AI: a scientific & technical change … or a cultural challenge?

ALBERTO ARRIBAS
Hd Informatics Lab - Prof. Univ. of Exeter - Turing Fellow

Thanks to …

Niall Robinson
Rachel Prudden
Sam Madge
Maria Athanassiadou
.
.

Suman Ravuri
Karel Lenc
Shakir Mohamed
Ellen Clancy
Matt Wilson
.
.

Dmitry Kangin

Some references:

What is next for National Met Services? (https://arxiv.org/abs/2005.01425)

A review of radar-based nowcasting (https://arxiv.org/abs/2005.04988)

ML Nowcasting (working title …
sorry, I cannot tell you much about this one yet!)

Established in 2015

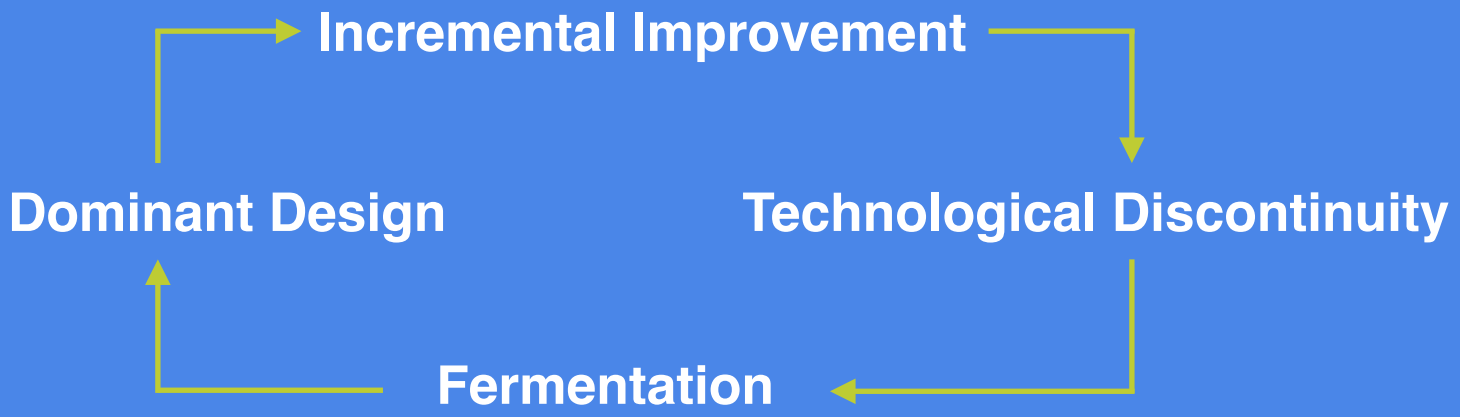
Focus on making environmental science and data useful

Multi-disciplinary (science/tech/design) department at Met Office

Extensive collaborations across academia, government and IT industry (DeepMind; Microsoft; Amazon and others)

And, concluding ….

AI is a cultural challenge



Technology Lifecycle Model
(Tushman and Anderson, 2004)

“Incremental” framing

“Discontinuity” framing

“Incremental” framing

Doing the same better with new tools

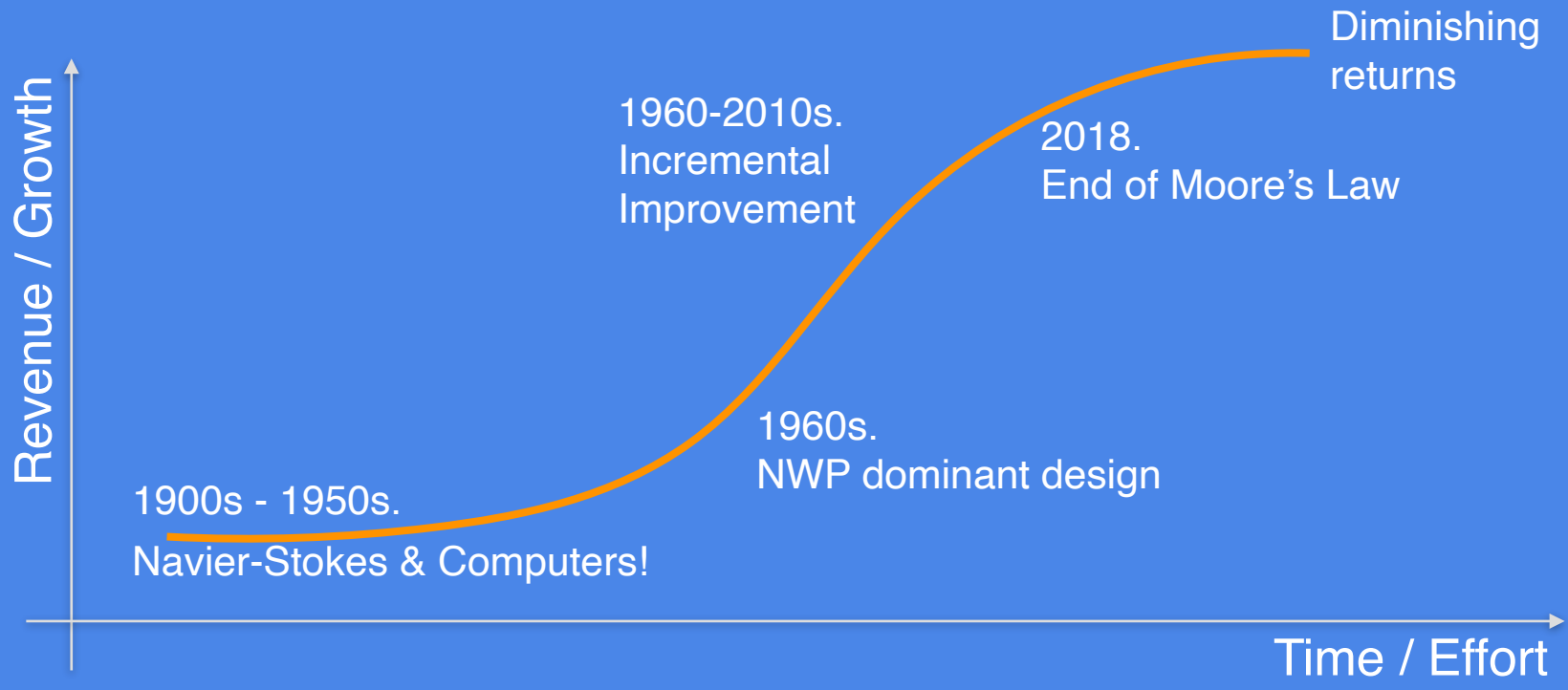
vs.

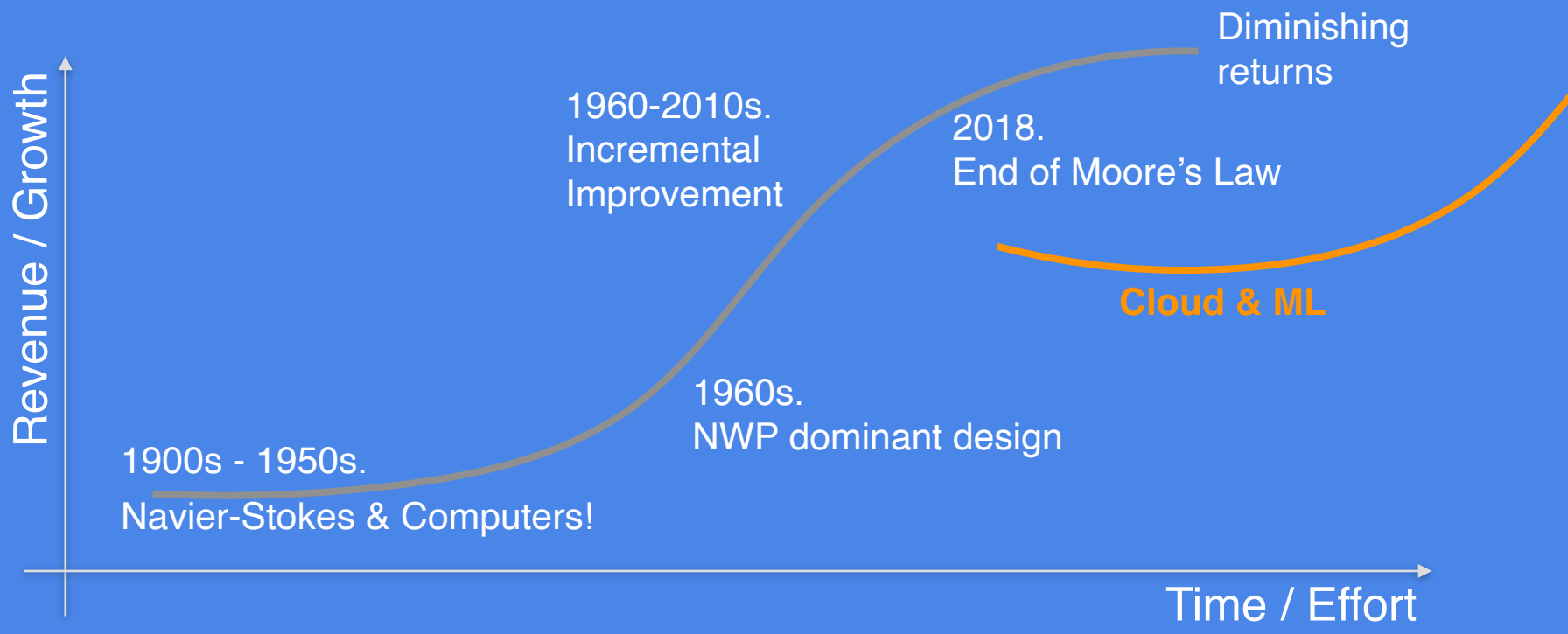
“Discontinuity” framing

Doing something different
that we could not do before
(i.e. cultural challenge)

Companies are more likely to survive a recession than to adapt to a tech. discontinuity

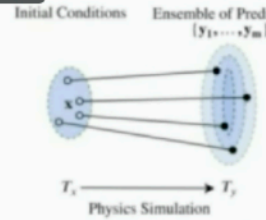
Anderson and Tushman (1991)

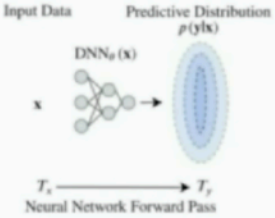

Revenue / Growth
Time / Effort
1900s - 1950s.
Navier-Stokes & Computers!
1960s.
NWP dominant design
1960-2010s.
Incremental
Improvement
2018.
End of Moore's Law
Diminishing
returns


Revenue / Growth
Time / Effort
1900s - 1950s.
Navier-Stokes & Computers!
1960s.
NWP dominant design
1960-2010s.
Incremental
Improvement
2018.
End of Moore's Law
Diminishing
returns
Cloud & ML

“Incremental” framing

LIVE 113





Numerical Weather Prediction	Neural Weather Models
Uncertainty through ensembling	Uncertainty from direct probabilistic model
Relies on equations of physics	Relies on generic learnable transformations (and requires data)
Latency of prediction is generally linear as a function of lead time	Latency of prediction is constant (or logarithmic) as a function of lead time. MetNet takes under a second for any lead time of up to 8 hours
Accuracy depends on underlying resolution. Doubling resolution requires ~ 8 times more computation.	Accuracy does not depend on underlying resolution. Doubling resolution requires ~ 4 times more computation.

Google

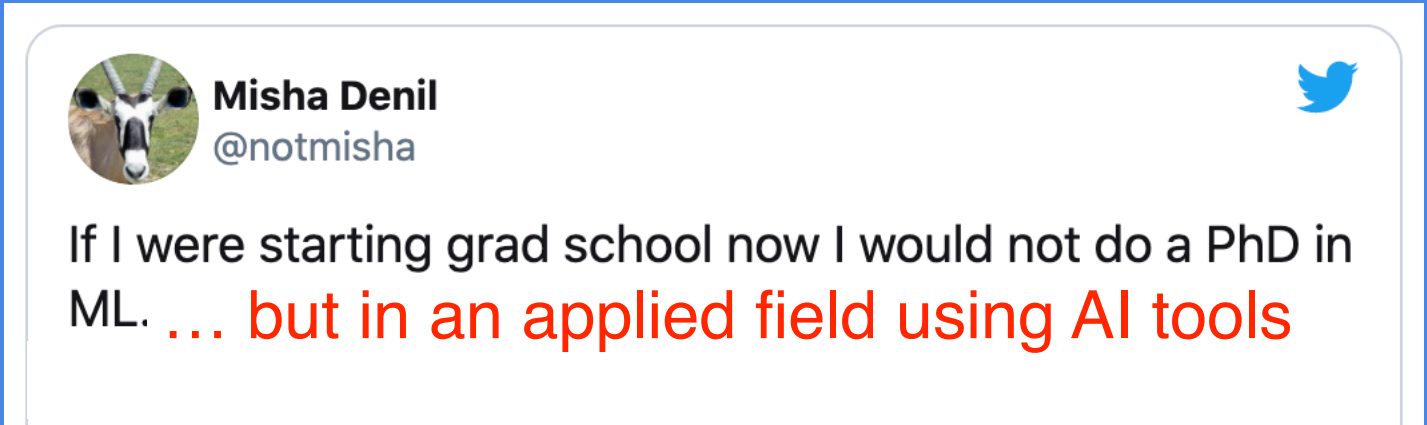
ECMWF

“Discontinuity” framing

NOT a one-way street

… but a crucial approach to improve AI

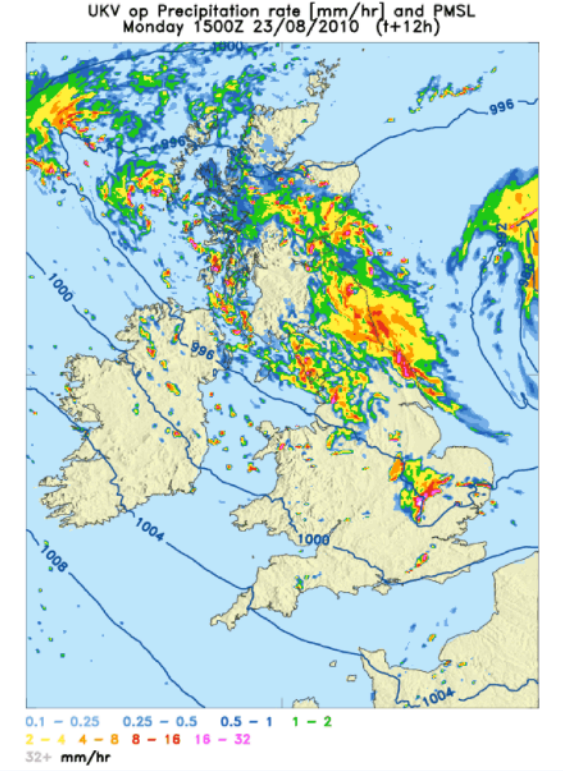
Tweet from Misha Denil (DeepMind)

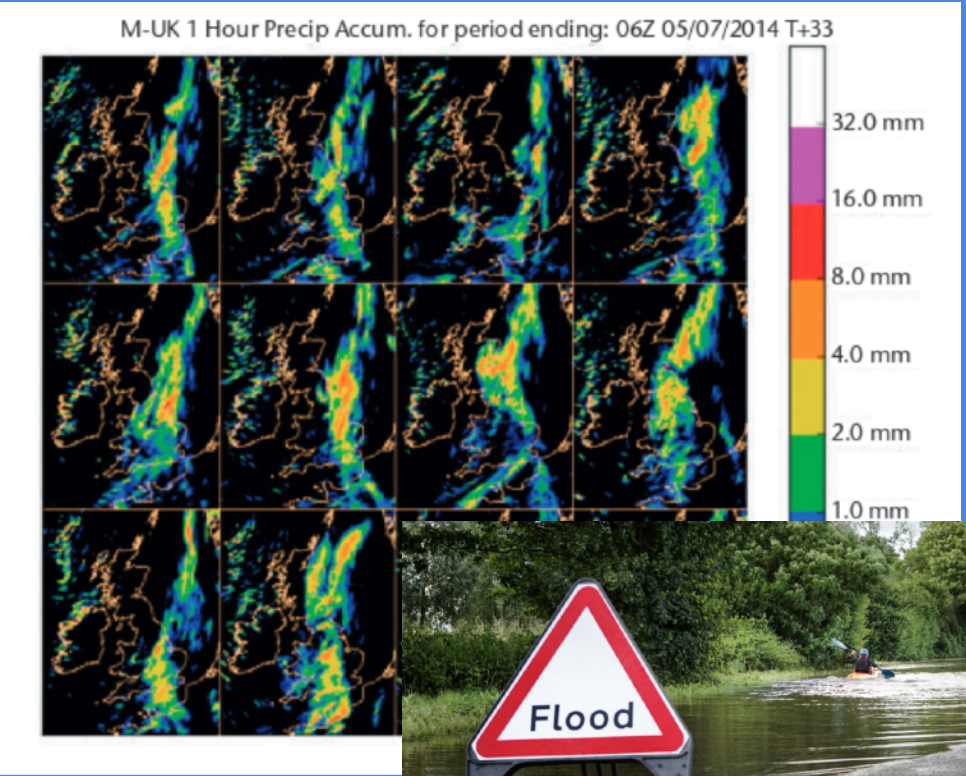


DeepMind & Informatics Lab

ML Nowcasting

Over two years working together and still the beginning of a long journey


UKV op Precipitation rate [mm/hr] and PMSL
Monday 1500Z 23/08/2010 (t+12h)
1000
996
1004
1008
0.1 – 0.25 0.25 – 0.5 0.5 – 1 1 – 2
2 – 4 4 – 8 8 – 16 16 – 32
32+ mm/hr


M-UK 1 Hour Precip Accum. for period ending: 06Z 05/07/2014 T+33
32.0 mm
16.0 mm
8.0 mm
4.0 mm
2.0 mm
1.0 mm

Flood

Point is not to “solve nowcasting” but …

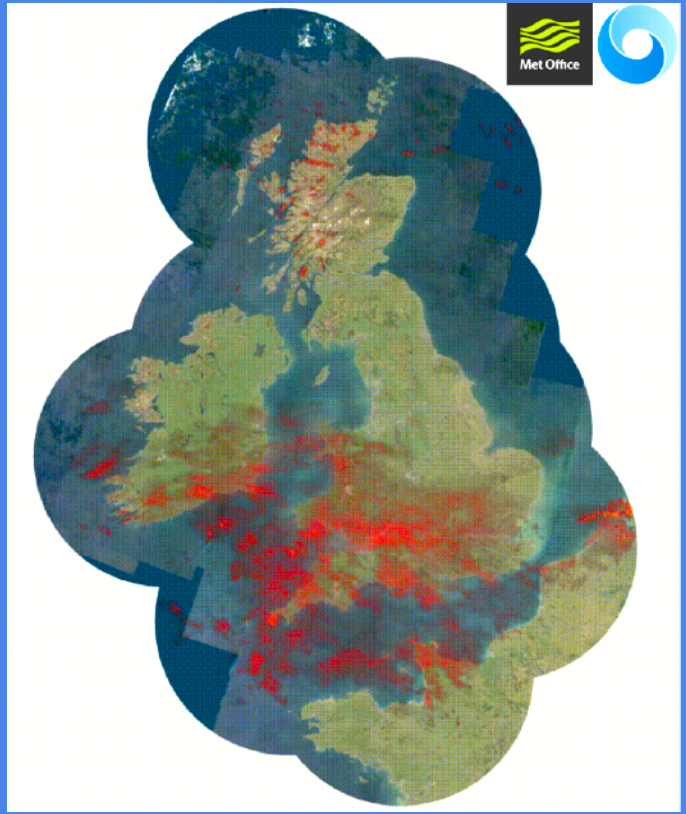
Physics-driven ML,
Personalised on-demand predictions,

.

.

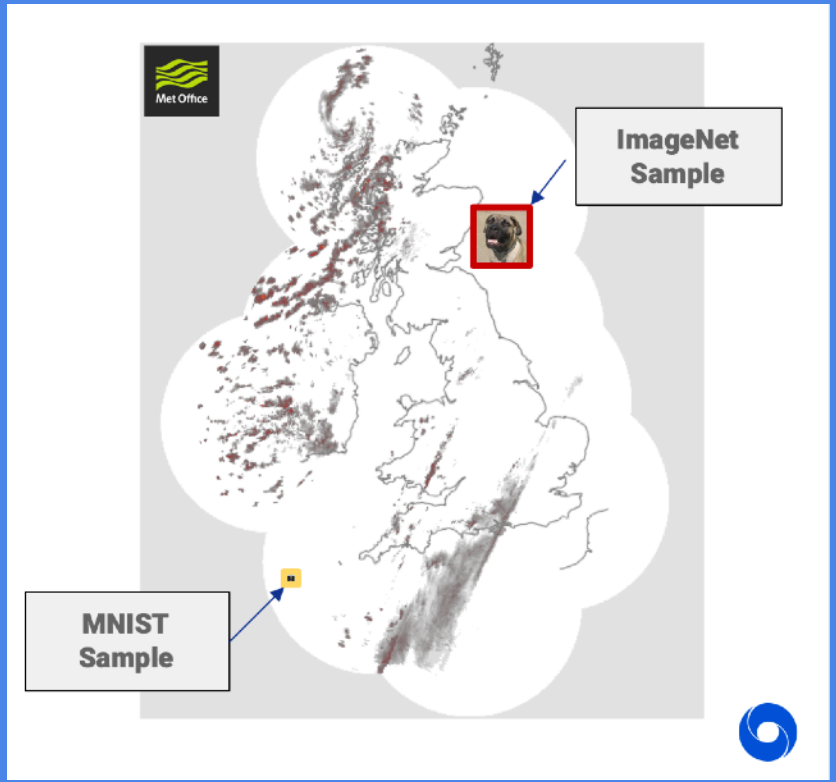
Challenges from **Nowcasting**:

1.- Multiple spatio-temporal scales (& regime changes!)
2.- Dense prediction task
3.- Extremes and out-of-sample events are critical
4.- Probabilistic prediction (and verification issues)



Radar images:
3072 x 2560 pixels
every 5 min

Far bigger than
“standard”
ML Image problems

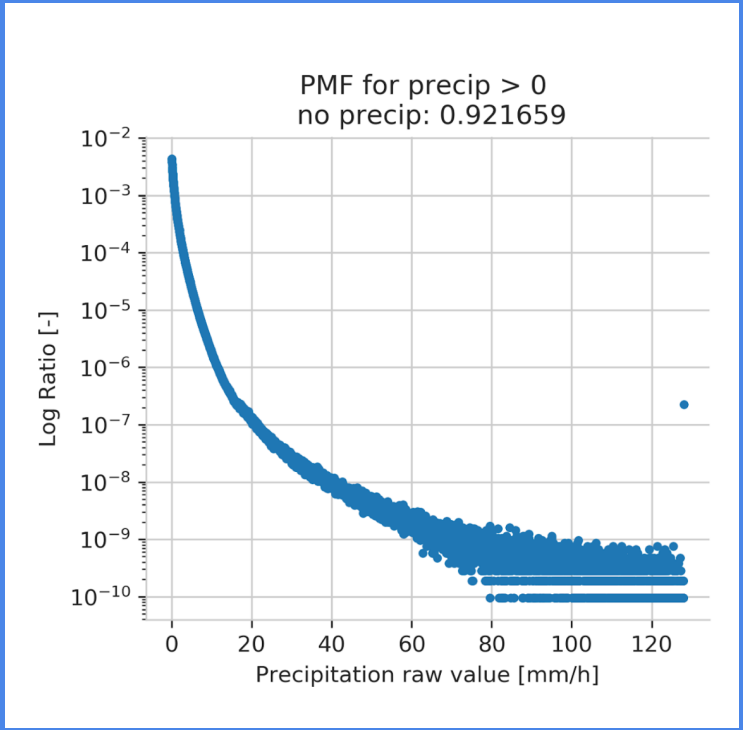


Data engineering

workflow, not GCM, is the centre of gravity

Data policy, licensing, benchmarking

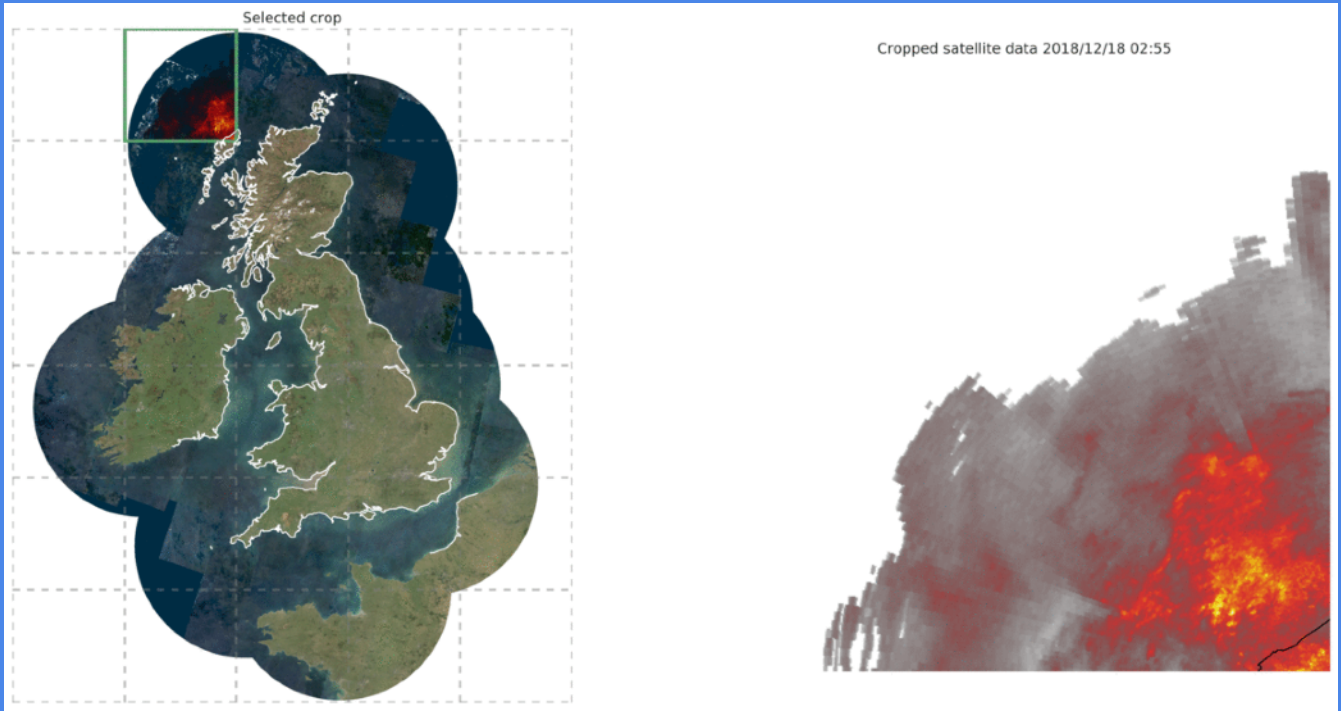
Opening Data to the community



Most of the data frames contain little precipitation

This bias ML model towards learning to predict low prec

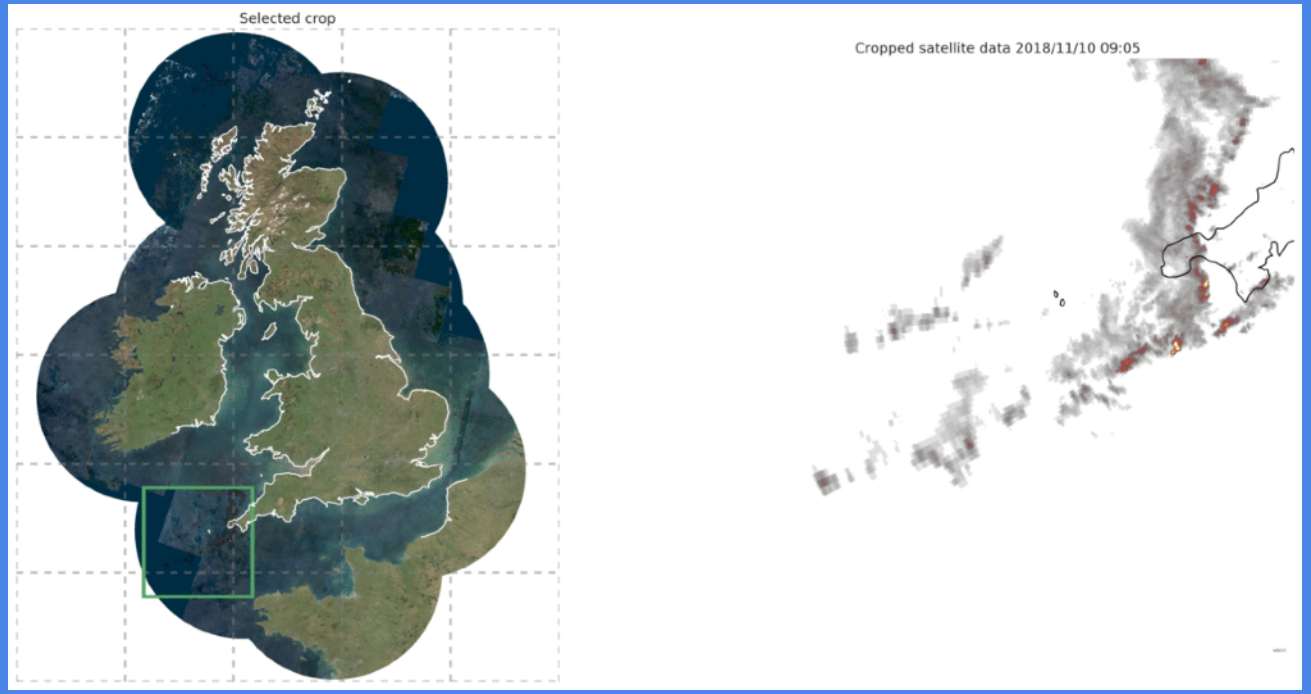
Validation Dataset



22 Fixed tiles
512x512

Importance Sampling: sample tiles heavy prec with higher prob

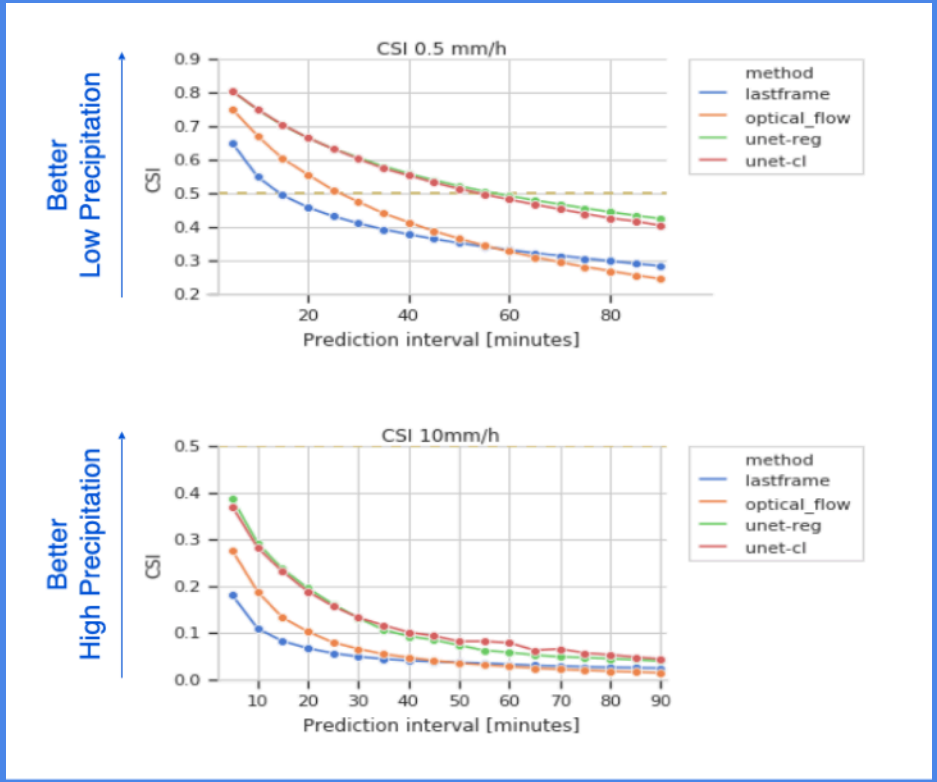
Training Dataset



Random Tiles

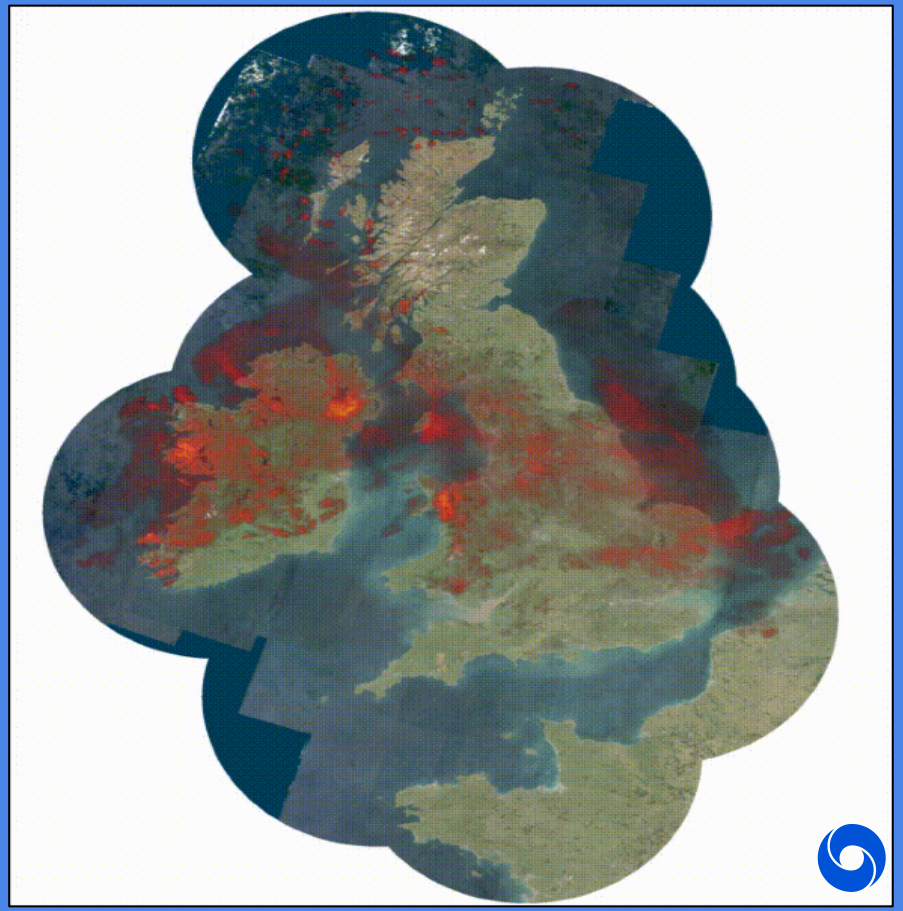
Largest possible Dataset desired

Verification … and the tyranny of the Leaderboard

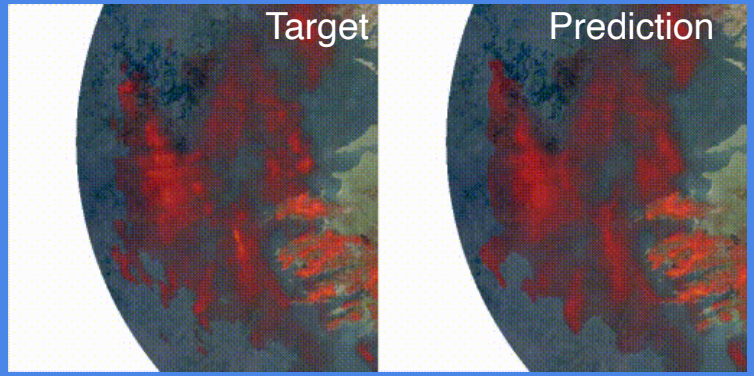


U-net looking great compared to baseline

But …



U-Net CL



Parachuting ML

Real-world question vs. Toy problem
Complex verification vs. Leaderboard

… you may end doing the same worse!

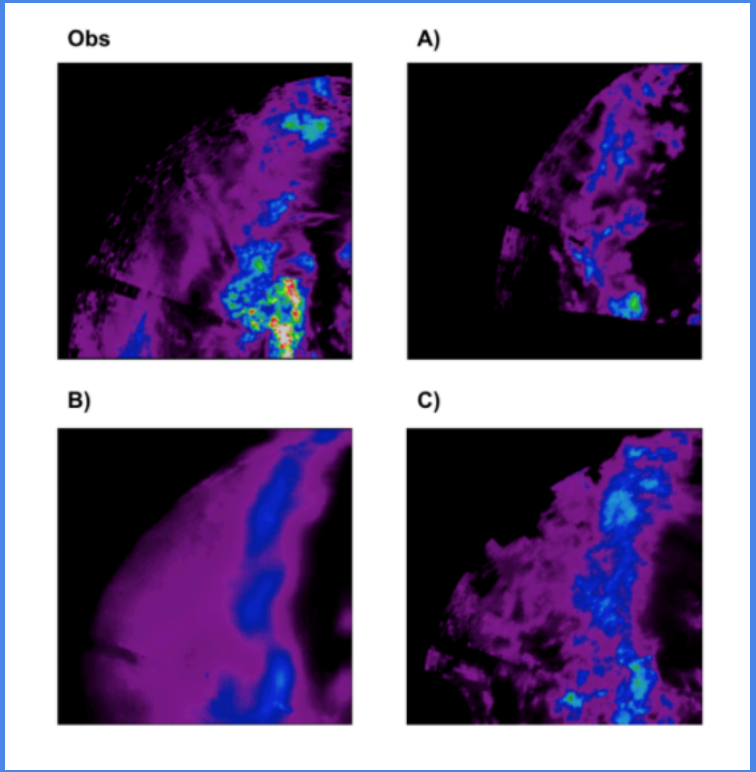
Non-parachuting ML

Complex basket of verification
“Behavioural verification”

.

.

Forecaster agency
Human - ML teaming

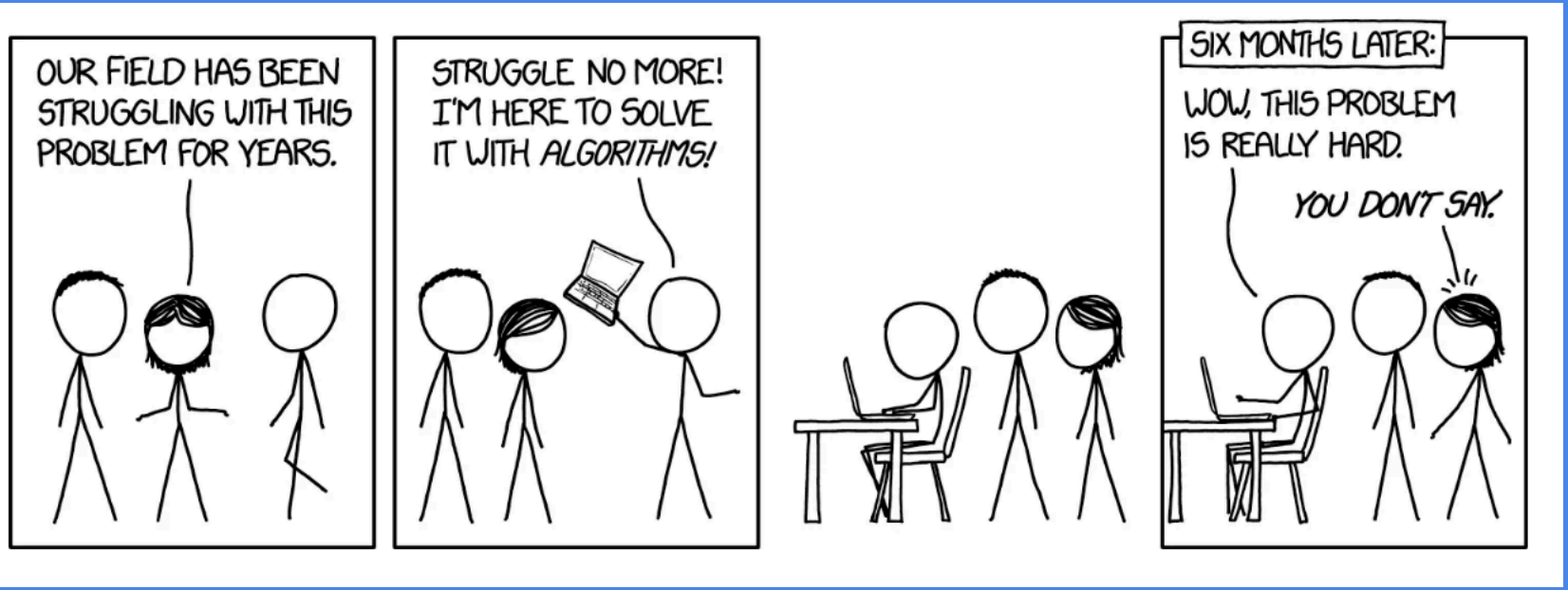


Non-parachuting ML

Discontinuity framing & Fermentation

… opportunity to start doing what you couldn’t do before!

It is hard work …



So, concluding again …

AI is a cultural challenge

(but do not underestimate the sci/tech complexity)

Lessons learnt and personal views

1. Enable cultural change
   & Discontinuity framing

2. Address Data and Engineering

… then you can identify and solve the interesting scientific questions

Cultural change is deceptively difficult

Cultural change is deceptively difficult

As per Kahneman's …

A bat and a ball cost $1.10
The bat costs $1.00 more than the ball
How much does the ball cost?

Cultural change is deceptively difficult

And remember this fact …

Companies are more likely
to survive a recession than
to adapt to a tech. discontinuity

THANKS!
QUESTIONS?

www.informaticslab.co.uk / alberto.arribas@informaticslab.co.uk