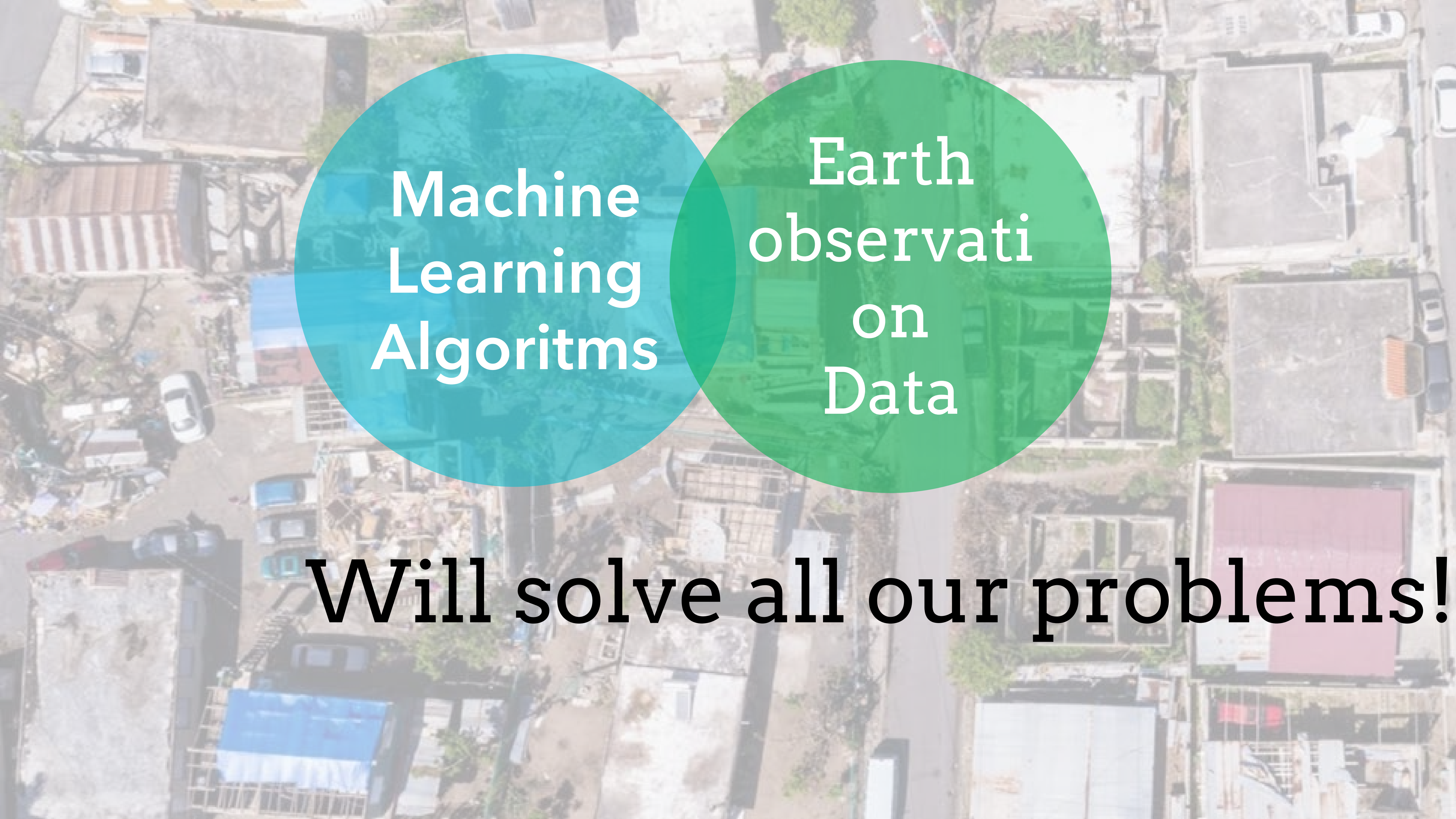
The (un)limited potential of machine learning for Earth Observation

Stef Lhermitte

AN INITIATIVE OF THE NETHERLANDS RED CROSS

@steflhermitte

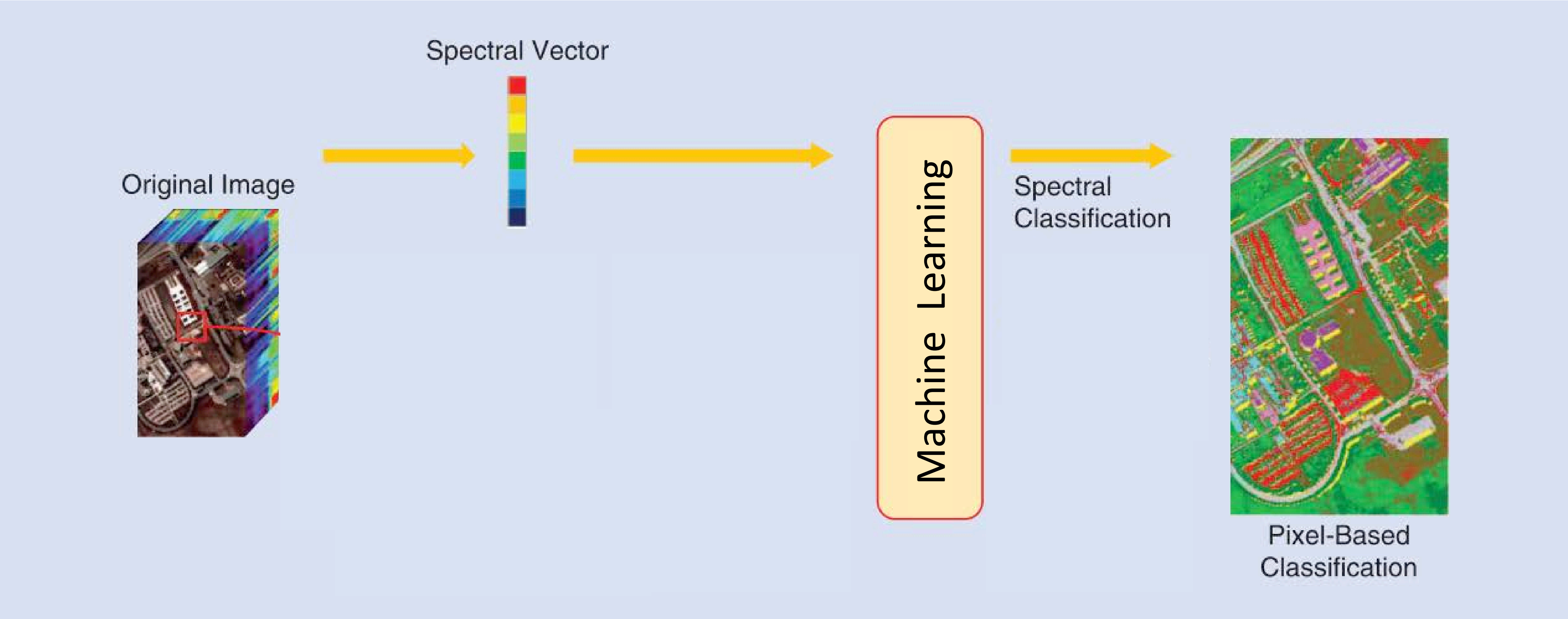


Machine Learning Algoritms

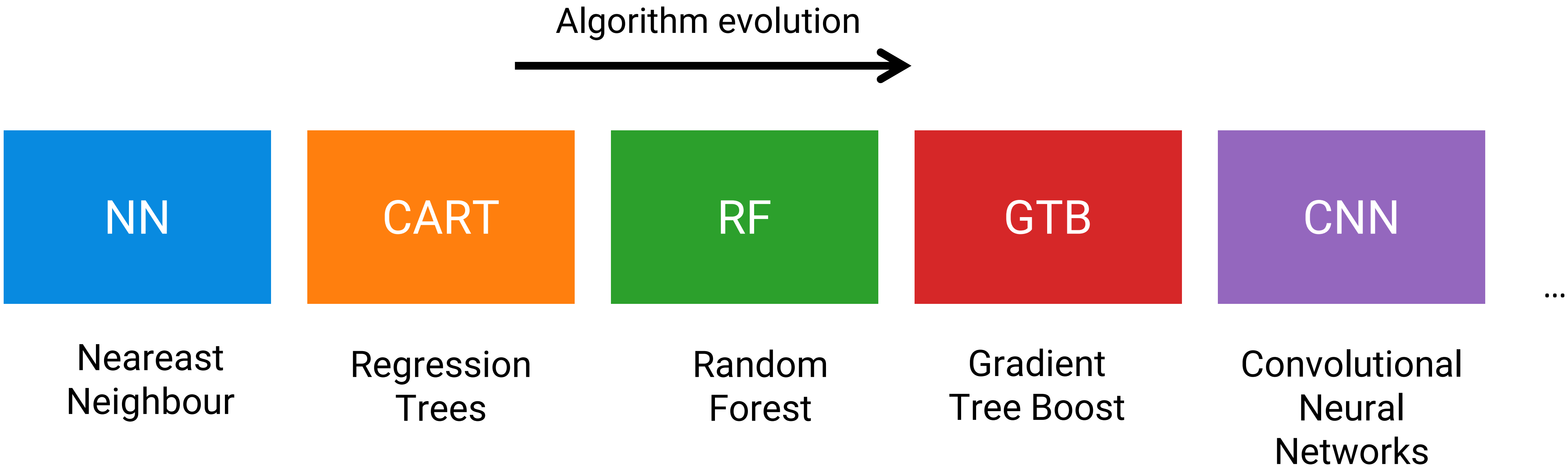
Earth observati on Data

Will solve all our problems!

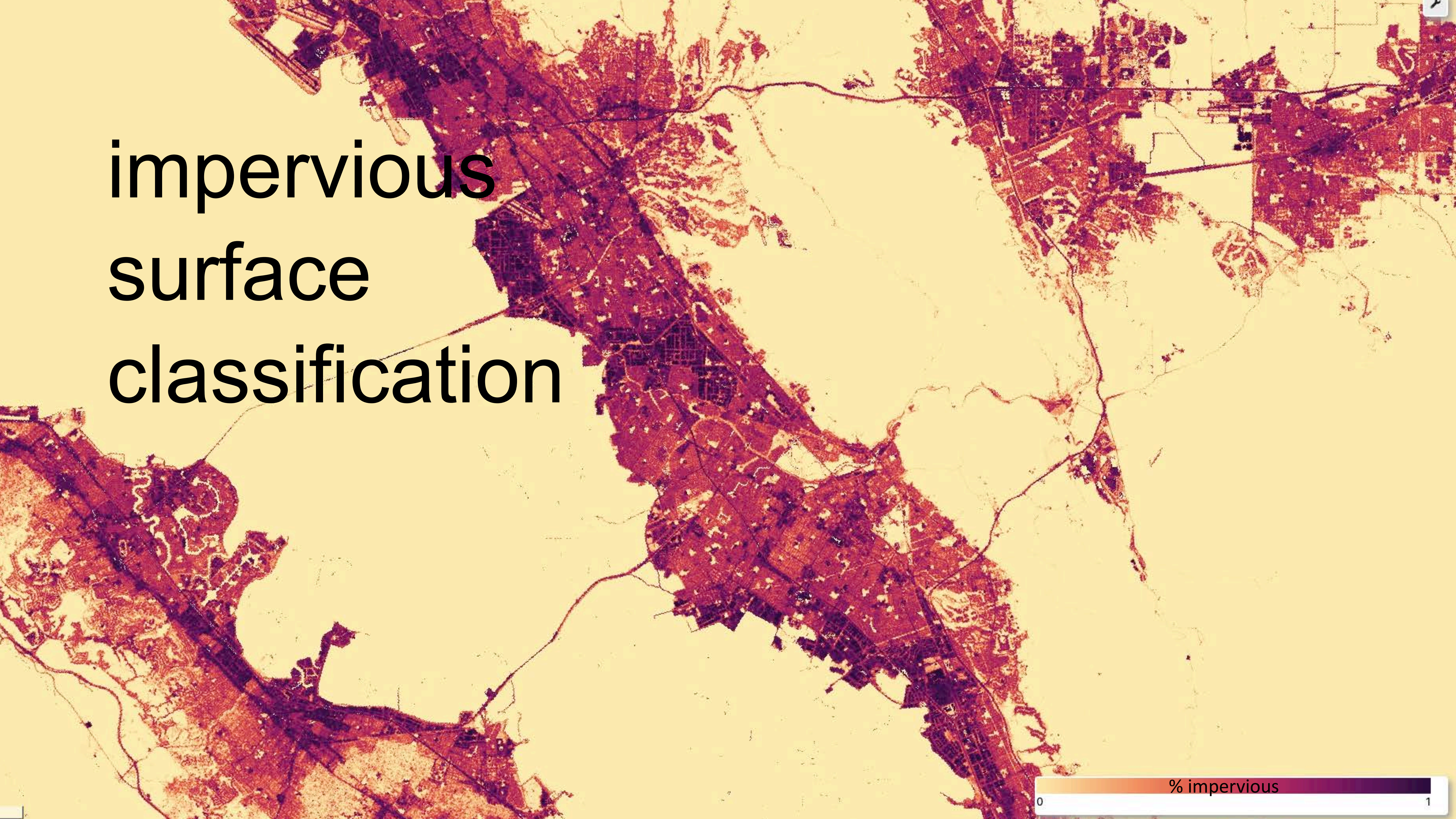
Hooray for the unlimited potential of ML methods



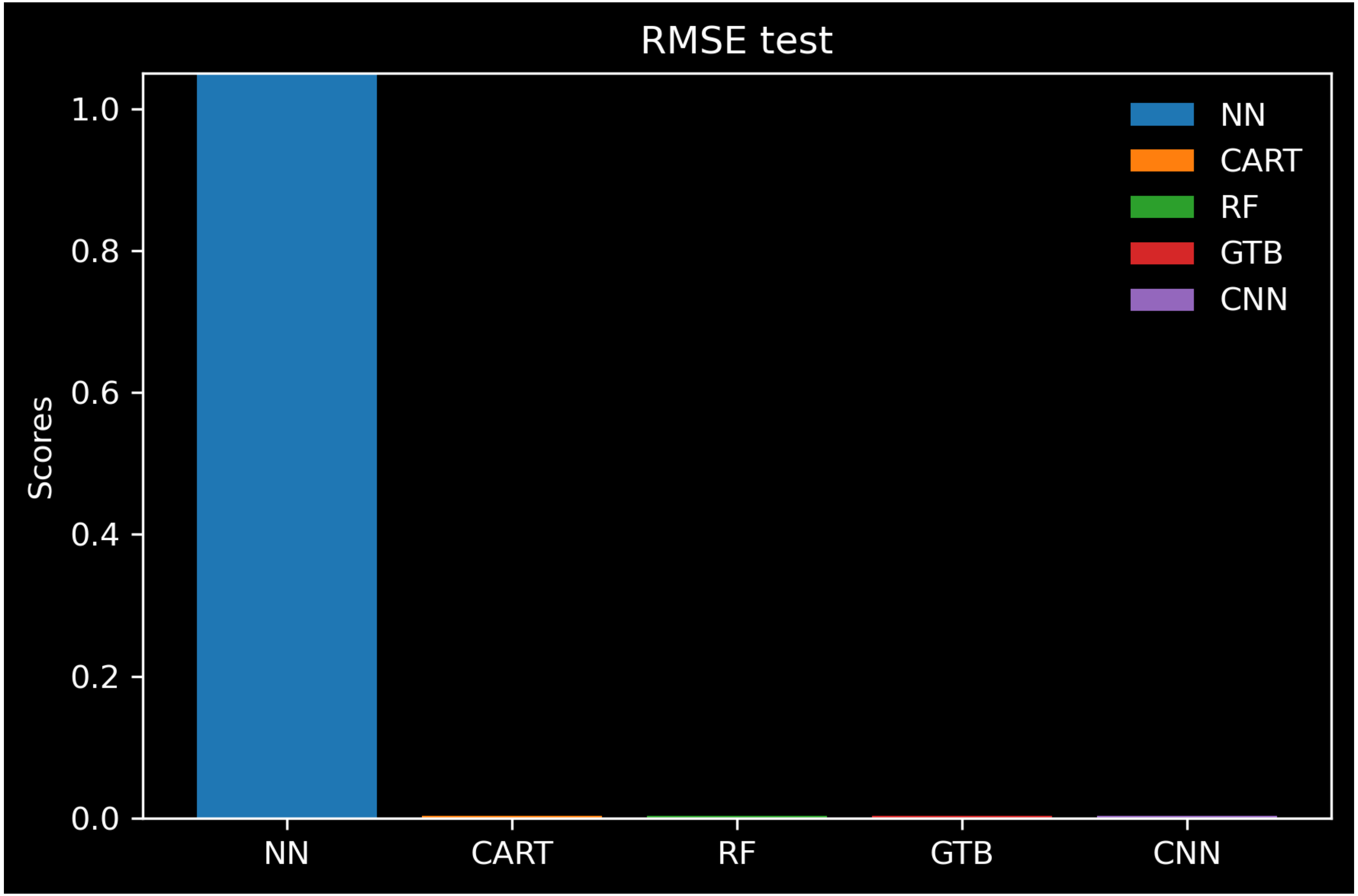
ML method evolution

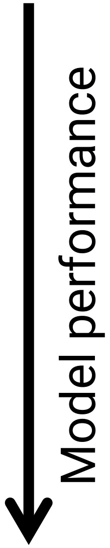


impervious surface classification

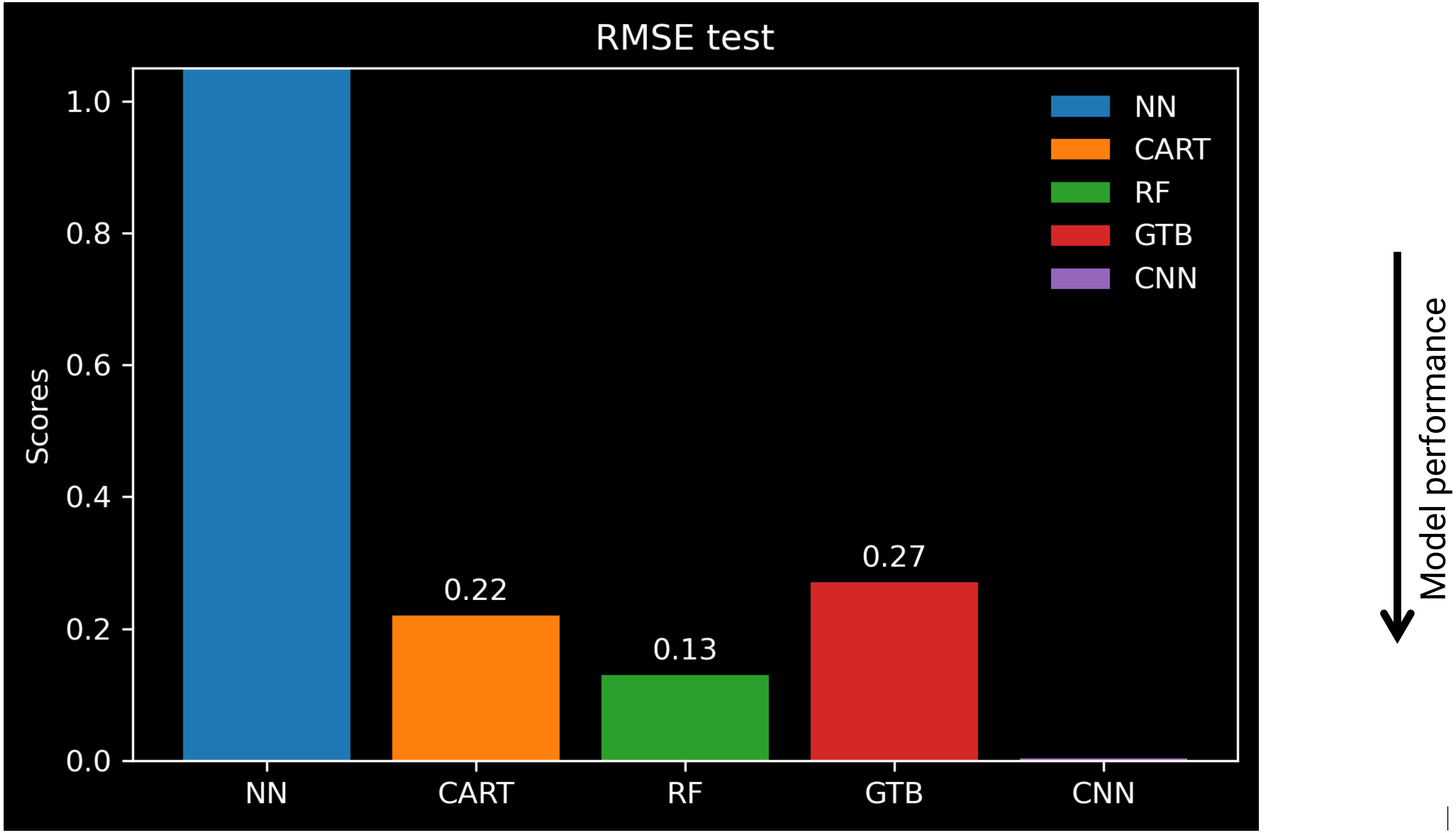


How good are the ML algorithms?

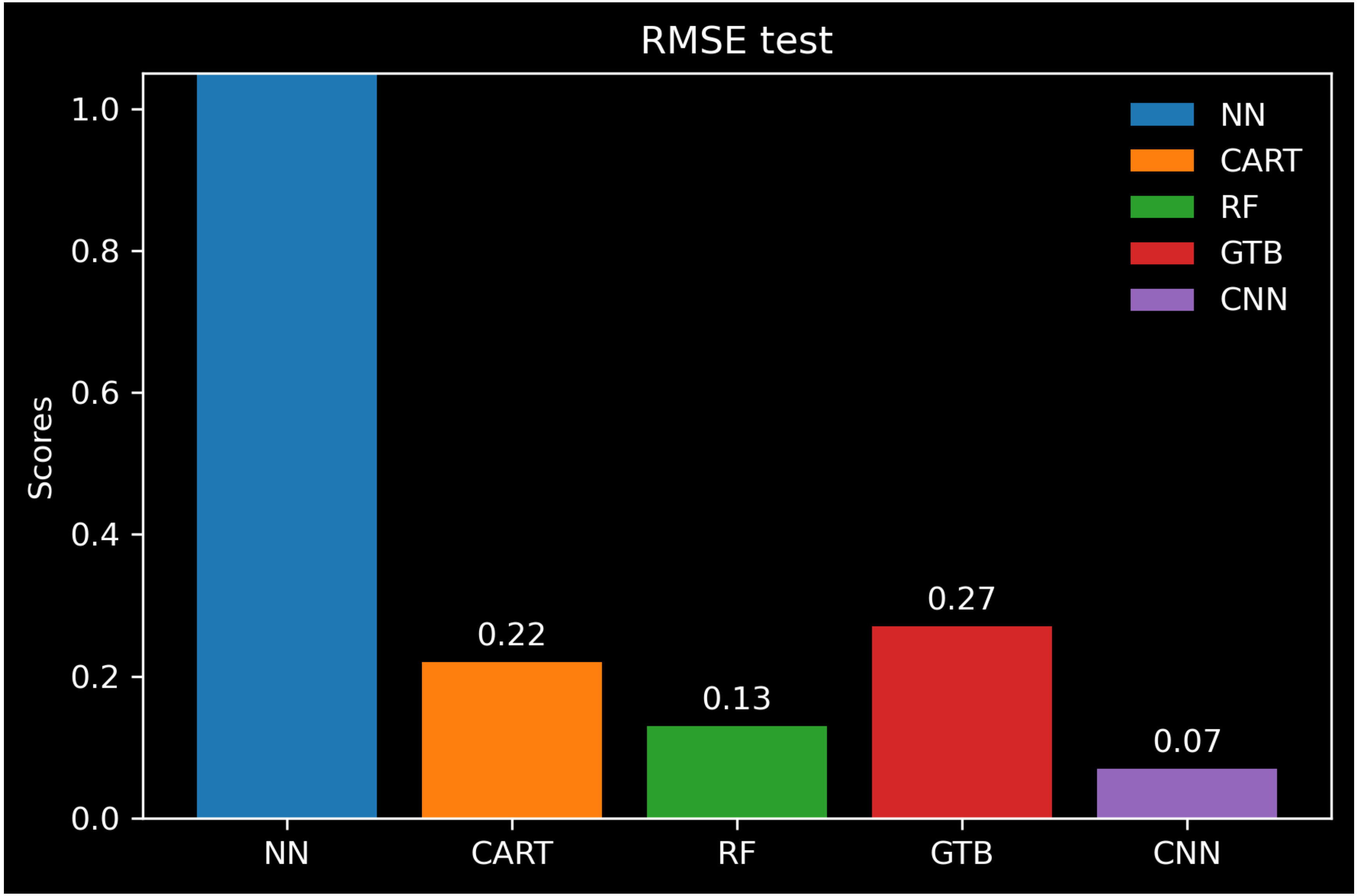


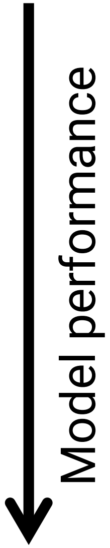


How good are the ML algorithms?



How good are the ML algorithms?






□et□ods are constantl□ improving
□
□nlimited potential

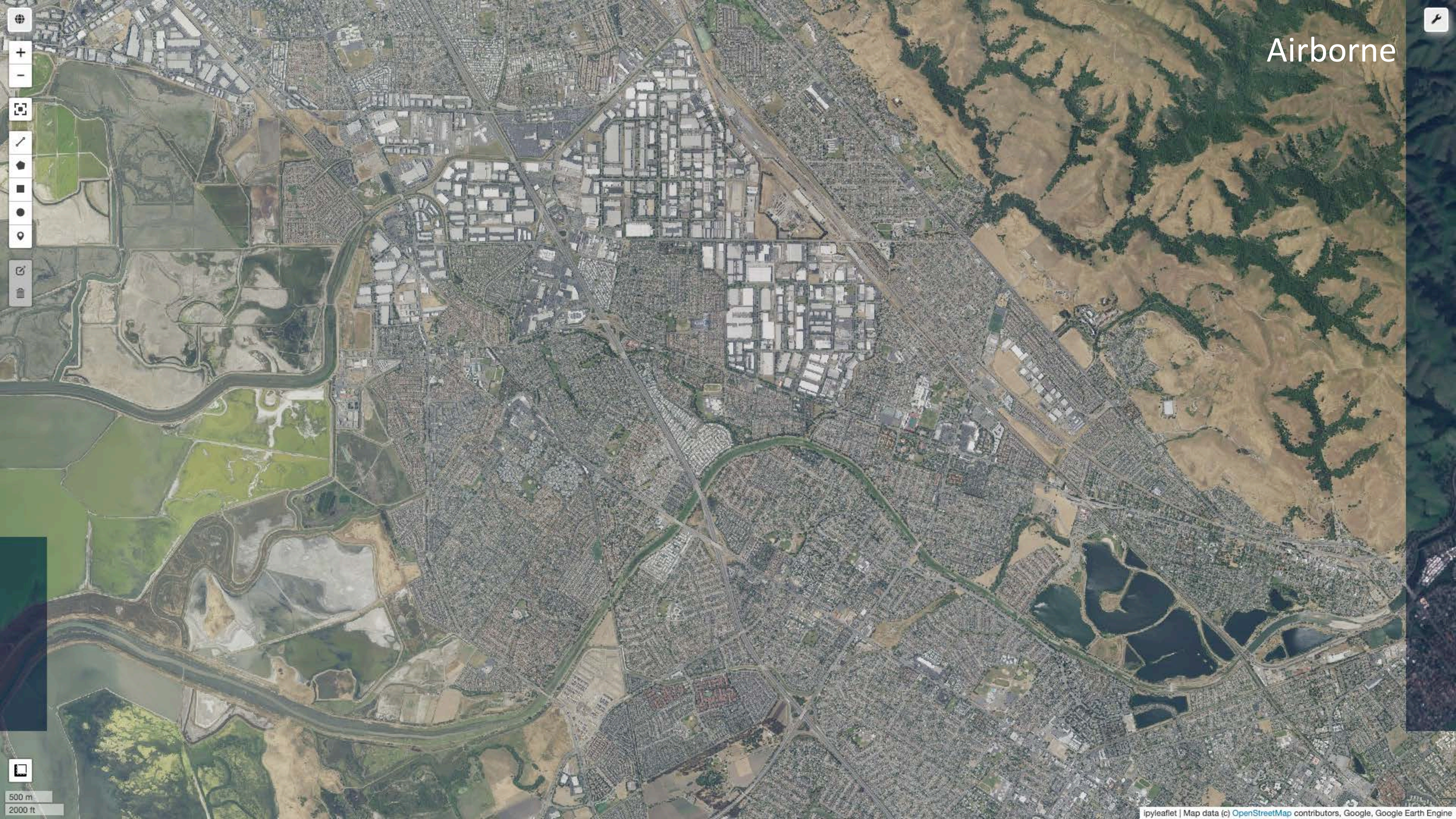
But what about the effect of data quality on machine learning performance?

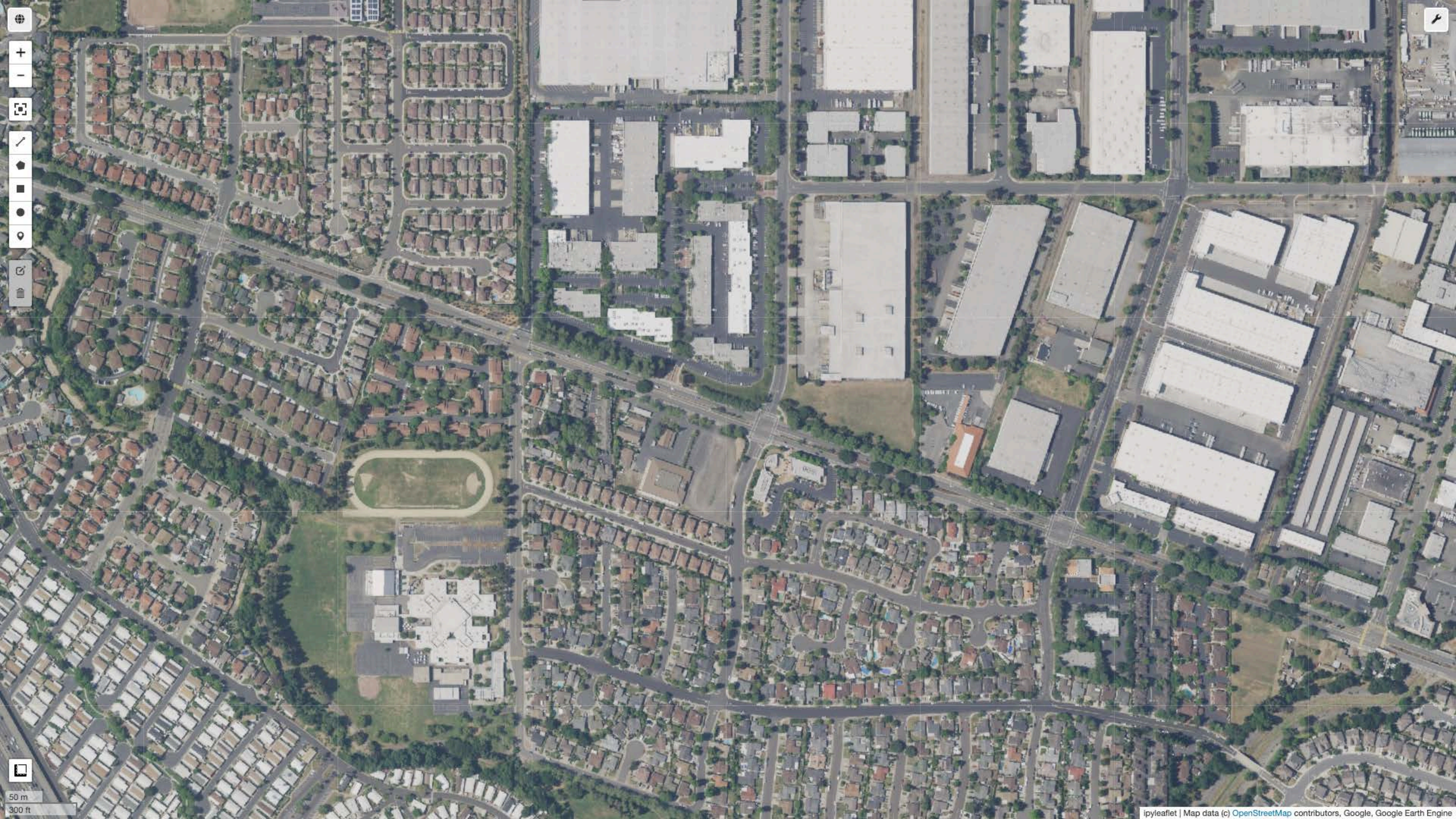

Ne□ Data
□
ne□ problems

New data, new problems: e.g. shadows

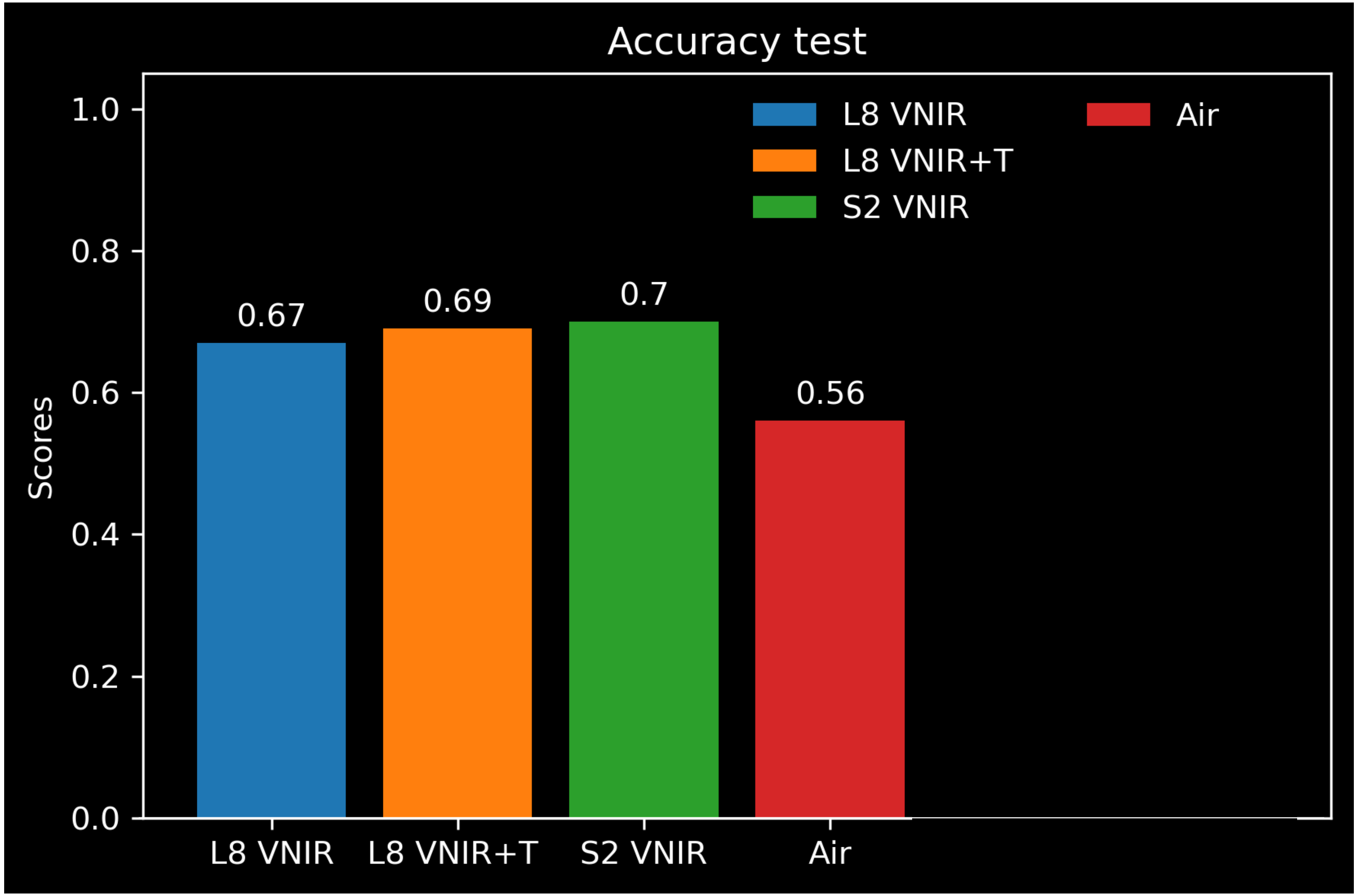

Landsat-8
500 m
2000 ft
ipyleaflet | Map data (c) OpenStreetMap contributors, Google, Google Earth Engine

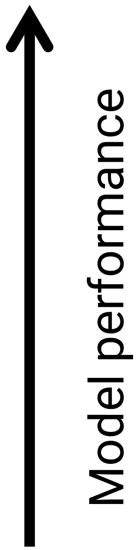

Sentinel-2
500 m
2000 ft
ipyleaflet | Map data (c) OpenStreetMap contributors, Google, Google Earth Engine


Airborne
500 m
2000 ft
ipyleaflet | Map data (c) OpenStreetMap contributors, Google, Google Earth Engine

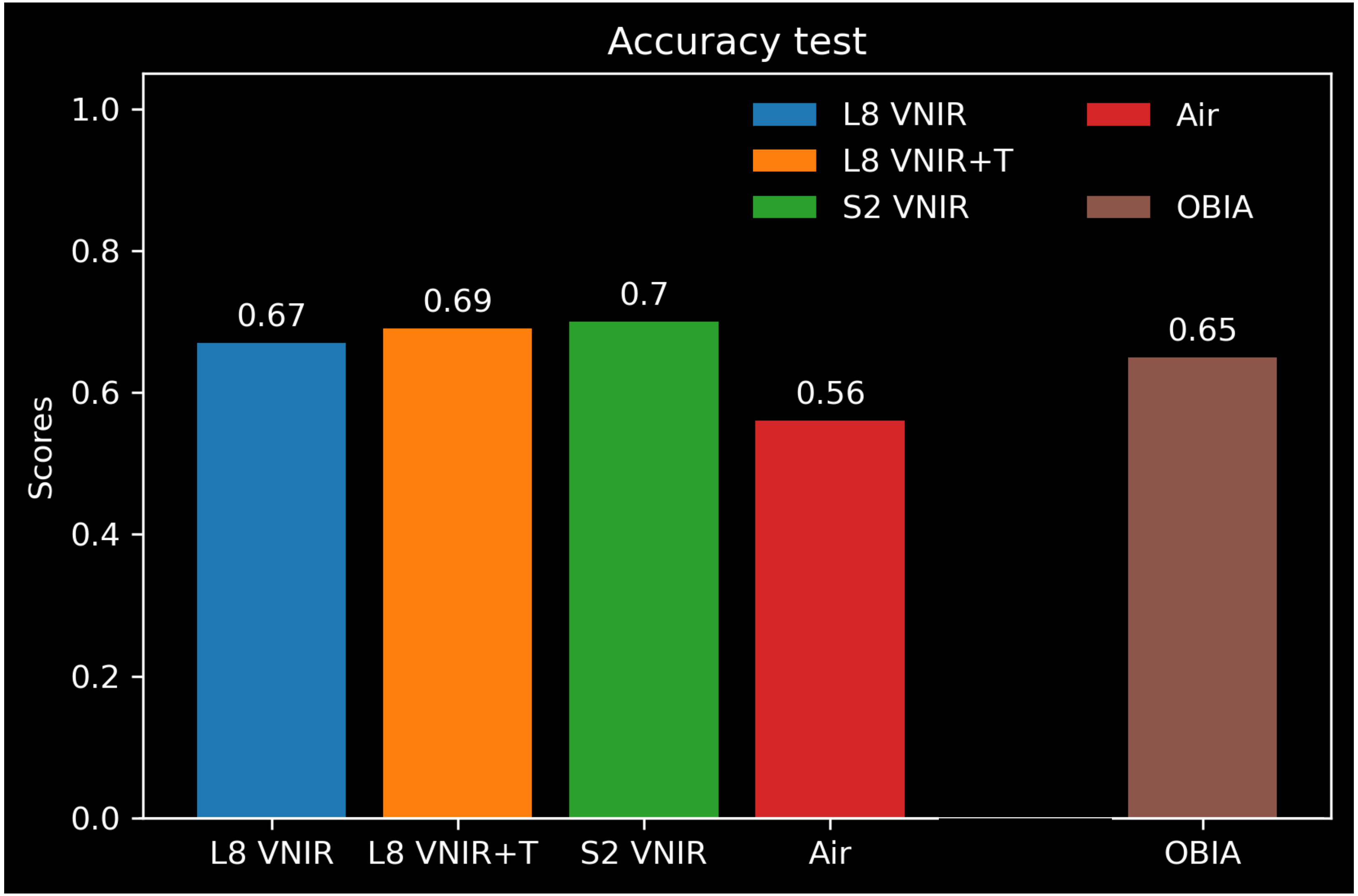

50 m
300 ft
ipyleaflet | Map data (c) OpenStreetMap contributors, Google, Google Earth Engine

New data problems: e.g. shadows





New data problems: e.g. shadows



Model performance

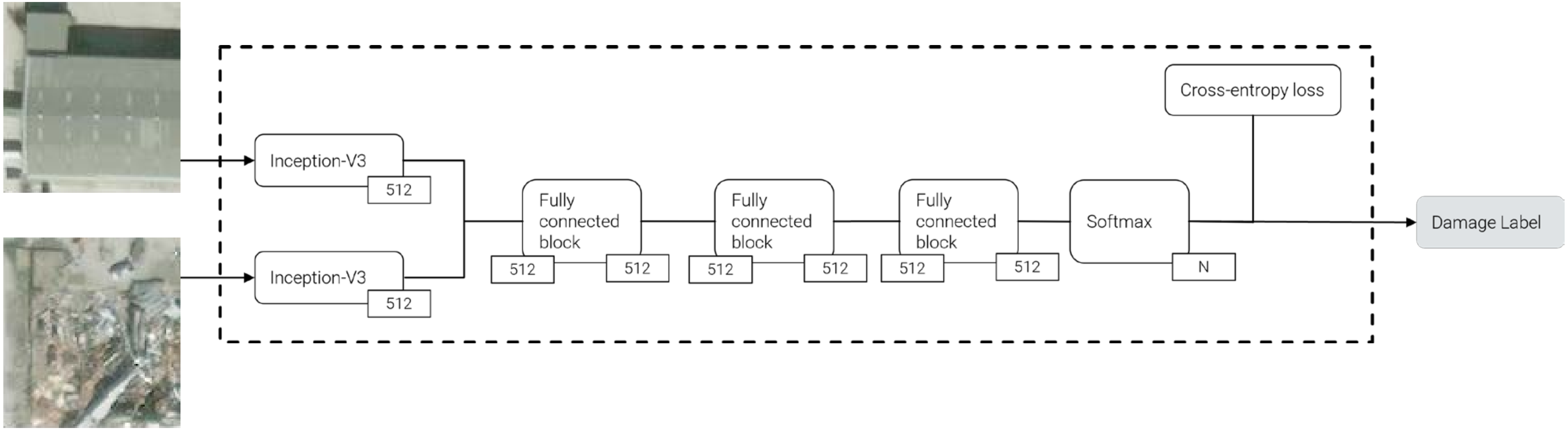
New data problems: change detection for disaster management

Pre-event

Post-event

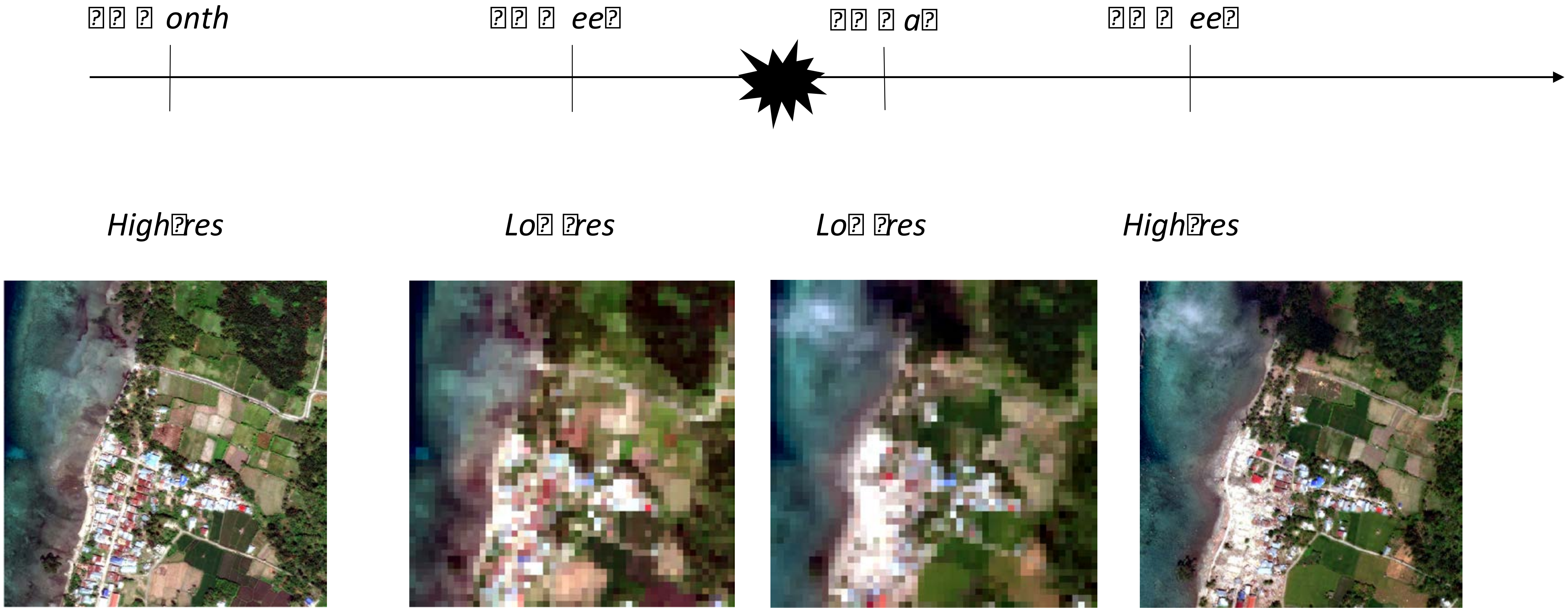
Verschoor et. al. (2022)

The 510 Caladrius model



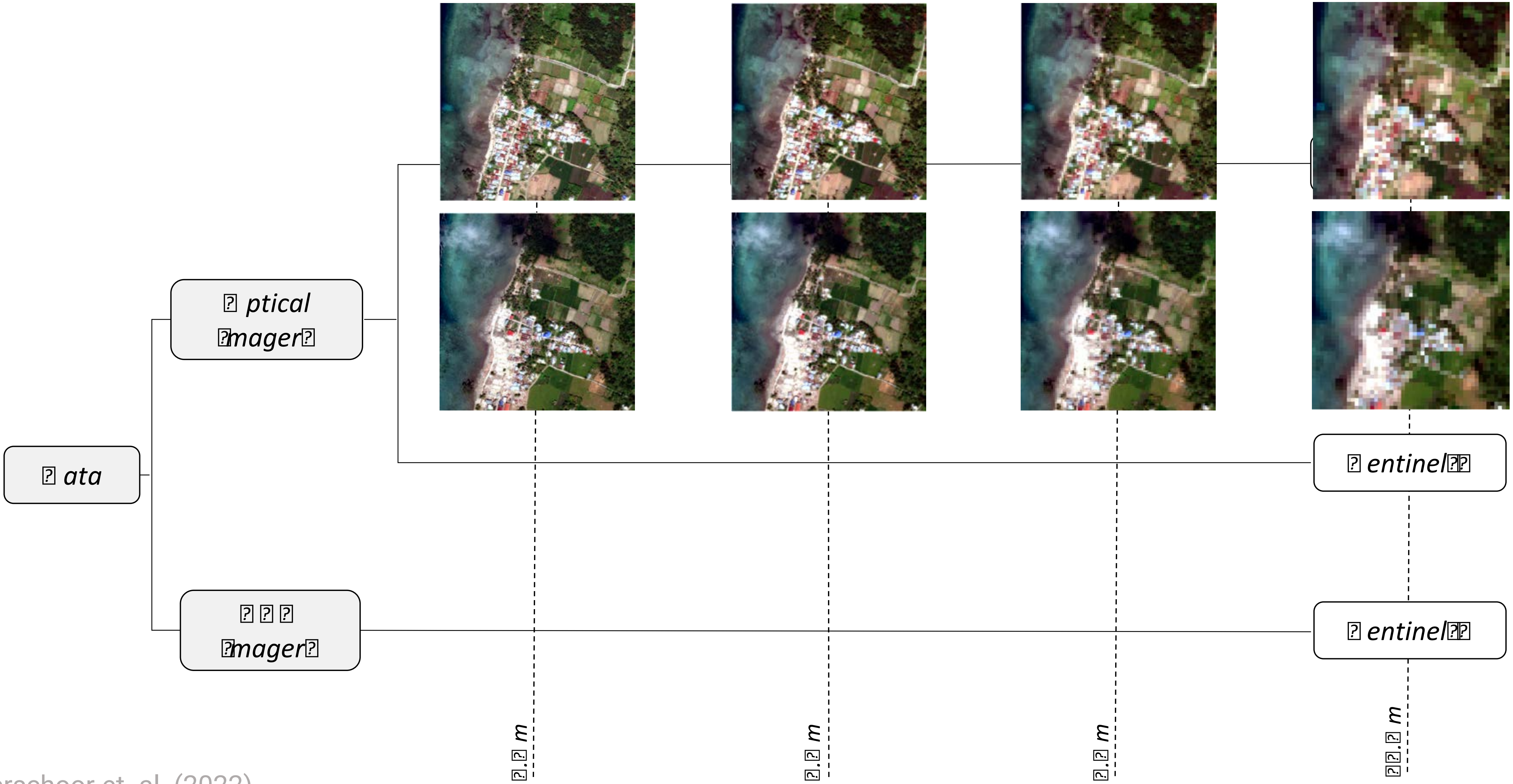
Verschoor et. al. (2022)

New data problems: time urgency



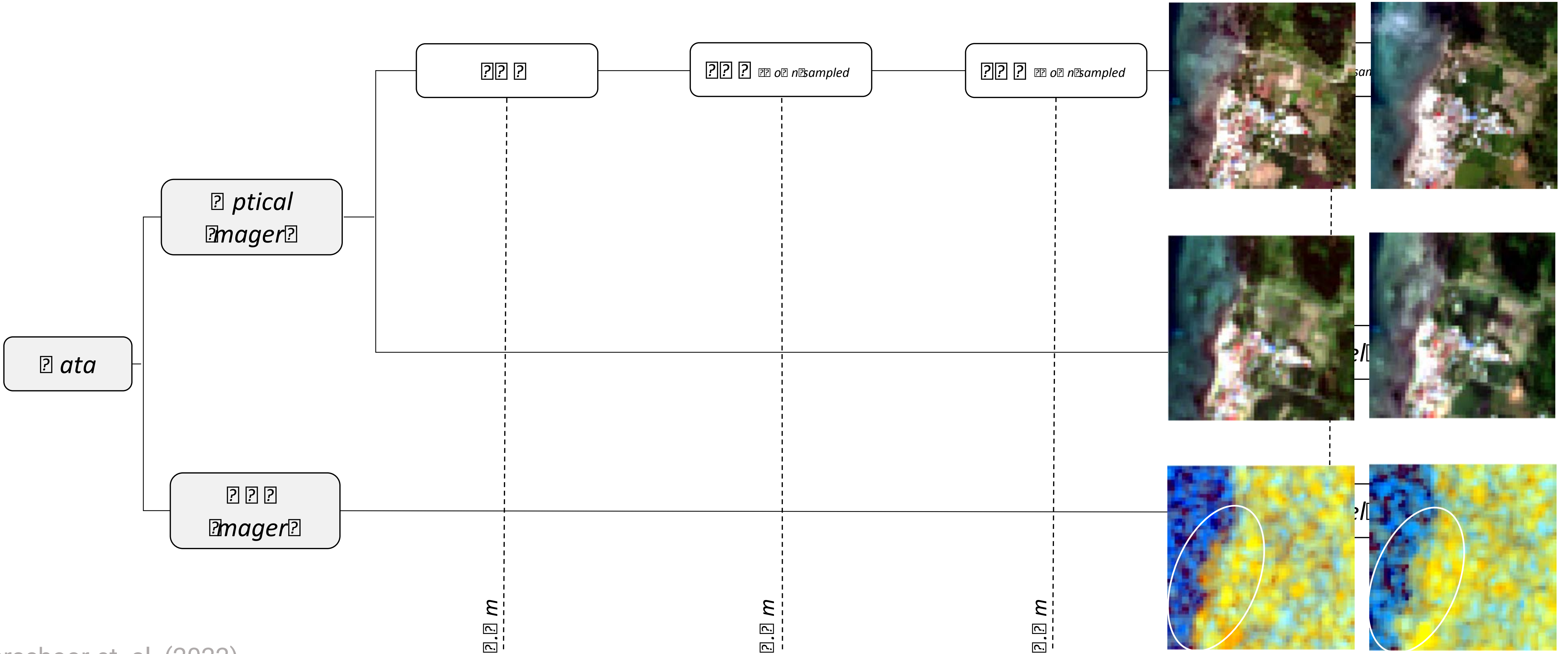
Verschoor et. al. (2022)

The input data is originated from the xBD, Sentinel-2 and Sentinel-1 dataset



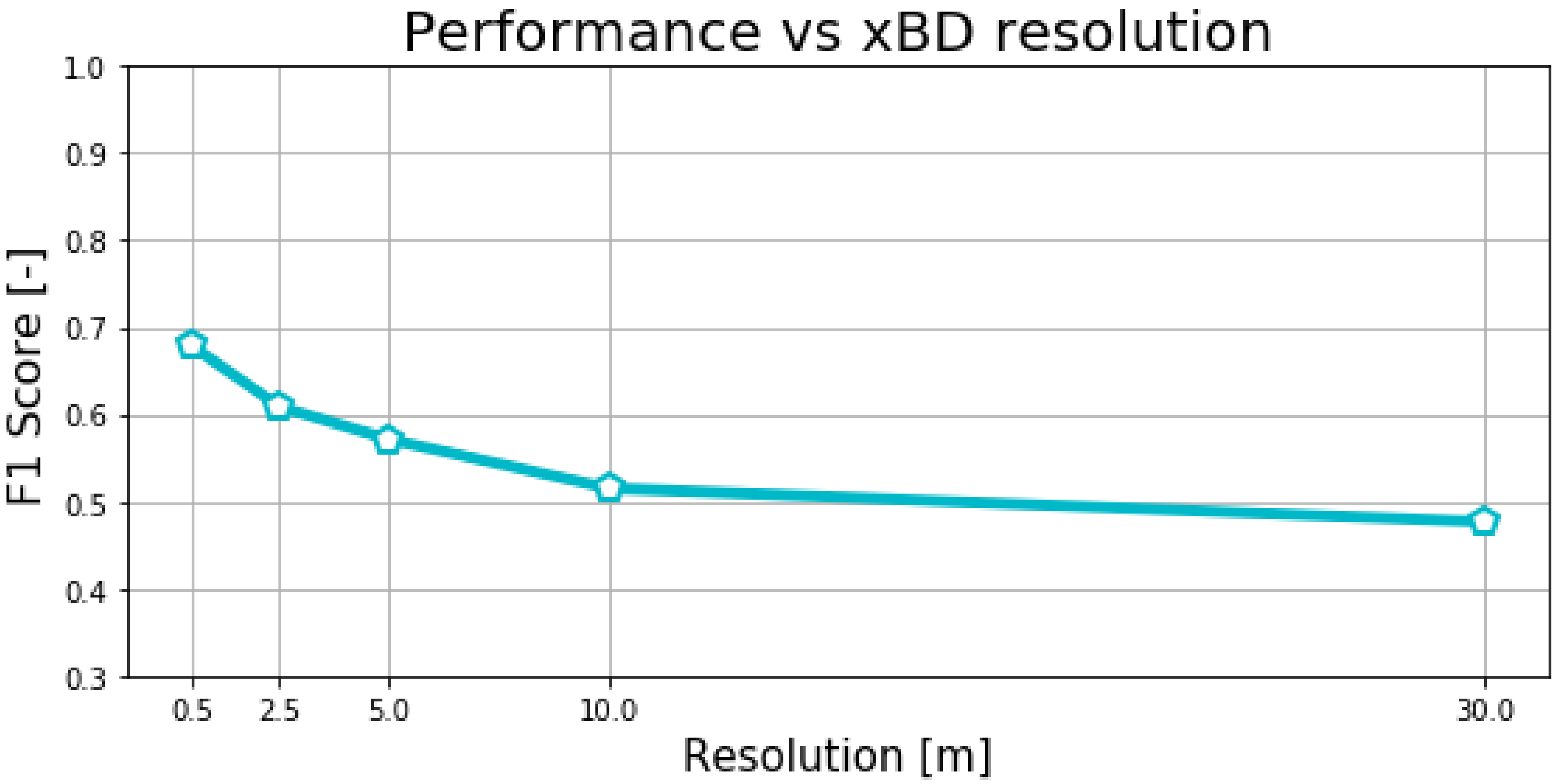
Verschoor et. al. (2022)

The input data is originated from the xBD, Sentinel-2 and Sentinel-1 dataset



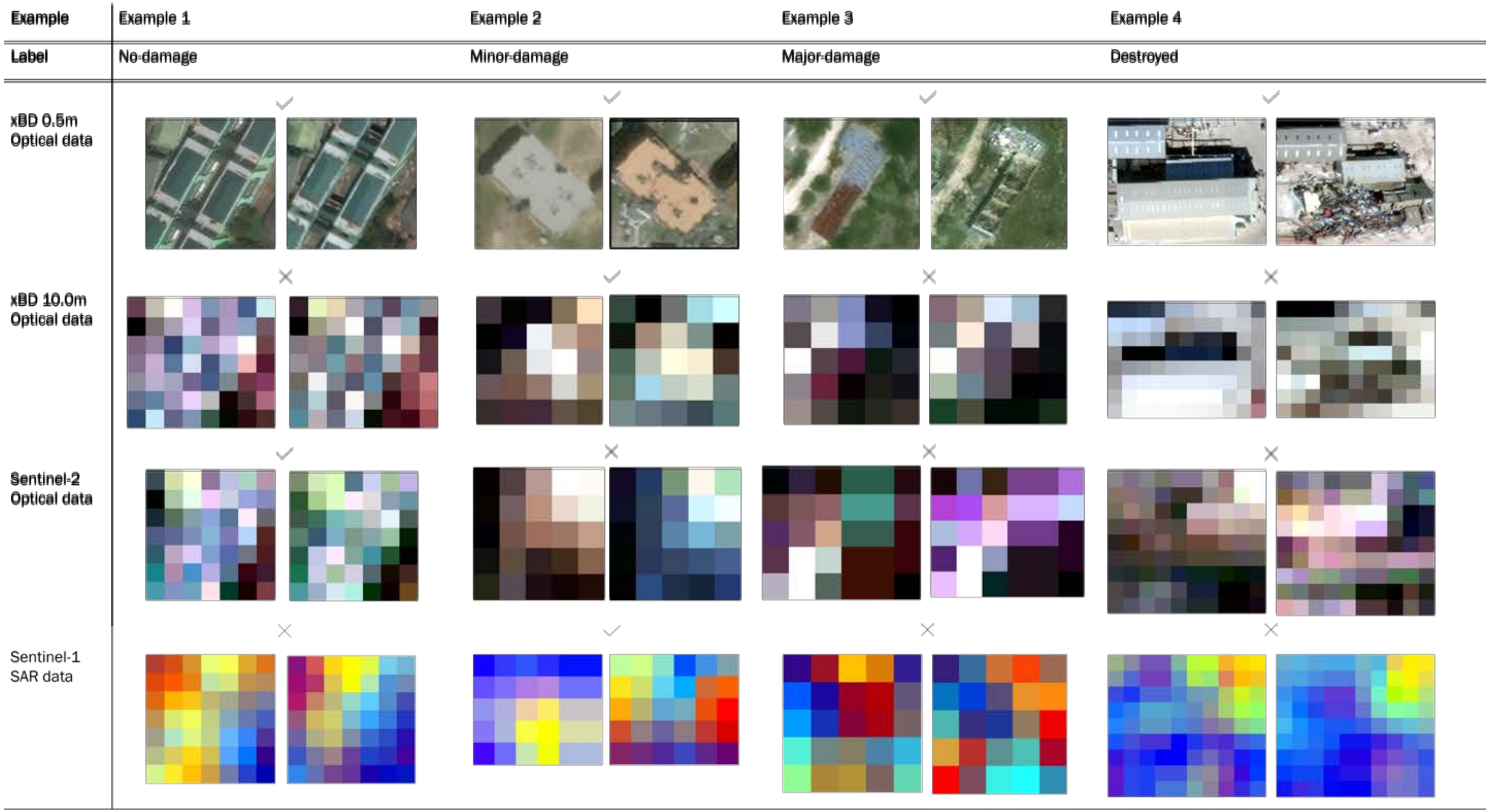
Verschoor et. al. (2022)

Data resolution matters …



Verschoor et. al. (2022)

Data type matters ....



0.68

0.52

0.44

0.32

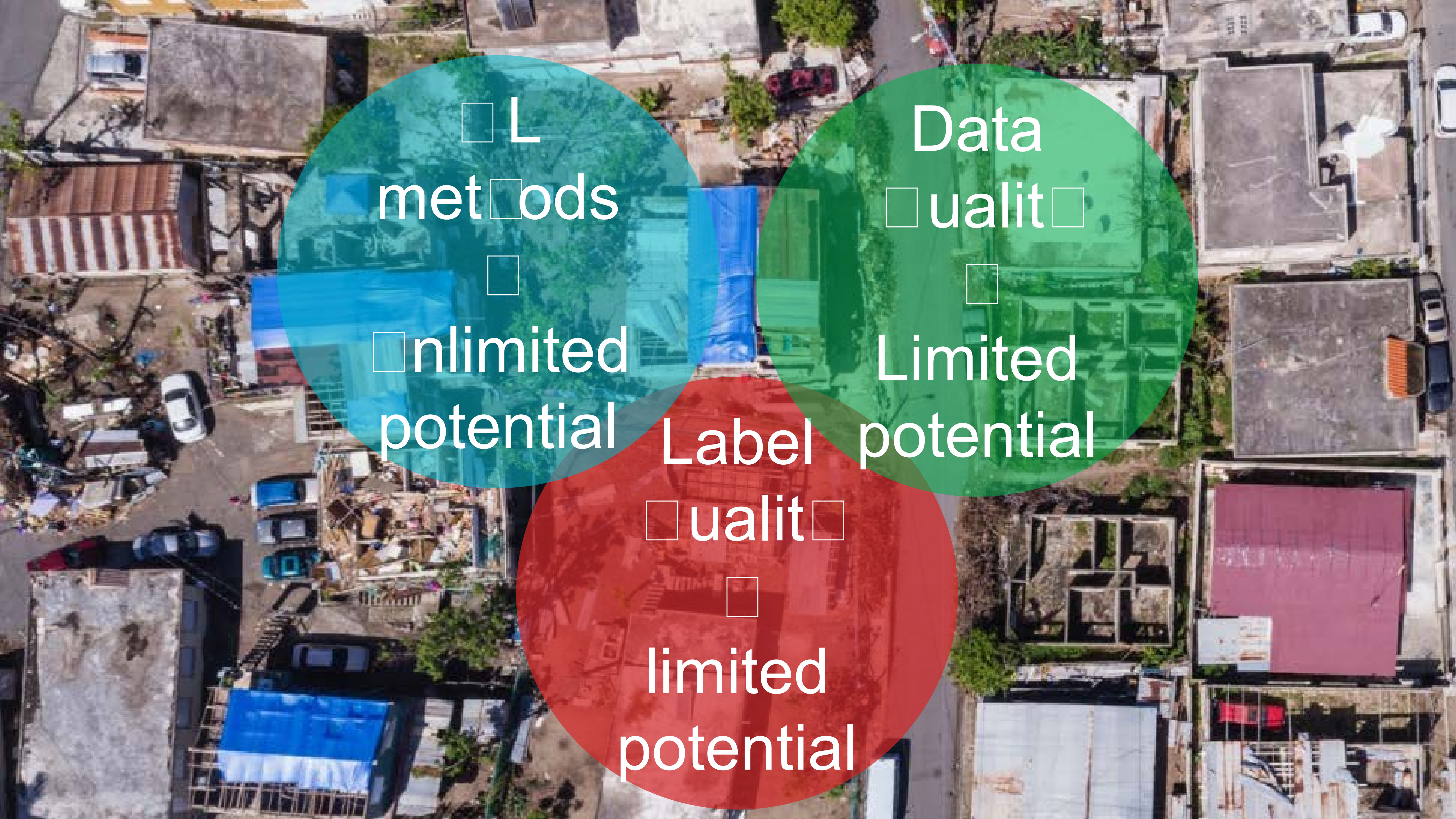
correct incorrect

Verschoor et. al. (2022)

… but not always that much

Pre-Event \ Post-Event	0.5m	2.5m	5.0m	10.0m
xBD 0.5m	0.65	0.63	0.61	0.57
xBD 2.5m	X	0.61	0.58	0.57
xBD 5.0m	X	X	0.57	0.53
xBD 10.0m	X	X	X	0.52

Verschoor et. al. (2022)


□L
met□ods
□
□nlimited
potential
Data
□ualit□
□
Limited
potential
Label
□ualit□
□
limited
potential

Conclusions

ML methods offer promising methodological opportunities
but are no magical solutions.

ML methods require application/data/label-specific development

Even the best ML methods suffer from "*garbage in = garbage out*"