# The Stunning Success Story of Machine Learning for Weather Prediction

Prof. Dr. Thomas Ludwig
German Climate Computing Center (DKRZ)
University of Hamburg, Department for Computer Science (UHH/FBI)

# Agenda

- History
- Weather Prediction 101
- Recent Integration of Machine Learning
- Google DeepMind´s GraphCast
- Status at ECMWF
- The ICON Model

# History

- Goal: predict conditions of the atmosphere for a given location and time
- People attempted informally for millennia and formally since the 19$^{th}$ century
- Method: collect quantitative data about the current state of the atmosphere, land, and ocean …
- … and use meteorology to project changes at a location
- First time published: The Times, on August 1, 1861
  - Weather maps followed in the same year
  - First transmission via telegraph a bit earlier

# Lewis Fry Richardson (1881-1953)

- Vilhelm Bjerknes (1862-1951) asks for a numerical weather prediction in 1904 as the weather formation follows laws of physics
  - Can not specify a practical way to do this
- In 1922 **Lewis Fry Richardson** publishes his book *Weather Prediction by Numerical Process*

Lewis Fry Richardson

# Richardson‘s Vision: The First Weather Prediction

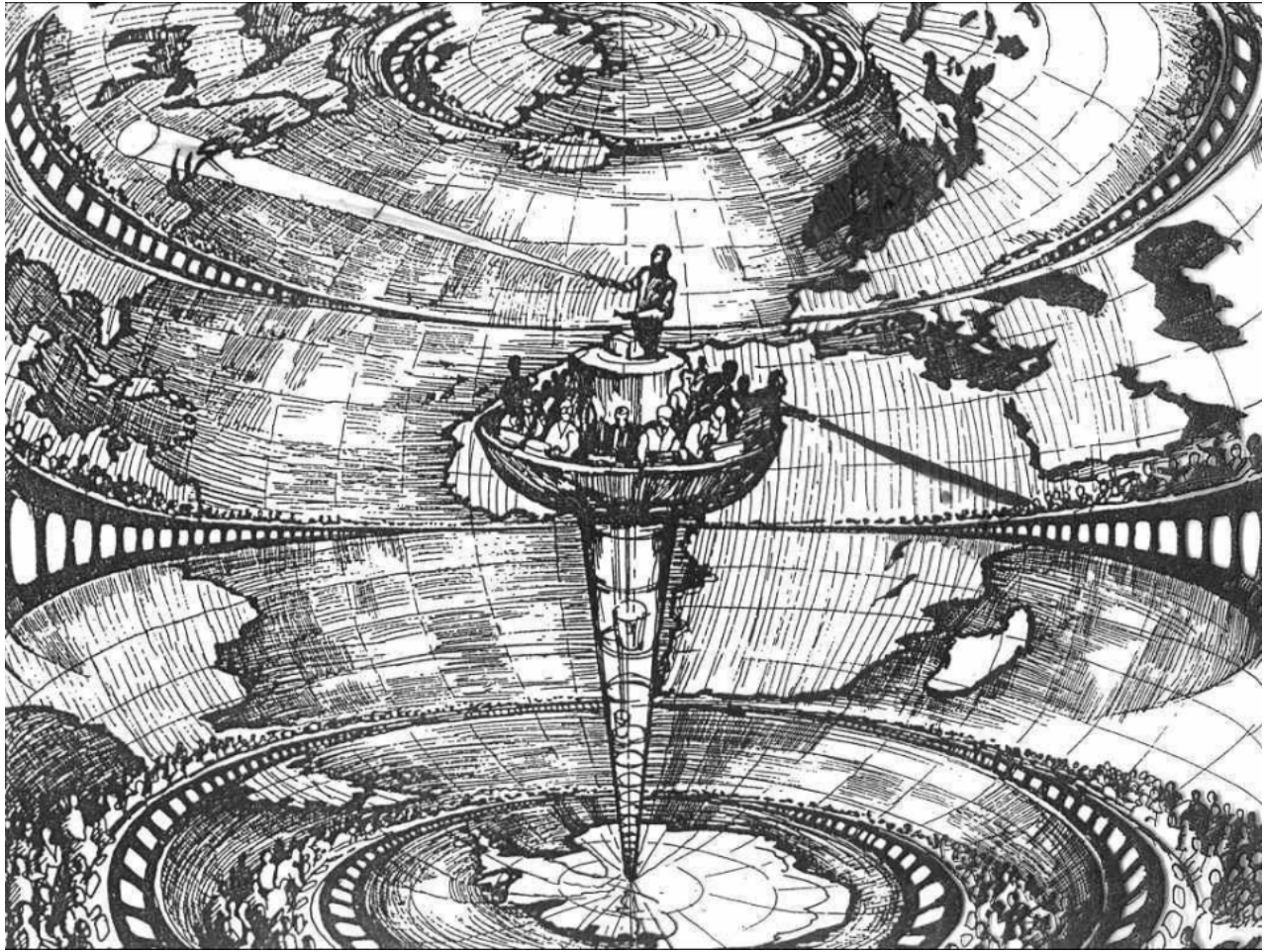
*Perhaps some day in the dim future it will be possible to advance computations faster than the weather advances and at a cost less than the saving to mankind due to the information gained*

[Richardson, WPNP]

Weather forecasting is a part of the economy. For example in 2009, the US spent approximately $5.8 billion on it, producing benefits estimated at six times as much.

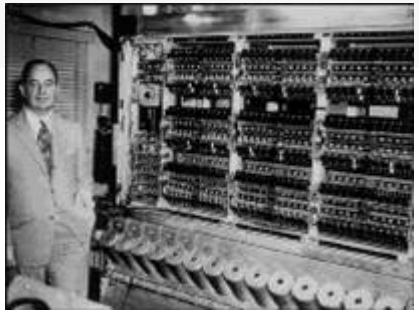
[Wikipedia “Weather Forecasting” [1]]

# Richardson´s Forecast Factory



- 64,000 human computer
- kind of message passing
- also load balancing

# History of Numerical Weather Predicition

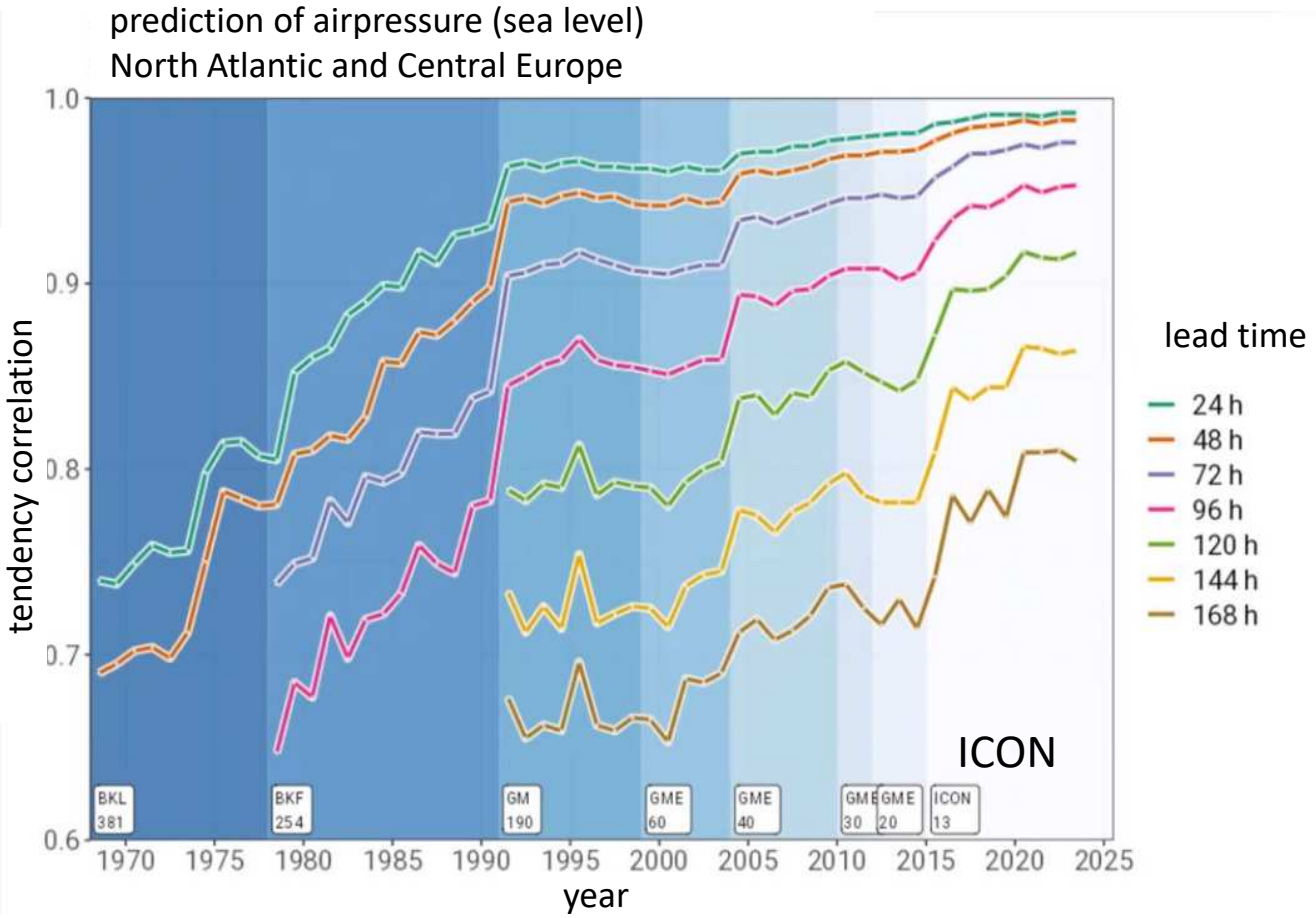


- First weather prediction on ENIAC in 1950
  - Team included Jule Charney, John von Neumann (mathematician), Klara Dan von Neumann (programmer) and others
  - First prediction took them 24 hours to produce and was for the next 24 hours.
    - Most work were manual operations, though
  - ENIAC was at Univ. of Pennsylvania, had about 500 FLOPS

# From 1950 to 2019

- Better computers, of course
- More investment in HPC for weather (and climate)
  - E.g. Met Office (UK) announced deal with Microsoft in FEB 2020: £1.2B (about €1.4B) for 10 years
  - “It is expected that the investment will result in financial benefits totalling up to £13 billion (or £9:1) for the UK over its 10-year lifespan.” (press release)
- Global forecast models, global climate models
- High resolution regional models
- Tropical cyclone models, ocean models
- Ensemble forecast: set of forecasts for one situation
  - evaluate uncertainty aspects

# Improvements over Time (One Example)




- y-axis: quality measure
- x-axis: year
- plots for different lead times (i.e. predicted intervals)

[data by Roland Potthast (DWD)]

# Peter Bauer´s Vision (ECMWF) (ISC 2019)

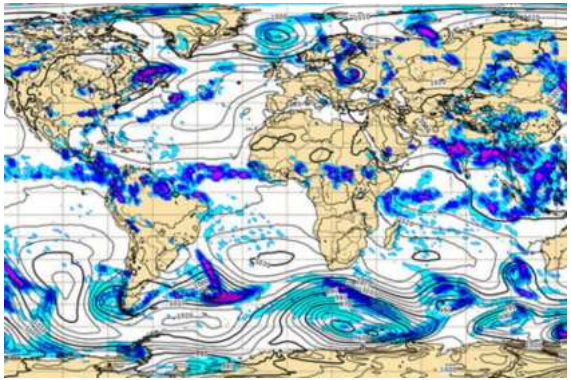
Title of the talk: *“Exascale Systems Present a Vision for Weather and Climate Prediction - Can we Meet the Challenges?”*

*“While artificial intelligence methods cannot overcome the main bottlenecks of efficient computing they can help alleviate algorithmic cost and support information extraction from both observational and simulated data.”*

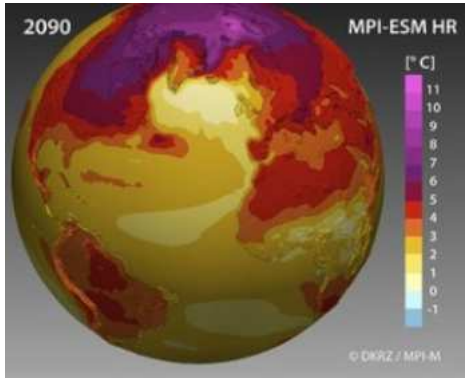
# Forecasting 101 [5]

ML-based success stories

main DKRZ users



also ML-based success stories




minute — hour — day — week — month — year

nowcasting

medium-range weather forecasting

sub-seasonal forecasting

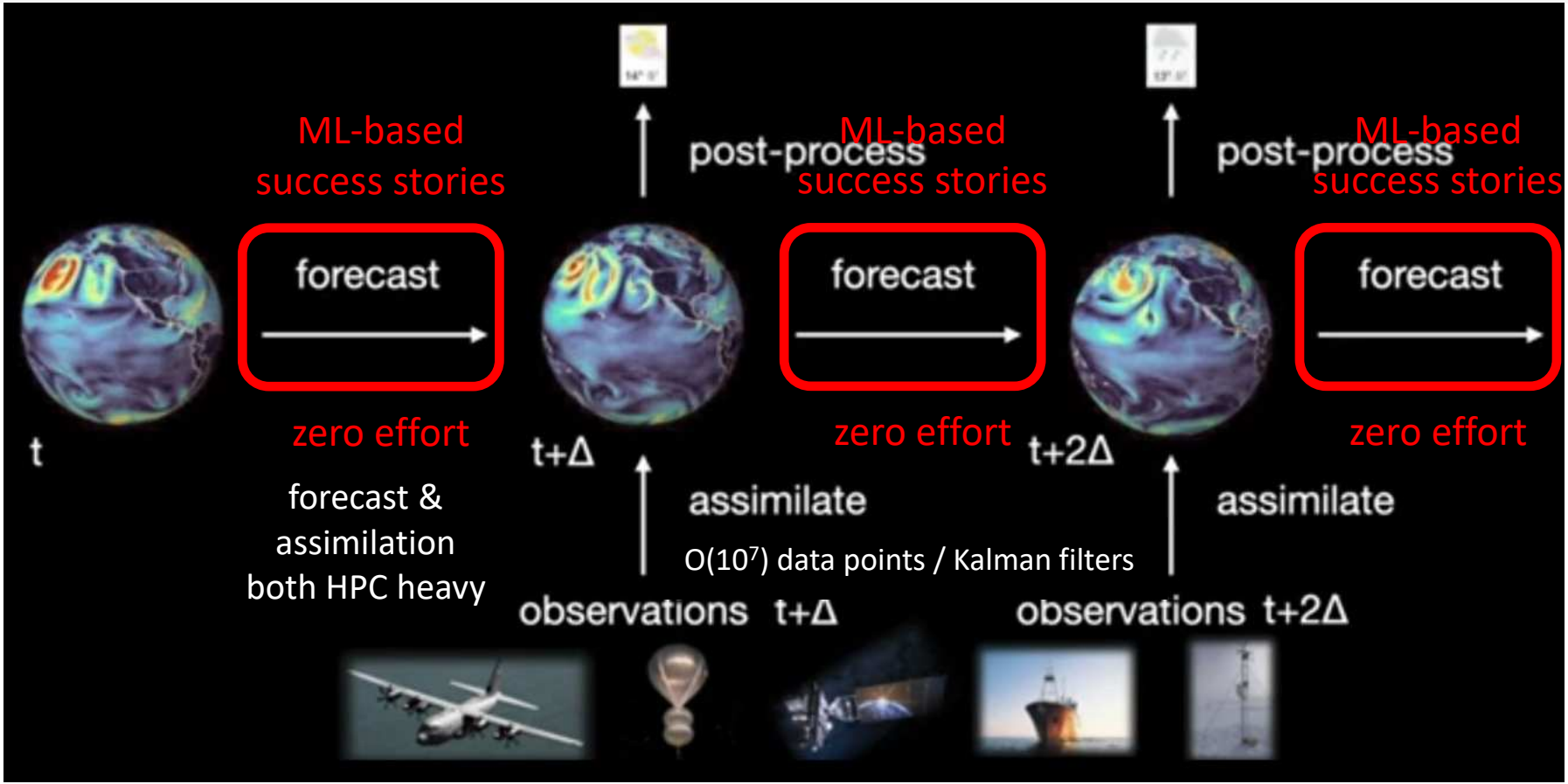
seasonal & climate

deterministic local atmosphere-driven dynamics

global

stochastic ocean-driven dynamics

# Medium-Range Weather Forecasting Process




© Richard Turner [5]

# Integration of Machine Learning into Forecasting

- Wikipedia “History of numerical weather prediction”
  - Updated until recently
    - no mentioning of ML concepts (March 24)
    - now extended and ML is mentioned (December 24)
- DWD: started in 2018
- ECMWF: started in 2018
- Quick adoption of ML methods
- Three excellent results in 2023

# Three Breakthroughs in 2023

June: Nvidia

FourCastNet

FourCastNet: Accelerating Global High-Resolution Weather Forecasting Using Adaptive Fourier Neural Operators

**Authors:** Thorsten Kurth, Shashank Subramanian, Peter Harrington, Jaideep Pathak, Morteza Mardani, David Hall, Andrea Miele, Karthik Kashinath, Anima Anandkumar Authors Info & Claims

July: Huawei

Pangu

**Accurate medium-range global weather forecasting with 3D neural networks**

Kaifeng Bi, Lingxi Xie, Hengheng Zhang, Xin Chen, Xiaotao Gu & Qi Tian

Nov: Google DeepMind

GraphCast

**Learning skillful medium-range global weather forecasting**

REMI LAM, ALVARO SANCHEZ-GONZALEZ, MATTHEW WILLSON, PETER WIRNSBERGER, MEIRE FORTUNATO, FERRAN ALET, SUMAN RAVURI, TIMO EWALDS, ZACH EATON-ROSEN, [...], AND PETER BATTAGLIA +8 authors Authors Info & Affiliations

# Foundation for ML Success: Algorithms and Data

- **Algorithms**: transformers and graph neural networks
  - FourCastNet: transformer based
  - Pangu: transformer based
  - GraphCast: graph neural network

- **Data**: weather and climate science hold vast volumes of well curated data
  - E.g. so-called ERA data from ECMWF

# Data: ERA5

- ECMWF produces reanalysis data from observational data by using mathematical methods
  - Ground sensors and satellite data
  - But also aircrafts, balloons, ships, buoyes, etc.
- ERA5 spans 1979 to present with a global 31km grid on 137 levels with a 1 hour time interval
- Full dataset is about 1.5 PB
- Freely available for others
- Basis for all three ML-based methods
- Now also available at DKRZ [9]

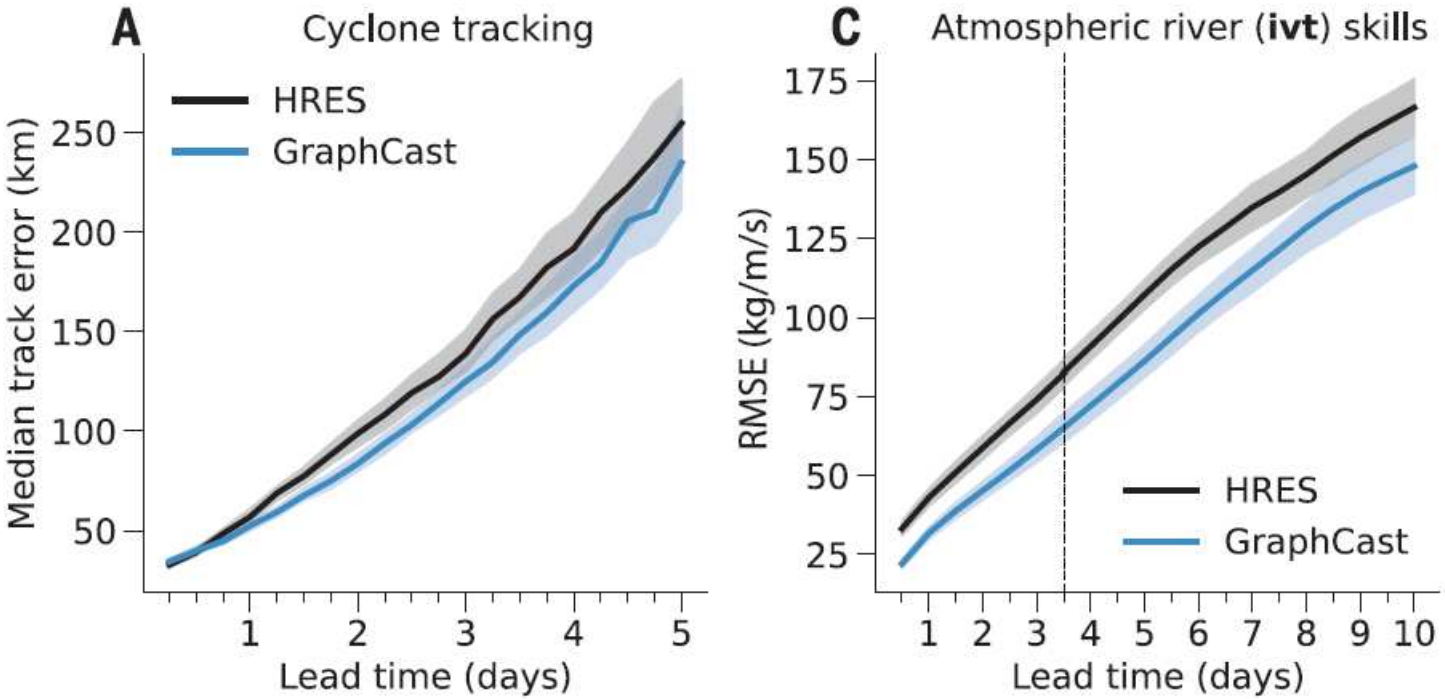
# Google DeepMind´s GraphCast [6]

- Graph neural network architecture
  - 36.7 million parameters, i.e. small
- Trained to minimize training objective using gradient descent
  - About 4 weeks on 32 Cloud TPU v4 devices with batch parallelism
- Forecast skill compared to ECMWF´s High-Resolution Forecast (HRES) for many variables, levels and lead times
  - Metrics: root mean square error (RSME) and anomaly correlation coefficient (ACC)
- Forecast needs PC not HPC ☺

# Results 1

- GraphCast outperformed HRES on 90.3% of the 1380 targets and significantly outperformed HRES on 89.9% of targets
- Also compared GraphCast's performance to the top competing ML-based weather model, Pangu-Weather and found that GraphCast outperformed it on 99.2% of the 252 targets presented
- ☺ See supplementary materials (132 pages)

# Results 2: Cyclone Tracking and Atmospheric Rivers




Better results than HRES although NOT specifically trained for extreme weather events

# Discussion: GraphCast and ML-based Forecast Technology

- Can it predict local events?
  - Was trained on 31km grid
- Can it predict seldom events?
  - Perhaps not enough information in ERA5
  - However, cyclone prediction is promising
- Many more issues …

spring 2024

- What will be the influence on DWD´s and ECMWF´s future HPC installations?
  - Challenge for scientists [7]
  - Relief for politicians: we will in future finance you a desktop computer

- DWD, ECMWF, and all others already develop own ML approaches

# More to Come… (Update November 2024)

Peter Dueben (ECMWF) at Supercomputing 24 [10]

Invited Talk „The Digital Revolution of Earth System Modelling“

General questions and their answers

- Can machine learning models avoid the smearing out for long predictions? Yes
- Can machine learning models learn uncertainties? Yes
- Can machine learning models represent extreme events? Yes
- Can machine learning models represent physical consistency? Yes
- **Can machine learning models do data assimilation? Almost**

Next step: Machine learning based climate simulations

# The Race is on… - Always Check the Date of an Article

- Google Deepmind´s GenCast (Dec 04, 2024)
  - For 15 days lead time better than ECMWF´s forecast
  - Needs only 1(!) 5th generation TPU from Google
    - Prediction takes around 8 minutes
  - Plan to provide realtime production mode for integration into other prognostic systems

- ECMWF AIFS (Feb 25, 2025)
  - Presents Artificial Intelligence Forecasting System (AIFS)
  - New operational model / provides data to public
  - 800 million observations processed on a daily basis
  - Outperforms many state-of-the-art physics-based models
  - Targets 50 members ensembles in future
  - Very low power consumption (1/1000x ?)

# AI/ML @ DKRZ for Climatology

- Department for AI/ML lead by Christopher Kadow
- Projects
  - Infilling of missing weather data in historic data sets [13]
  - Large language model-driven scientific assistant for climate research and data analysis [14]
  - GPT-4o and GPT4o-mini for climate data analysis [14]
  - …

# Short Announcement…

EXCLAIM Symposium

**“Is AI the Future of Weather and Climate Modeling?”**

ETH Zurich (in collaboration with MeteoSwiss and the Swiss National Supercomputing Centre (CSCS))

June 2-4, 2025 in Zurich, Switzerland.

Session A - Creating the foundations of AI-based modeling of weather and climate

Session B - Harnessing the power of AI in modeling weather and climate

Session C - Harnessing the power of physics-based modeling of weather and climate

Session D - Merging AI and physics-based modeling of weather and climate

# Back to ICON – the (still) Physics-Based Weather Prediction

- Leading model code for weather and climate
  - https://www.icon-model.org/ - now open source
- Developed by Deutscher Wetterdienst, Max-Planck-Institut für Meteorologie, Deutsches Klimarechenzentrum, Karlsruher Institut für Technology, Center for Climate System Modelling (Switzerland)
- Used by DWD for operational weather prediction

# Marbellous Cloud Computing 1972/2022


NASA Blue Marble photo

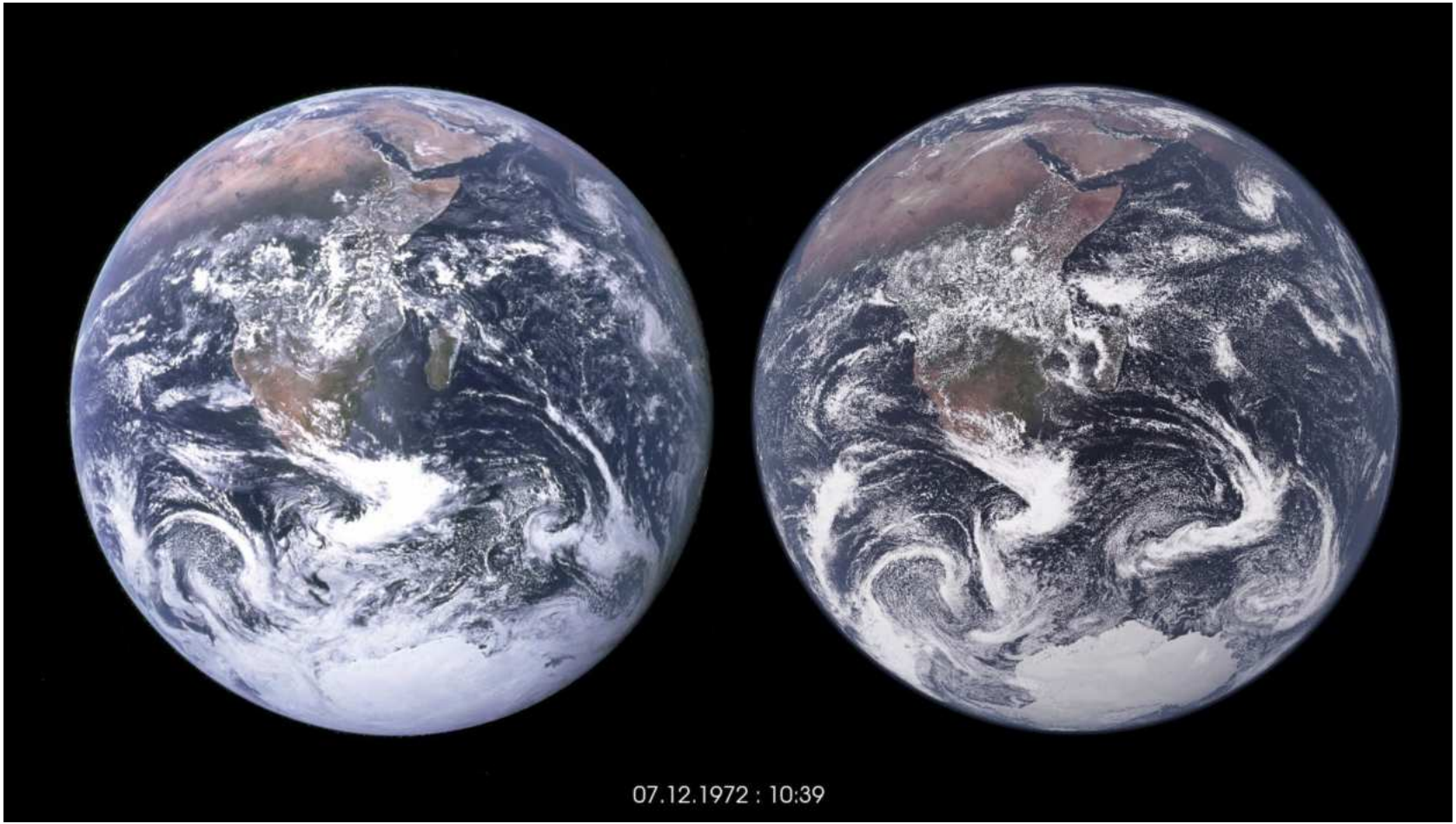
Dec, 7th, 1972 – 10:39 UTC

Apollo 17, last crewed lunar mission

# Blue Marble Recomputed

- Max Planck Institute for Meteorology succeeded in October 2022 to simulate a full global model on a kilometre scale
  - Two decades of development of the ICON model
- Decided to simulate the weather of Dec 7th, 1972
  - Together with NVIDIA and DKRZ
- Started the simulation 2 days before the photo had been taken, thus it is a two day forecast [8]
  - Ocean simulation was started 4 years earlier
  - Reanalysis data from Dec 5th, 1972 was used for atmosphere
  - 7,200 node hours on Levante per simulated day

# Blue Marble: Photo (left) and Recomputed (right)



# References


[1] Fostering Innovation, Creating Jobs, Driving Better Decisions: The Value of Government Data. Economics and Statistics Administration Office of the Chief Economist. July 2014. p. 15.

[2] https://en.wikipedia.org/wiki/History_of_numerical_weather_prediction

[3] https://www.ecmwf.int/en/about/media-centre/science-blog/2023/rise-machine-learning-weather-forecasting

[4] Blue Marble simulation https://www.youtube.com/watch?v=Nict8tBs4lY

[5] Richard Turner – The Quiet AI Revolution in Weather Forecasting https://www.youtube.com/watch?v=JGn18WH0d6s

[6] Learning skillful medium-range global weather forecasting https://www.science.org/doi/10.1126/science.adi2336 [GraphCast, 6 pages]

[7] https://www.sueddeutsche.de/wissen/wetter-ki-vorhersage-1.6342702

[8] https://mpimet.mpg.de/en/communication/detail-view-news-ii

[9] https://www.dkrz.de/de/kommunikation/aktuelles/era5-auf-levante

[10] The Digital Revolution of Earth System Modelling https://sc24.conference-program.com/presentation/?id=inv106&sess=sess493

[11] Weather forecasting takes big step forward... https://www.ft.com/content/5642ef4d-5f42-4987-9034-92a4500b807c?utm_source=chatgpt.com

[12] Ein Tausendstel des Strombedarfs... https://www.heise.de/news/Europaeischer-Wetterdienst-Neues-KI-Wettermodell-braucht-1000-Mal-weniger-Strom-10294362.html

[13] Artificial intelligence reconstructs… https://www.nature.com/articles/s41561-020-0582-5

[14] EGU25 contributions https://meetingorganizer.copernicus.org/EGU25/session/54186