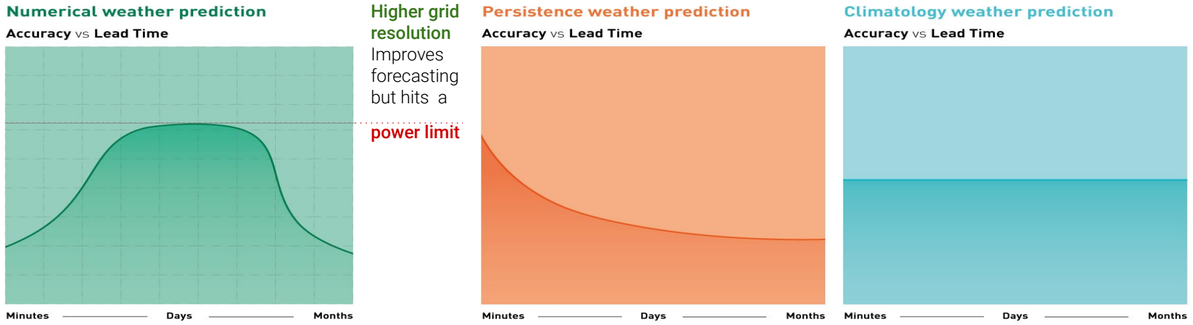
RainBench: Towards Global Precipitation Forecasting from Satellite Imagery

Christian Schroeder de Witt* - **Catherine Tong*** - Valentina Zantedeschi - Daniele De Martini
: Piotr Biliński - Matthew Chantry - Freddie Kalaitzis - Duncan Watson-Parris

Why do we need a digital twin for Earth?

Neural networks can **learn physics** from data.
Potential to **predict weather**, with less compute.



Numerical forecasting is best for the timescale of a couple of days.

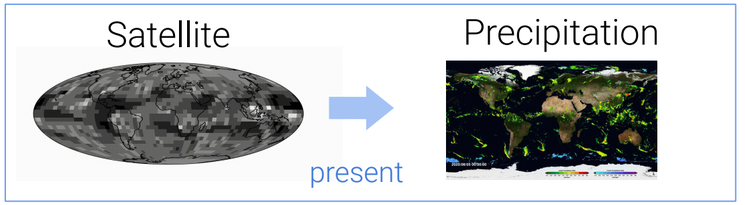
While other methods don't deliver accuracy...

Challenges

- Forecast representation
- Optimal time-window?
- Precipitation representation
- Sparsity & imbalance
- Global view to local predictions
- Useful variables?
- Land, sea, mountain behavior
- How to introduce prior knowledge
- Trusted & explainable
- Integration with existing tools
- Stochasticity representation

Predict precipitation up to days, **globally.**

Machine Learning for learning **data-driven weather forecasting models** with **lightning-fast inference.**



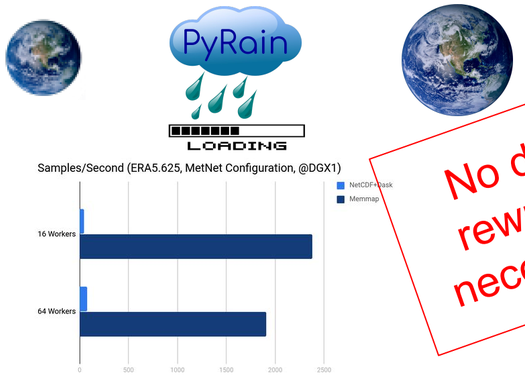
RainBench:
3 datasets
End-to-end
Neural baselines

Inputs	1-day	3-day	5-day
Persistence	0.6249	0.6460	0.6492
Climatology	0.4492 (1979-2017)		
Climatology (weekly)	**0.4447** (1979-2017)		
SimSat	0.4610	0.4678	0.4691
ERA	0.4562	0.4655	0.4677
SimSat + ERA	**0.4557**	**0.4655**	**0.4675**
ERA (1979-2017)	0.4485	0.4670	0.4699

(b) Predicting Precipitation from IMERG

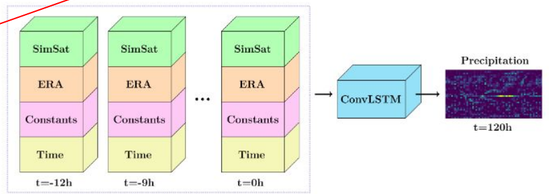
Inputs	1-day	3-day	5-day
Persistence	1.1321	1.1497	1.1518
Climatology	0.7696 (2000-2017)		
Climatology (weekly)	**0.7687** (2000-2017)		
SimSat	0.8166	0.8201	0.8198
ERA	0.8182	0.8224	0.8215
SimSat + ERA	**0.8134**	**0.8185**	**0.8185**
ERA (2000-2017)	0.8085	0.8194	0.8214

PyRain: Lightning-fast dataloading and processing



No data rewriting necessary!

https://github.com/FrontierDevelopmentLab/PyRain



First small steps to multi-day precipitation forecasts...

- Increase spatial resolution
- Physics-inspired models
- Class balancing
- Calibrate probabilistic predictions