

Automated instance generation for physics-aware machine learning

Laurens Arp

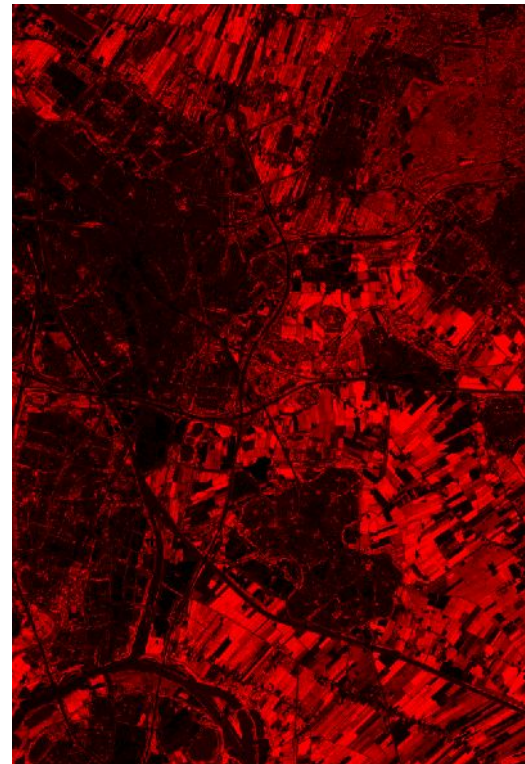
Supervisors: Mitra Baratchi, Holger Hoos, Peter van Bodegom

Collaborating: Alistair Francis (ESA), Nguyen Dang (Saint Andrews University)



**Universiteit
Leiden**
The Netherlands

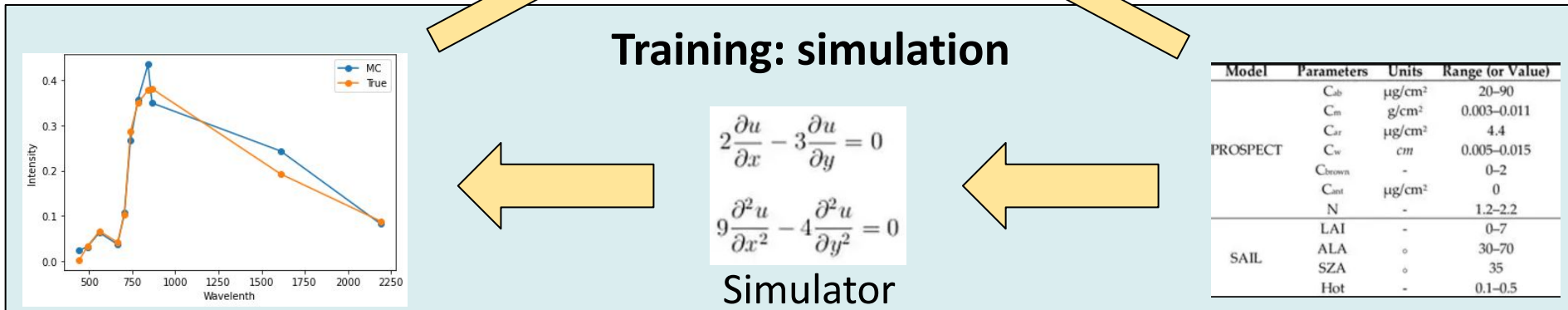
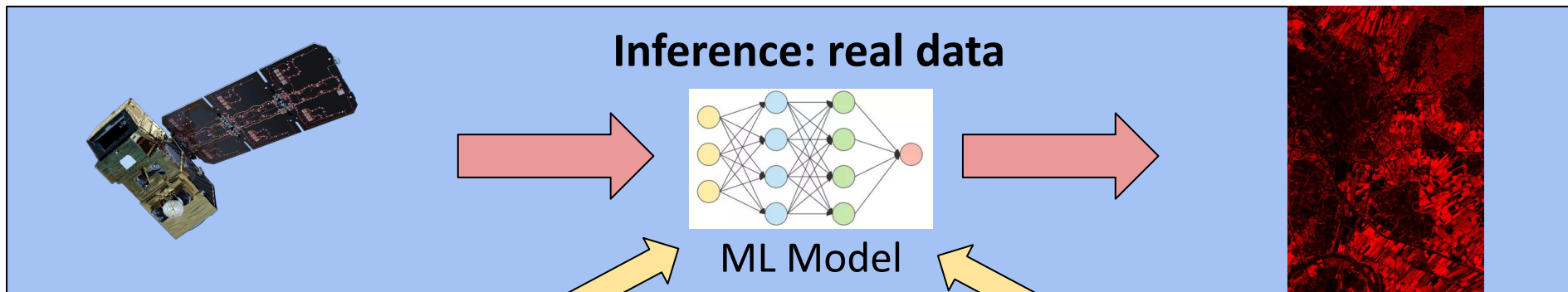
Some background



Problem: ground truth (in-situ)

- Physical parameters are expensive to measure
- Earth is incredibly diverse
 - Issues for transferability and generalisability
- Typical solution: use simulators
 - Physical model inversion

Model inversion



State-of-the-art

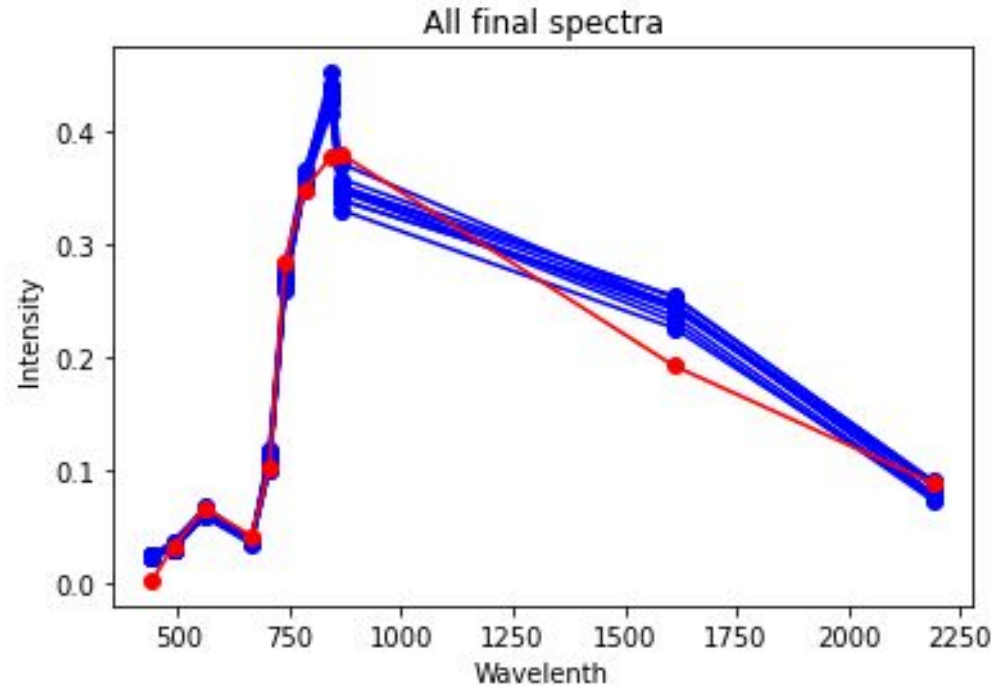
- Latin hypercube sampling (LHS)
- Active learning (AL)
 - Heuristics-based
 - Too general, no guarantees
 - Validation loss-based
 - Ground truth availability, transferability issues

Our proposed solution (AutoAL*?)

- Leverage availability of unlabelled satellite data
 - Compute distribution of true reflectance
 - Generate lookup-table
 - Use optimisation techniques to match distribution to true reflectance
 - Specific to study area
 - Can be generated *without* ground truth data

*ADA seal of approval for this nice catchy cheesy name?

Simulated output VS real spectrum



Some selected preliminary results

- AutoAL much better than random search
 - 0.84 — 4.92
- AutoAL better than post-hoc selection
 - 0.84 — 2.96
- Pure ML much better than AutoAL
 - 0.07 — 0.84

(results table, check later if interested)

Method	MAE
Random search	4.92
Random + post-hoc instance selection	2.96
ML (similar ecosystem training)	0.07
ML (dissimilar ecosystem training)	0.18
AutoAL	0.84
AutoAL selected parameters	1.15

Open problems

- Ill-posedness
- Performance compared to ‘vanilla’ ML
- Optimising population (to match distribution)
- Simulator configuration
- Simulator inaccuracies
- Advances in the AL field

Automated instance generation for PA-ML

- Physical models can be combined with ML
 - Physical model creates ML training data
- Existing approaches try to be too general
- Our proposed method generates data specific to study area
- WIP, stay tuned for more!

Laurens Arp

LIACS – Leiden University

l.r.arp@liacs.leidenuniv.nl