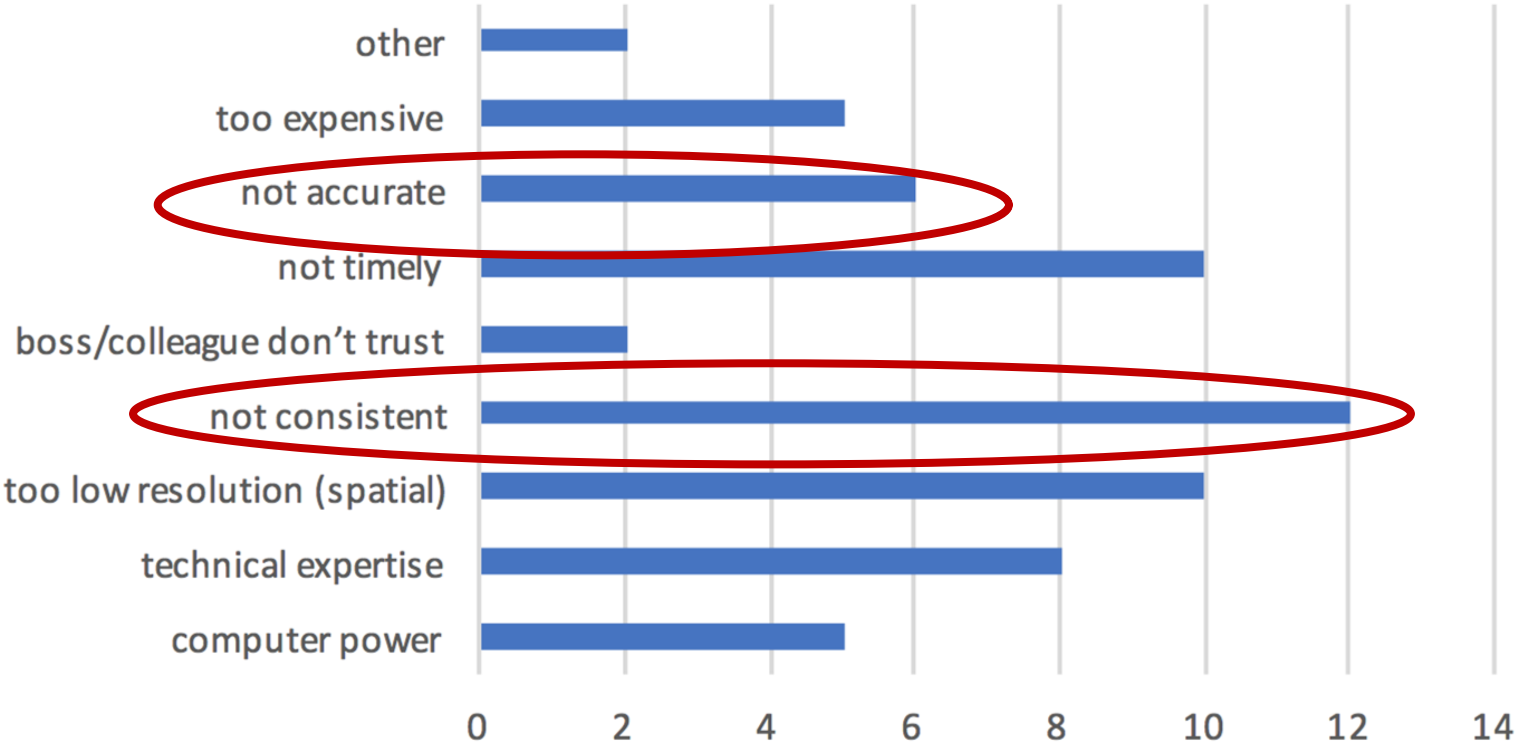
For whom are we (or should be!) measuring accuracy?



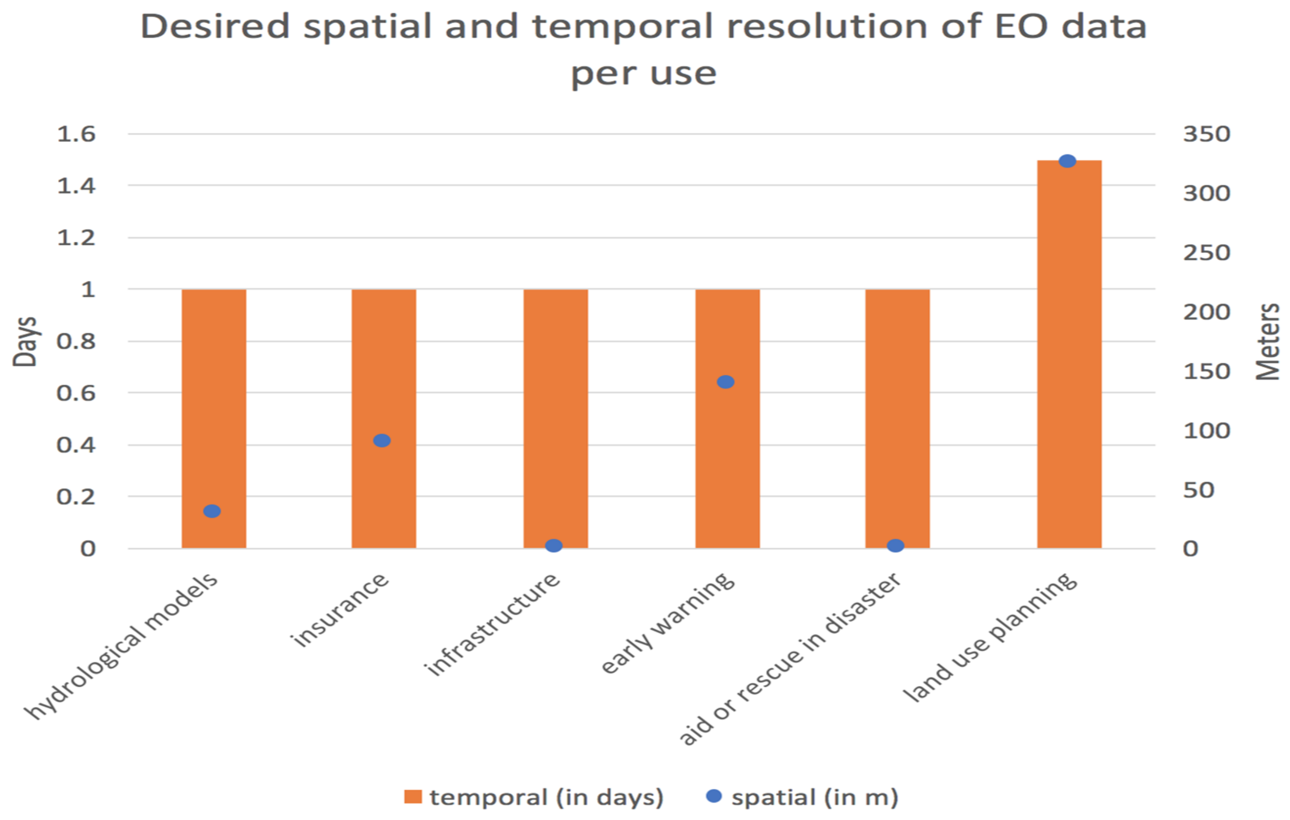
Kettner, A.J., Schumann, G.J.-P., Tellman, B., 2019. The push toward local flood risk assessment at a global scale, Eos, 100, DOI:10.1029/2019EO113857.

2018 NASA Flood Risk Meeting

Main Barriers to Use of EO



Users want- daily data, but require different spatial resolutions



Disaster cycle to decision horizon

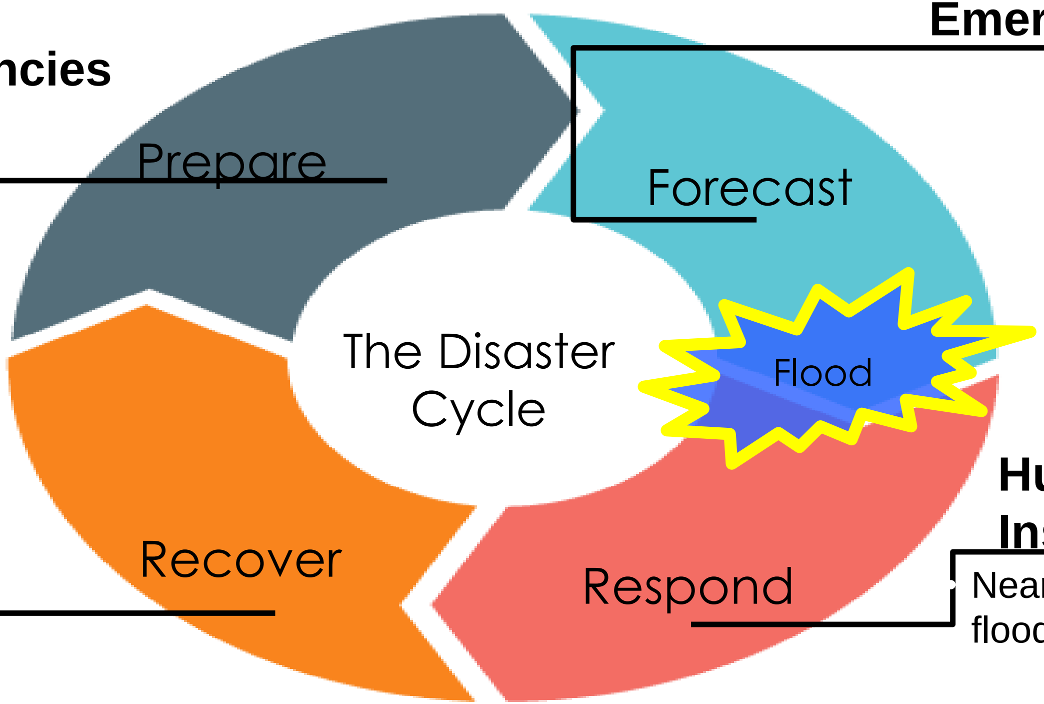
Insurers & Development Agencies

Risk mapping, affordable catastrophe insurance

Emergency Managers

Predict the size and damage of a flood

Early warning and evacuation



Governments

Map of communities hardest hit

Target recovery programs

Humanitarians, Insurers

Near-real time map of floods

Release aid in 24 hours

Disaster cycle to decision horizon

5 qualities of flood maps

- **Event accuracy**
- **Temporal consistency**
- Spatial resolution
- Spatial completeness
- Speed

Users:

Recovery personnel **(respond)**

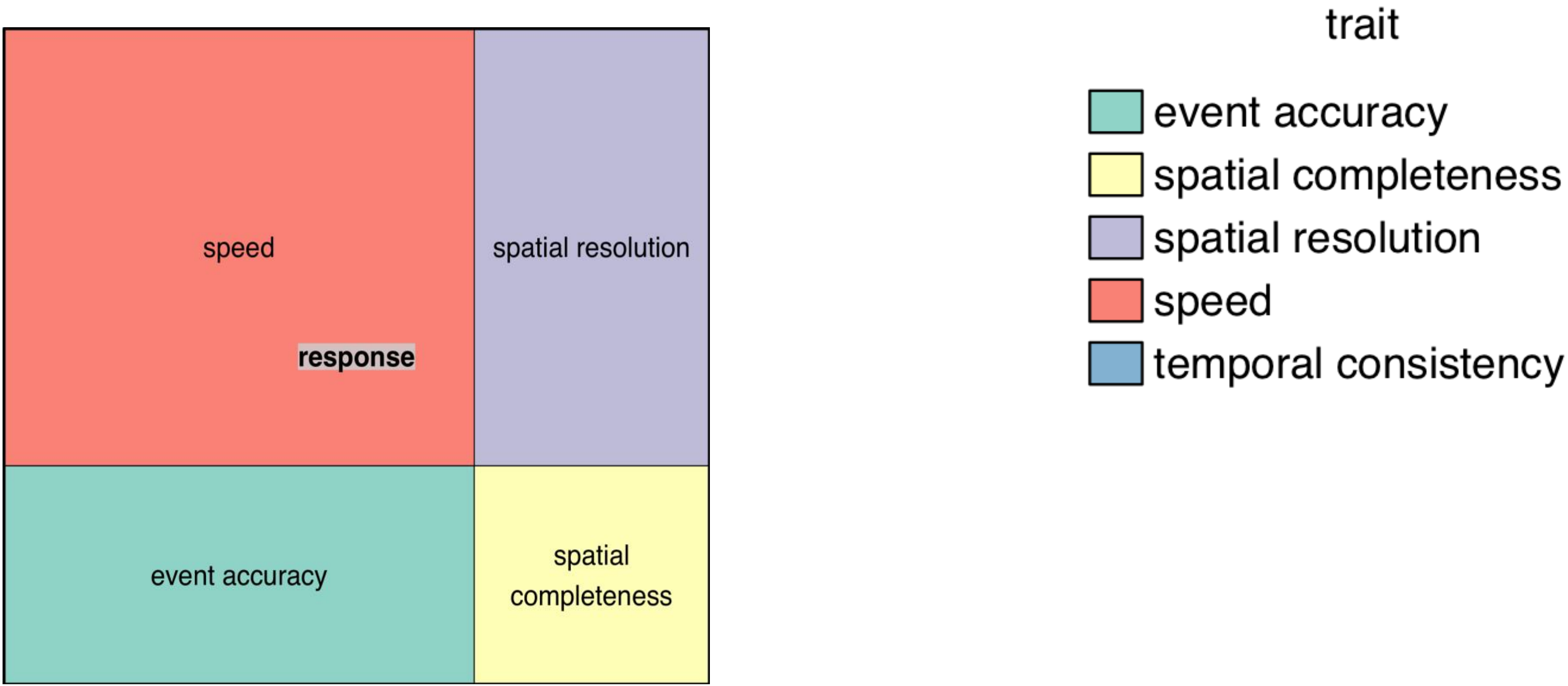
Land use planners, engineers **(prepare)**

Insurers **(prepare, recovery)**

Emergency managers **(forecast)**

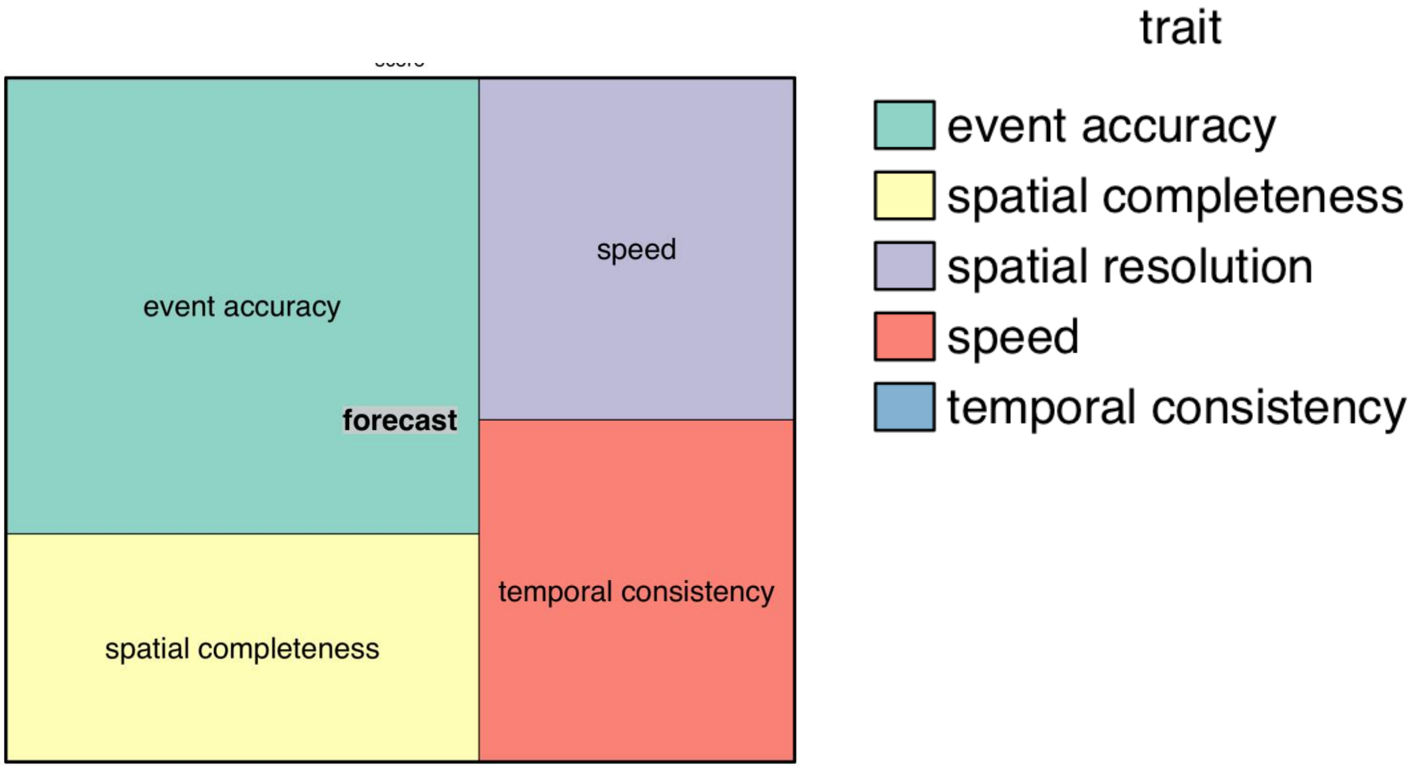
Citizens **(respond, prepare, recover)**

Scientists **(Model/calibrate)**



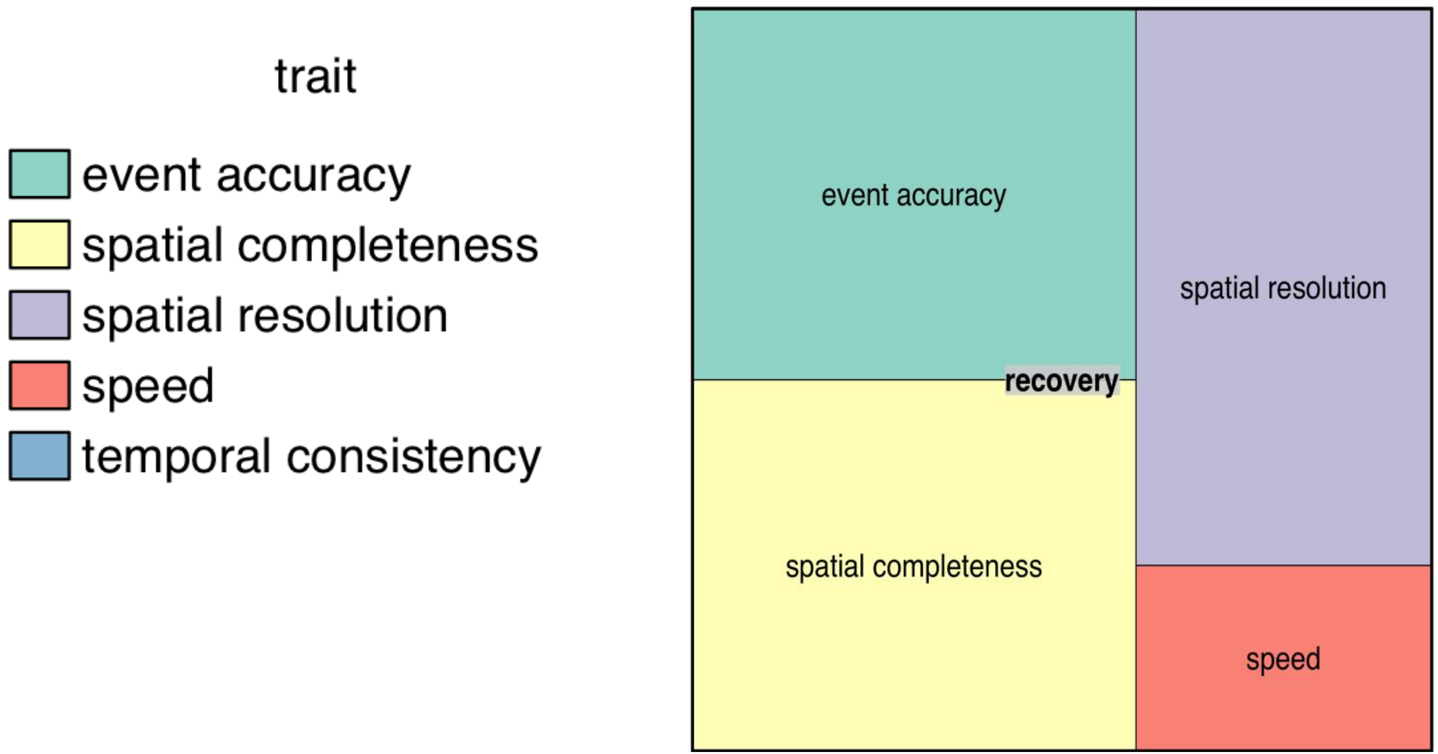
Respond **Forecast** **Recover** **Prepare** **Model/calibrate**

days months **TIME** years

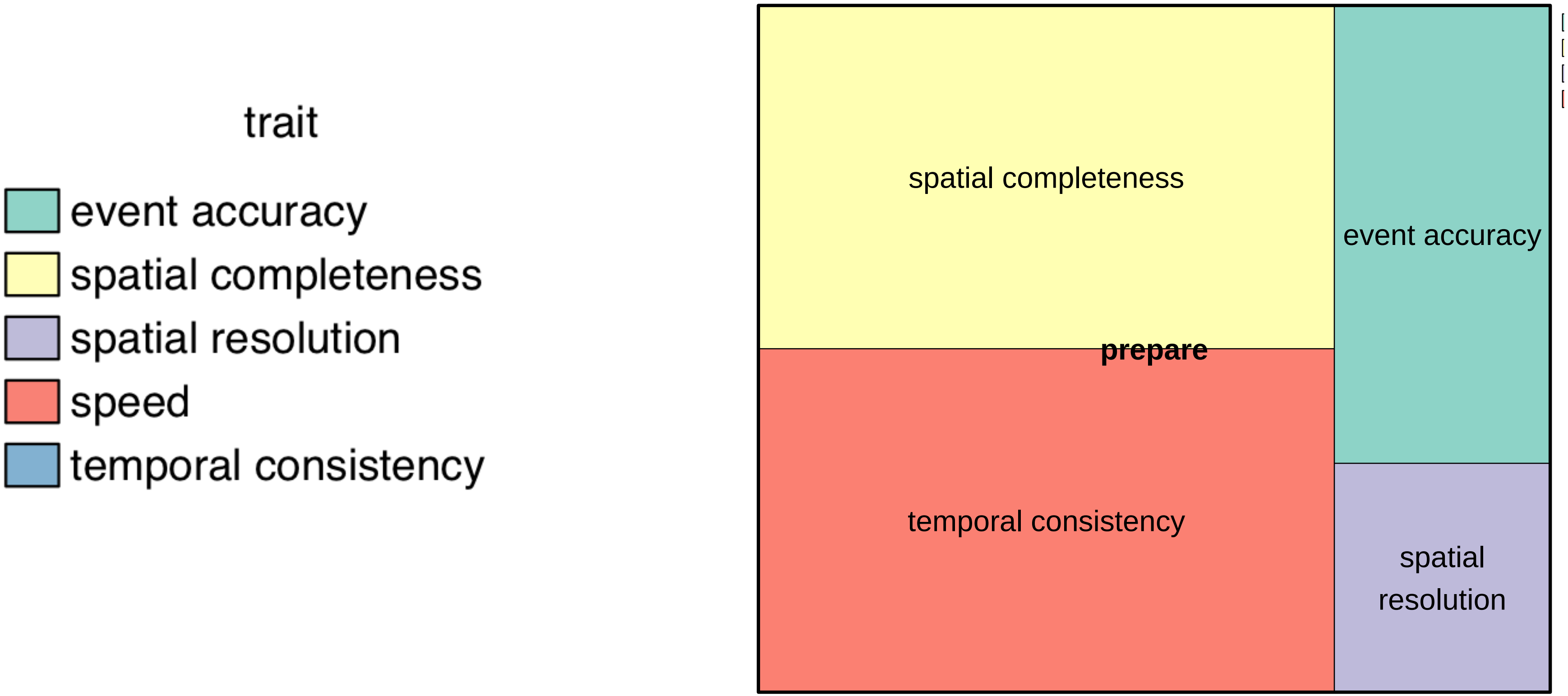


Respond **Forecast** **Recover** **Prepare** **Model/calibrate**

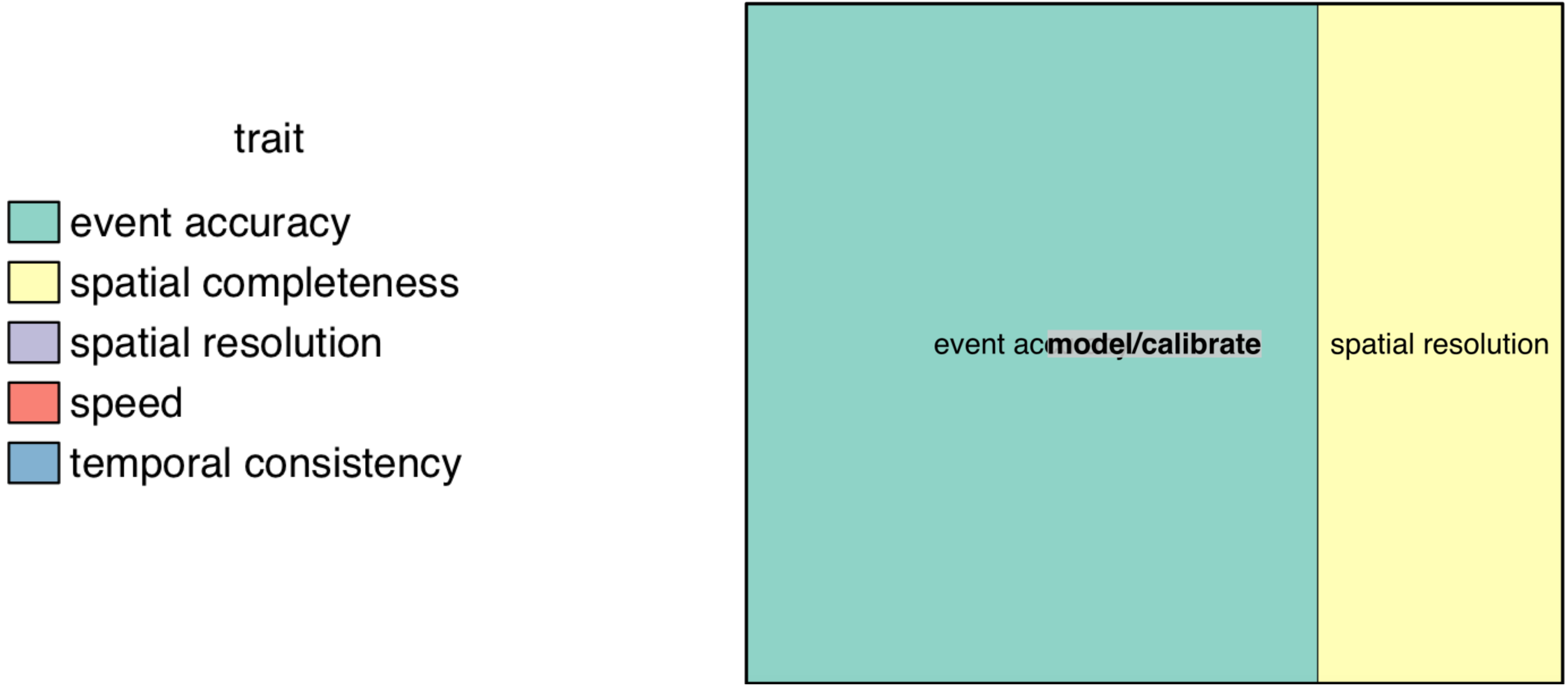
days months **TIME** years


trait
event accuracy
spatial completeness
spatial resolution
speed
temporal consistency
event accuracy
recovery
spatial resolution
spatial completeness
speed


Respond
Forecast
Recover
Prepare
Model/calibrate
days
months
TIME
years

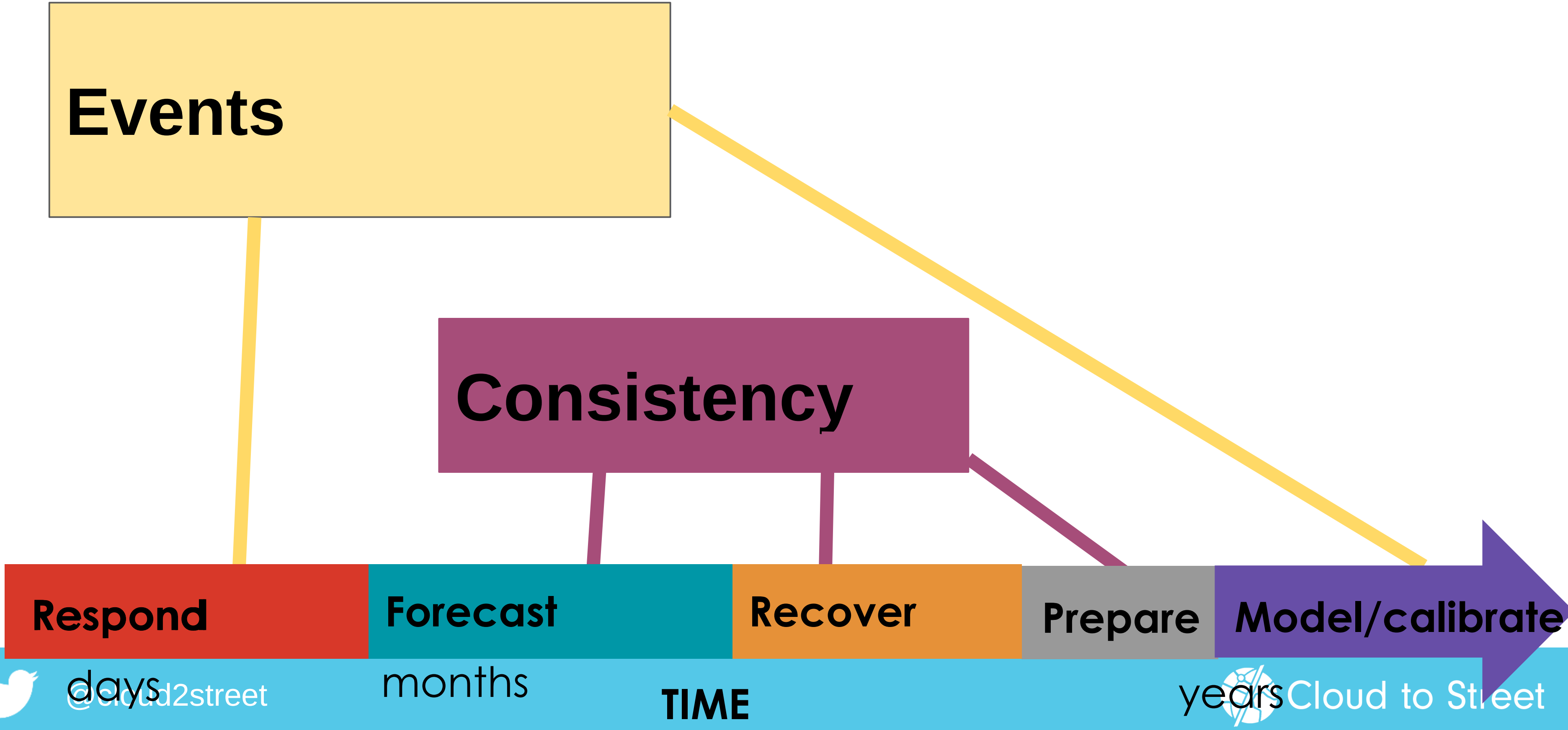


days months TIME years


trait
event accuracy
spatial completeness
spatial resolution
speed
temporal consistency
event ac
model/calibrate
spatial resolution


Respond
Forecast
Recover
Prepare
Model/calibrate
days
months
TIME
years

Two main types of accuracy mapped onto decision time horizon/users



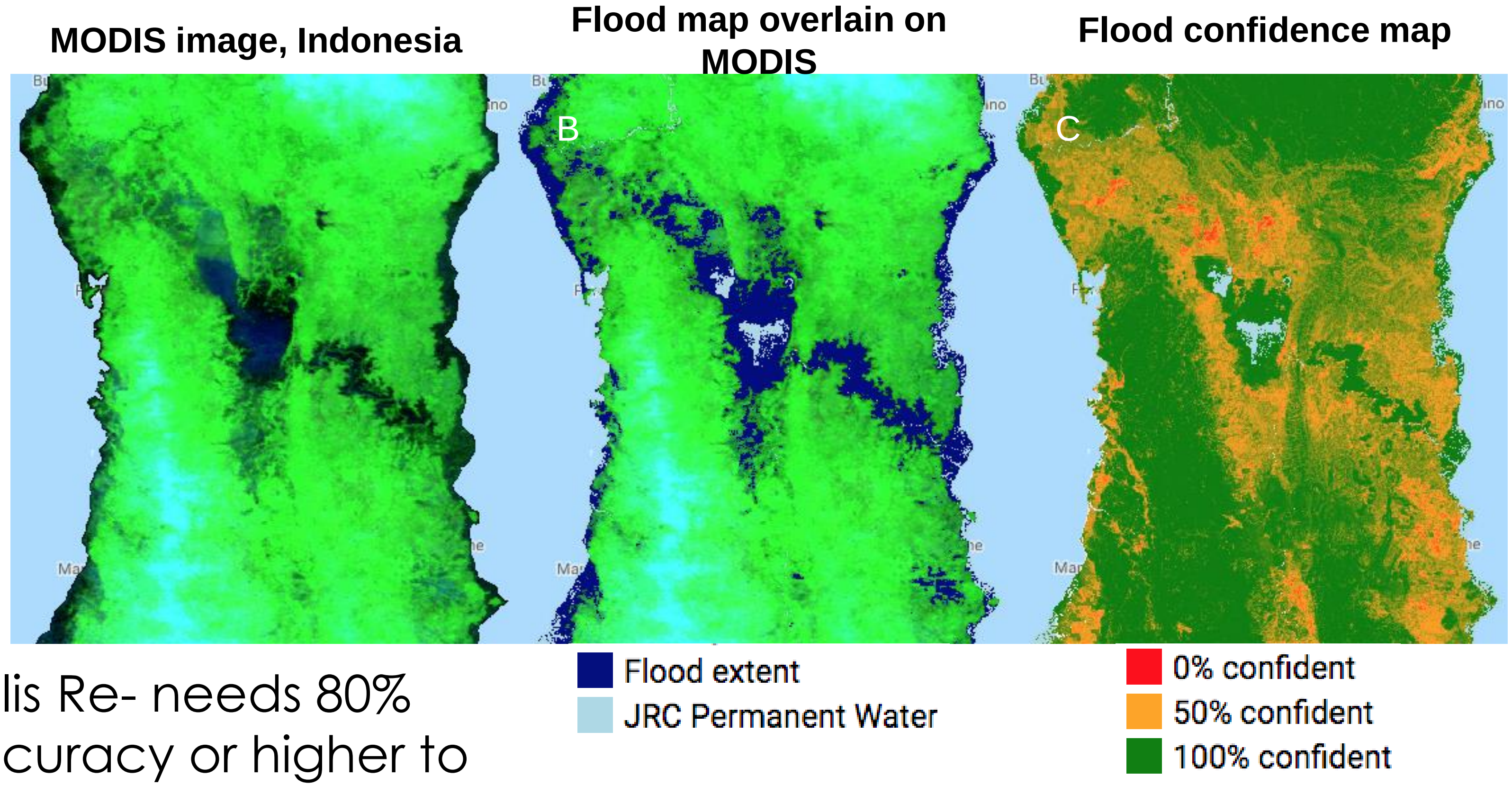
Single “event” accuracy

1.Go beyond weighted stratified random sample,CSI,

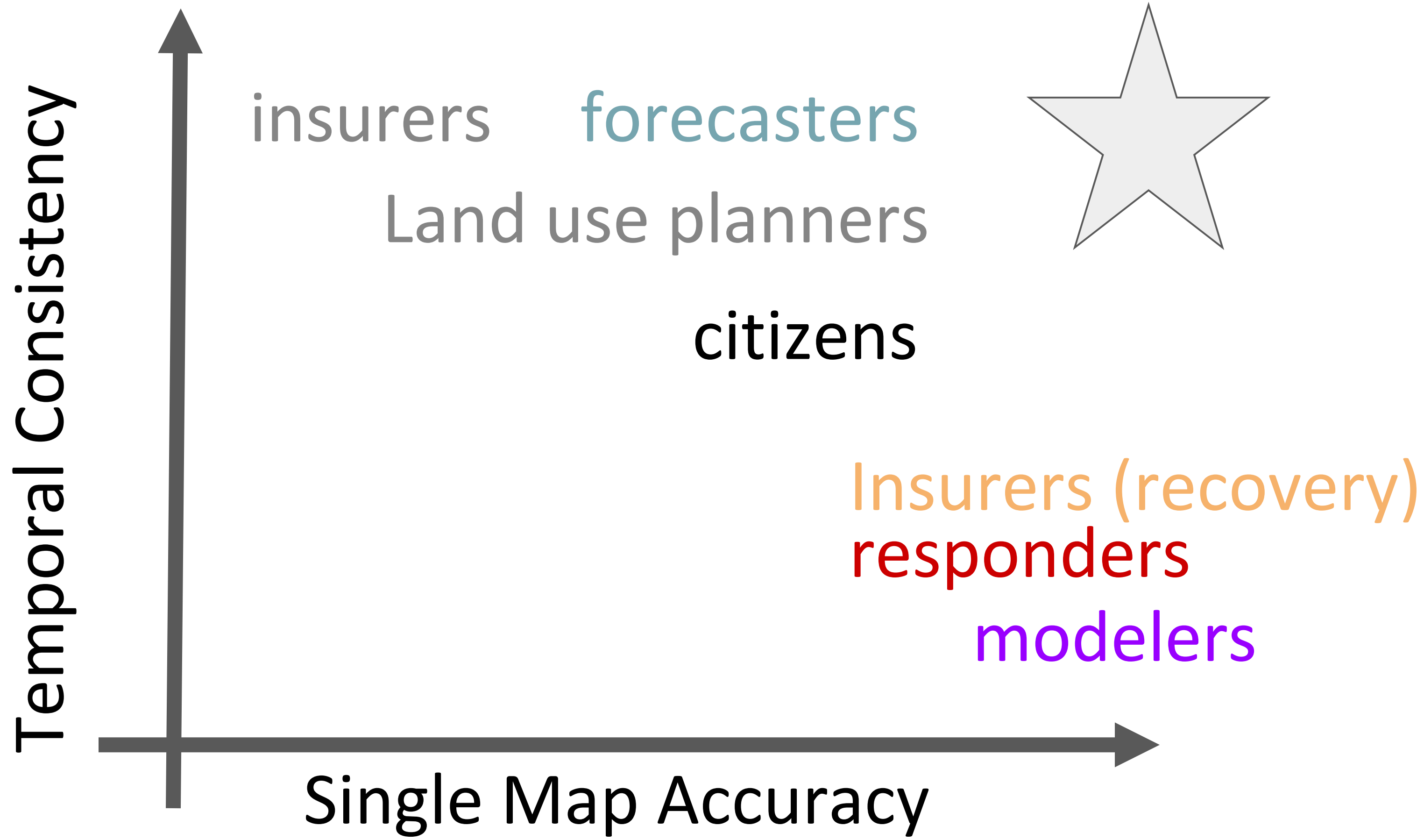
2. Focus on CRITICAL OBJECTS for users: (crops, assets, population centers, roads) and report their accuracy

3. Assess representativeness of “peak” flood uncertainty based on sensor visibility and known issues (e.g. flooded vegetation in SAR-blind spots)

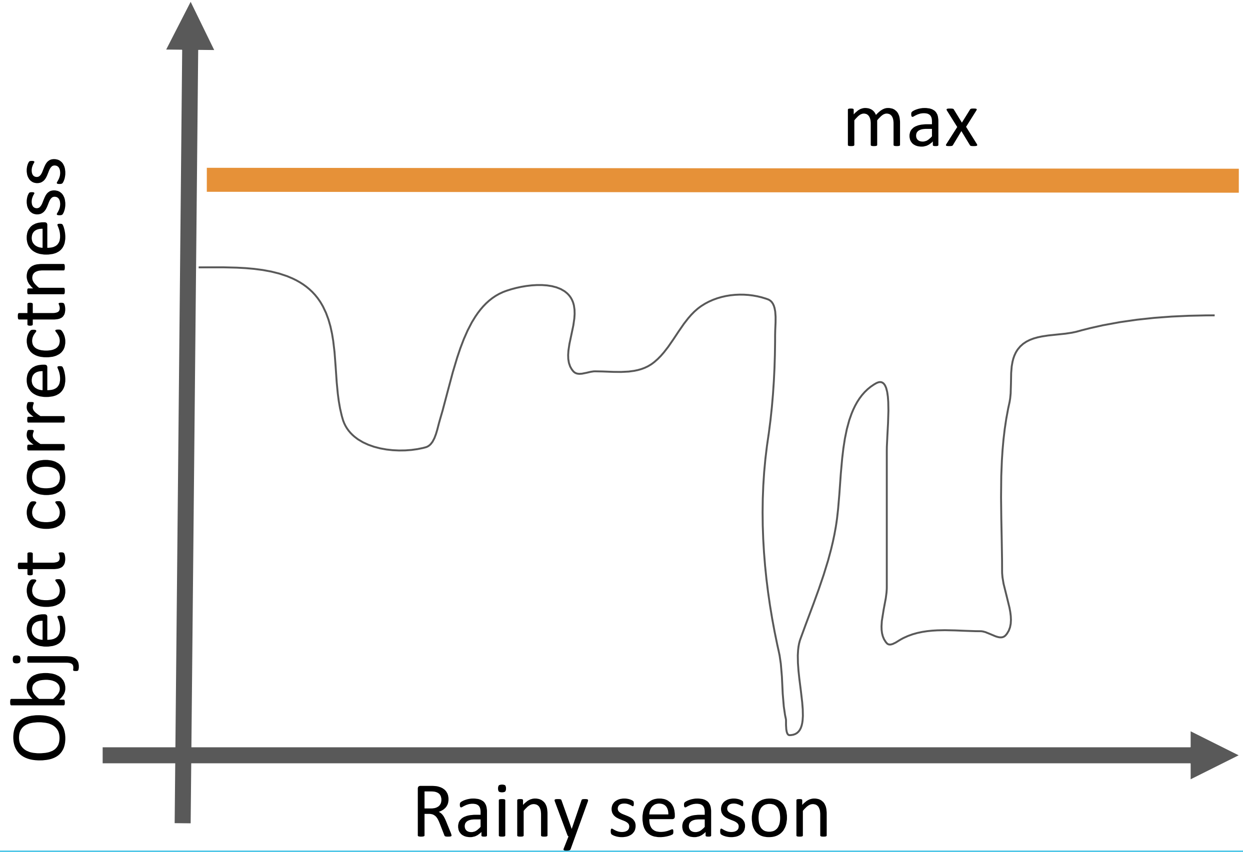
Assess if "peak event" is captured



Willis Re- needs 80% accuracy or higher to calibrate their models

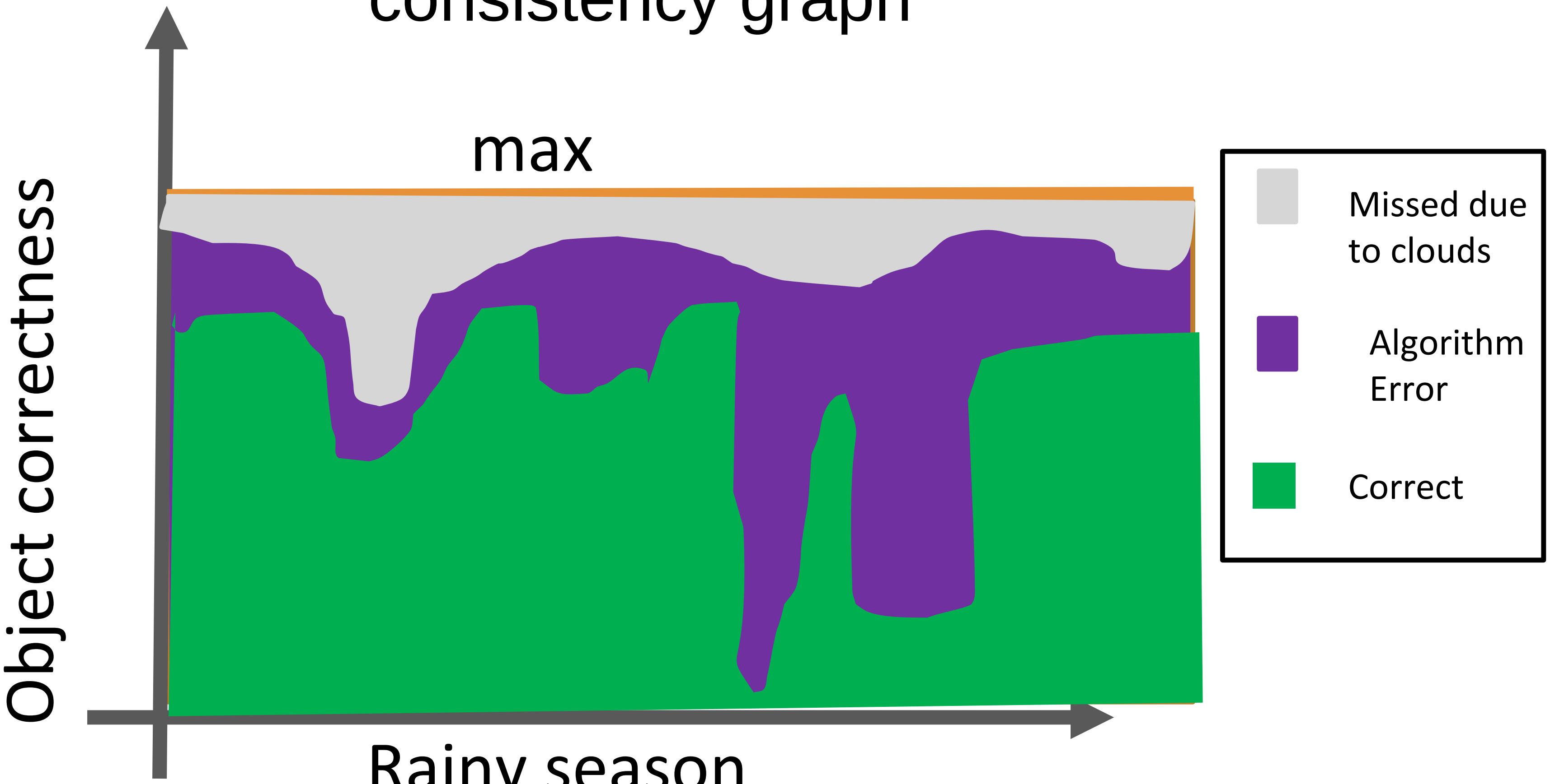

Temporal Consistency
insurers
forecasters
Land use planners
citizens
Insurers (recovery)
responders
modelers
Single Map Accuracy

consistency graph



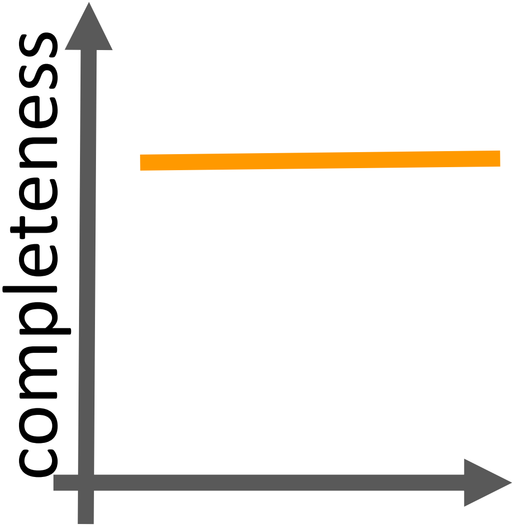
1. Select 50-100 critical floodable objects
2. For each object, determine "floodability"

consistency graph

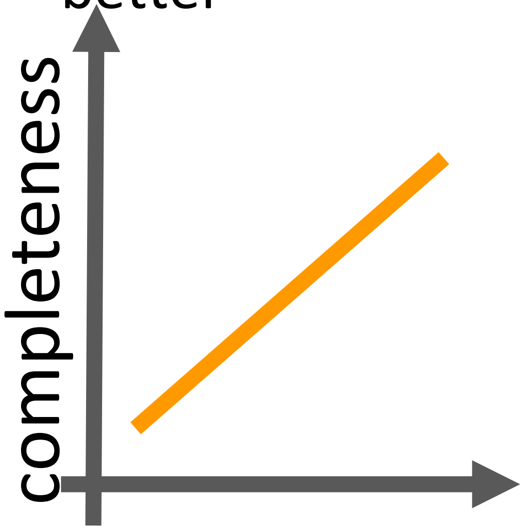


Spatial Completeness for Events

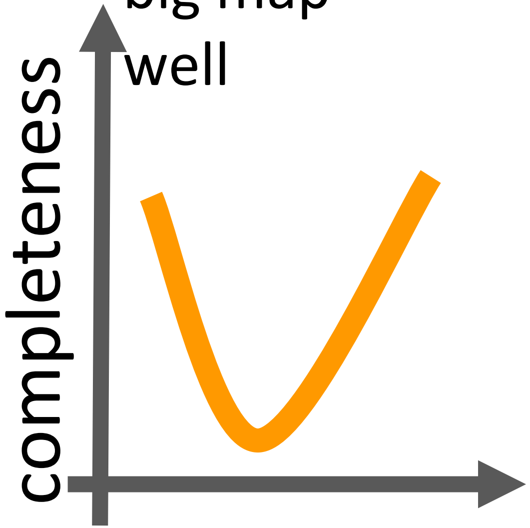
All events map well



Bigger events map better



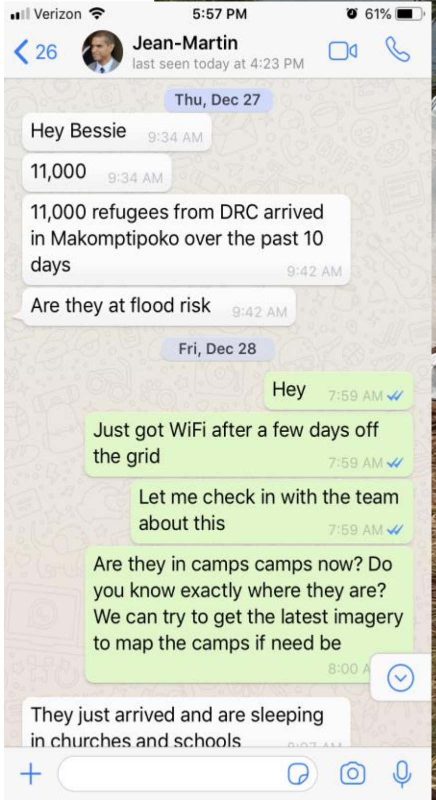
Only small and only big map well



Congo refugee relocation

Sometimes there is no magic metric
when expert opinion is the only option

Flood risk concern at new refugee sites

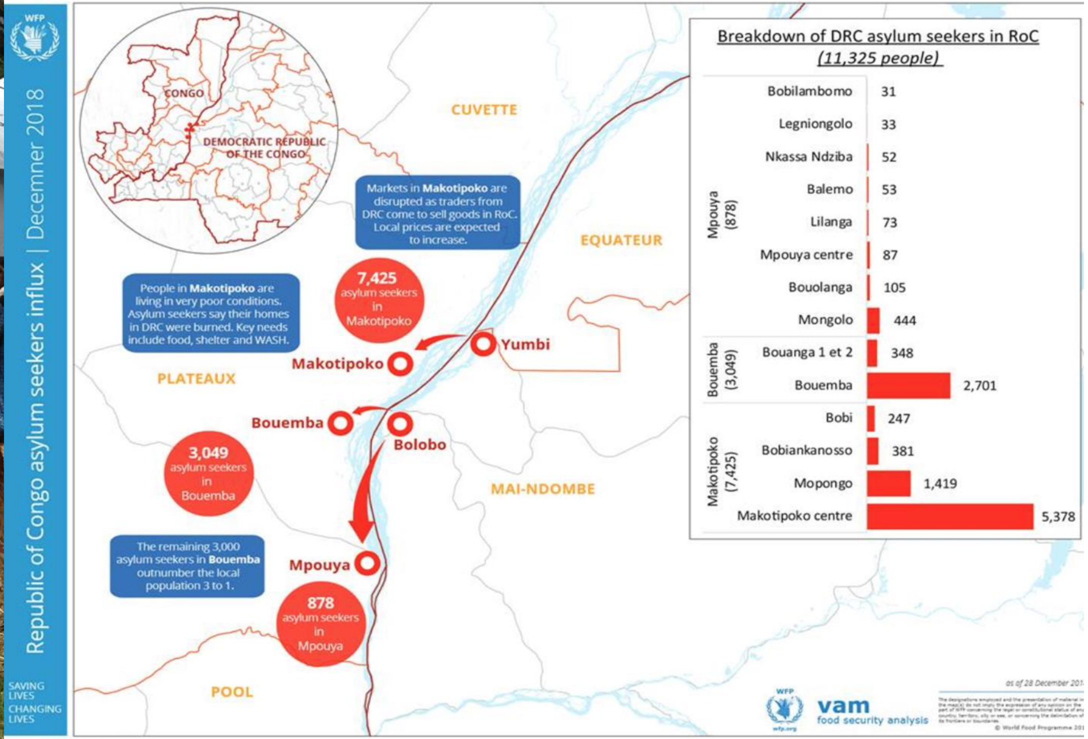


DRC refugees

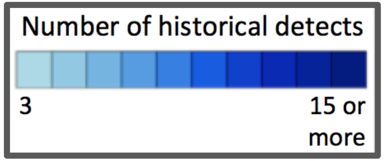
To: Bessie Schwarz, Bessie Schwarz, Cc: William VU

Details

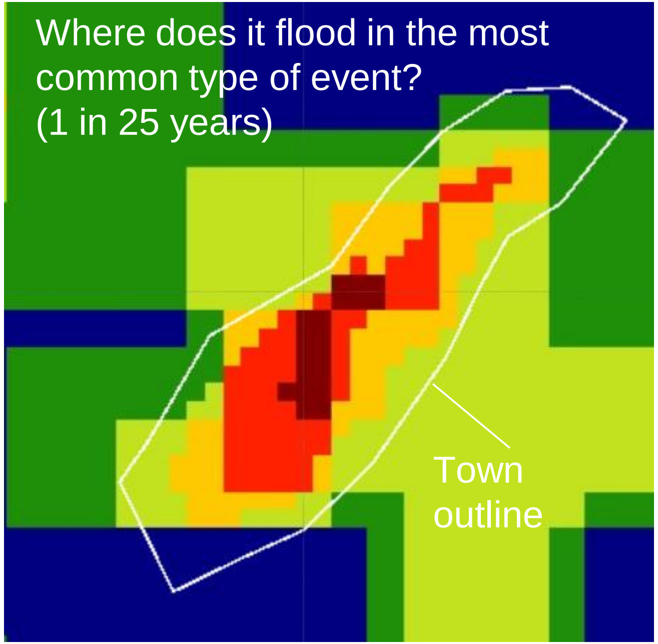
Hi, here is the situation at present. The sites called Makomptipoko and Mopongo are already waterlogged and we're wondering about relocating these refugees. Thanks



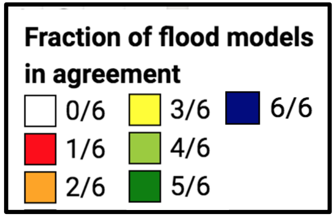
Makotipoko: Historical risk and modeled flood risk



Areas of Makotipoko have tended to flood in the last 30 years.



There's also high risk based on data we have from six flood models(Trigg et al., 2016), and also high certainty of this risk (i.e., multiple models agree).

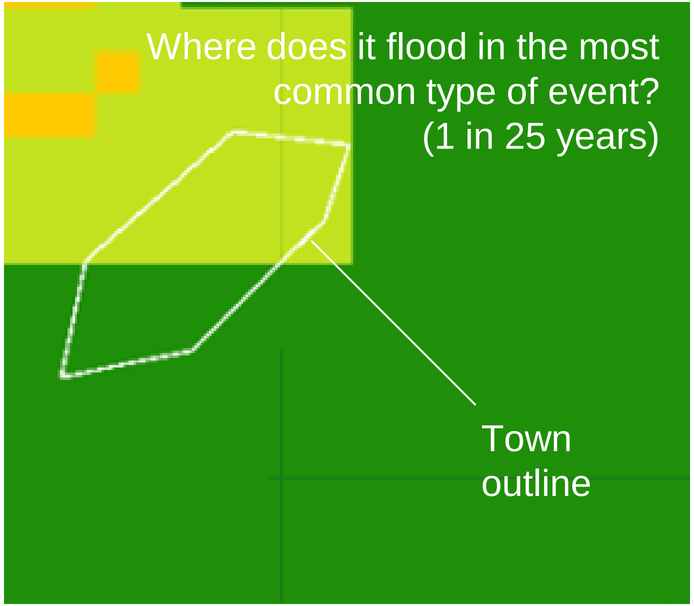


Mopongo: Historical risk and modeled flood risk

Number of historical detects

3 15 or more

We did not observe historical flooding in Mopongo.



However, the flood models indicate medium risk and medium certainty of that risk.

Fraction of flood models in agreement

0/6 1/6 2/6 3/6 4/6 5/6 6/6

Relocating refugees with flood maps/models

possible

4. **Bouemba:** results are too uncertain to recommend moving the asylum-seekers

Risk	Certainty
	n
	um
	um
Medium to Low	Low

The Global Flood models we are using may identify areas that are likely to flood, but they could miss other areas and so are not useful for identifying "safe" areas. Unfortunately, this problem is largest in places like Republic of the Congo where elevation data is poor and dense forest vegetation influences model results. Therefore, we cannot provide a recommendation as to which areas would be safe for them to move. Dr. Mark Trigg, who has worked on this reach of the Congo river, said local knowledge of past flooding will be most useful for determining safer zones for each location and that communities can usually identify those areas.

Conclusions

1. We can do better than what remote sensing gives us for accuracy assessment information

2. Focus on critical objects and features (events vs. consistency) the user cares about and their decision timeline
    1. Events= peak flood uncertainty, objects
    2. Consistency= measure over time and event magnitude

@Cloud2Street
Come see us in NYC
(ps we are hiring…)

www.cloudtostreet.info

Dr. Beth Tellman
@pazjusticiavida

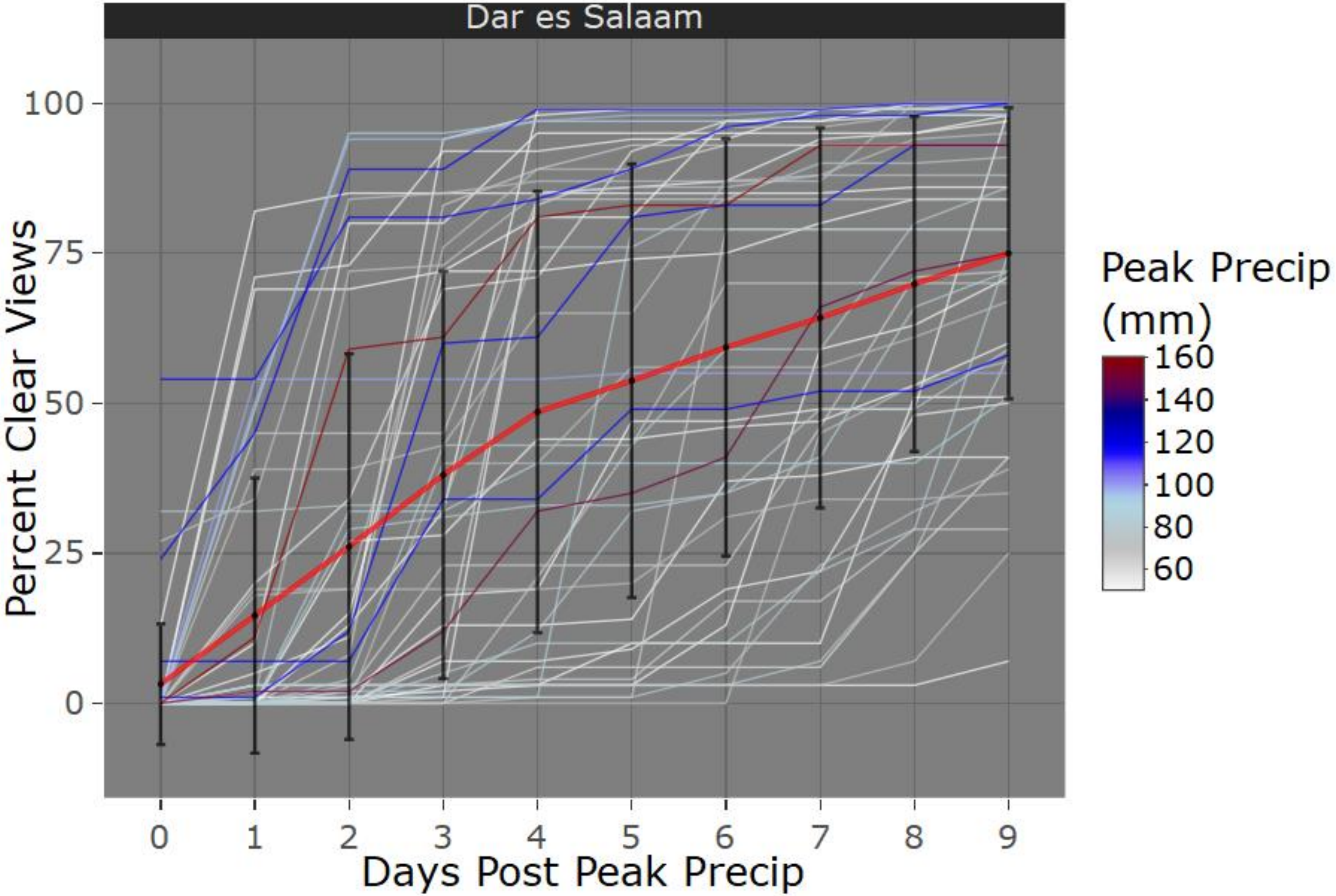
Back up and other

Temporal Consistency and Spatial Completeness

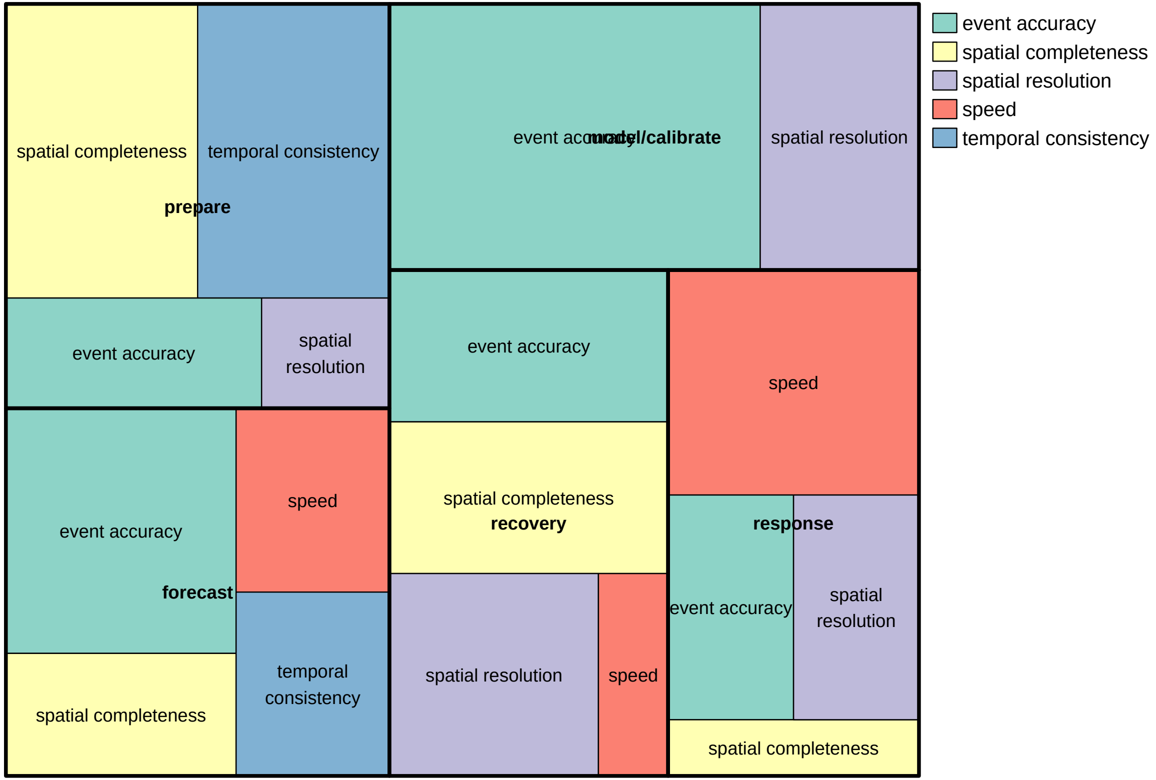
1. Select 50-100 critical floodable objects
2. For each object, determine "floodability"
3. Determine start/end of rainy season
4. Calculate and graph number of objects visible daily and number correct (flooded or not)
5. Accuracy= (area under curve/total area)

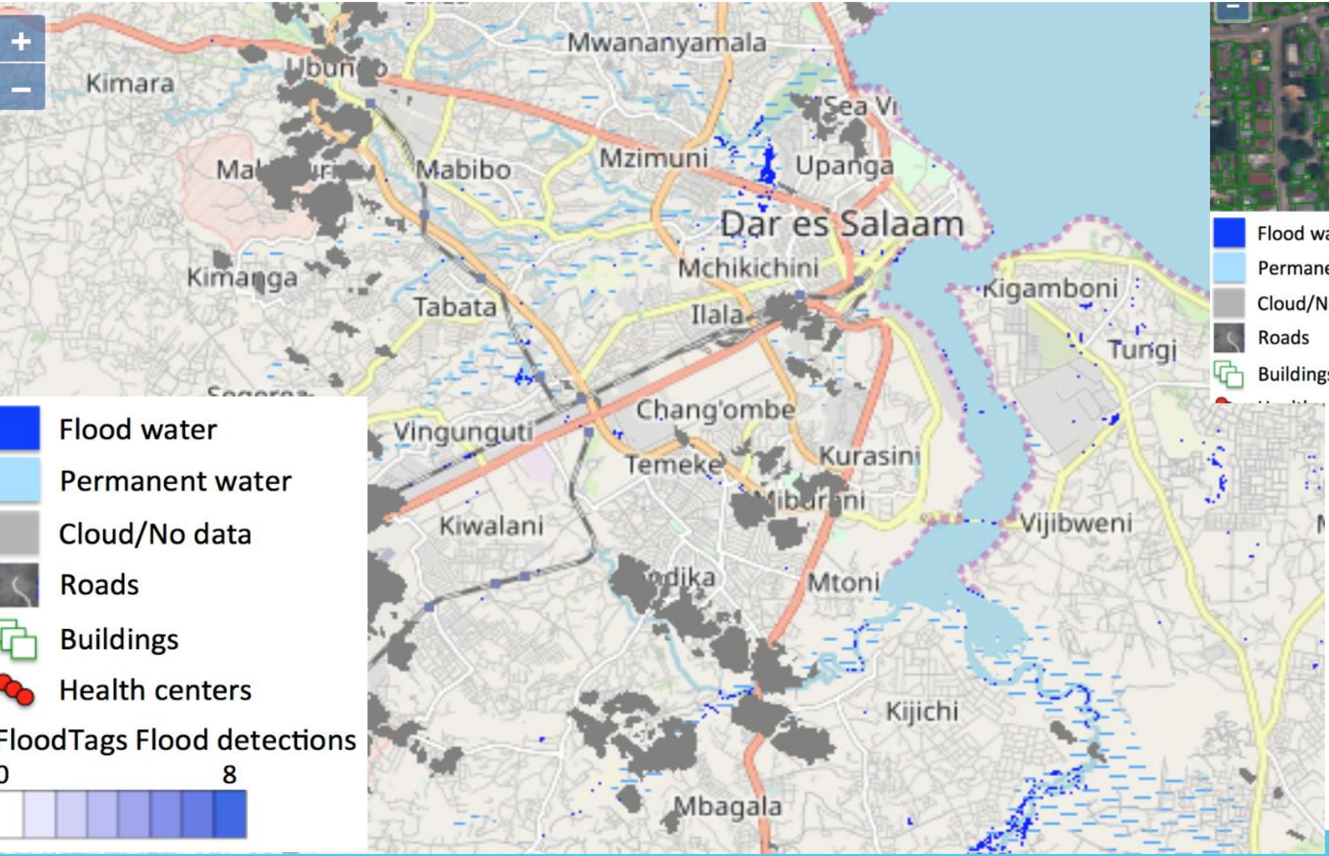
Skysat Beira Comparison

- Water extent higher from July 2018, but the structural damage from March 2019 means it was much worse.

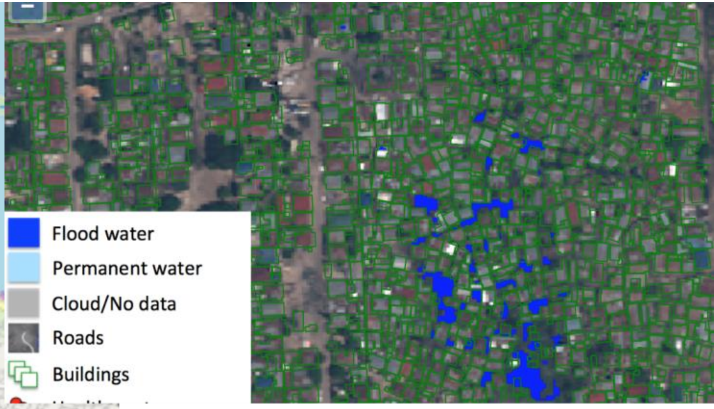

Dar es Salaam
Percent Clear Views
100
75
50
25
0
Days Post Peak Precip
0
1
2
3
4
5
6
7
8
9
Peak Precip (mm)
160
140
120
100
80
60

Most important flood map feature HYPOTHETICAL





Mwananyamala
Kimara
Mzimuni
Mabibo
Upanga
Dar es Salaam
Mchikichini
Kimanga
Tabata
Ilala
Kigamboni
Tungi
Chang'ombe
Vingunguti
Temeke
Kurasini
Kiwalani
Vijibweni
Mtoni
Kijichi
Mbagala
Flood water
Permanent water
Cloud/No data
Roads
Buildings
Health centers
FloodTags Flood detections
0
8



Flood water
Permanent water
Cloud/No data
Roads
Buildings

Temporal Consistency and Spatial Completeness

1. Select 50-100 critical floodable objects. If they are points, buffer then by some amount (~30m)
2. For each object, determine its average floodability (using distance from a place that has ever flooded using C2S recurrence, a model, or the HAND index). Since floodability is by pixel, you will area weight the object for its per pixel floodability to get the average score.
3. Determine the rainy season for the watershed or country of interest
4. Every day, calculate the number of objects visible (more than 50%). For the visible objects, record if the satellite of the day objects, correctly identified significant flooding in the object (using your eyes)- binary yes or no.

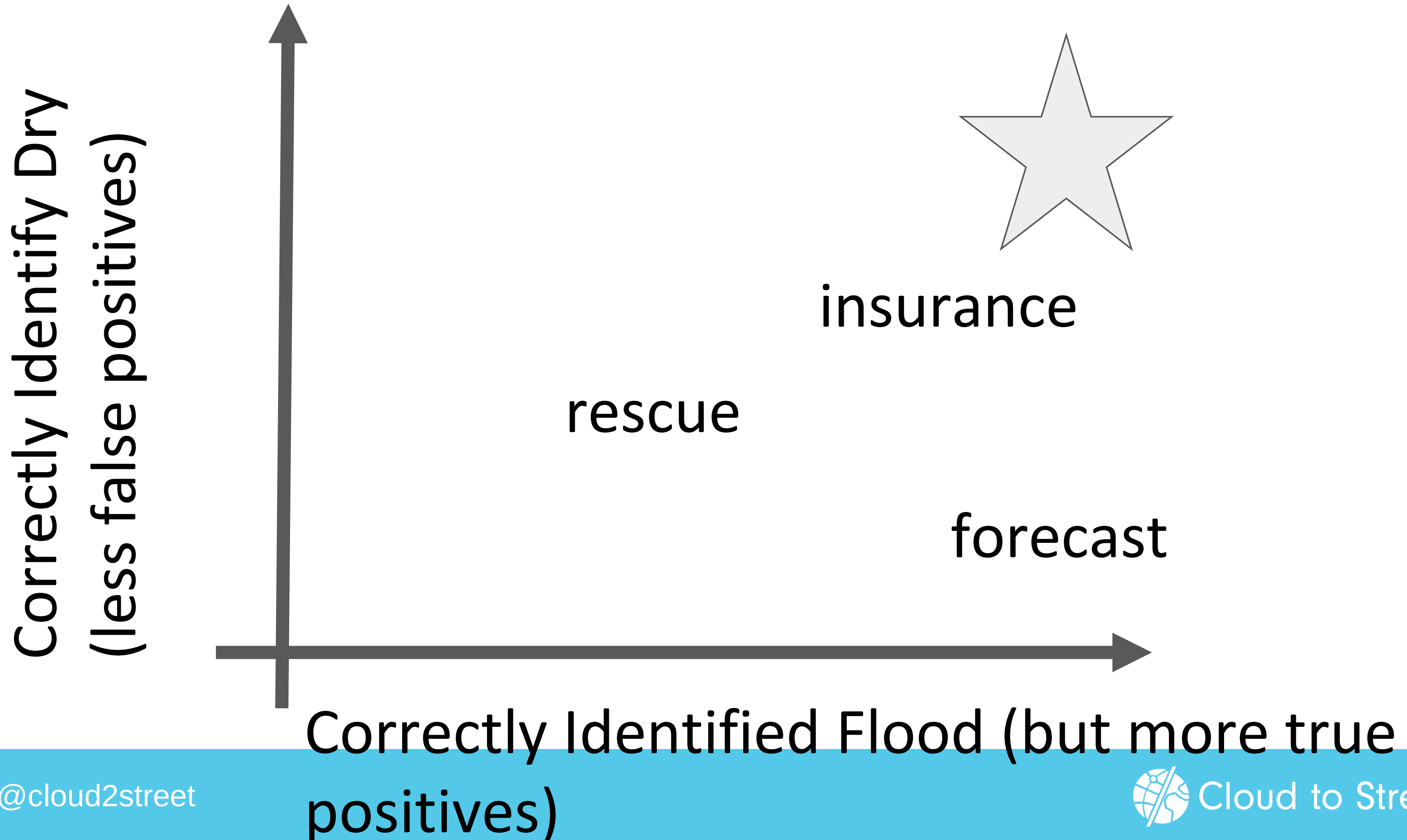
5. Graph over an entire rainy season the daily score by summing the object scores that were identified.

6. 1-Ratio under the curve is the consistency metric

This can also be mapped- by summing objects. Hotspots of 1s and hotspots of 0 should pop out and a getis-ord score can be generated (hotspot analysis)

7. This can be done in the past, but I suggest parsing it up by chunks of years given satellite variability

8. This can be done in the future, by using the average cloudiness (from a typical or series of


Correctly Identify Dry
(less false positives)
insurance
rescue
forecast
Correctly Identified Flood (but more true positives)

Cloud to Street

The Global High Resolution Flood Mapping and Monitoring System

Designed to protect the most vulnerable and enable resilience worldwide

www.cloudtostreet.info

Flood map science to decisions

Algorithm Developed

Flood protection decision from flood map

Impfondo, Congo, November 2017, 5,500 people need food aid

Aid took 3 weeks- because impact was unknown

But high res optical (1.8m) imagery identified this event

Gambia Story

- Picked up lots of seaonsal flooding (great!)
- But nothing that the government cared about (where people live)