# How AI is revolutionizing climate modeling

Nikolay Koldunov and colleagues from AWI and ECMWF

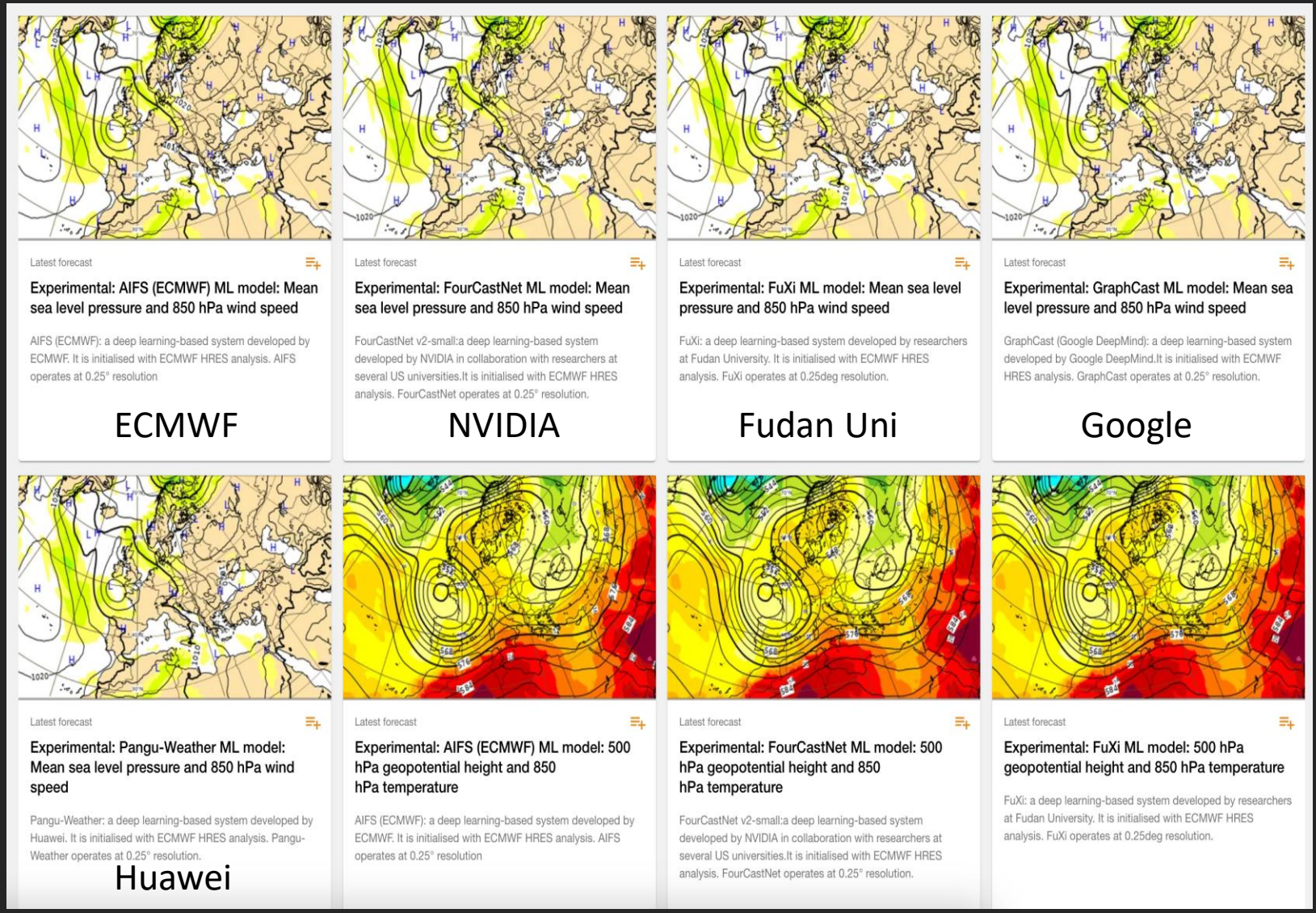
FESOM2 – The Finite-volumE Sea ice–Ocean Model – AWI
Visualization – DLR

# ML/AI for weather and climate applications

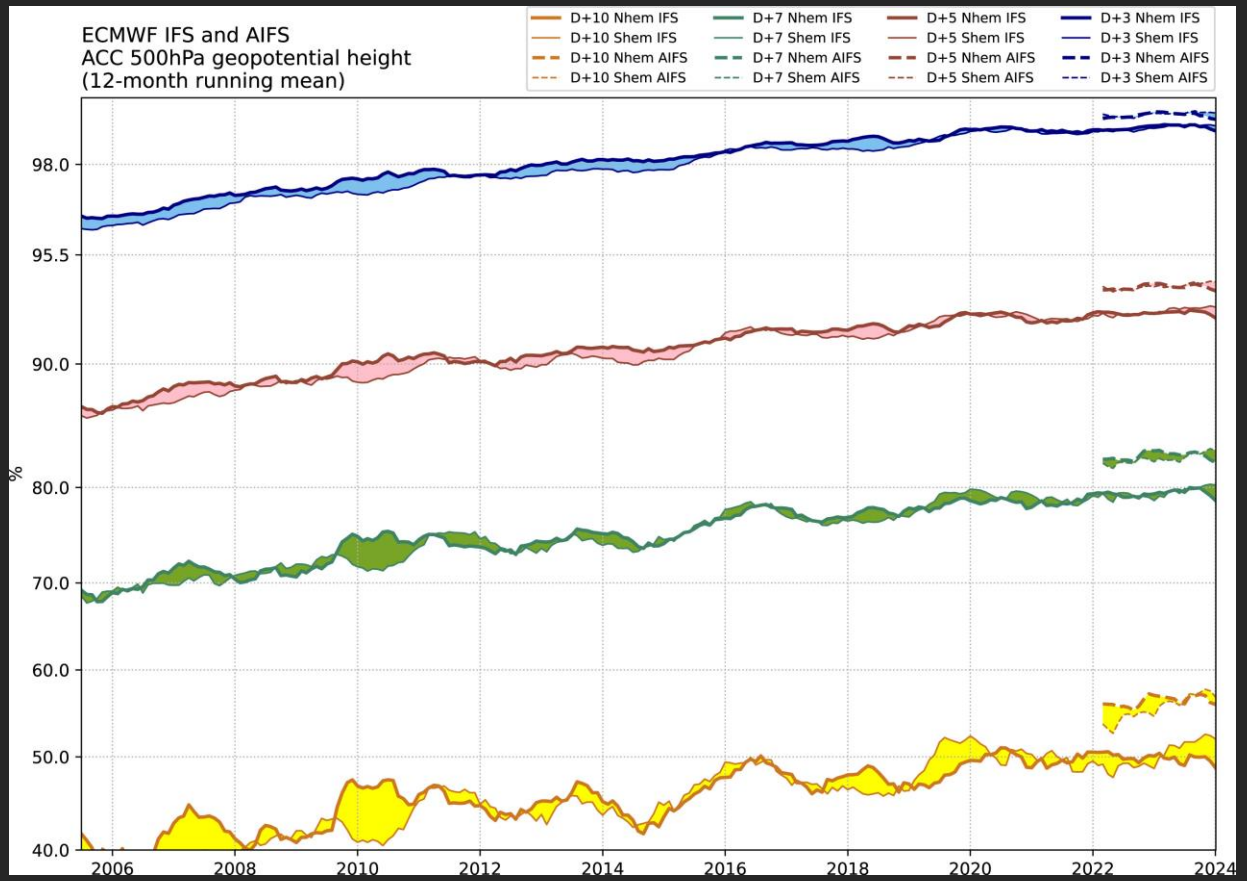
- Feature extraction
- Compression
- Sub-grid scale parameterizations
- Equation discovery
- Hybrid emulators
- Full emulators
- Data processing
- Climate change communication

# Weather prediction

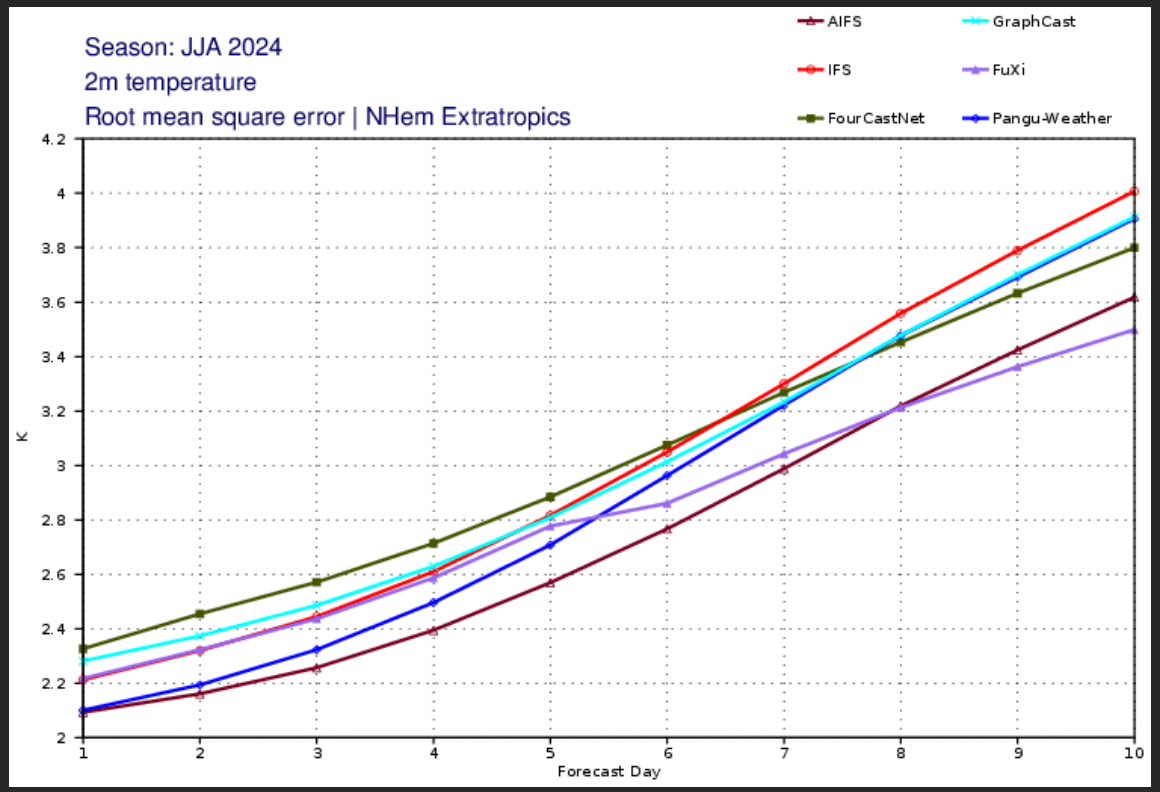
ECMWF forecast page



# Weather prediction




Bauer, Peter. "What if? Numerical weather prediction at the crossroads." *Journal of the European Meteorological Society* 1 (2024): 100002.



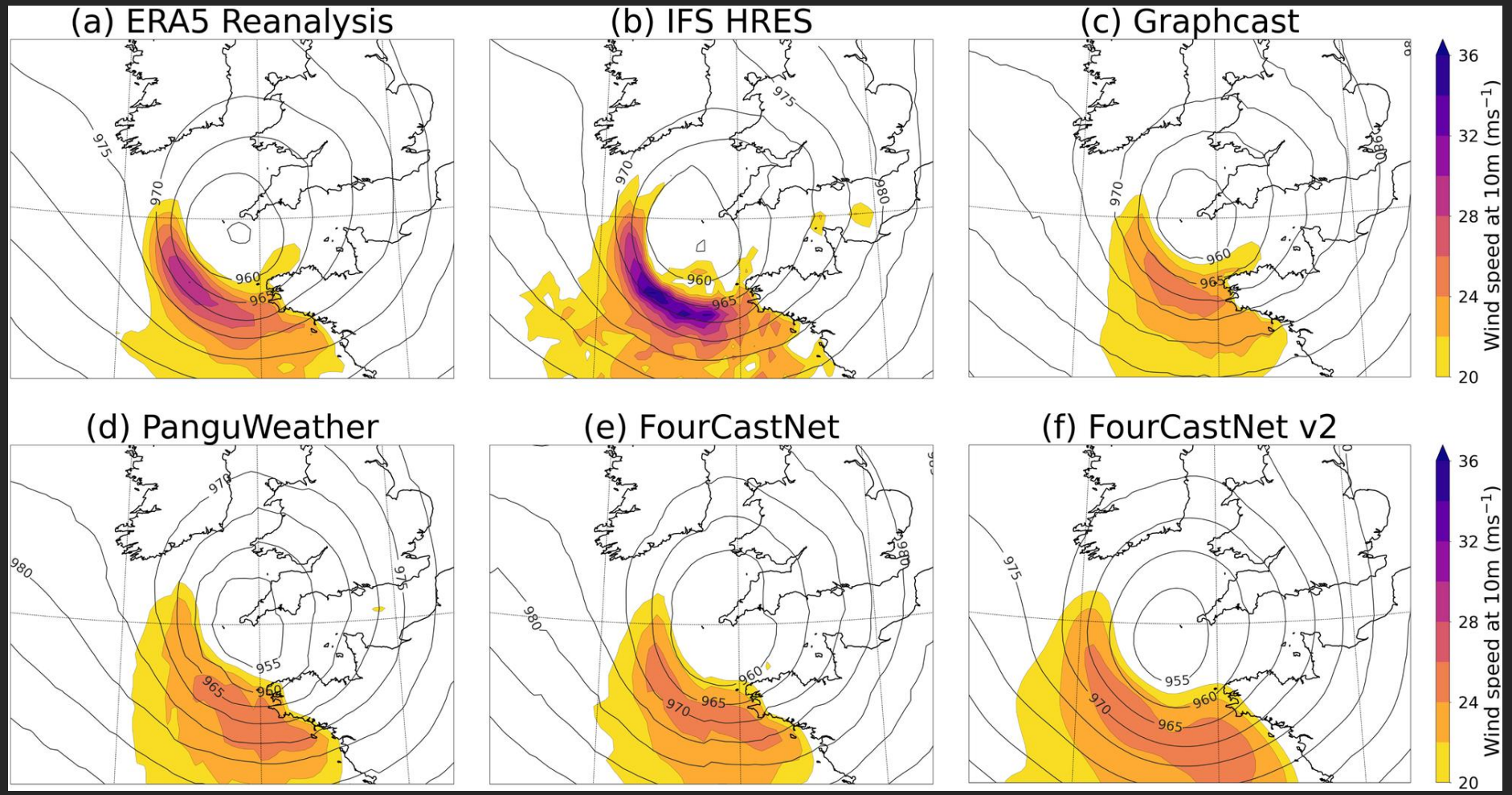

https://charts.ecmwf.int/products/plwww_3m_fc_aimodels_surface_mean

- Different architectures
- Trained on ERA5 reanalysis

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

# Weather prediction

10-m wind speed (shading) and MSLP (contours) for **Storm Ciarán**




Charlton-Perez, Andrew J., Helen F. Dacre, Simon Driscoll, Suzanne L. Gray, Ben Harvey, Natalie J. Harvey, Kieran MR Hunt et al. "Do AI models produce better weather forecasts than physics-based models? A quantitative evaluation case study of Storm Ciarán." *npj Climate and Atmospheric Science* 7, no. 1 (2024): 93.

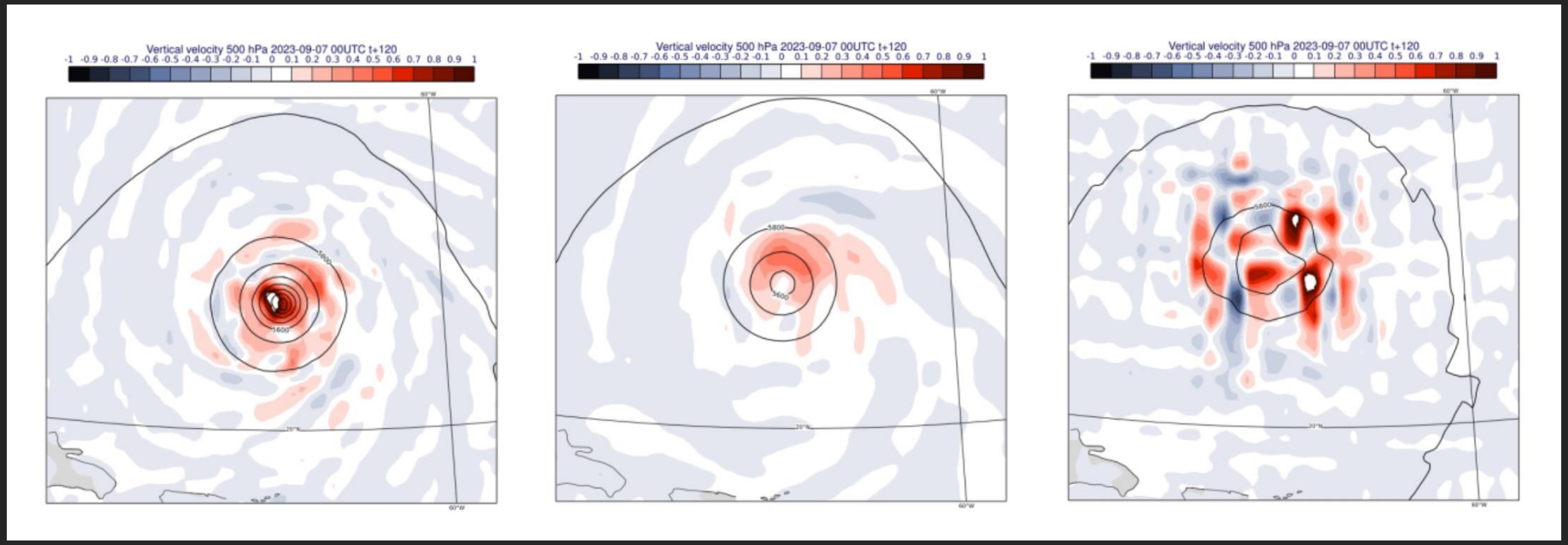
# Weather prediction

## t+120 hours forecast vertical velocities at 500 hPa

IFS

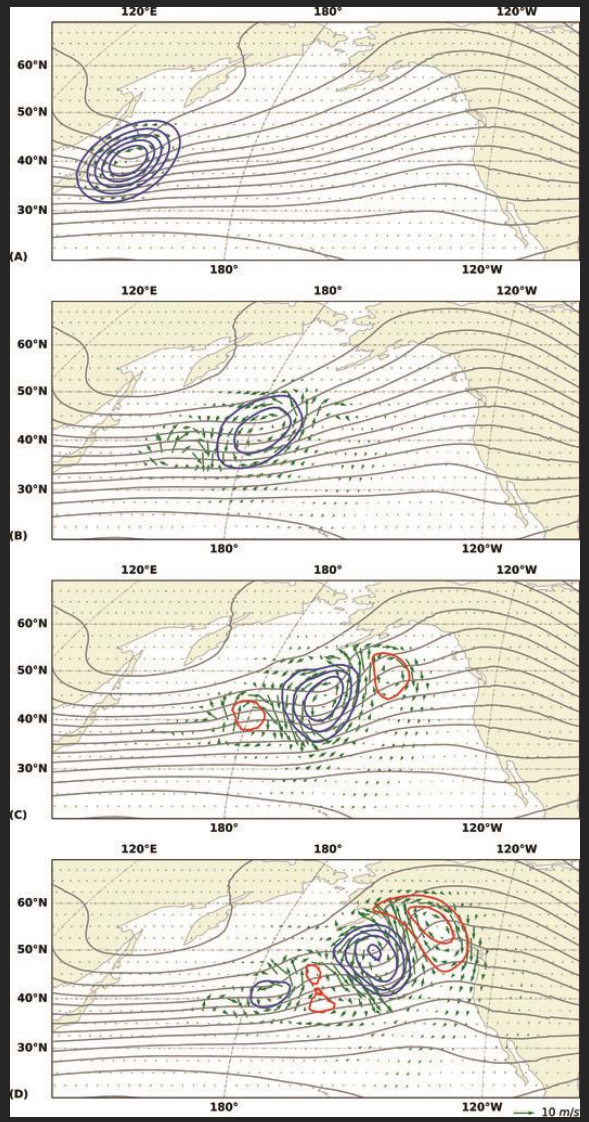
ERA5

Pangu-Weather (diagnosed)




Bonavita, Massimo. "On the limitations of data-driven weather forecasting models." *arXiv preprint arXiv:2309.08473* (2023).

# Weather prediction




Dynamical Tests of a Deep Learning Weather Prediction Model

- Steady tropical heating
- Extratropical cyclone development
- Geostrophic and hydrostatic adjustment
- Atlantic hurricane development

“Observations and idealized modeling results show the development of a localized extratropical cyclone, with subsequent cyclones appearing downstream as the disturbance evolves into a spreading wave packet”

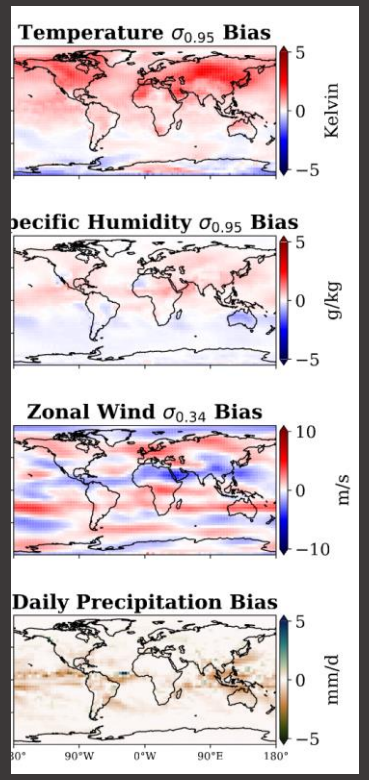
“We conclude that the model encodes realistic physics in all experiments and suggest that it can be used as a tool for rapidly testing a wide range of hypotheses.”

Hakim, Gregory J., and Sanjit Masanam. "Dynamical tests of a deep-learning weather prediction model." *Artificial Intelligence for the Earth Systems* (2024).

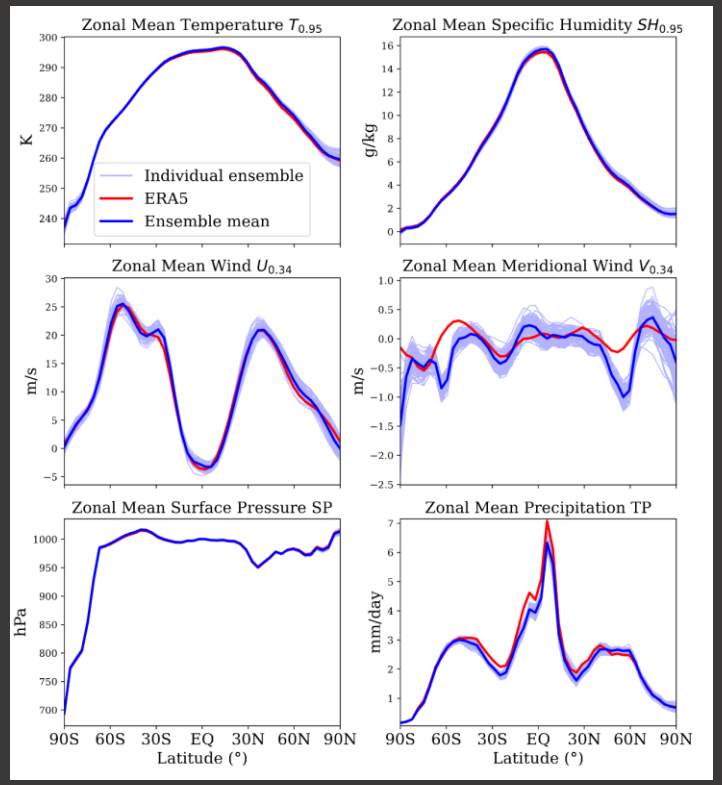
# Climate simulations

## LUCIE

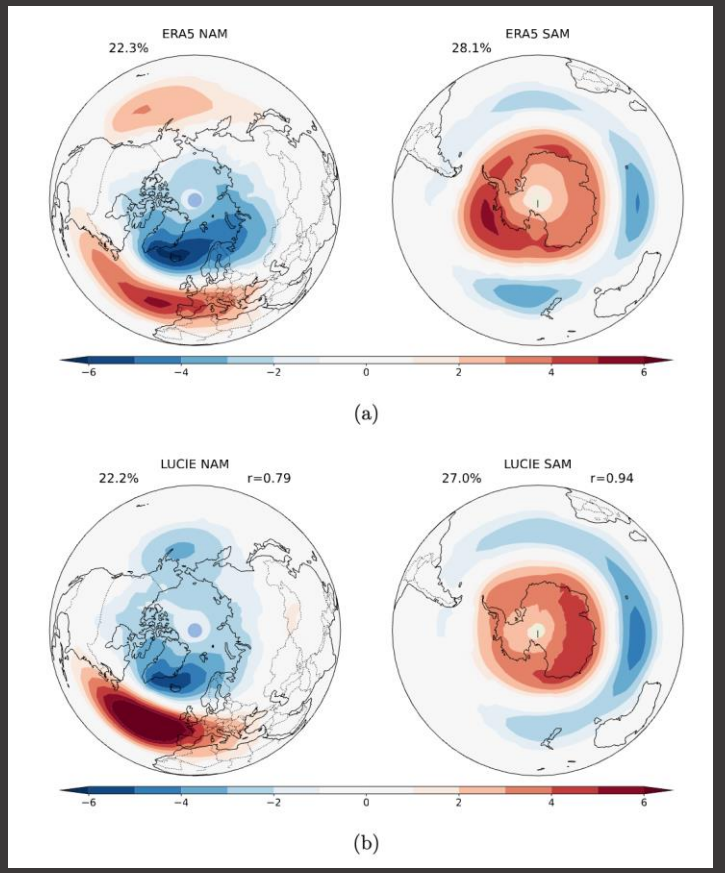
### Ensemble mean annual climatology bias




### Zonal mean climatology




### Northern Hemisphere Annular Mode (NAM), Southern Hemisphere Annular Mode (SAM)



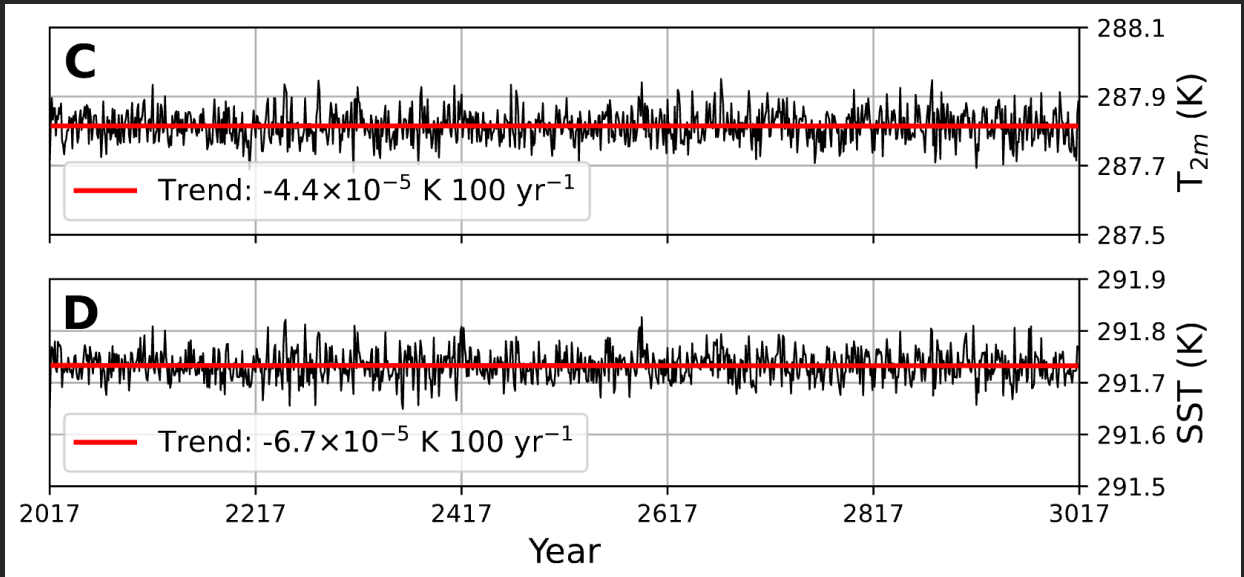

Atmosphere only, very low (400km) resolution.
Can be trained on as low as 2 years of 6-hourly ERA5 data
Forcing: Total incoming solar radiation, orography (good for paleo?).
Can reproduce present climate, and generate huge ensembles.
Trained on  in 2.4 hour on single GPU


Guan, Haiwen, Troy Arcomano, Ashesh Chattopadhyay, and Romit Maulik. "LUCIE: A Lightweight Uncoupled ClImate Emulator with long-term stability and physical consistency for O (1000)-member ensembles." *arXiv preprint arXiv:2405.16297* (2024).

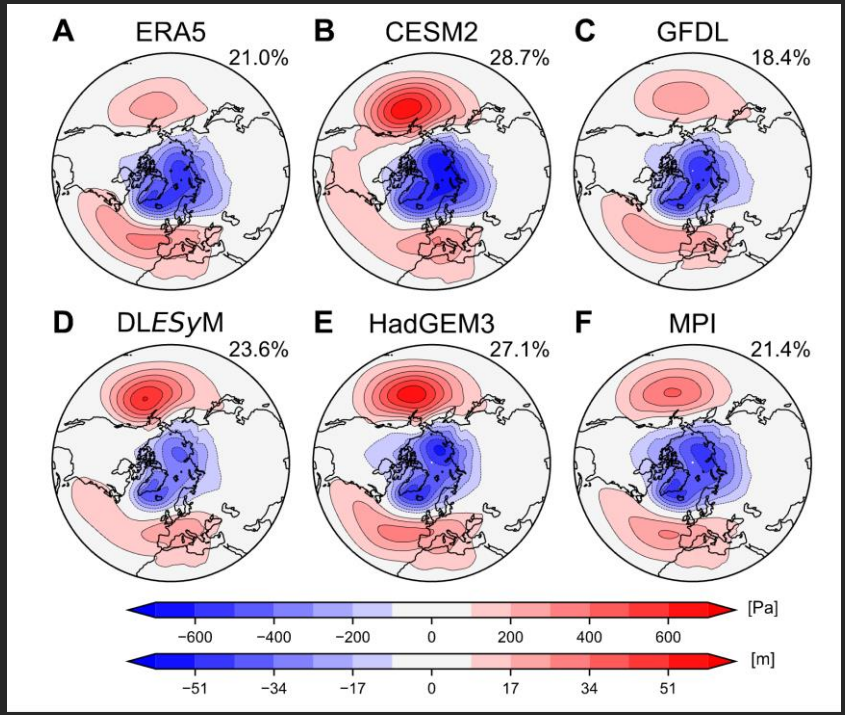
# Climate simulations

DLESyM

## Stable 1000 years rollout

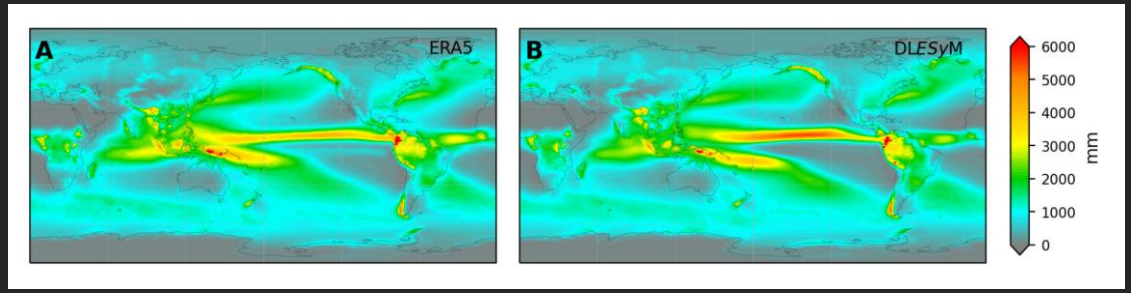



## Leading mode of NH variability




- Coupled deep learning atmosphere and ocean models (110km HEALPix resolution)
- Trained on ERA5 and some satellites, no antropogenic forcing.
- Can do stable 1000 years of present day climate
- Takes 24 h to train, and 12 h to do a 1000 year simulation.

## Climatology of global annual precipitation (diagnosed)



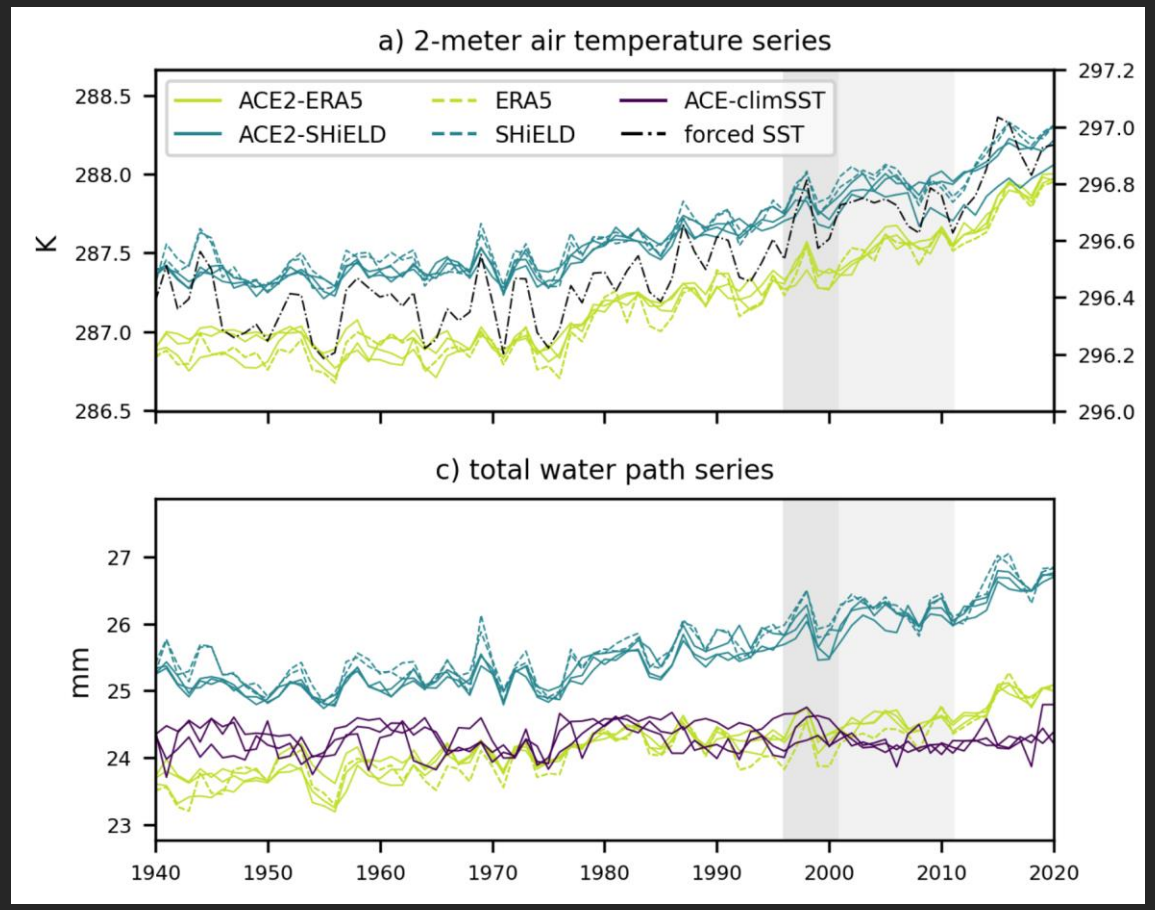


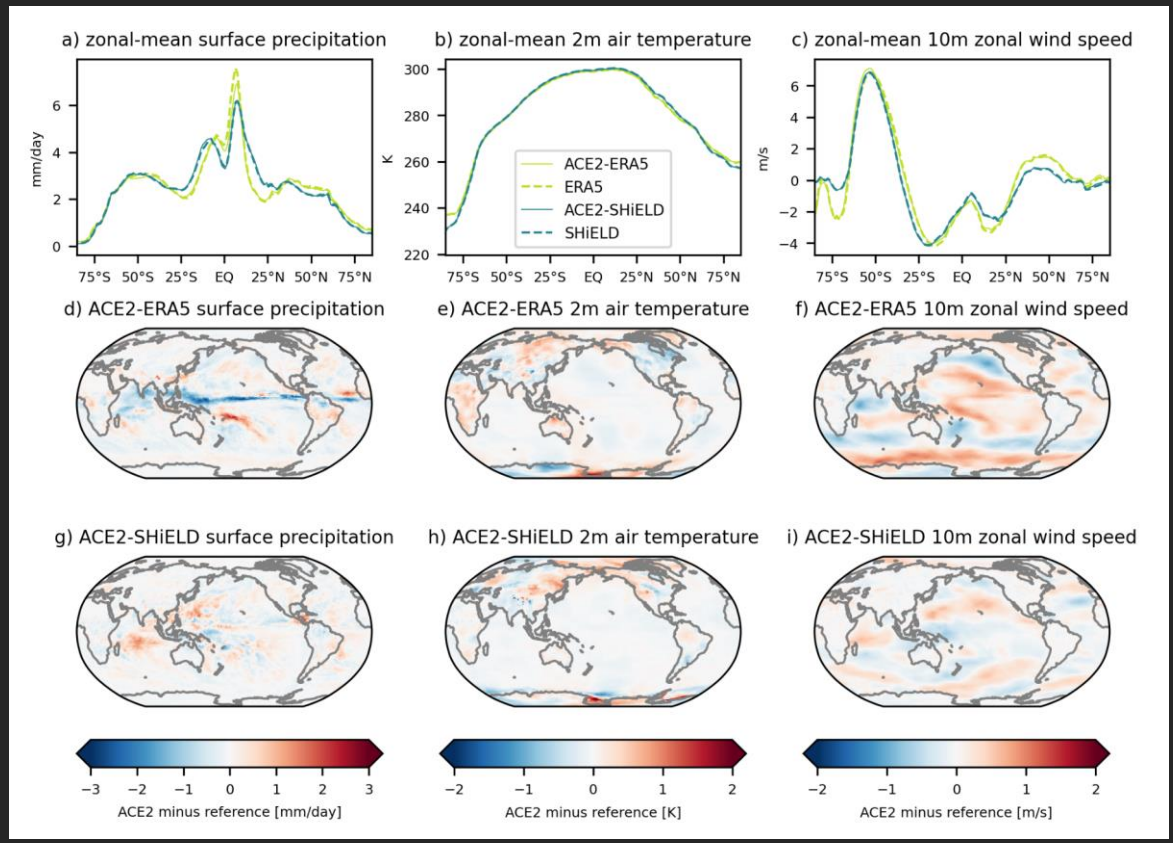
Cresswell-Clay, Nathaniel, Bowen Liu, Dale Durran, Andy Liu, Zachary I. Espinosa, Raul Moreno, and Matthias Karlbauer. "A deep learning earth system model for stable and efficient simulation of the current climate." *arXiv preprint arXiv:2409.16247* (2024).


AWI HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

# Climate simulations

## ACE2 (traind on SHiELD and ERA5)




Atmosphere only + prescribed SST, 1 degree reolution

Trained on ERA5 and SHiELD. SST and CO2 as boundary conditions. Conserves global dry air mass and moisture.

Reproduce present day climate, and can run stably for 1000s of years under climatological mean SST and stable CO2.

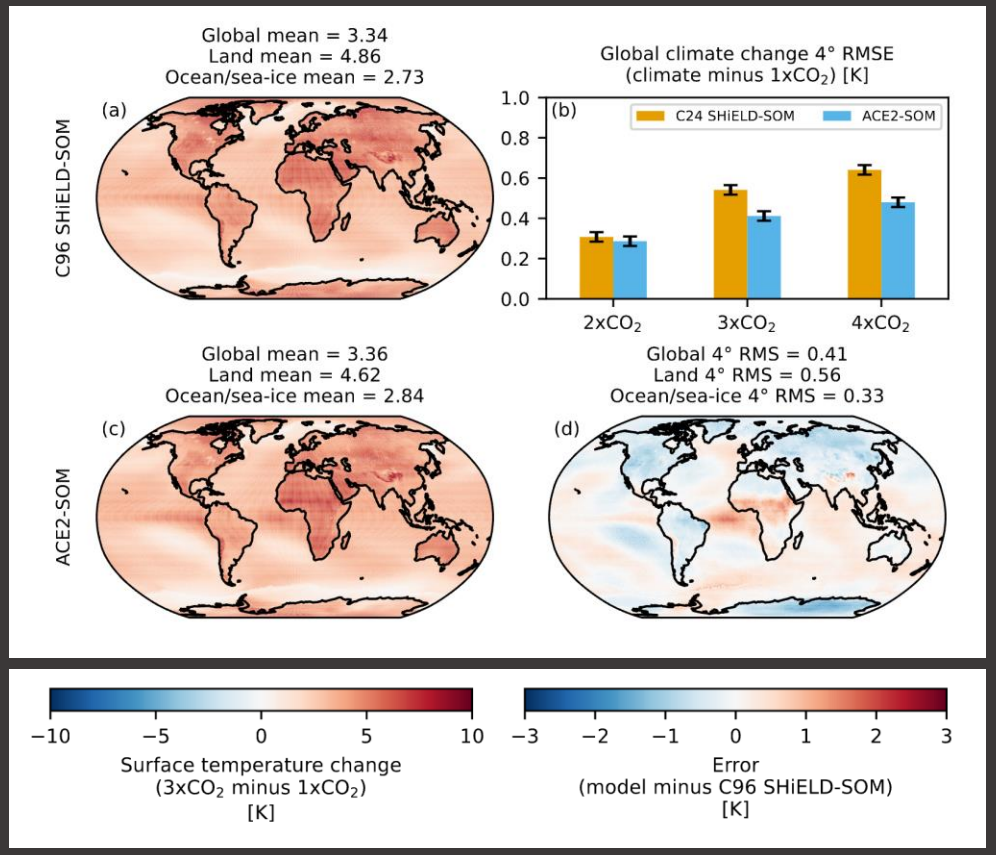
Training: 4.5 days on 8 x H100, Inference: 1500 SYPD on 1 x H100


Watt-Meyer, Oliver, Brian Henn, Jeremy McGibbon, Spencer K. Clark, Anna Kwa, W. Andre Perkins, Elynn Wu, Lucas Harris, and Christopher S. Bretherton. "ACE2: Accurately learning subseasonal to decadal atmospheric variability and forced responses." *arXiv preprint arXiv:2411.11268* (2024).

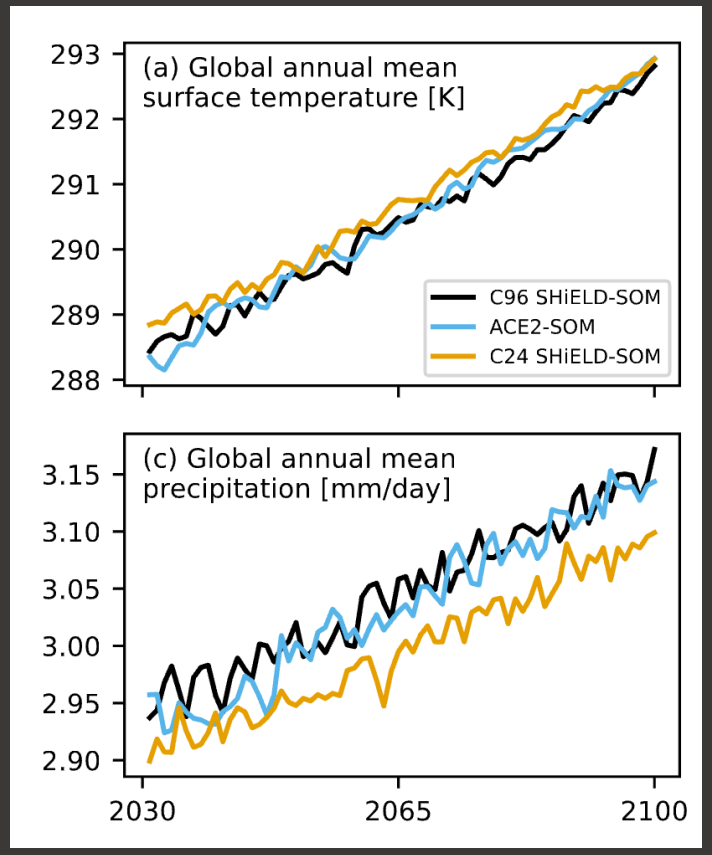
# Climate simulations

## ACE2-SOM (with slab ocean)

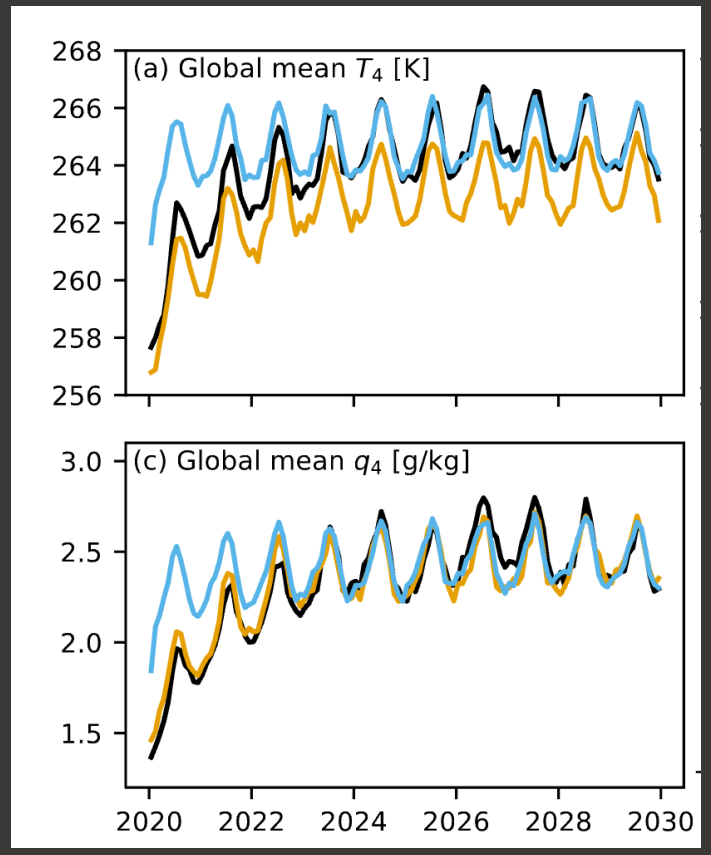
### Climate change pattern (3xCO2)




### Gradually increasing CO2




### Abrupt CO2 increase (4xCO2)



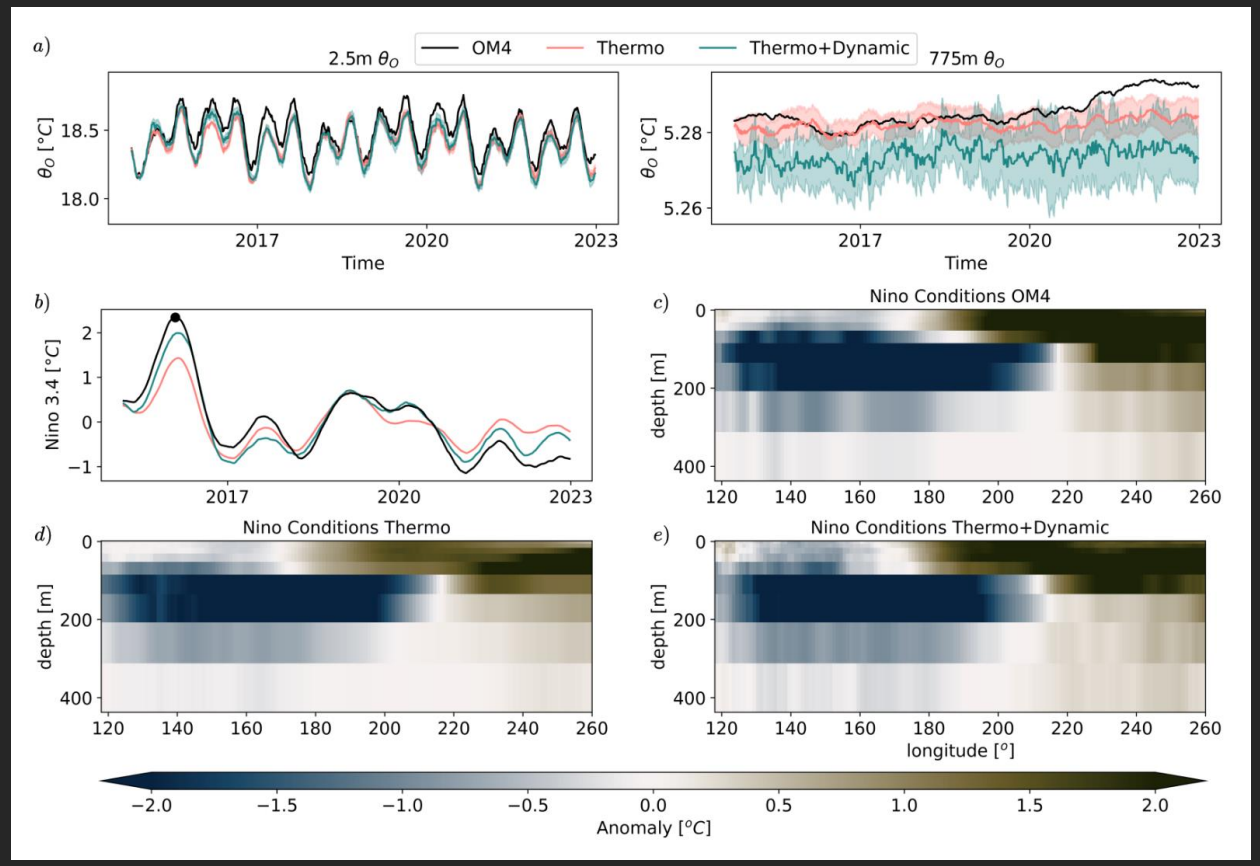

Atmosphere ML + slab ocean, trained on 1 degree SHiELD simulations (also with slab ocean).
Trained only on 1xCO2, 2xCO2, and 4xCO2 equilibrium climates.
Reproduce 3xCO2 equilibrium climate, largely reproduce CO2 increasing at a rate of 2 %/year, have problems with abrupt 4xCO2 (warms too fast).


Clark, Spencer K., Oliver Watt-Meyer, Anna Kwa, Jeremy McGibbon, Brian Henn, W. Andre Perkins, Elynn Wu, Christopher S. Bretherton, and Lucas M. Harris. "ACE2-SOM: Coupling to a slab ocean and learning the sensitivity of climate to changes in CO $ _2$." *arXiv preprint arXiv:2412.04418* (2024).

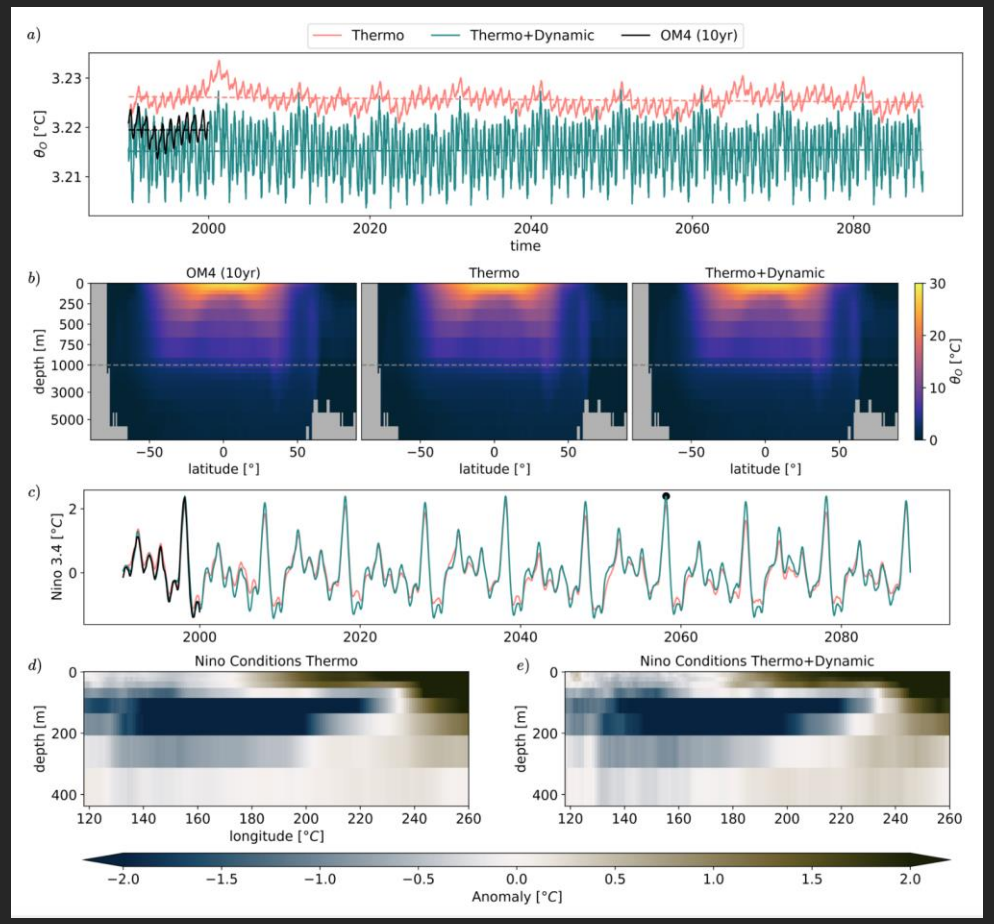
# Ocean climate simulations

## Samudra

8 year test (JRA55 forcing)




1000 year test ( repeated 1990-2000 forcing)




Full depth emulator of ocean model (OM4)

Prognostic: T,S,U,V, SSH, 19 levels, 5-day average, 1 degree resolution (from original ¼ degree).

Forcing: wind stress, net heat flux and it's anomalies.

Trained on 4 A100 GPUs in less than a day. Rollout: ~1800 SYPD on one GPU.


Dheeshjith, Surya, Adam Subel, Alistair Adcroft, Julius Busecke, Carlos Fernandez-Granda, Shubham Gupta, and Laure Zanna. "Samudra: An AI Global Ocean Emulator for Climate." *arXiv preprint arXiv:2412.03795* (2024).


AWI HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

# Climate simulations

- Currenlty mostly AMIP type
- Emulators of ERA5 or existing models
- Low resolution (1 degree or couarser)
- Very fast for training and inference
- Moving towards coupled and scenario simulations

# What we do in AWI?

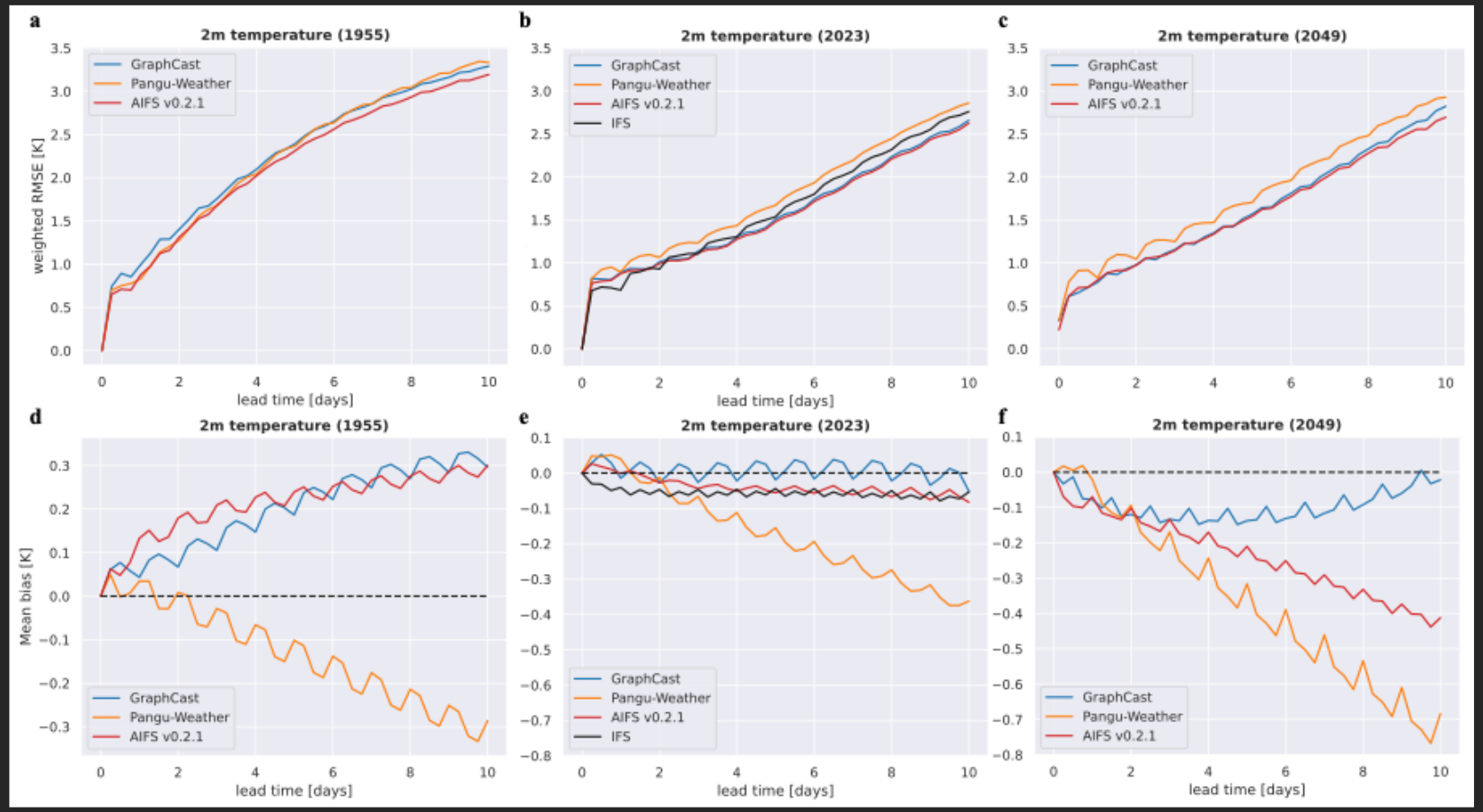
# Let’s have some fun with those models

Models are trained on ERA5 reanalysis (present day, ¼ degree).

- How will they behave in “out of sample” climate?

- How will they behave if initialised with low resolution data? What if those data come from climate model?

# Use AI-NWP models to reproduce weather of climate models




- 3 AI-NWP models
- Initial conditions from 3 different out-of-distribution years

AI-NWP models can reproduce weather of the climate model.

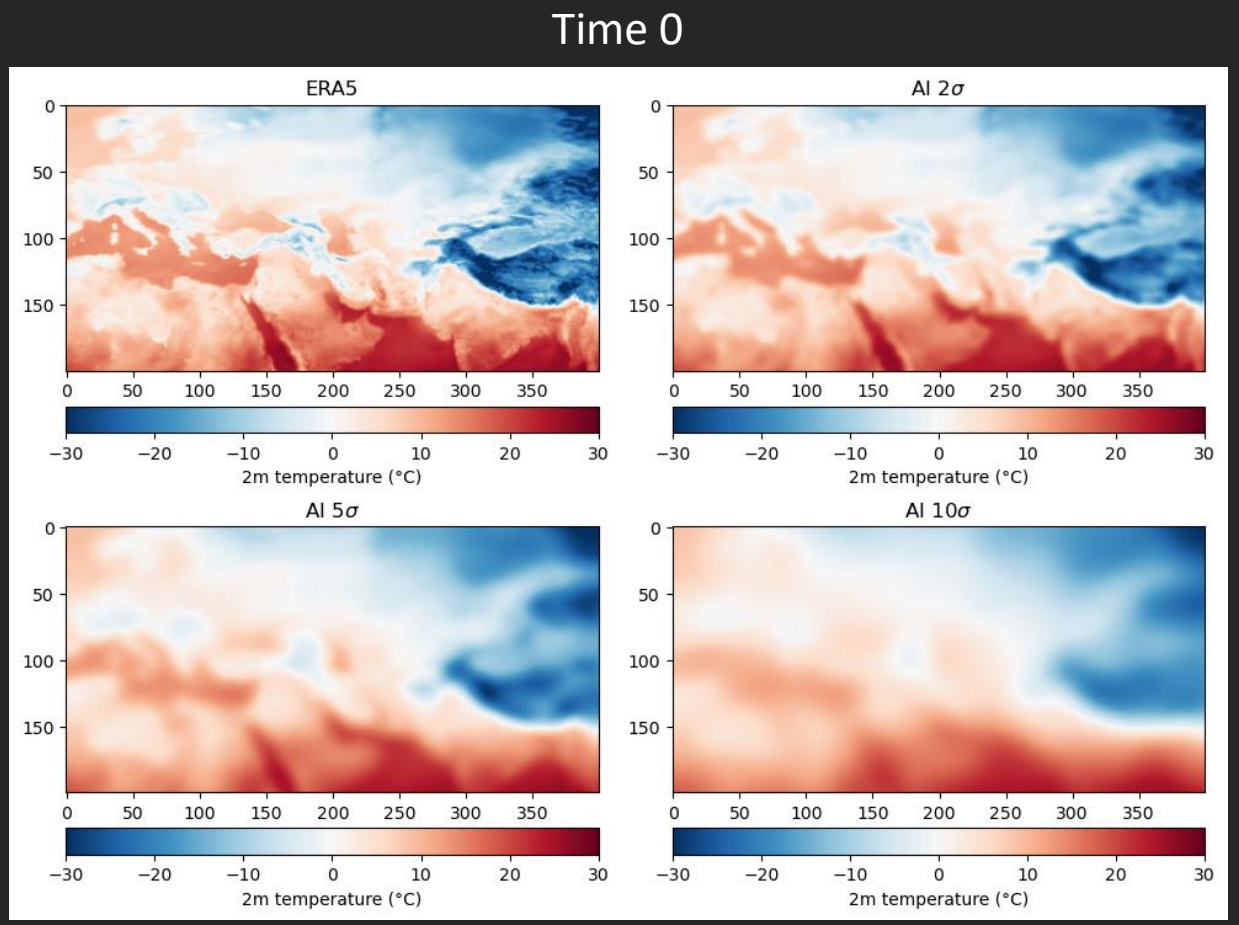
Temporal interpolation

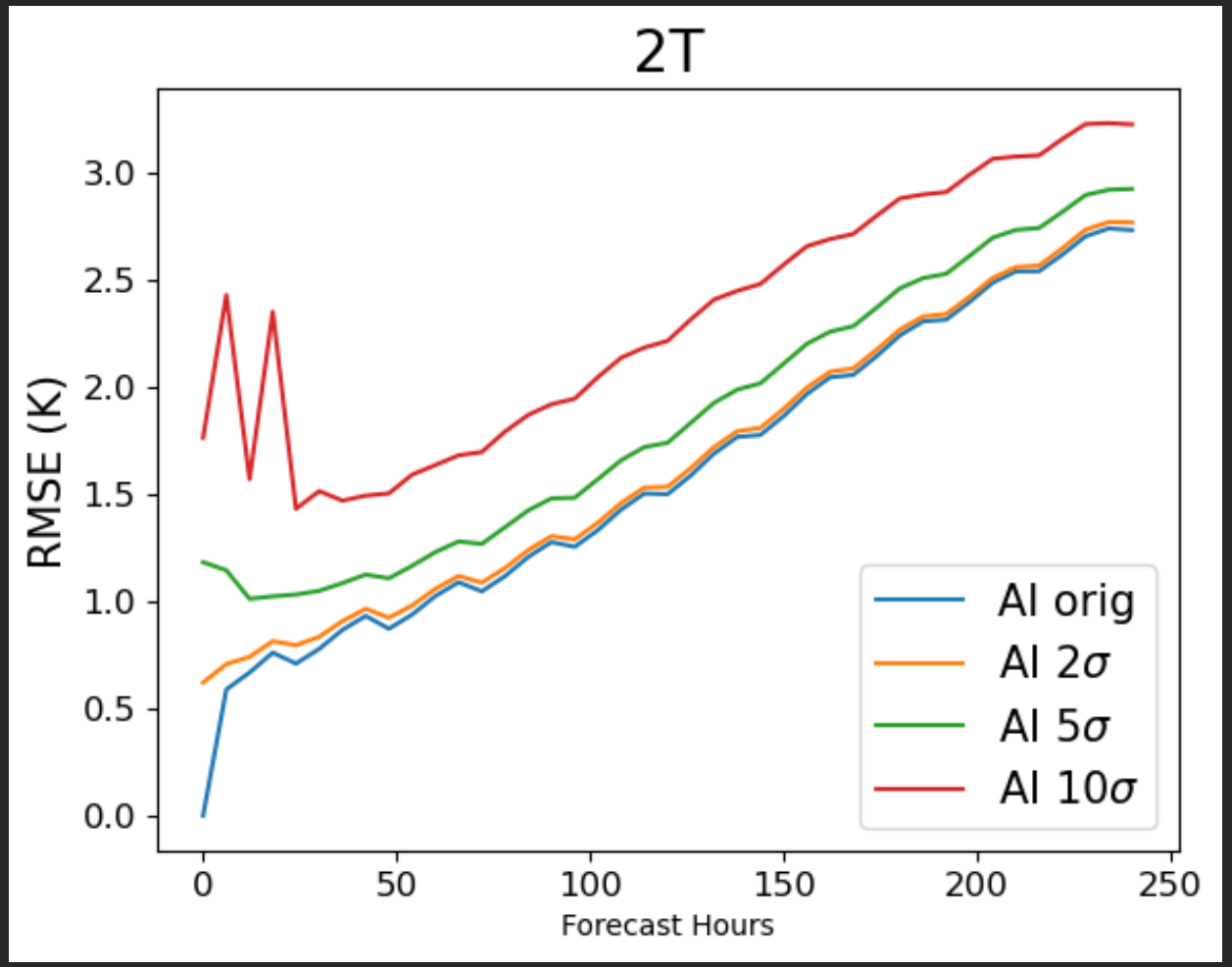
Uncertainty estimates

Rackow et al., 2024 Robustness of AI-based weather forecasts in a changing climate, arXiv

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

# AI-NWP models as downscaling tools



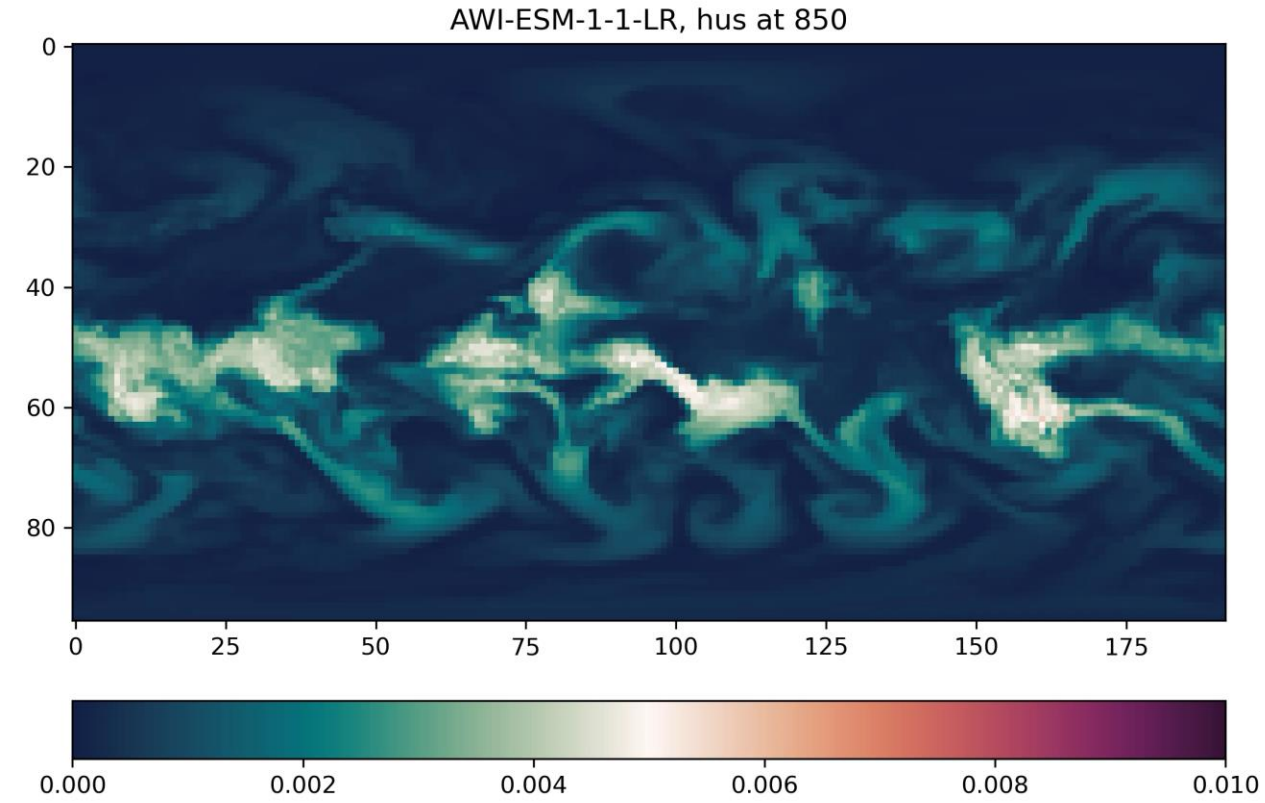



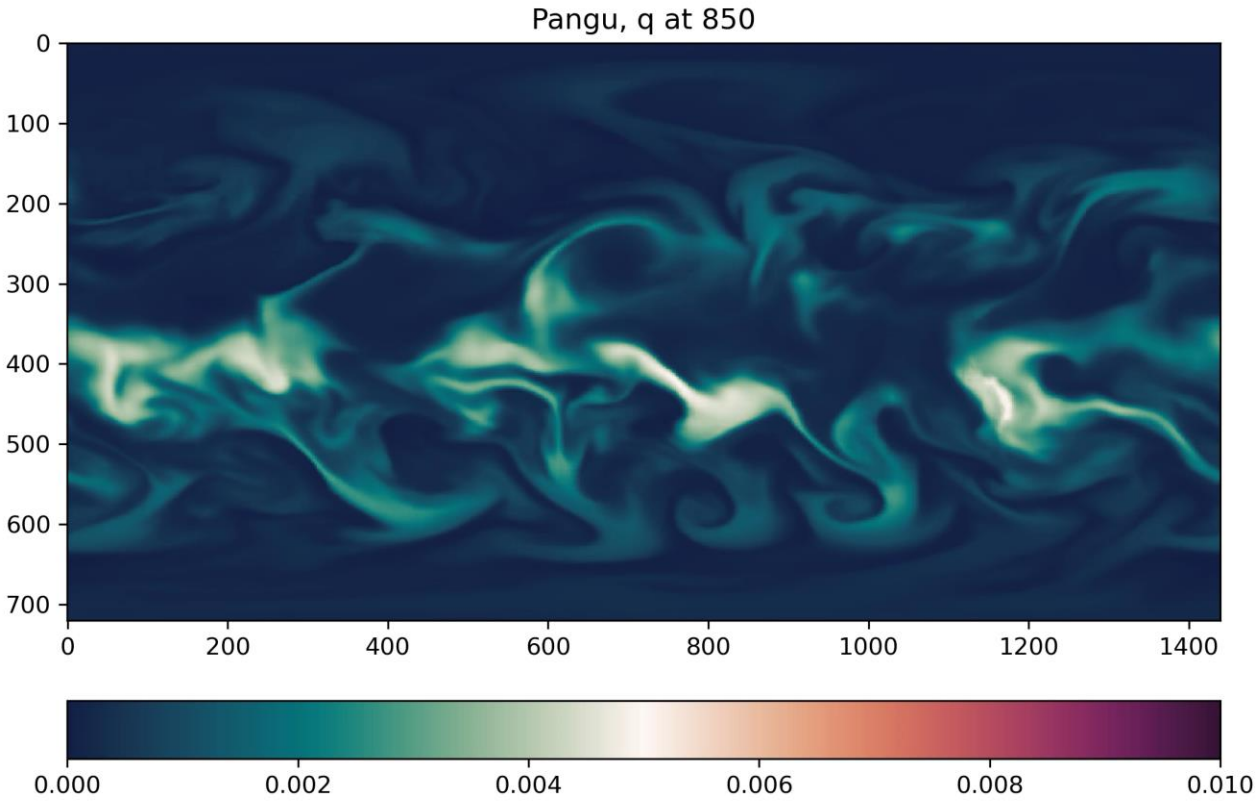

Koldunov, N., et al., 2024. Emerging AI-based weather prediction models as downscaling tools. *arXiv*

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

# AI-NWP models as downscaling tools

## Reproduction of global Q850 evolution from 2 day forecasts

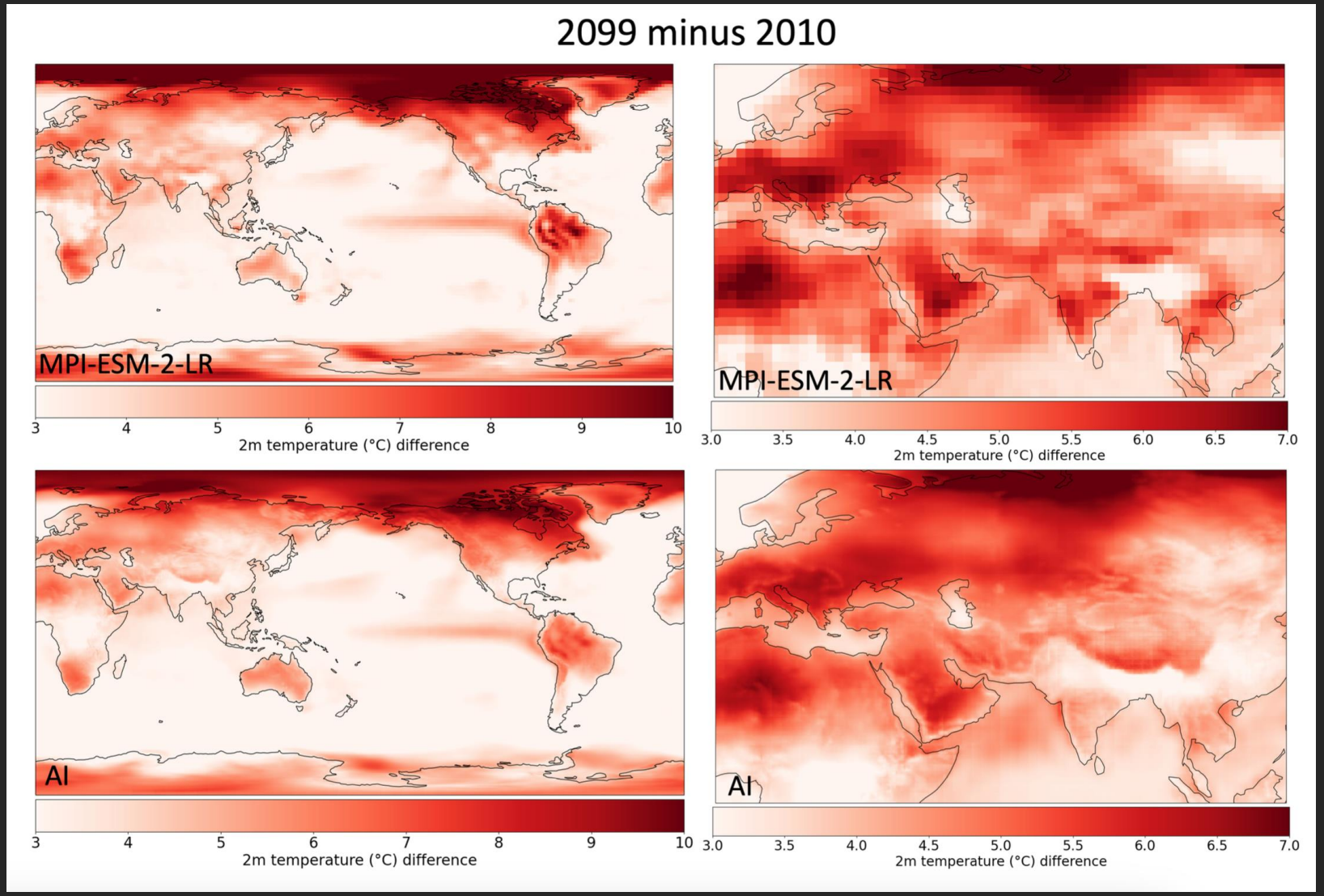



HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

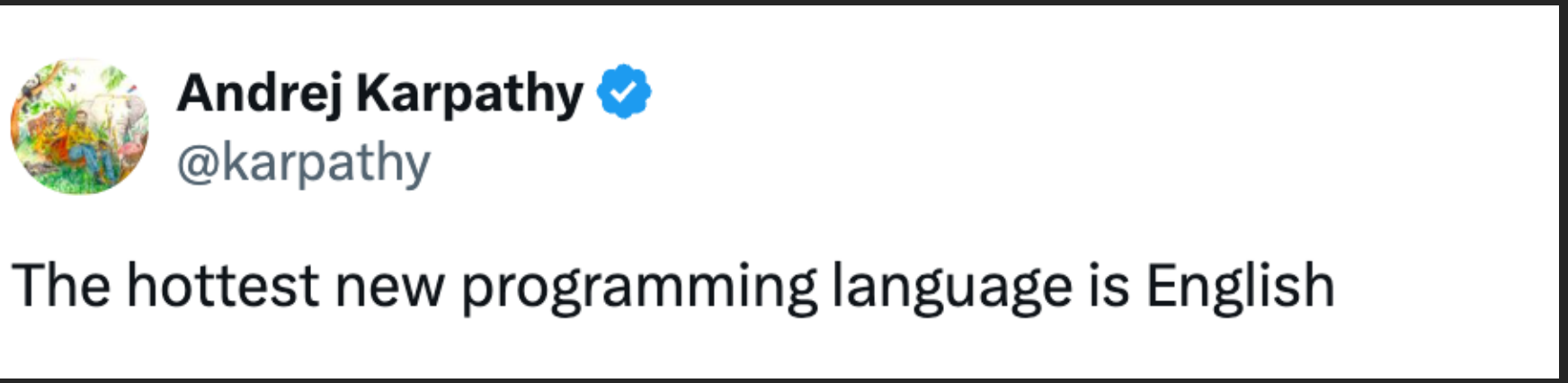
# AI-NWP models as downscaling tools

## Climate chance signal




And now for something completelly different!

# Data processing




Is natural language the next hot thing in data interaction?

# Imagine…

What oceanographic data available for the Fram Strait region in 2000-2002?

There are 10 dedicated oceanographic expeditions, 20 mooring sites and several expeditions of opportunity. What you are interested in?

Show me the map of measurement locations

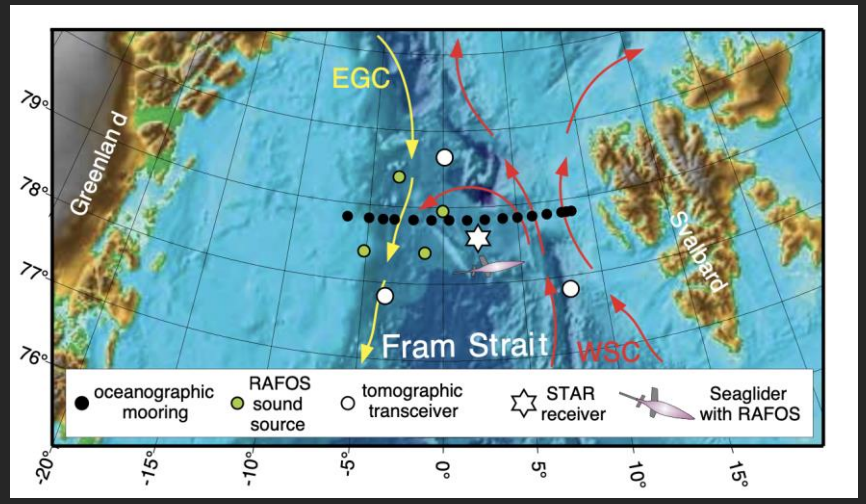



Image source: Beszczynska-Möller et al,, 2011

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

# Imagine…

Show mean temperature profile through Fram strait for May 2001

Unfortunately, there are only a few measurements for this period, so the profile may not be robust. Would you like me to proceed anyway, or should we use data from June 2001, where more data is available?

Do it for June

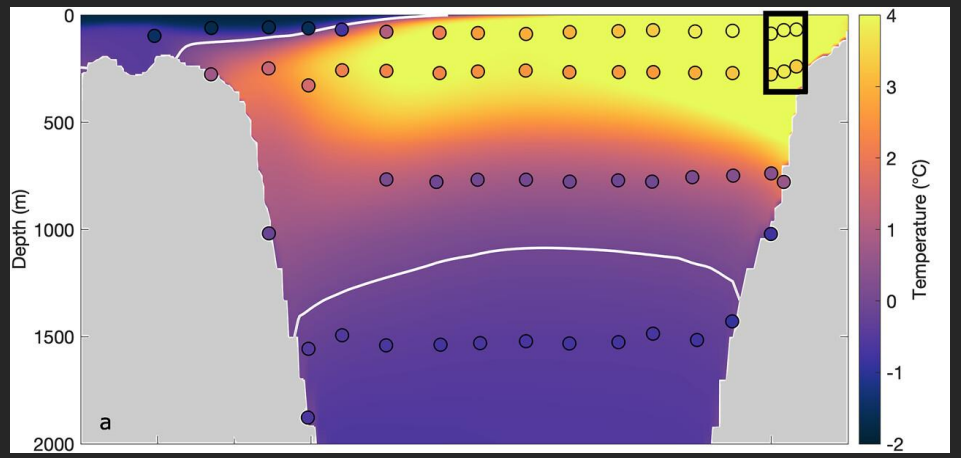



Image source:McPherson et al., 2023

# Many LLM-based systems are working on your problem

## Agents


**ClimDev** stands as a **virtual climate service company** that operates through various **intelligent agents** holding different roles, including Chief Scientist, Programmer, Climatologist, Economist, Geographer, Policy expert, art designer.

https://github.com/OpenBMB/ChatDev

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

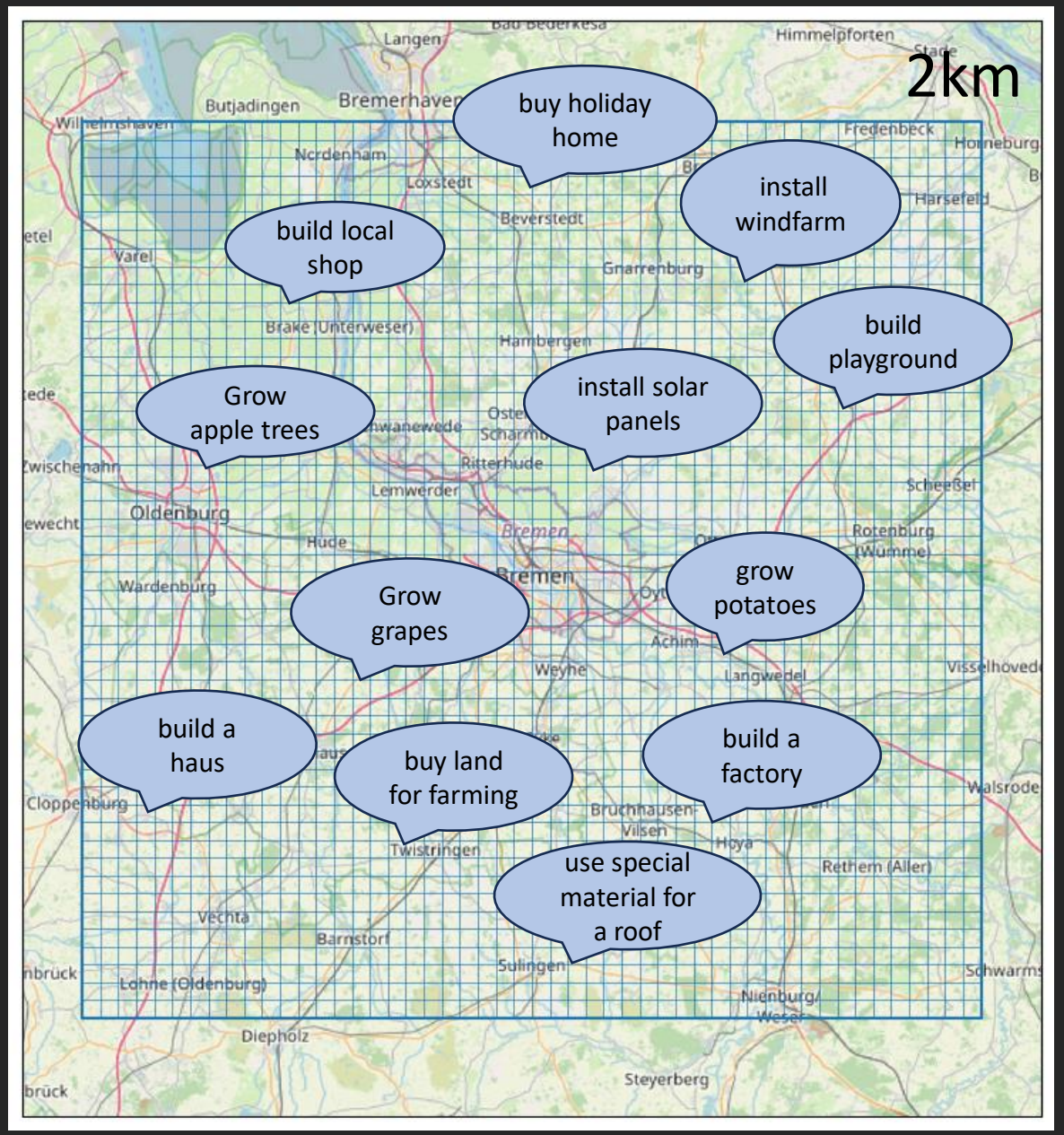
Deploy

# PANGAEA GPT

## Dataset Explorer

Enter your query:

# How climate change will affect my plans if I want to…

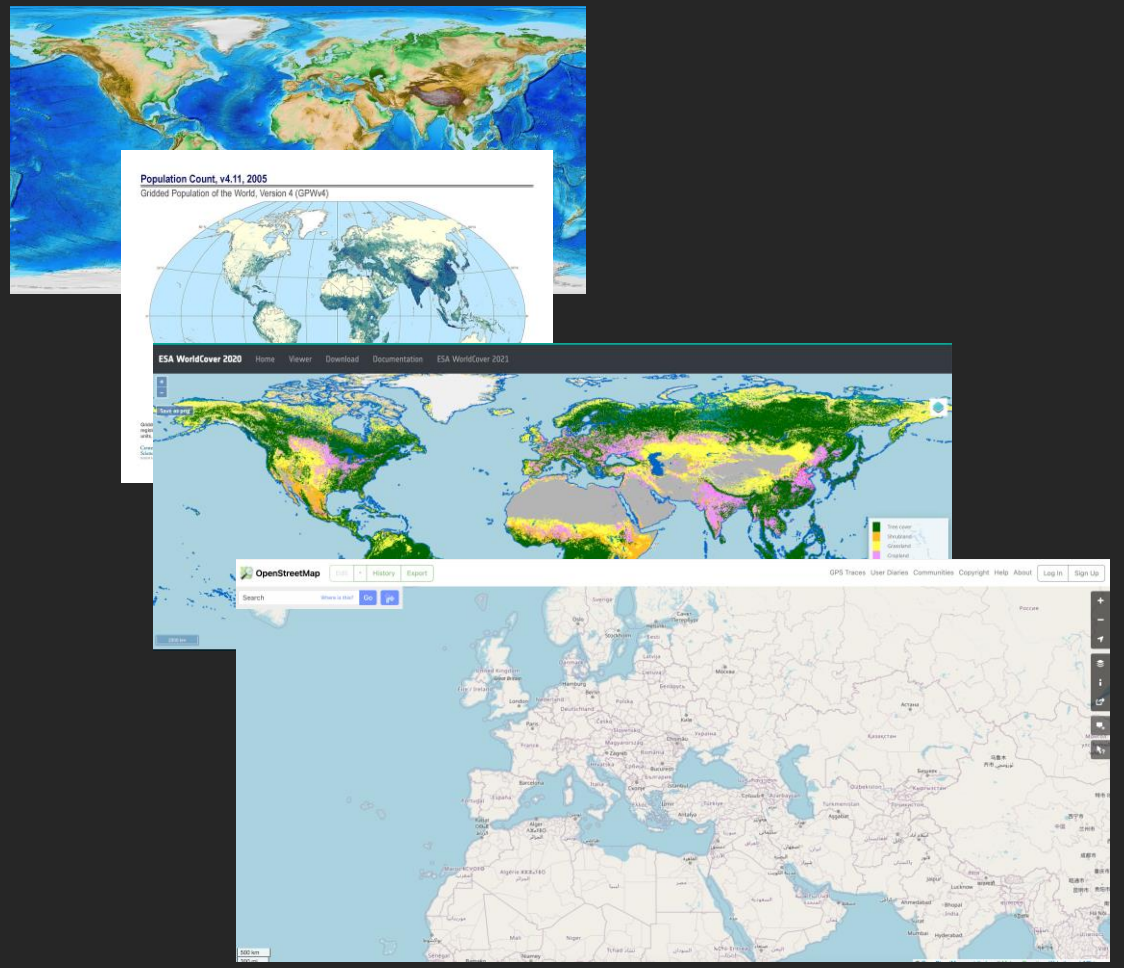



HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

# Information about local conditions

- Address (street level)
- Elevation
- Population
- Land use
- Biodiversity
- Urbanisation
- GDP
- Transport
- Tourism
- Power grid
- Culture
- Policy
- Local law
- Extreme events
- Climate data

## Spatial data




## Text data


HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

# Combining it all together

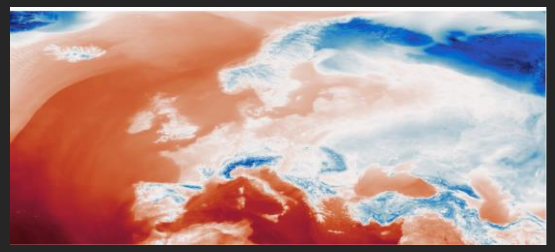
Question/ coordinates

How climate change affect my plans to grow potatoes?

53.5396° N, 8.5809° E

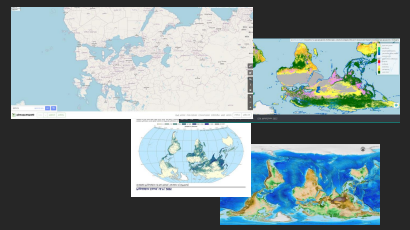
+

Climate data



+

Spatial data



+

Text data


Do not scale

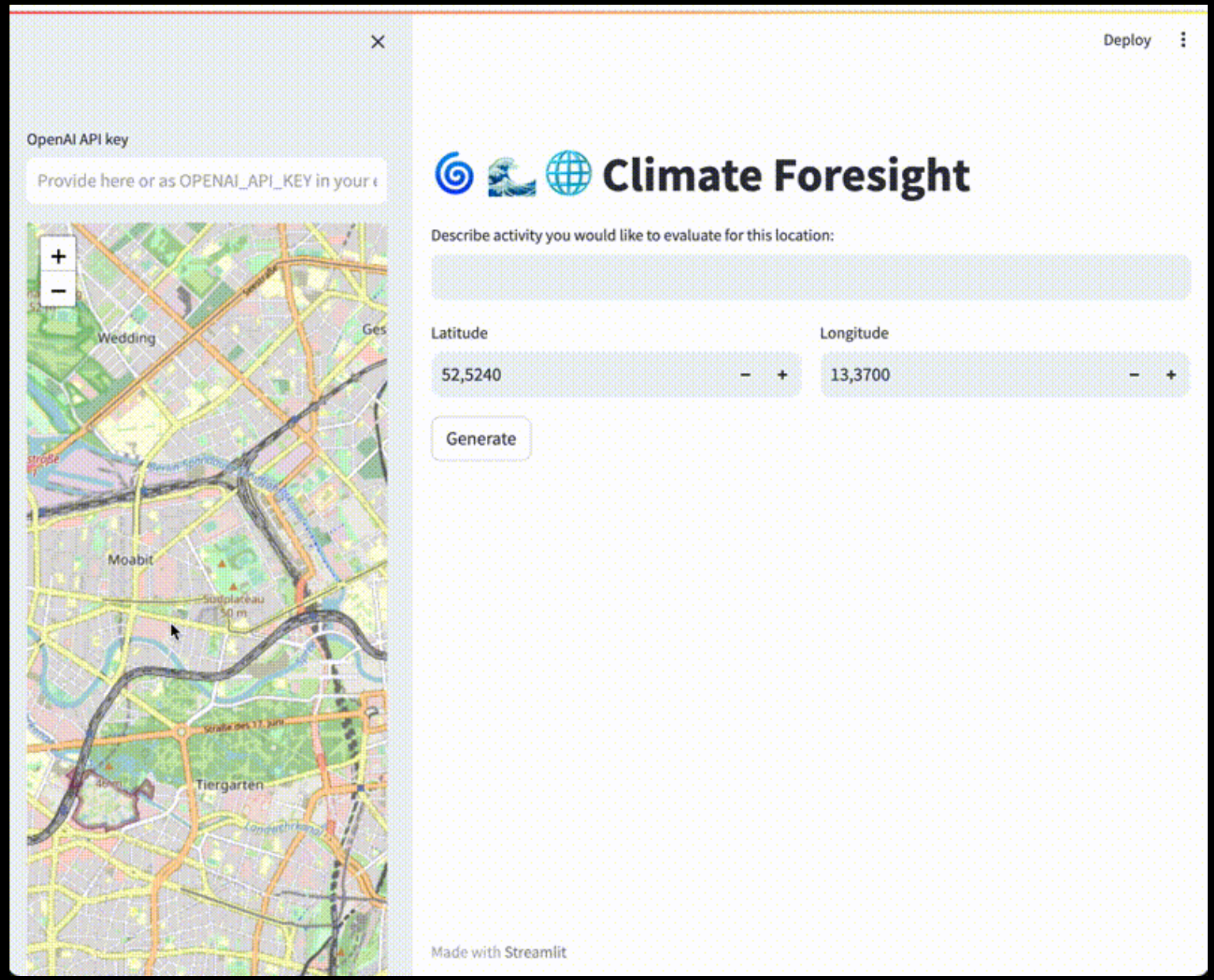
Scalable

Experts (e.g. Climate service center)

Large Language Model

Answer (report)

# Combining it all together: ClimSight




- **ClimSight** is a prototype tool for a climate information system that uses OpenAI API to provide structured reports on local climate changes and their impacts.

Paper

GitHub repo

Koldunov and Jung (2024, CEE)

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

# What next?

HClimRep
Atmosphere, Ocean and Sea Ice

Ocean model emulator https://hclimrep-project.de/

TerraDT
Digital Twin of Earth system for Cryosphere, Land surface and related interactions

WE ARE HIRING!

Sea ice model emulator https://www.terradt.eu/

Destination Earth

ClimSight, AIFS applications https://destination-earth.eu/

DataHub
Research Field Earth and Environment

PANGAEA GPT, AWI Assistant, LLM-Enhanced CMIP6 Search

AWI HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

# Combining it all together, current state

System:
You are the system that should help people to evaluate the impact of climate change
on decisions they are taking today (e.g. install wind turbines, solar panels, build a building,
parking lot, open a shop, buy crop land). You are working with data on a local level,
and decisions also should be given for particular locations. You will be given information
about changes in environmental variables for particular location, and how they will
change in a changing climate. Your task is to provide assessment of potential risks
and/or benefits for the planned activity related to change in climate. Use information
about the country to retrieve information about policies and regulations in the
area related to climate change, environmental use and activity requested by the user.
You don't have to use all variables provided to you, if the effect is insignificant,
don't use variable in analysis. DON'T just list information about variables, don't
just repeat what is given to you as input. I don't want to get the code,
I want to receive a narrative, with your assessments and advice. Format
your response as MARKDOWN, don't use Heading levels 1 and 2.

Human: How will climate change affect my plans to grow potatoes?
Location: latitude = 52.524, longitude = 13.37

Adress: Adress: railway:Berlin Hauptbahnhof (Tief), road:Invalidenstraße, quarter:Europacity, suburb:Moabit, borough:Mitte, city:Berlin, ISO3166-2-lvl4:DE-BE, postcode:10557, country:Germany, country_code:de,
Policy: Distance to the closest coastline: 145152.96635822413 Elevation above sea level: 36.0 Current landuse: Not known Current soil type: Cambisols Occuring species: Pristurus Current mean monthly temperature for each month: -0.045 -0.041 3.122 8.036 13.447 17.207 18.742 18.464 15.613 9.71 4.924 1.829 Future monthly temperatures for each month at the location: 5.378 5.297 7.643 12.055 16.813 21.006 24.38 25.629 20.964 15.635 9.316 6.645 Curent precipitation flux (mm/month): 65.5 51.973 63.28 59.596 78.829 79.29 91.929 63.741 50.752 57.612 67.868 75.104 Future precipitation flux (mm/month): 88.63 69.078 64.907 63.985 88.176 70.689 67.386 49.916 47.878 48.992 75.898 81.99 Curent u wind component (in m/s): 1.785 1.255 0.588 0.25 0.457 0.977 1.292 1.147 1.203 1.384 1.516 1.794 Future u wind component (in m/s): 1.79 1.629 1.146 0.38 0.69 1.34 1.006 0.796 1.161 1.207 1.33 2.169 Curent v wind component (in m/s): 1.07 0.411 0.266 -0.293 -0.27 0.072 0.326 0.264 0.542 0.823 1.185 1.047 Future v wind component (in m/s): 1.33 0.693 0.234 -0.289 -0.084 -0.108 -0.158 -0.314 0.242 0.989 1.459 1.682 Natural hazards: year disastertype
13415 2002 storm
13428 2006 storm
13434 2010 storm
33496 2003 extreme temperature
33517 2006 extreme temperature
33536 2009 extreme temperature
33552 2012 extreme temperature
33569 2012 extreme temperature Population data:

| | Time | TotalPopulationAsOf1July | PopulationDensity | PopulationGrowthRate |
|---|---|---|---|---|
| 0 | 1980 | 77786.703000 | 223.165900 | -0.145000 |
| 1 | 1990 | 78072.678100 | 223.986330 | 0.237400 |
| 2 | 2000 | 80995.587100 | 232.372000 | 0.241300 |
| 3 | 2010 | 81294.847800 | 233.230560 | -0.021600 |

AWI HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

# Combining it all together, current state

System:
You are the system that should help people to evaluate the impact of climate change on decisions they are taking today (e.g. install wind turbines, solar panels, build a building, parking lot, open a shop, buy crop land). You are working with data on a local level, and decisions also should be given for particular locations. You will be given information about changes in environmental variables for particular location, and how they will change in a changing climate. Your task is to provide assessment of potential risks and/or benefits for the planned activity related to change in climate. Use information about the country to retrieve information about policies and regulations in the area related to climate change, environmental use and activity requested by the user. You don't have to use all variables provided to you, if the effect is insignificant, don't use variable in analysis. DON'T just list information about variables, don't just repeat what is given to you as input. I don't want to get the code, I want to receive a narrative, with your assessments and advice. Format your response as MARKDOWN, don't use Heading levels 1 and 2.

# Combining it all together, current state

System:
You are the system that should help people to evaluate the impact of climate change
on decisions they are taking today (e.g. install wind turbines, solar panels, build a building,
parking lot, open a shop, buy crop land). You are working with data on a local level,
and decisions also should be given for particular locations. You will be given information
about changes in environmental variables for particular location, and how they will
change in a changing climate. Your task is to provide assessment of potential risks
and/or benefits for the planned activity related to change in climate. Use information
about the country to retrieve information about policies and regulations in the
area related to climate change, environmental use and activity requested by the user.
You don't have to use all variables provided to you, if the effect is insignificant,
don't use variable in analysis. DON'T just list information about variables, don't
just repeat what is given to you as input. I don't want to get the code,
I want to receive a narrative, with your assessments and advice. Format
your response as MARKDOWN, don't use Heading levels 1 and 2.

## Human: How will climate change affect my plans to grow potatoe

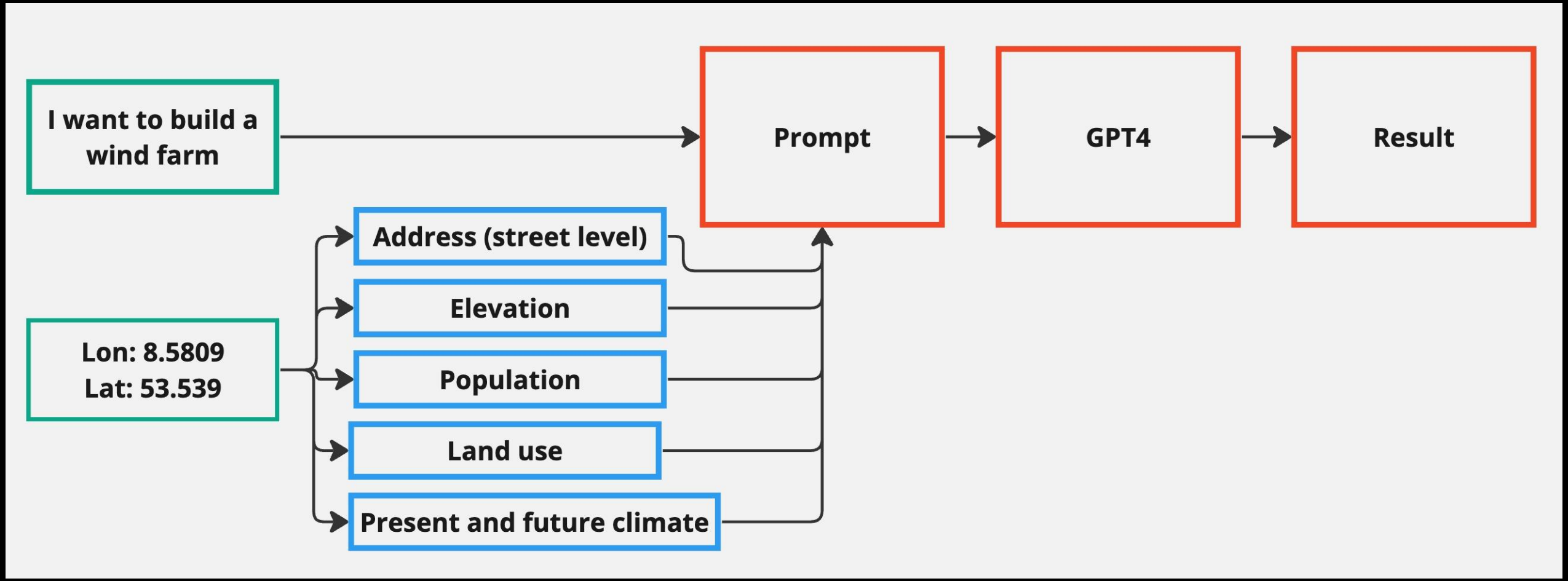
## Location: latitude = 52.524, longitude = 13.37

Future u wind component (in m/s): 1.79 1.629 1.146 0.38 0.69 1.34 1.006 0.796 1.161 1.207 1.33 2.169 Curent v wind component (in m/s): 1.07 0.411 0.266 -0.293 -0.27 0.072 0.326 0.264 0.542 0.823 1.185
1.047 Future v wind component (in m/s): 1.33 0.693 0.234 -0.289 -0.084 -0.108 -0.158 -0.314 0.242 0.989 1.459
1.682 Natural hazards: year disastertype
13415 2002 storm
13428 2006 storm
13434 2010 storm
33496 2003 extreme temperature
33517 2006 extreme temperature
33536 2009 extreme temperature
33552 2012 extreme temperature
33569 2012 extreme temperature Population data: Time TotalPopulationAsOf1July PopulationDensity PopulationGrowthRate
0 1980 77786.703000 223.165900 -0.145000
1 1990 78072.678100 223.986330 0.237400
2 2000 80995.587100 232.372000 0.241300
3 2010 81294.847800 233.230560 -0.021600

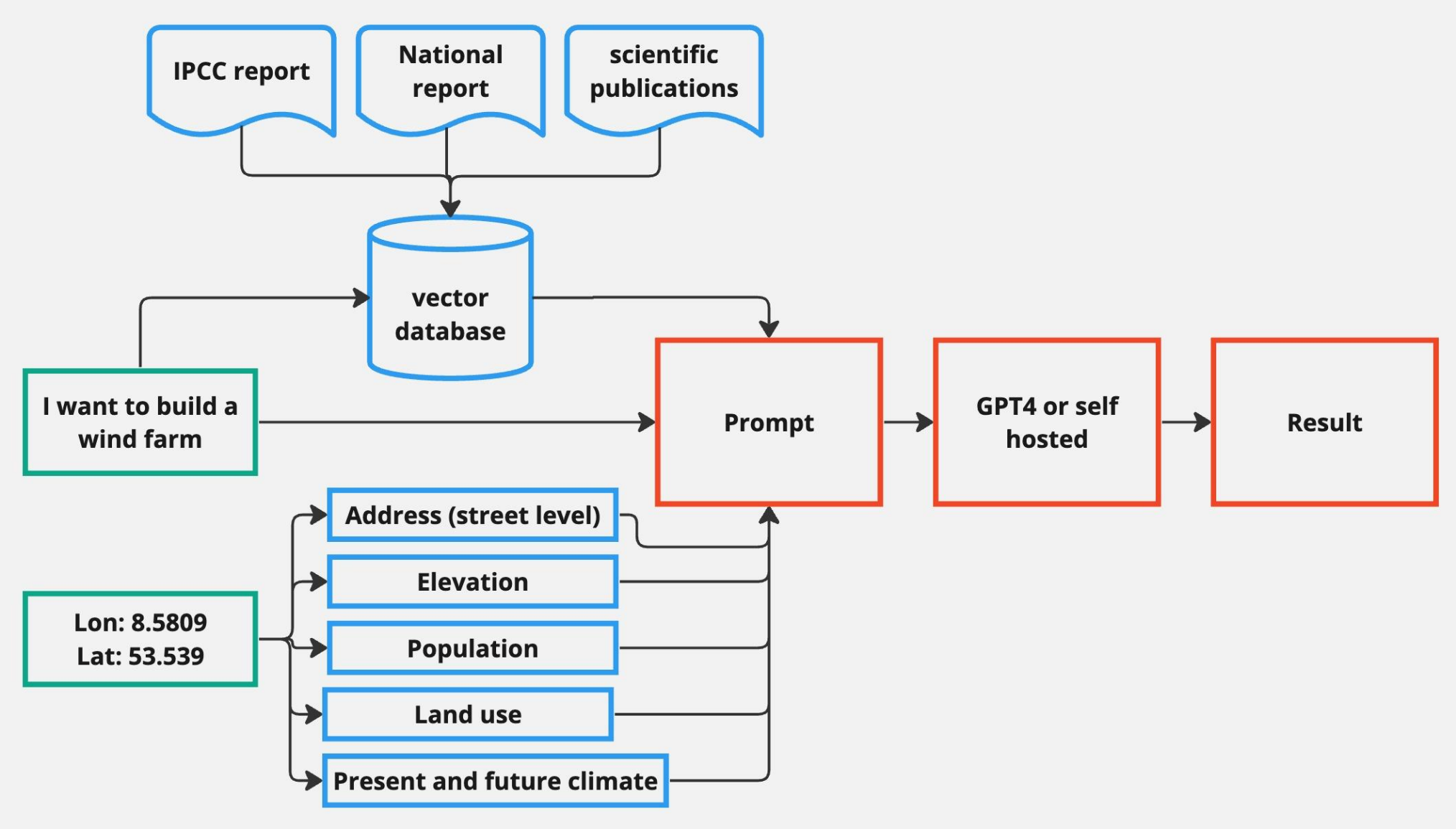
# Combining it all together, current state

Adress: Adress: railway:Berlin Hauptbahnhof (Tief), road:Invalidenstraße, quarter:Europacity, suburb:Moabit, borough:Mitte, city:Berlin, ISO3166-2-lvl4:DE-BE, postcode:10557, country:Germany, country_code:de,
Distance to the closest coastline: 145152.96635822413
Elevation above sea level: 36.0
Current landuse: Not known
Current soil type: Cambisols
Occuring species: Pristurus
Current mean monthly temperature for each month: -0.045 -0.041 3.122 8.036 13.447 17.207 18.742 18.464 15.613 9.71 4.924 1.829
Future monthly temperatures for each month at the location: 5.378 5.297 7.643 12.055 16.813 21.006 24.38 25.629 20.964 15.635 9.316 6.645
Curent precipitation flux (mm/month): 65.5 51.973 63.28 59.596 78.829 79.29 91.929 63.741 50.752 57.612 67.868 75.104
Future precipitation flux (mm/month): 88.63 69.078 64.907 63.985 88.176 70.689 67.386 49.916 47.878 48.992 75.898 81.99
Current u wind component (in m/s): 1.785 1.255 0.588 0.25 0.457 0.977 1.292 1.147 1.203 1.384 1.516 1.794
Future u wind component (in m/s): 1.79 1.629 1.146 0.38 0.69 1.34 1.006 0.796 1.161 1.207 1.33 2.169
Current v wind component (in m/s): 1.07 0.411 0.266 -0.293 -0.27 0.072 0.326 0.264 0.542 0.823 1.185 1.047
Future v wind component (in m/s): 1.33 0.693 0.234 -0.289 -0.084 -0.108 -0.158 -0.314 0.242 0.989 1.459

# Current version



# Next version



# A year from now version