Towards Data-Driven Physics-Informed Global Precipitation Forecasting from Satellite Imagery

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

FDL FRONTIER DEVELOPMENT LAB
esa Google Cloud NVIDIA · SCAN IBM AIRBUS planet. IWMI International Water Management Institute CATAPULT Satellite Applications UNIVERSITY OF OXFORD TRILLIUM ACCELERATED RESEARCH LAB

Data-driven Precipitation Forecasting

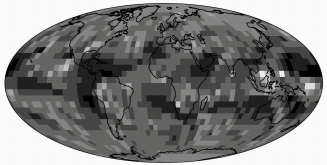
w.r.t. numerical models:

reduced data and compute requirements and inference time lag

We propose to forecast precipitation

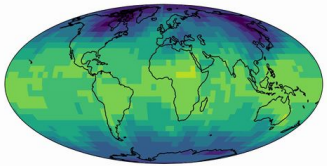
1. **globally**: to push the time horizon to **3+ days**
2. with a **probabilistic model**: to account for the stochasticity of the system
3. with implicit **physics** knowledge: to improve performance
4. from **Earth observations**: satellite imagery, ground measurements

Data from RainBench [1]



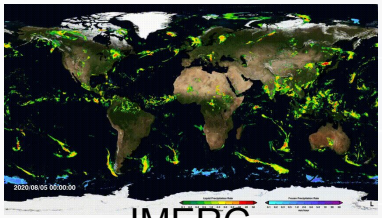
SimSat

- bands measuring global cloud cover and moisture features
- retrieved at resolution 5.625°, 3 hour frequency
- 2016 - present



ERA5

- estimated physical and atmospheric variables at different heights (e.g. humidity, temperature)
- retrieved at resolutions 5.625° and 0.25°, 3 hour frequency
- 1979 - present

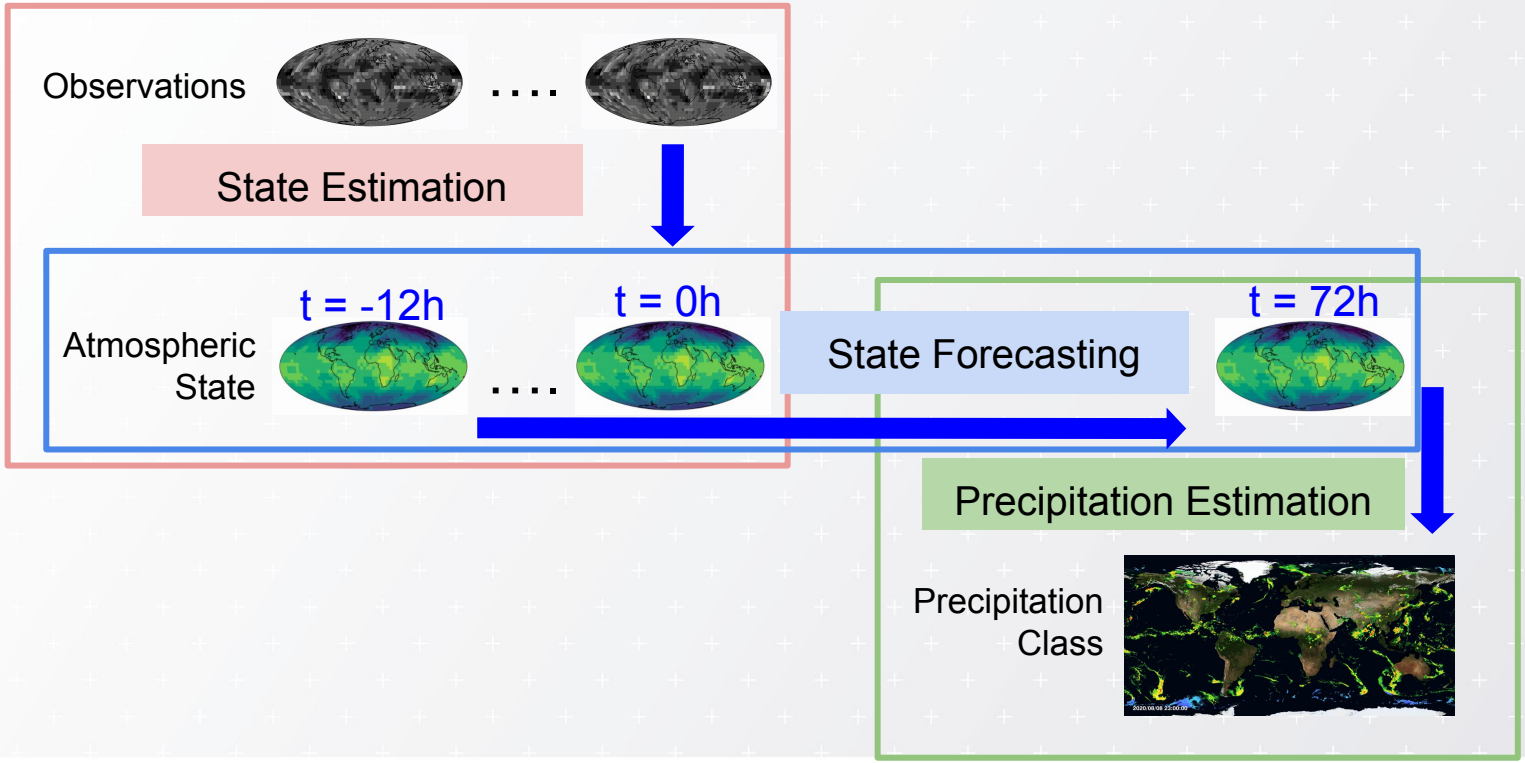


IMERG

- estimated precipitation intensity
- retrieved at resolution 0.25°, 3 hour frequency
- 2000 - present

[1] Schroeder de Witt, Tong, et al. "RainBench: Enabling Data-Driven Precipitation Forecasting on a Global Scale." *2020*

Three-Steps physics-informed and probabilistic approach

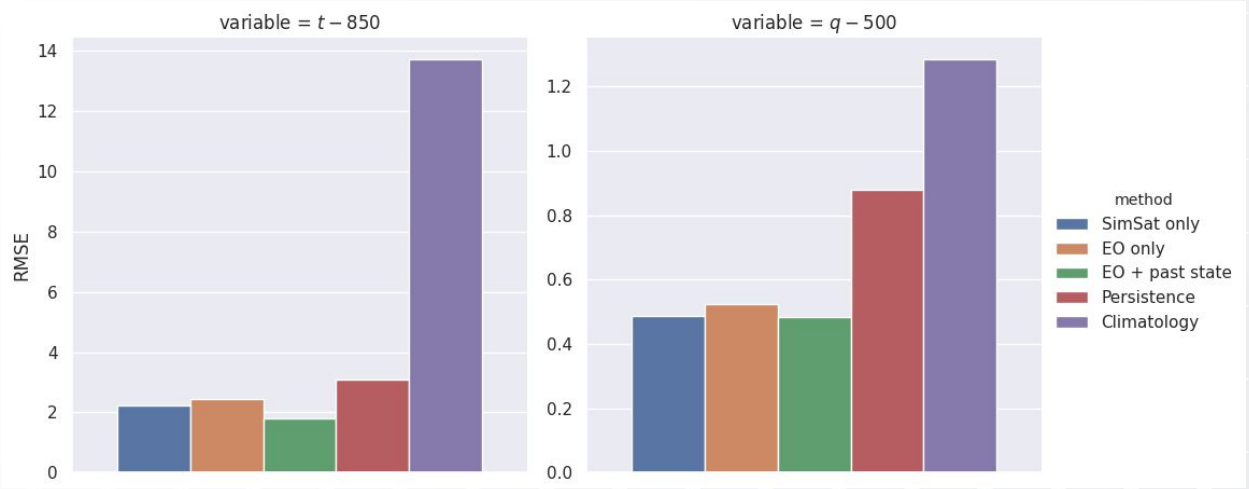


1. Present State Estimation

Model: Convolutional LSTM [1]

Input: stream of Earth Observations

Output: Atmospheric State (17 ERA5 variables)



Earth Observations up to t = 0h:

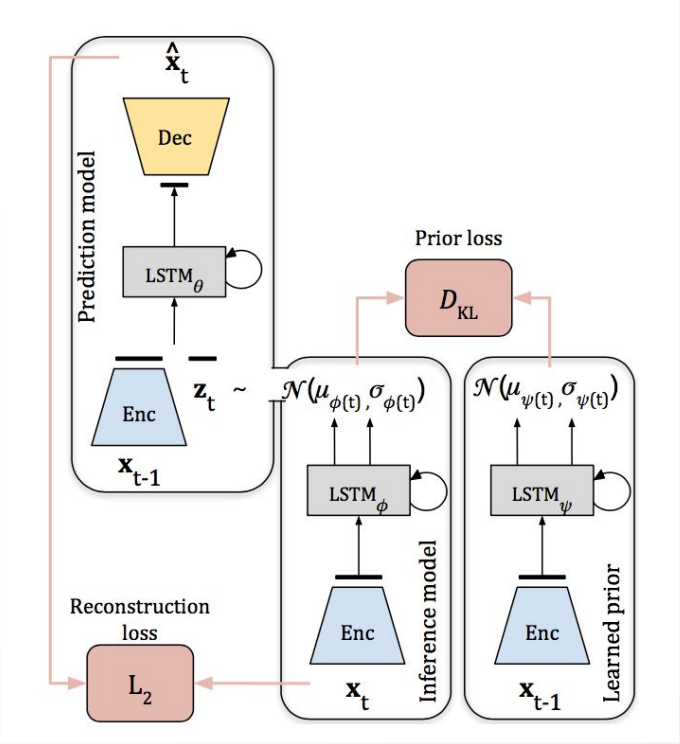
- SimSat
- Surface pressure
- Temperature 2m

Past State up to t = - 24h:

- ERA5 variables

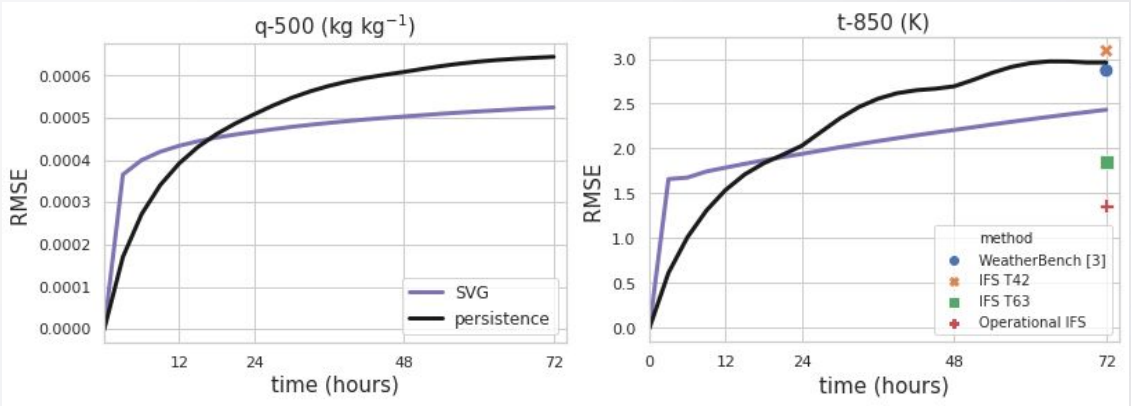
[1] Shi, Xingjian, et al. "Convolutional LSTM network: A machine learning approach for precipitation nowcasting." *NeurIPS 2015*

2. Future State Forecasting



Model: Stochastic Video Generation with Learned Prior [2] with Variational AutoEncoder and VGG Discriminator

Input-Output: Atmospheric State (17 ERA5 variables)



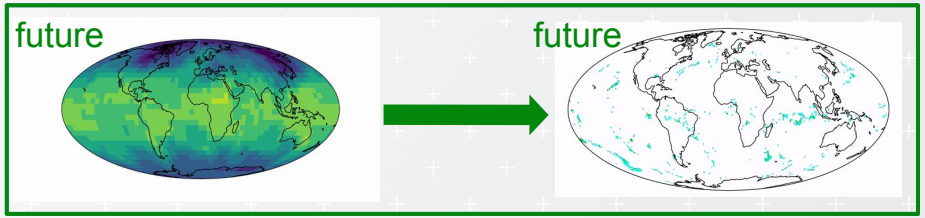
[2] Denton, Emily, and Rob Fergus. "Stochastic video generation with a learned prior." *JMLR 2018*
[3] Rasp, Stephan, et al. "WeatherBench: A benchmark dataset for data-driven weather forecasting." *2020.*

3. Precipitation Estimation

Model: Gridcell-wise Fully Connected Network

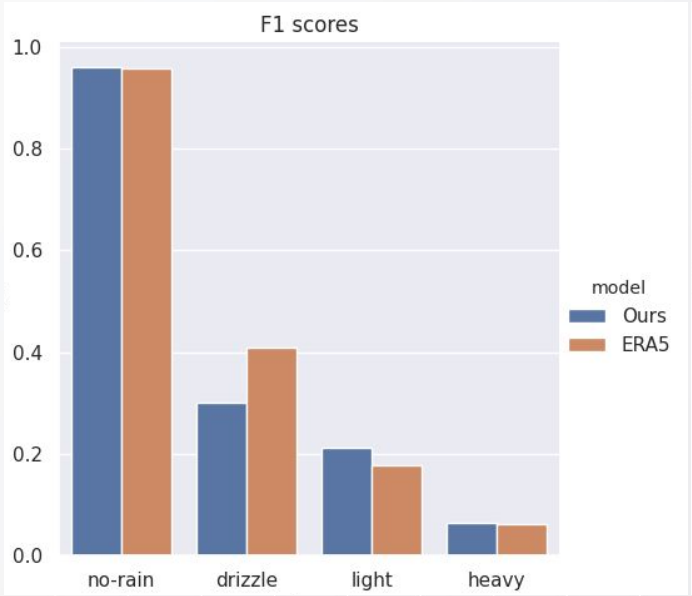
Input: Atmospheric State (17 ERA5 variables)

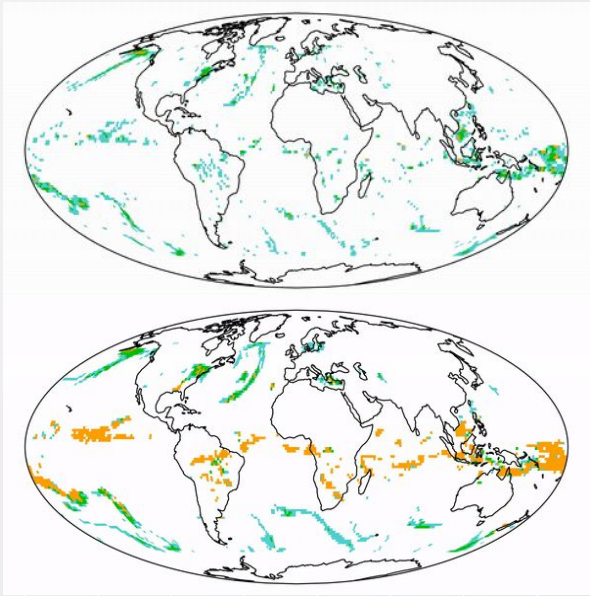
Output: Precipitation Class



Segmentation Task:
(mm over 3 hours)

No precip: < 1
Drizzle: [1, 7.5]
Light: [7.5, 22.8]
Heavy: > 22.8





Future work

1. End-to-end model
2. Calibrate model confidence
3. Regress precipitation intensity
4. Increase resolution

Thank you for your attention!