Paving the Way for AI?

Development of HPC and NWP in the Last Decades

The DWD Perspective

Ulrich Schättler
Department for Numerical Modeling (FE1)
Deutscher Wetterdienst, Germany

And many colleagues from DWD!

Contents

History Lesson

NWP - AI

Skipping the ancient period...

Milestones of Forecasting: NWP

- 1904: Vilhelm Bjerknes defined a program for rational weather forecasting. He showed how the laws of physics could be used to develop a procedure for atmospheric prediction.

  (1904: Das Problem der Wettervorhersage, betrachtet vom Standpunkt der Mechanik und Physik. – Meteorol. Zeitschrift, 21,1–7.)

Horizontal Momentum $\frac{d\mathbf{V}^H}{dt} + \frac{1}{\rho}\nabla p + f\mathbf{k}\times\mathbf{V}^H = \mathbf{F}^H$

Vertical Momentum $\frac{dw}{dt} + \frac{1}{\rho}\frac{\partial p}{\partial z} = -g + F^z$

Continuity $\frac{d\ln\rho}{dt} + \nabla.\mathbf{V} = 0$

Thermodynamic $\frac{d\theta}{dt} = F^\theta$

Moisture $\frac{dq}{dt} = F^q$

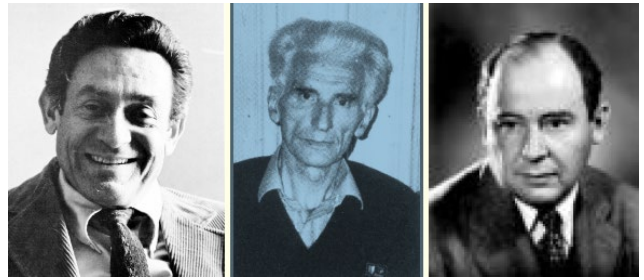
State $p = \rho RT \quad where \quad \theta = T\left(\frac{p_r}{p}\right)^k$

- 1916: Lewis Fry Richardson tried to apply Bjerknes' approach, but with discouraging results: the pressure changes forecasted were orders of magnitude higher than observed.

  (L.F. Richardson: „Weather Prediction by Numerical Process“, Cambridge University Press, London, 1922)

- 1950: John von Neumann et al. integrated a barotropic vorticity model on the ENIAC, one of the first multi-purpose electronic digital computers.

  (Charney, J.G., R. Fjørtoft and J. von Neumann: „Numerical Integration of the barotropic vorticity equation“, Tellus, 2, pp. 237-254)

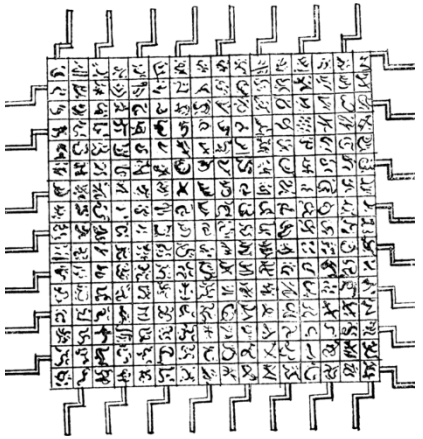


Highlights in the Development of AI

- 1726: Jonathan Swift described a wonderful machine, the *Engine*, located in the *Academy of Projectors* in the fictional city of *Lagado*:

  A frame filled with wooden squares, with words written on them. By turning handles, the words shift positions randomly, producing different combinations. Students then copy the meaningful arrangements and claim them as original work.“

  (1726: Gulliver‘s Travels, Part III, Visit to the Academy of Lagado)



„The Engine“
From: commons.wikimedia.org

- 1943: Warren McCulloch and Walter Pitts described networks of artificial neurons: mathematical functions serving as a model of a biological neuron. It is the elementary unit of an artificial neural network.

  (1943: A logical calculus of the ideas immanent in nervous activity, in: Bulletin of Mathematical Biophysics, 5(4), pp. 115-133, doi: 10.1007/BF02478259)

- 1957: Frank Rosenblatt developed the Perceptron: A system with three components: Sensory System, Association System, Response System. This can be considered as an early neural network

  (1957, The Perceptron – A Perceiving and Recognizing Automaton, Cornell Aeronautical Laboratory, Inc., Report No. 85-460-1)

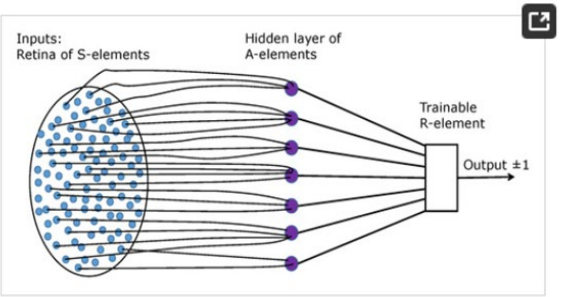


Figure 1. Rosenblatt's elementary perceptron (re-drawn from the Rosenblatt book [1]).

From the 50ies to the 90ies and Beyond

1950

Early Enthusiasm (until mid 70ies)

- Perceptron as an early neural network
- early chatbot ELIZA

1970

First AI Winter (until mid 80ies)

- AI hit performance and funding limits.
- Perceptrons were proven incapable of solving simple problems (like XOR).

Expert Systems (in the 80ies)

- popular in business, but hard to maintain

1990

Comeback of neural networks / machine learning

- But still could not work on large interesting problems.

With computational power available in the 90ies, neural networks could not demonstrate their real potential.

Quasi-geostrophic (filtered) equations (50ies)

- Baroclinic models, typically single-level.

Primitive equations (from early 60ies):

- Initially hemispheric models with hydrostatic assumption.
- Included physical processes in the late 60ies.
- Global models started in the 70ies.
- Special European highlight: ECMWF: 1st August 1979

Fully compressible, non-hydrostatic equations (90ies)

- Initially implemented in limited-area models.
- 1999: COSMO-Model by DWD and COSMO consortium.

For some decades NWP models have been able to utilize the maximum available computing power by increasing model complexity and resolution. This still holds for climate models.

NWP and HPC at the Deutscher Wetterdienst

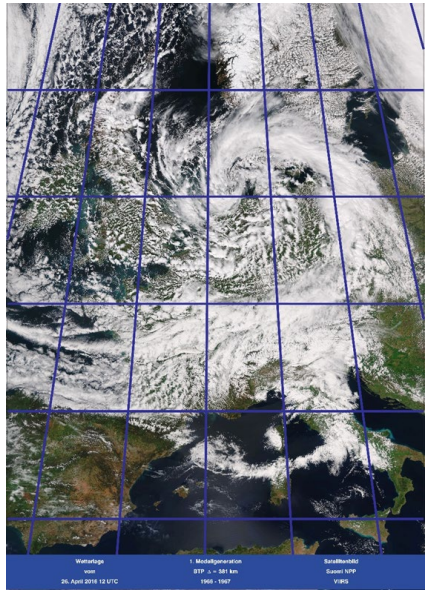
1966: Deutsches Meteorologisches Rechenzentrum im DWD (© DWD)

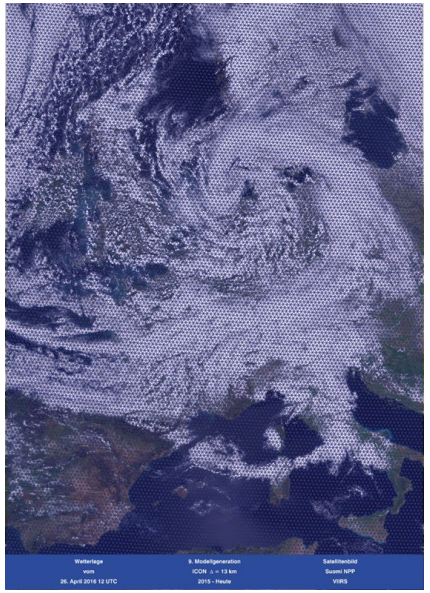
3 October 1966

Start of operational forecasting on a CDC3400 with a barotropic single-level model: **BTP**

NWP Generations of Global Models 1966-2025

Year	Model	Δ /Levels	
1966	BTP	381 / 1	barotropic
1967	BKL	381 / 5	baroclinic, hemispheric model, dry
1978	BKF	254 / 9	barocl., hemispheric, moist
1991	GM	190 / 19	global spectral model (from ECMWF)
1999	GME	60 / 31	global icosahedral / triangular grid
2004	GME	40 / 40	
2010	GME	30 / 60	
2012	GME	20 / 60	
2015	ICON	13 / 90	non-hydrostatic; developed with MPI-M
2022	ICON	13 /120	





© Bodo Ritter, Detlev Majewski

3-4 grid points for Germany
BTP, Δ = 381 km, 1 layer

~2000 grid points for Germany
ICON, Δ = 13 km, 120 layers

Since 3rd September 2025 supplemented by AI forecast: AICON

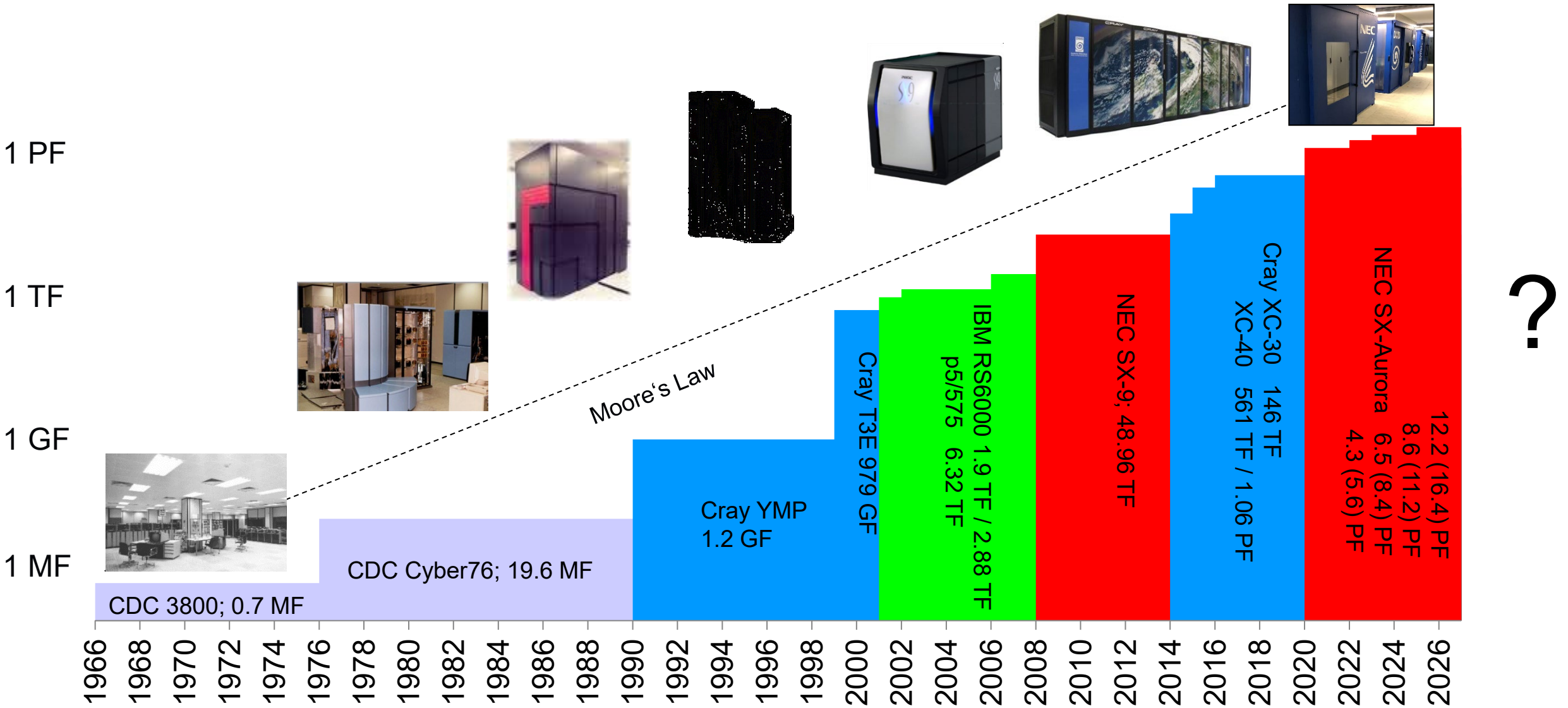
Based on ICON-DREAM Re-Analysis (13 km)

NWP Generation of Regional Models 1991-2025

Year	Model	Δ	Area	
1991	Europa	55	Europe	hydrostatic
1993	Deutschland	14	Germany	in addition to EM
1999	COSMO	7	Central Europe	non-hydrostatic
2002	COSMO-EU	7	Europe	
2007	COSMO-DE	2.8	Germany	in addition to COSMO-EU
2011	COSMO-DE-EPS	2.8	Germany	convection permitting EPS
2019	COSMO-D2(-EPS)	2.2	Central Europe	
2021	ICON-D2(-EPS)	2.1	Central Europe	
2024	ICON-RUC	2.1	Central Europe	ICON-D2: 24 forecasts per day
2025	ICON-D05	0.5	Central Europe	ICON-D2 with nesting to 500 m

The NWP generations are also characterized by increasing complexity of the physical parameterizations, numerical methods and software design.

DWD Computers 1966-2026



Hardware and Programming Models: Some (DWD) Facts

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Computer	Proc. Unit	Programming Model	NWP Model
CDC	special	Fortran (?)	BTP, BKL
YMP	Vector	Fortran + shared memory	GM, EM, DM
T3E	DEC Alpha	Fortran/C + MPI	GME, COSMO
SP4/5	Power	Fortran/C + MPI (+OpenMP)	
SX-9	Vector	Fortran/C + MPI	
XC30/40	x86	Fortran/C + MPI + OpenMP	ICON, COSMO
SX-Aurora	Vector	Fortran/C + MPI (+OpenMP)	ICON

- DWD models after 1990 always were written in vectorized way.
- No special implementations / optimizations for scalar processors.
- Major change for distributed memory parallel computers: implementation of message passing (MPI).

- OpenMP was never implemented in GME and COSMO. Tests with COSMO showed that it is not on all computers beneficial. But (to be honest) we also never had enough staff resources and time to implement it.
- It was implemented in ICON (together with MPI-M, DKRZ), but still not beneficial on all platforms.
- GPU support is implemented in COSMO and ICON, strongly pushed by our colleagues from MeteoSwiss (using OpenACC). But this is not yet used at DWD.

Future Programming Models will sure look rather different!

The Present Situation

HPC @ DWD: Current Systems

Vector based systems: NEC SX-Aurora TSUBASA

- ➔ Warm water cooling (35°C), free cooling possible ~350 days/yr
- ➔ 2 generations of vector processors: Aurora 10 (8 cores; Phase 0-2a) and 30B (16 cores; added for Phase 2b/3)

2 systems at 2 redundant locations

Phase	Operational		Research		Available
	Cores	Size EPS	Cores	Size EPS	
0	11392	49	14848	64	12/2019
1	14336	161	18688	210	12/2020
2a	21696	244	28160	316	09/2022
2b	28352	324	36864	421	09/2023
3	41920	484	54400	631	01/2025

Size EPS:
ensemble members to check performance upgrade factor.
Committed factors are:

Phase 0:	1
Phase 1:	~3.3
Phase 2b:	~6.6
Phase 3:	~10

Small GPU Partition: 4 nodes with

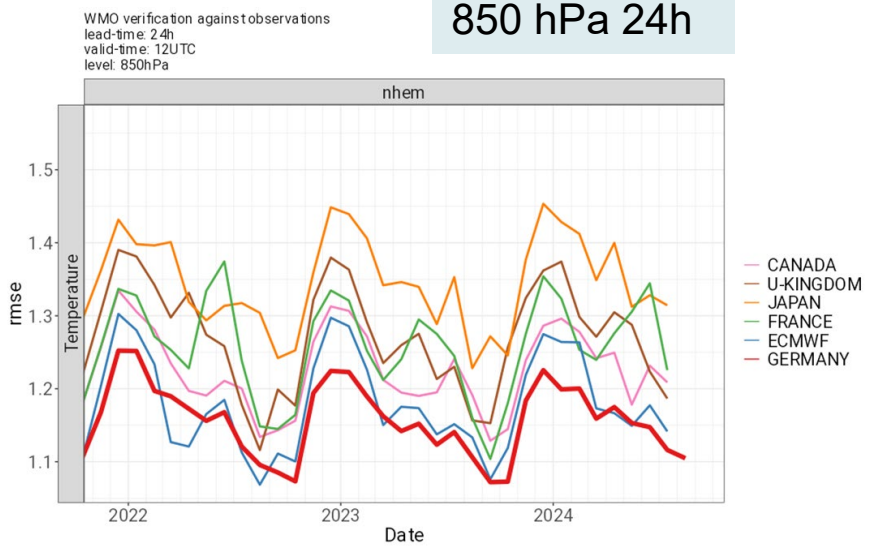
- ➔ 2x AMD EPYC, 64 cores
- ➔ 8x Nvidia A100 SXM4, 80 GB HBM2e

Top 500 Ranking

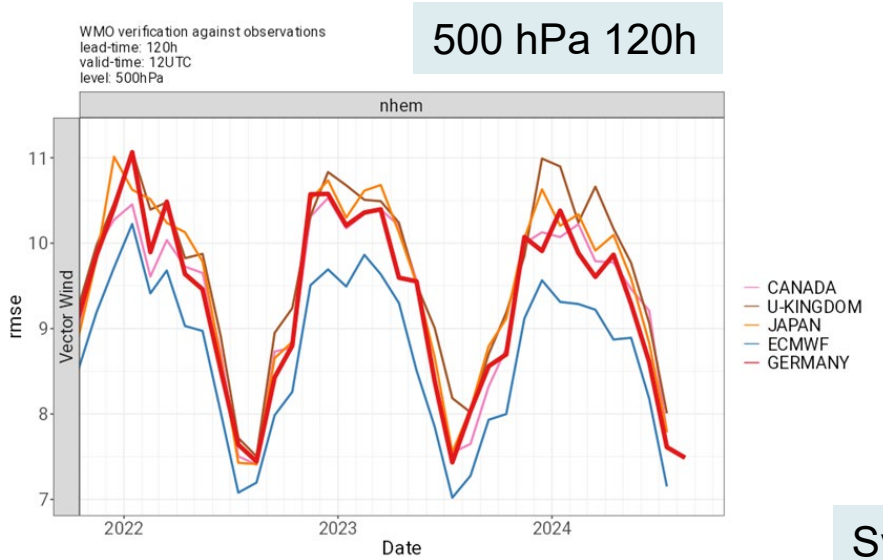
Rank in Top 500	Rank in Green 500	Rank in HPCG	#cores	Rmax (PFlop/s)	HPCG (TFlop/s)	Rpeak (PFlop/s)	Power (kW)	Efficiency (GFlops/W)
Phase 1, November 2021								
119	54	35	18,688	3.87	213.77	5.61	648	5.972
141	55	37	14,336	3.25	203.00	4.28	565	5.752
Phase 2b, June 2024								
101	90	24	36,864	9.36	517.24	11.20	1.213	7.717
136	92	35	28,352	6.73	399.10	8.58	942	7.144
Phase 3, June 2025								
113	125	21	54,400	13.58	754.45	16.43	1,692	8.035
156	127	27	41,920	9.20	585.75	12.22	1,164	7.905

Quality of ICON for NWP

850 hPa 24h



500 hPa 120h

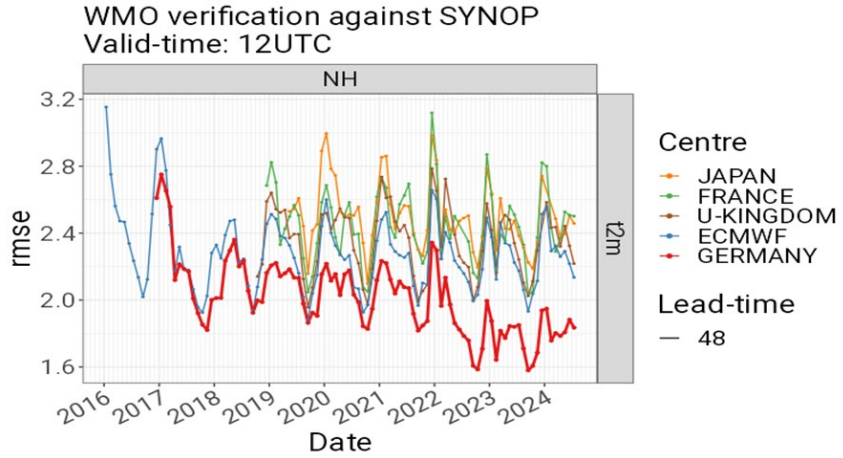


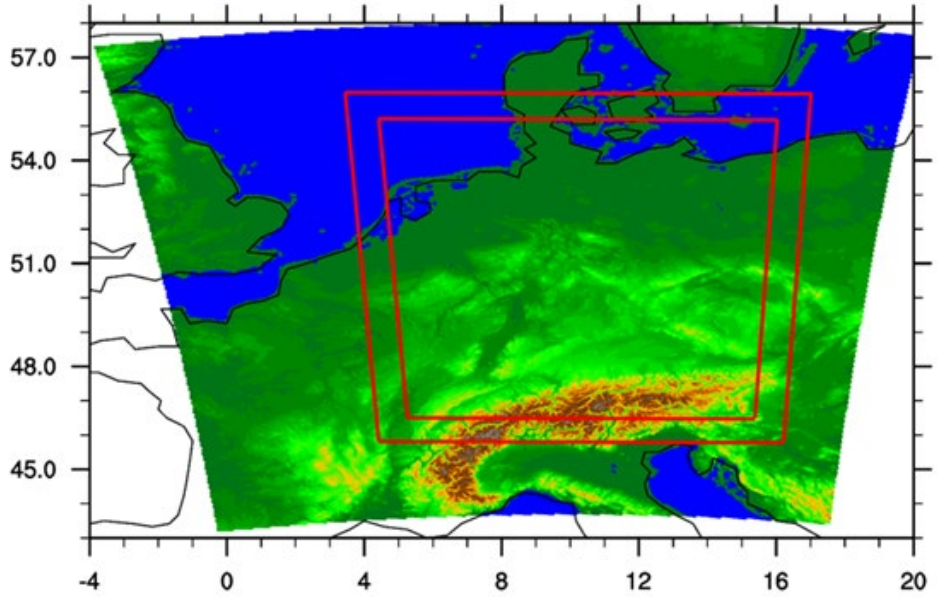
ICON + Ensemble Data Assimilation

14 NWP Centres with global Forecasts,
ICON well positioned in Top Group

All scores for the northern hemisphere.

Synop T2M 48h





Future Plans of DWD

You never know what might happen

The Wish List

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

ICON Global Non-Hydrostatic	ICON-EU 2-way-nest	ICON-D2 LAM Convective Scale	ICON-D05 LAM 2x 2-way-nest	ICON-D2 RUC Convective Scale	ICON-D05 RUC Convective Scale	AICON
Det 13 (10) km EPS 26 (13) km 120 level 40 member	Det 6.5 (5) km EPS 13 (6.5) km 74 level 40 member	Det 2 km EPS 2 km 65 level 20 (40) member	Det 2 | 1 | 0.5 km EPS 2 | 1 | 0.5 km 65 level 20 member	Det 2 km EPS 2 km 65 level 20 (40) member	Det 2 | 1 | 0.5 km EPS 2 | 1 | 0.5 km 65 level 20 member	AI Forecasts
Analysis (3h)	**Analysis (3h)**	**Analysis (1h)**	**ICON-D2 LAM**	**Analysis (1h)**	**ICON-D2 RUC**	
EnVAR + LETKF 40 member	EnVAR + LETKF 40 member	KENDA / LETKF 40 member	Nests are started during forecast	KENDA / LETKF 40 member	Nests are started during forecast	
Forecasts	**Forecasts**	**Forecasts**	**Forecasts**	**Forecasts**	**Forecasts**	**Forecasts**
180h: 00,12 120h: 06,18 51h: 03,09,15,21	180h: 00,12 120h: 06,18 51h: 03,09,15,21	48h: 00,03,06,09,12, 15,18,21	48h: 00,03,06,09, 12,15,18,21	14h: 00,01,…,23	14h: 00,01,…,23	180h: 00,12 120h: 06,18 51h: 03,09,15,21

Black text: Major parts of current operational system

Red text: Possible upgrades for the new system.

Compute intensive and challenging.

All systems will be supplemented by AI-based predictions.

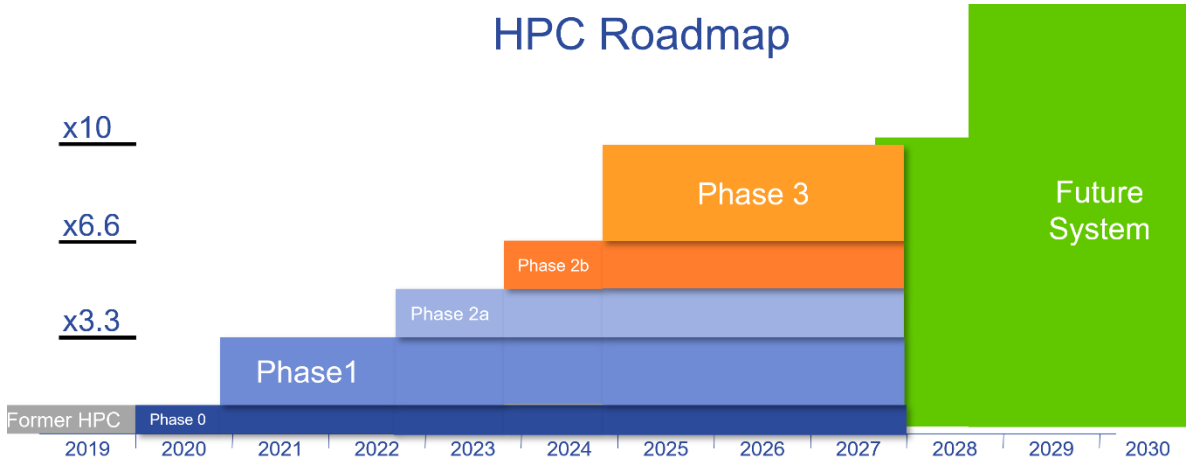
The Reality

- Original plans for a procurement in 2025.
- Result of a market research in 2024: Moore´s law definitely is dead.

  (ok: everybody knew before)
- In other words: Original plans and original budget do not match.

- Request to increase the budget, but:
  - Nov 24 – May 25: lack of government.
  - Since then (still) lack of a budget law.

- Hope dies last:
  - We start to prepare for a procurement in 2026.
  - (Hoped-for) growth in performance: Current HPC x 1.5 + AI partition



© Manuel Reiter, DWD

Conclusions

Did NWP Pave the Way for AI?

Question reformulated:

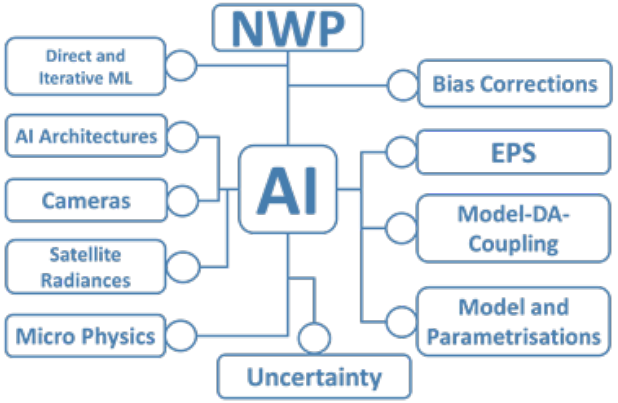
- ➔ Was NWP a driver for HPC development?
- ➔ There are other (more powerful / financially stronger) groups influencing HPC (or computer) development.

Not really!

Personal opinion!

Current situation:

- ➔ Computational power and scientific knowledge have reached a level to support many interesting AI systems and applications.
- ➔ Computational power has reached a financial level that NWP centers can hardly afford anymore.



© Roland Potthast, DWD

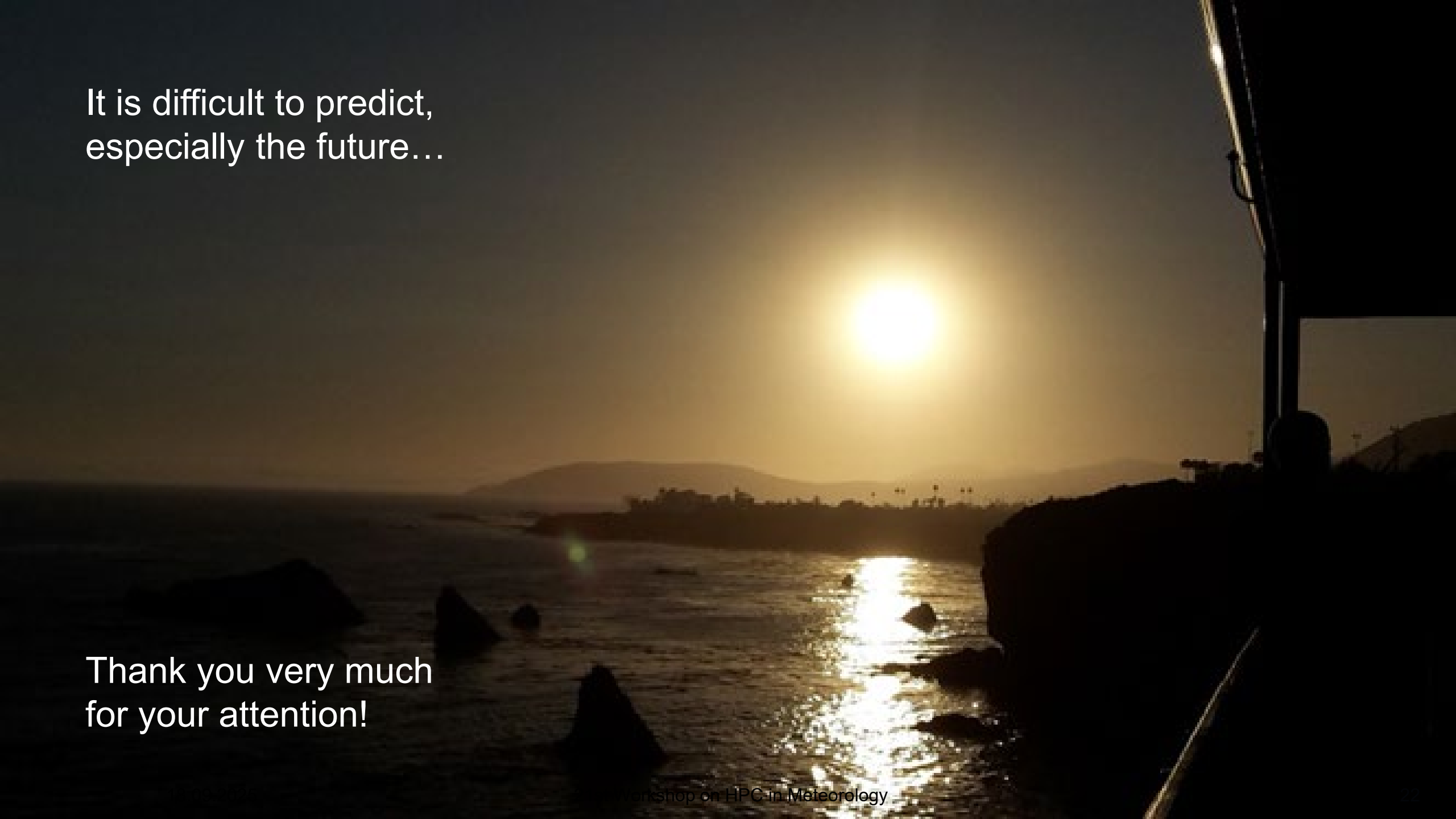
But weather forecasting provides a lot of interesting problems that now could be tackled by AI.

The future starts now!

Additional Literature and Web Sites Visited

- https://en.wikipedia.org/wiki/History_of_artificial_intelligence
- https://www.ibm.com/think/topics/history-of-artificial-intelligence
- https://en.wikipedia.org/wiki/History_of_numerical_weather_prediction

- AI: The Tumultuous History of the Search for Artificial Intelligence, Daniel Crevier, January 1993, Publisher: BasicBooks
- The origins of computer weather prediction and climate modeling, Peter Lynch, Journal of Computational Physics, 227(7), 3431-44. Bibcode: 2008JCoPh.227.3431L; doi: 10.1016/j.jcp.2007.02.034.
- 50 years numerical weather prediction. 50th Anniversary of Numerical Weather Prediction Commemorative Symposium, Wissenschaftspark Albert Einstein, Telegrafenberg, Potsdam, 9-10 March, 2000. Book of Lectures.


It is difficult to predict,
especially the future…
Thank you very much
for your attention!