tomorrow.io

Machine Learning:
Shifting the Forecasting Paradigm

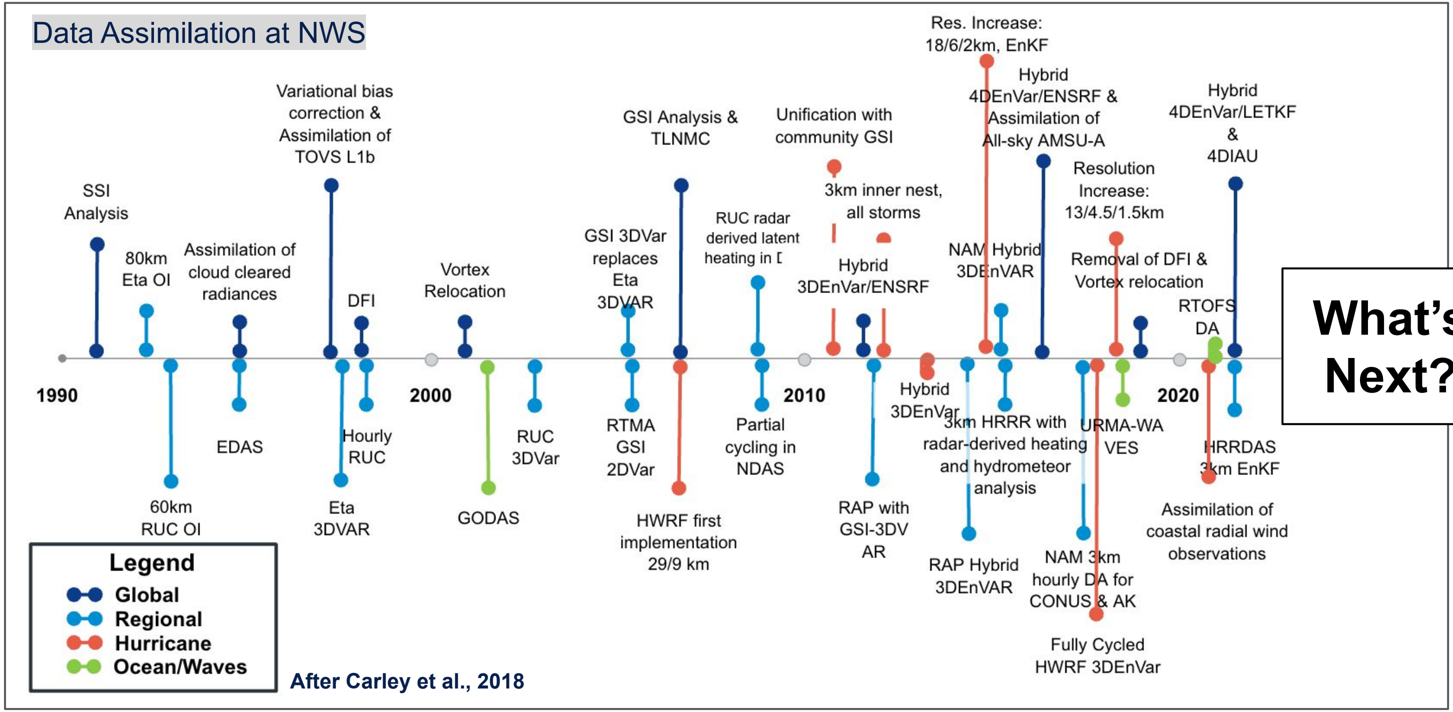
Stelios Flampouris - stylianos.flampouris@tomorrow.io

Steven Brey, Ashley Payne - Tomorrow.io

July 22, 2022

Introduction

Looking for the Next Breakthrough in Modeling and Model initialization…



What's Next?

Flampouris et al., 2020

tomorrow.io

Upgrading the NWP forecast and analysis systems

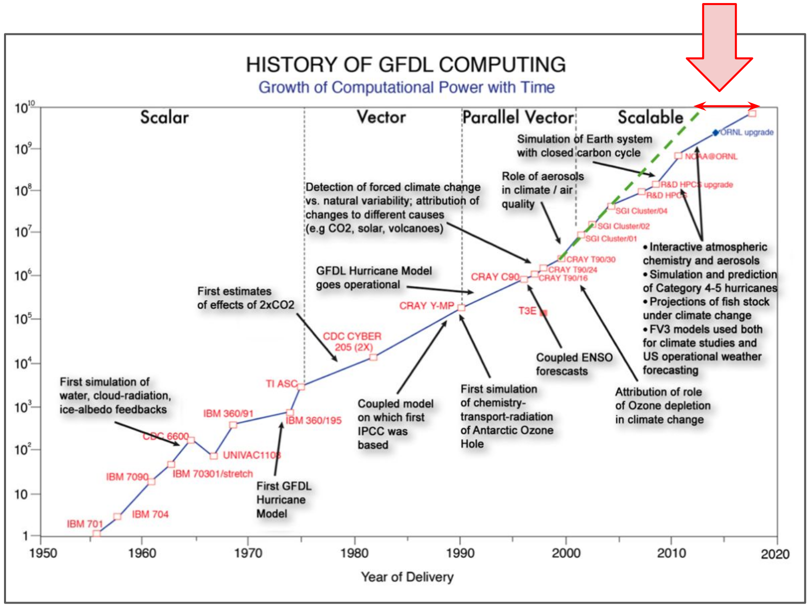
Constraints and Challenges

Innovating within the Modus Operandi

- Incremental scientific advancements
- Investment in the infrastructure
- Well-defined development path
- Operational Stability
- Optimization of cost

Overcoming ongoing Challenges to Innovation:

- Significant increase of data, model and observations (e.g., Tomorrow.io and others)
- Unresolved (coupled) physical processes
- Complexity of the models
- Prohibitive computational cost for the required resolution (Tolmann et al., 2022)



Balaji, 2020

Is Data Science (ML/AI) an Alternative/Supplemental Development Path?

Machine Learning and Earth Modeling

Machine learning: The study of computer algorithms that improve automatically through learning from data by using mathematics and the scientific process.

Why Now:

1. Unprecedented data volume
    - Our current analytical capabilities restrict the discovery

2. Daily maturation of the AI/ML capabilities/algorithms
    - Hundreds of applications are published daily

3. Evolution of hardware allows efficient ML models training
    - NVIDIA, AMD, Graphcore, Atmo, and more

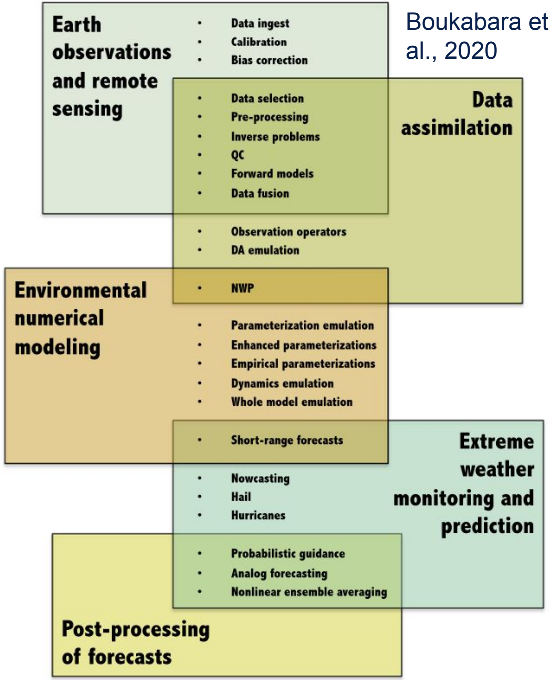
4. Availability of open-source, well-documented, extensively tested software frameworks
    - TensorFlow, PyTorch, Keras, Scikit-Learn and more

5. The cost of running ML models is orders of magnitude lower than for NWP models

Driven by the (Weather) Industry, the ML solutions have been adopted!

Adoption of ML/AI from Governmental Institutions

NOAA / NWS (Boukabara et al., 2020)



Boukabara et al., 2020

ESA/ECMWF (Schneider et al., 2022)

Standardization of ML applications within the NWP:

1. Enhancing Satellite Observation with ML
   a. Earth monitoring, biomass and volcanic plumes
   b. Radar backscatter and optical images
2. Hybrid Data Assimilation - ML approaches
   a. Approximation of nonlinear systems and extracting meaningful features from high-dimensional data
   b. Replacement of physically based model
   c. Application error corrections
3. Geophysical Forecasting with ML and Hybrid Models
   a. Speed up complex and time-consuming processing
   b. Diagnostics
   c. Model improvements
4. ML for Post-Processing and Dissemination
   a. Post-processing and optimization forecast outputs
   b. Downscaling

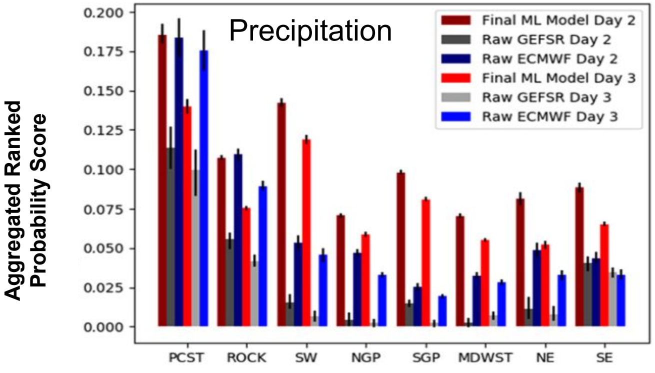
Intention: Integration of ML in the current forecast paradigm.

ML-Based Innovation in the Weather Industry

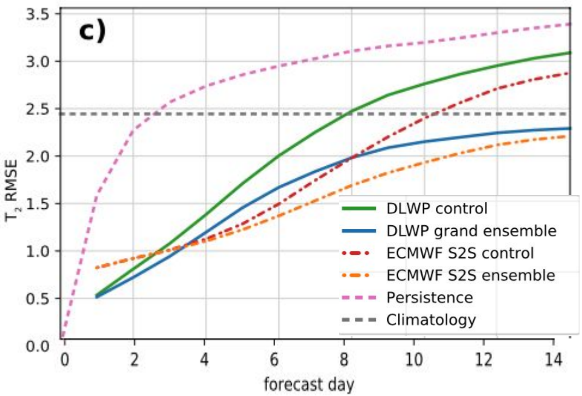
Broad spectrum of ML applications

Schneider et al., 2022: ESA-ECMWF Report on recent progress and research directions in machine learning for Earth System observation and prediction.

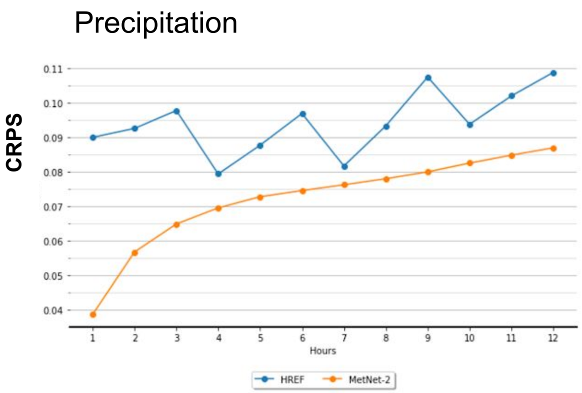
- *Herman & Schumacher, 2018 - Extreme **precipitation** forecasts with Random Forests
- Rasp & Lerch, 2018 - Neural network post processing of temperature
- Brey & Eckel, 2020, Dai and Hemri 2021- **Ensemble** ML-Prediction for Cloud Cover
- *Weyn et al., 2021- **NWP-free ML-based** weather forecast
- *Sønderby et al., 2020 - Neural-net based precipitation nowcasting



Herman & Schumacher, 2018



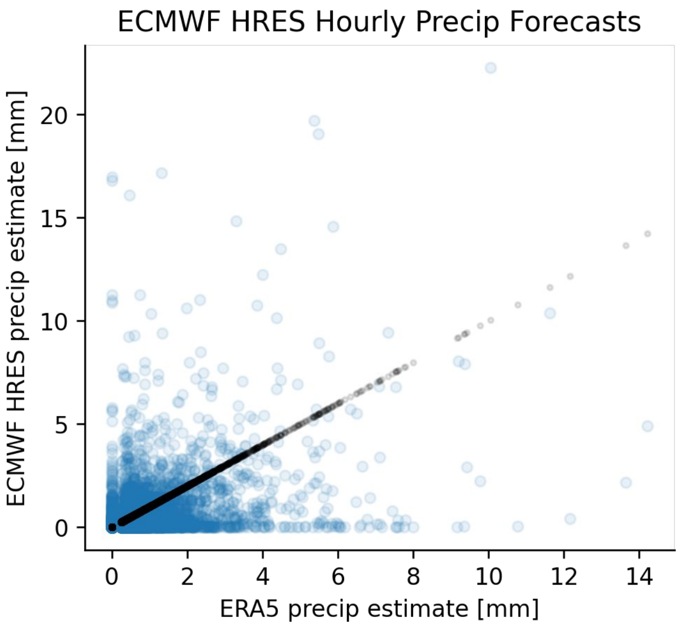
Weyn et al., 2021

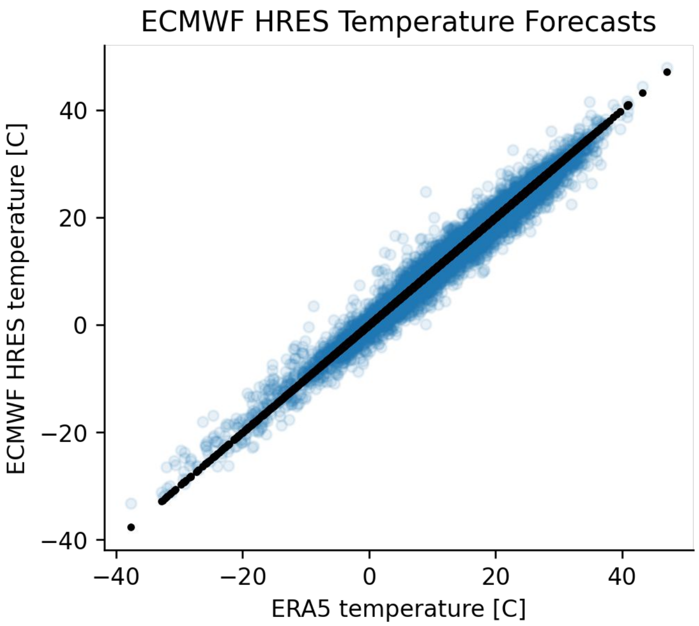


Sønderby et al., 2020

ML in post-processing at Tomorrow.io

Seamlessly merging multiple weather models into one **intelligent** and **accurate** forecast



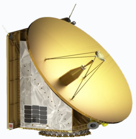


Focusing on the grand challenges of forecast, e.g., Precipitation Prediction

ML in Action at Tomorrow.io

ML is used to learn and correct the errors of traditional physics-driven forecasts

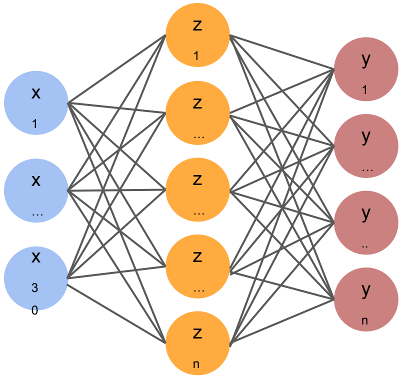
Input Predictors



Tomorrow.io NWP & Obs

ECMWF

Any physically relevant observation



The exact model, configuration, features, training strategy, *etc.*, vary prediction-to-prediction.

Predictions

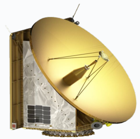
Deterministic or Ensemble:

Precipitation
Temperature
RH%
…
Any desired prediction

ML in Action at Tomorrow.io

ML is used to learn and correct the errors of traditional physics-driven forecasts and can provide uncertainty

Input Predictors



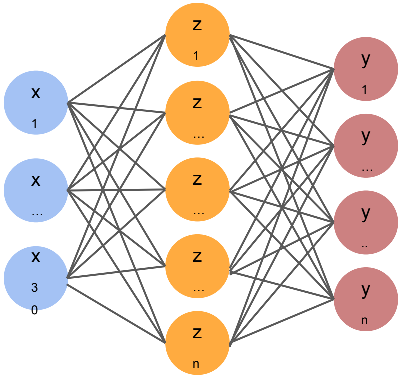
CBAM

Tomorrow.io NWP & Obs

ECMWF

NOAA

Any physically relevant observation



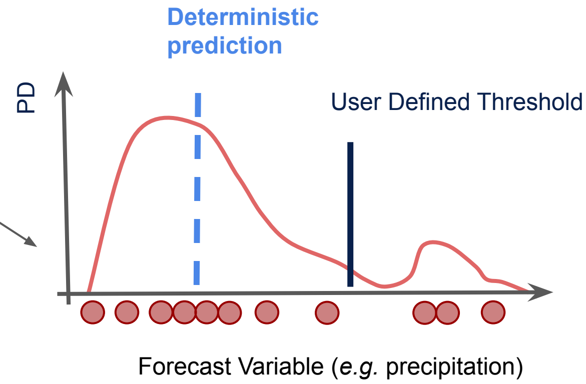
CRPS-Net based model which makes n equally likely predictions and the CRPS as its loss function.

$$CRPS(\hat{Y}, y) = E_F|\hat{Y} - y| - 1/2 \times E_F|\hat{Y} - \hat{Y}'|$$

Predictions

Ensemble:

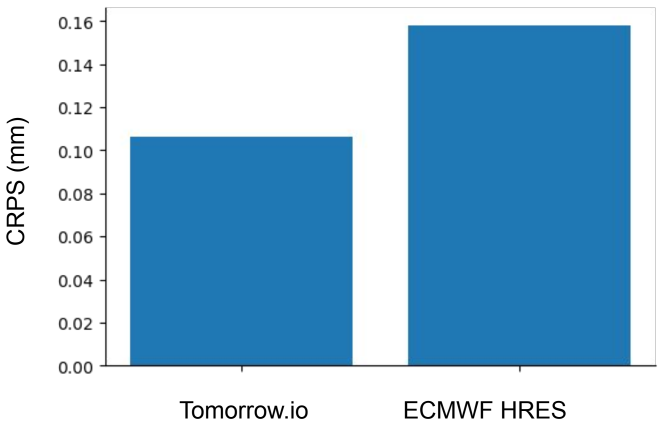
Precipitation
Temperature
RH%
…
Any desired prediction



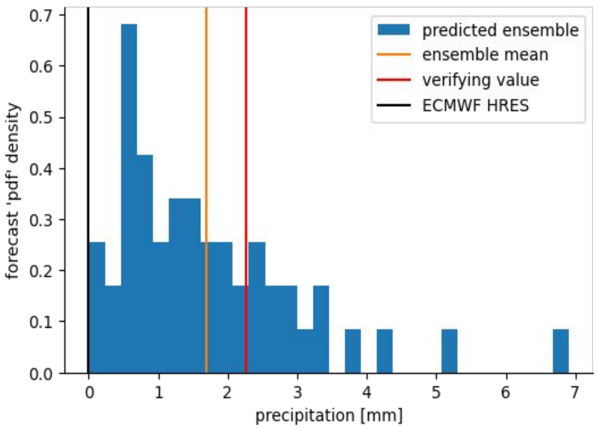
Brey & Eckel, 2020

Forecasting Precipitation

CONUS Precipitation Verification 1yr



51-Member Ensemble



Summary

- Data driven models have numerous applications in Weather Industry, for instanc Tomorrow.io has a suite of applications (DA, Nowcasting, Postprocessing)

- ML models are not the solution to all the problems

- The adoption of the ML approaches is changing the NWP: From modeling physical processes and initialization to data driven models and their training

- The typical development cycle, R2O2R, for NWP is extremely long for the Machine Learning common practices → Operational Innovation

- Considering the increasing number of observations and preliminary results, if and when purely ML weather predictions based on observations will be possible?

Thank you!