Next-generation climate modeling with new tools: leveraging the AI revolution

Pierre Gentine
Columbia Engineering & Climate School
Director, LEAP Center
USMILE ERC Synergy Grant co-PI

Outline

1. Introduction: climate uncertainties and the role of subgrid processes

2. AI: from climate process emulation to new discoveries

3. Solutions and challenges for next-generation climate models

4. A vision for next-generation climate models

Current gap: Climate Adaptation is Needed


Climate change is fueling deadly heat waves in India. It's putting the country's development at risk, study says


The New York Times

Spain Bakes in Summer-Like Heat, and Worries About What Comes Next

The April temperatures, over 100 degrees Fahrenheit in some places, come on top of a long-running drought that has depleted reservoirs and dried up fields.

But adaptation requires:

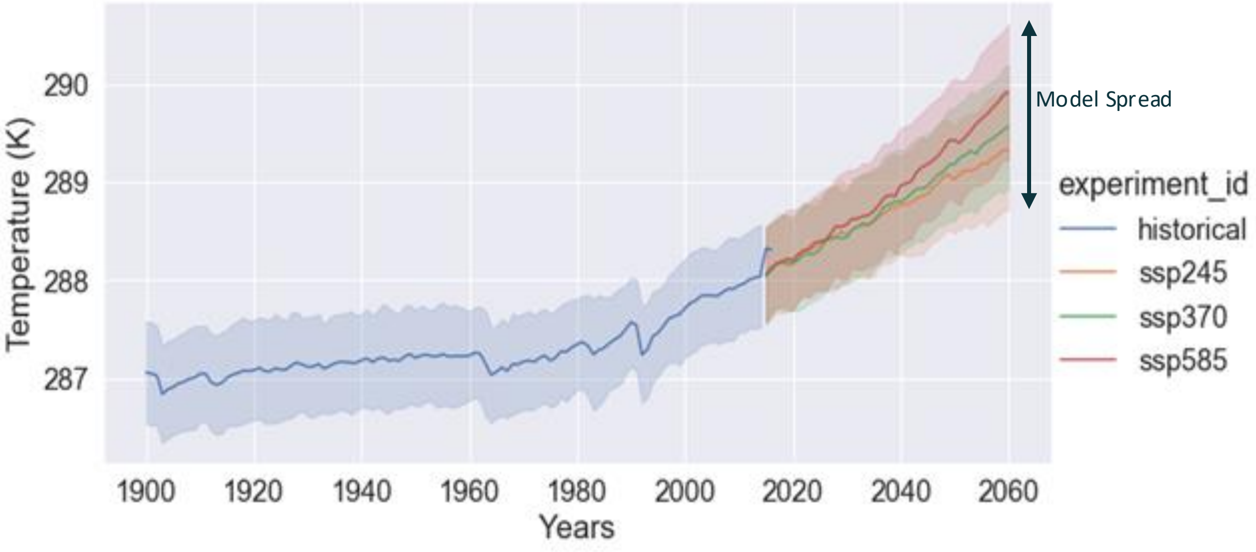
- Accurate information (projection) about the future using climate models
- Broadly and easily accessible climate projection information (outreach, LEAP-Pangeo platform not covered today)

Flood Fort Lauderdale 2023, Extreme heatwave in India and Spain (April 2023), Canadian wildfire smoke NYC (June 2023)

Yet, current projections are still too uncertain

Even for (simple) global metrics such as surface temperature

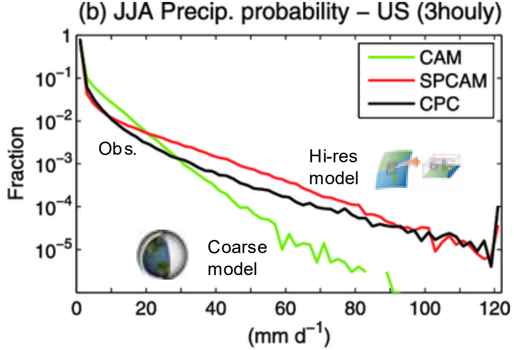
Global Surface Air Temperature - CMIP6 (New Generation)



*note it takes 5 minutes to plot this with modern cloud data infrastructure (LEAP-Pangeo)

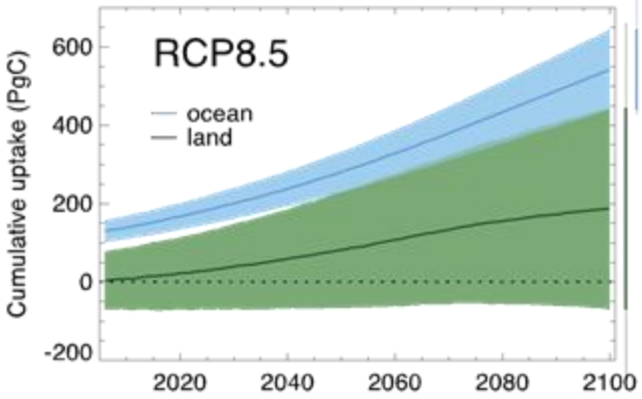
Regional climate and carbon cycles are even more uncertain

- Regional precipitation is quite inaccurate, especially for extremes



standard coarse atmospheric model
superparameterized CAM (kilometer-scale model resolving deep clouds)
observations

- Future of carbon cycle

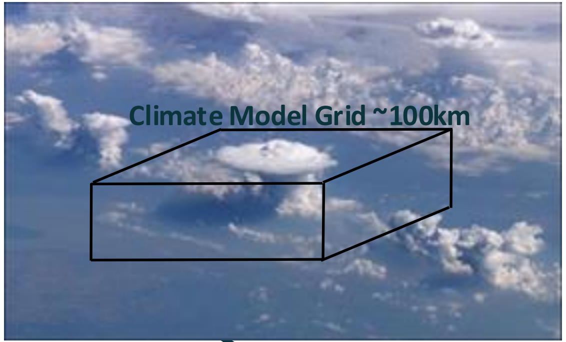


Collins et al., 2011 JAMES; Friedlingstein et al., 2014 J Climate

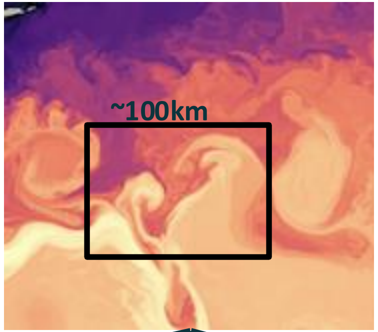
Unresolved or Unknown processes Require "**Parameterizations**"
Causing Projection Uncertainties

Physical
+ Biological
Processes

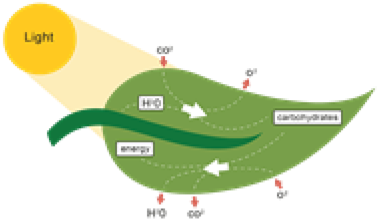
Clouds

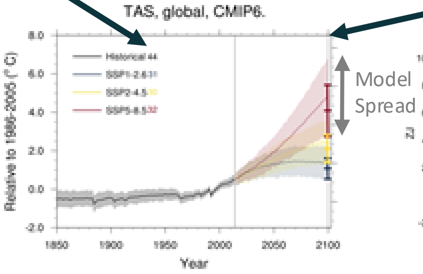


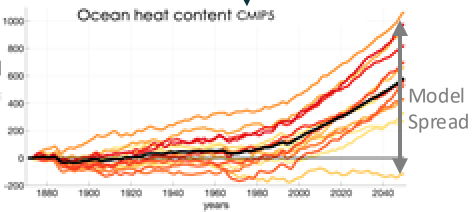
Ocean Eddies

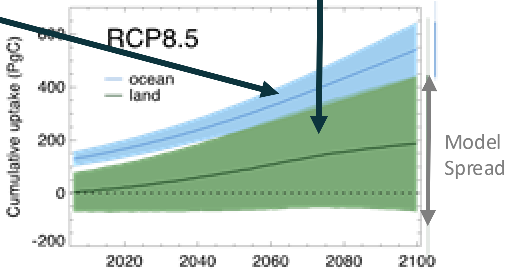


Photosynthesis









- Model errors dominate (>50%) uncertainties <50 years

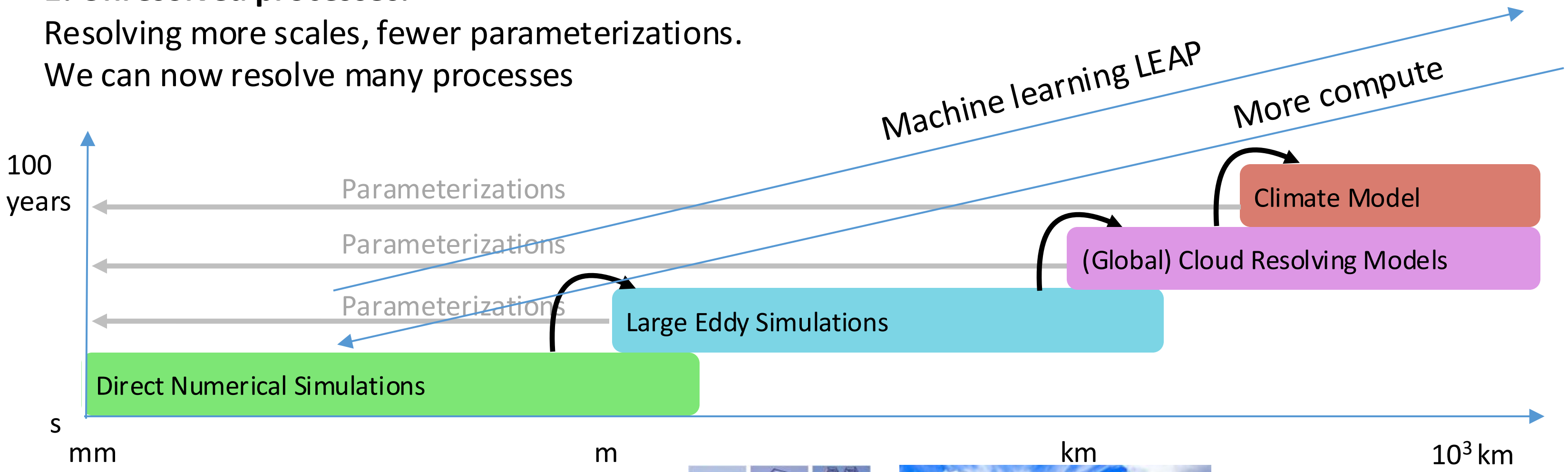
Gentine et al., 2018 Geo Res Lett, Newsom, Zanna et al., 2020 Geo Res Lett; Friedlingstein et al. 2006 J Climate; Green, Seneviratne, Gentine, Nature 2019

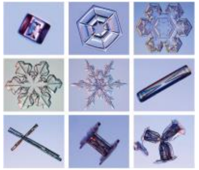
Strategies to improve climate modeling

1. Unresolved processes:

Resolving more scales, fewer parameterizations.

We can now resolve many processes





2. Unknown processes:

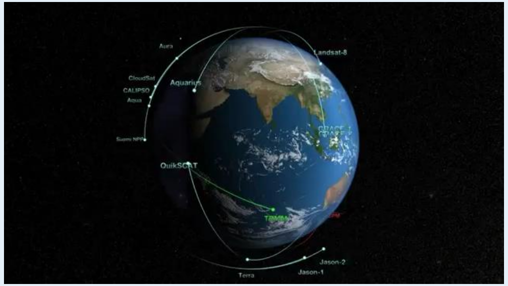
Many processes cannot be simulated: microphysics, biogeochemistry

→ Use observations (in situ, remote sensing) to learn processes

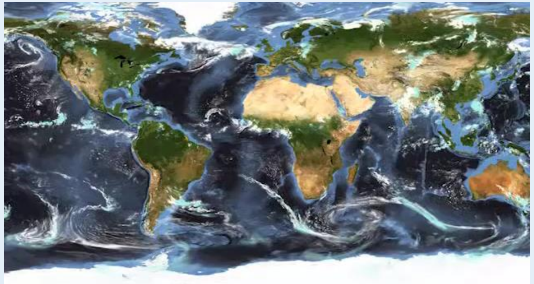
Gentine et al. 2021; Zanna et al., 2021 (Deep Learning for the Earth Sciences Linking Physics and Deep Learning Models)

Harvesting the Data Revolution to improve parameterizations

1. Massive data from Earth observation



2. High-resolution simulations



3. Progress in machine learning

Gentine et al. 2021; Zanna et al., 2021 (Deep Learning for the Earth Sciences Linking Physics and Deep Learning Models)

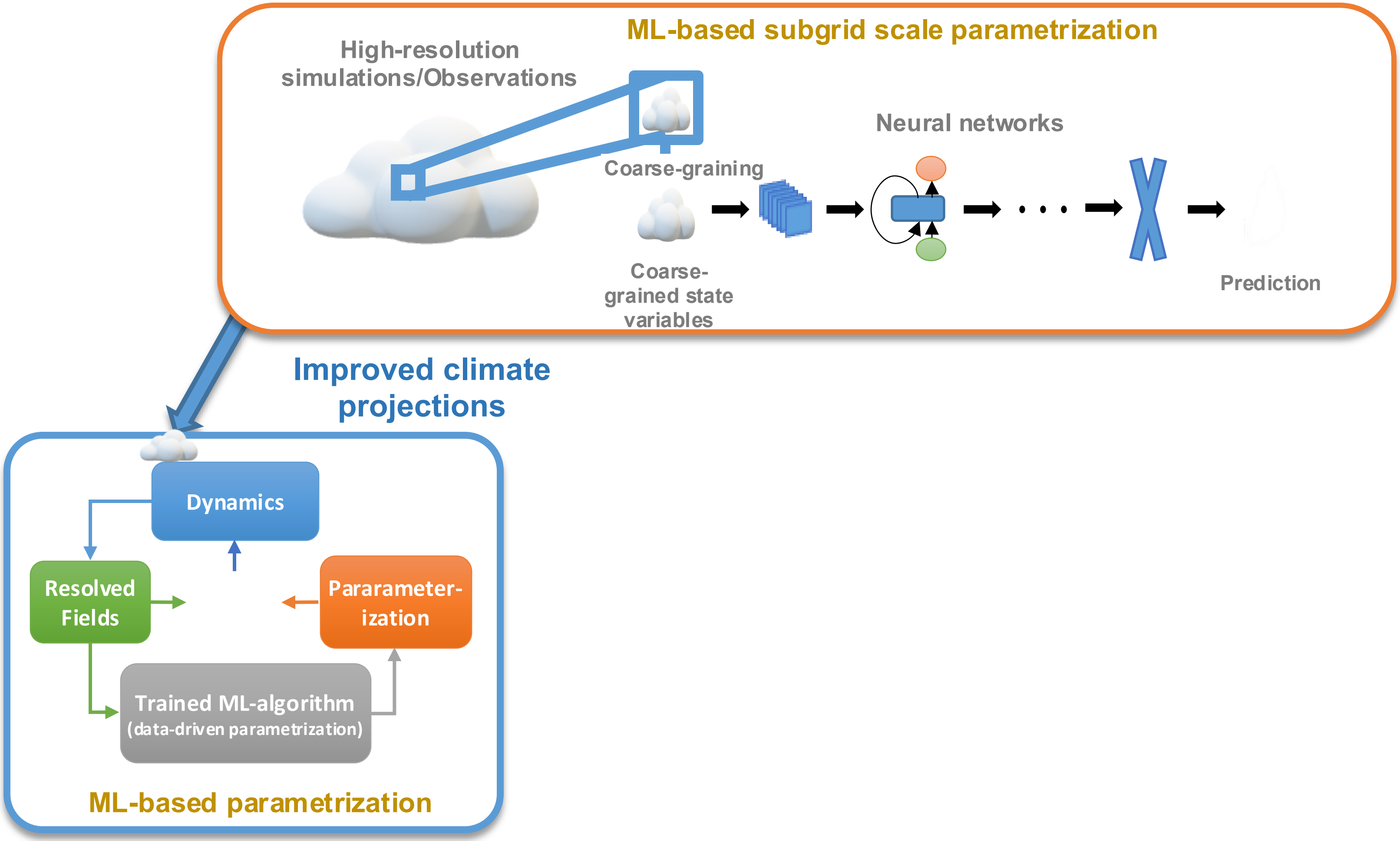
Outline

1. Introduction: climate uncertainties and the role of subgrid processes

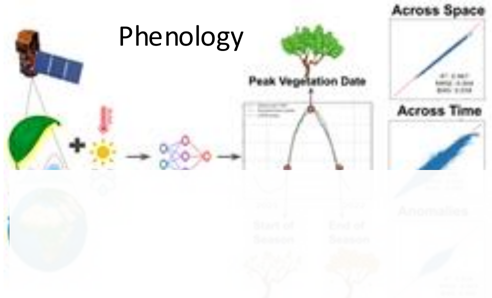
2. AI: from climate process emulation
to new discoveries

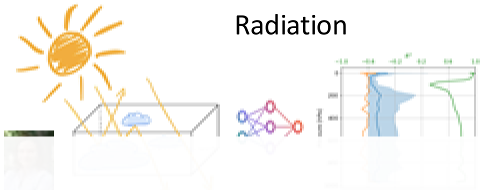
3. Solutions and challenges for next-generation climate models

4. A vision for next-generation climate models


ML-based subgrid scale parametrization
High-resolution simulations/Observations
Coarse-graining
Neural networks
Coarse-grained state variables
Prediction
Improved climate projections
Dynamics
Resolved Fields
Pararameter-ization
Trained ML-algorithm
(data-driven parametrization)
ML-based parametrization

Modeling the Climate System

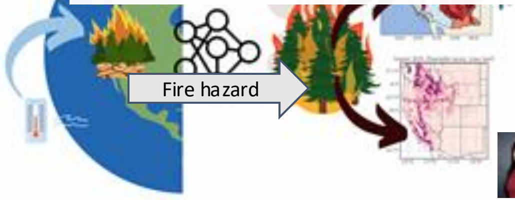


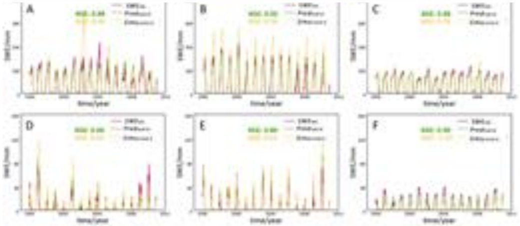


None of those processes

(besides deep convection)

are resolved in high-resolution cloud-resolving models!

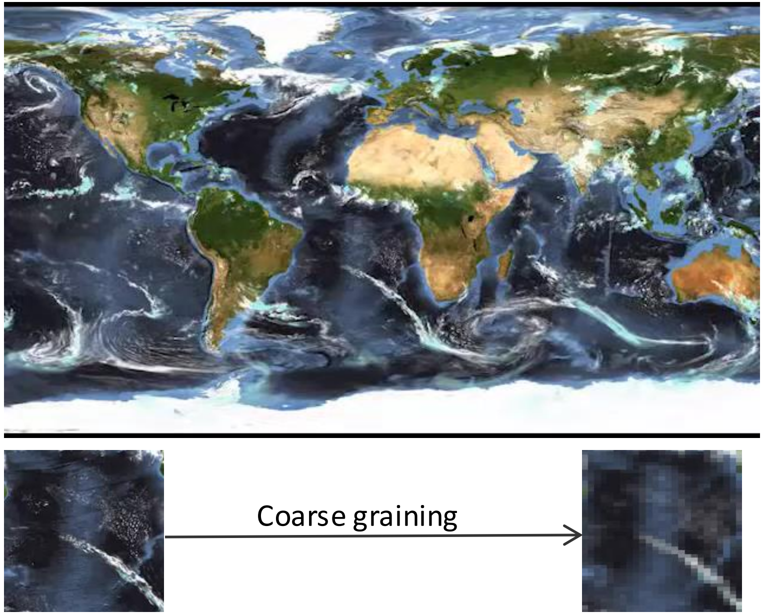




Zhao et al. 2019 GRL; El Ghawi et al 2023 ERL; Buch et al. 2023 GMD; Shamekh et a; JAMES in review; Lahlou et al. in prep

Learn subgrid convection parameterization with machine learning

- Main strategy: use high-resolution cloud-resolving model (few kms) data as “training” for closure development (supervised learning) at coarse resolution
- Multi-institution, inter-disciplinary, international efforts: LEAP/USMILE ERC/M^{2}LInES

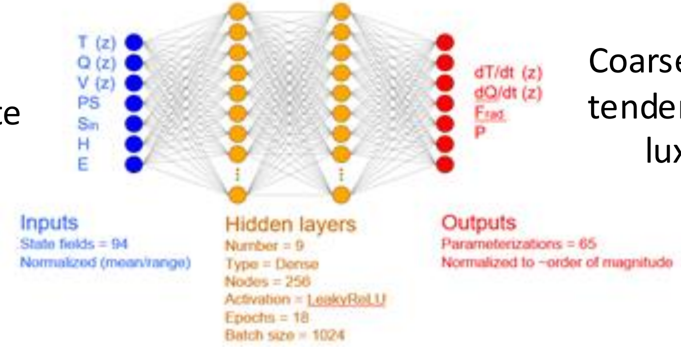


Gentine et al 2018 GRL; Rasp et al 2018 PNAS; Gentine et al. 2021; Zanna et al., 2021 (Deep Learning for the Earth Sciences Linking Physics and Deep Learning Models)

Learn subgrid convection closures with machine learning

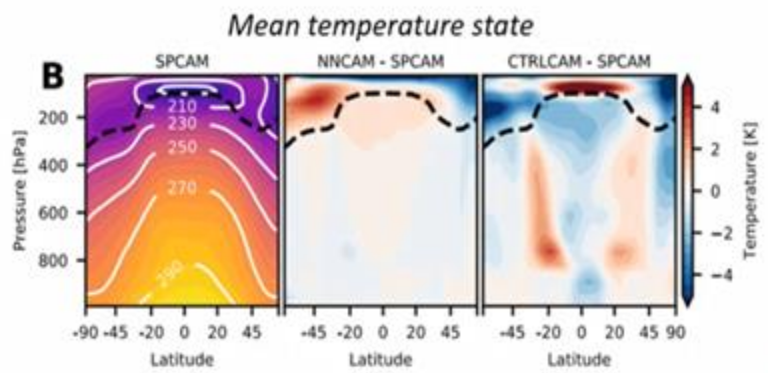
- Main strategy: use high-resolution cloud-resolving model (few kms) data as “training” for closure development (supervised learning) at coarse resolution

Coarse model state



Coarse-scale tendencies/fluxes

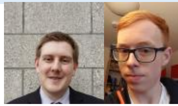
Online Aquaplanet (NCAR CAM)



Temperature profiles

Gentine et al 2018 GRL; Rasp et al 2018 PNAS; Gentine et al. 2021; Zanna et al., 2021 (Deep Learning for the Earth Sciences Linking Physics and Deep Learning Models)

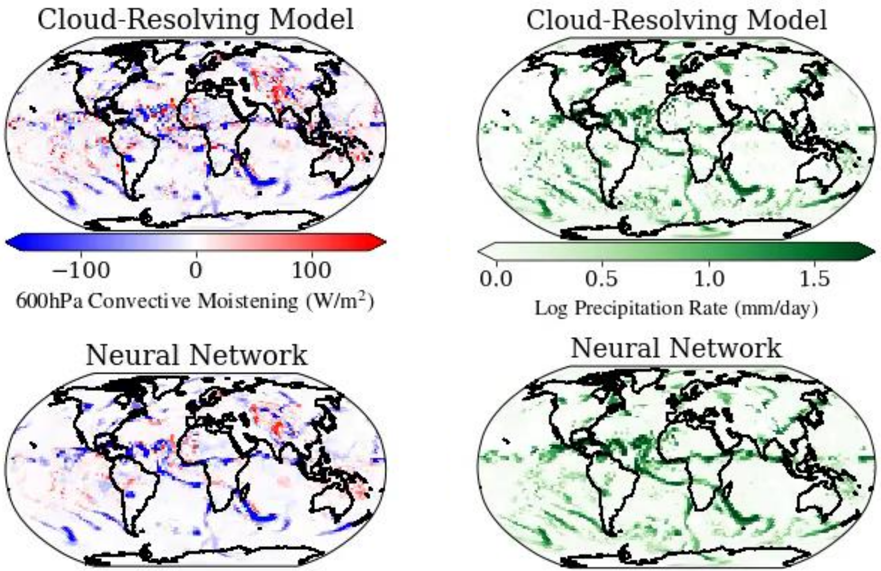
Successful Emulation of Convection on “*real geography*”

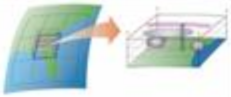


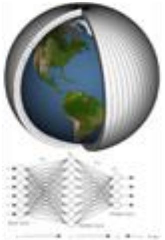
Real geography results in *Offline* Community Atmosphere Model (CAM)
Online in ICON with Max Planck/DLR

Truth
Super-param.
(SPCAM)
simulation

Prediction
NN







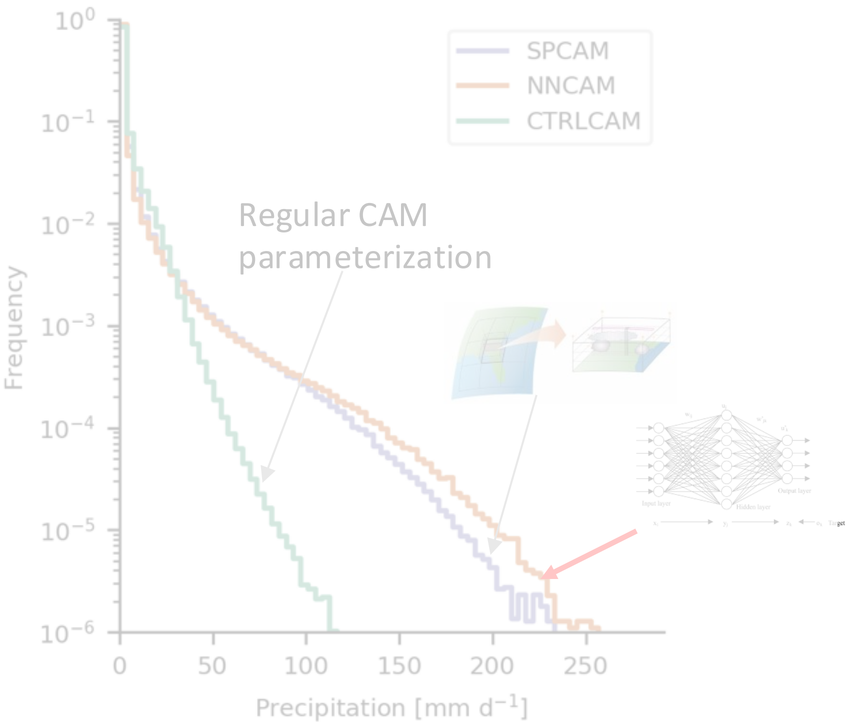
Skillful convection emulation with continents

Mooers et al. (2021) *JAMES;* Behrens et al *JAMES 2022, 2023;* Heuer et al. 2023 submitted

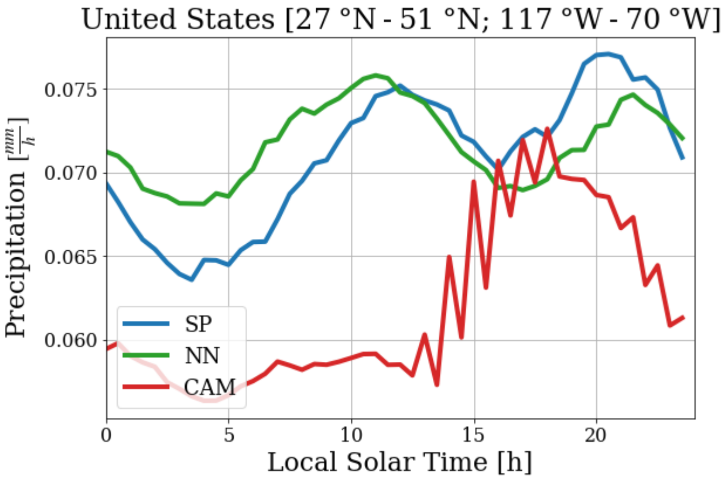
Inclusion of ML parameterizations leads to **Step-change** in climate model quality

Online global simulations 🌎

Precipitation distribution (CAM)

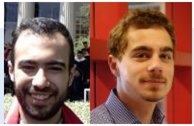


Precipitation diurnal cycle (CAM+ICON)



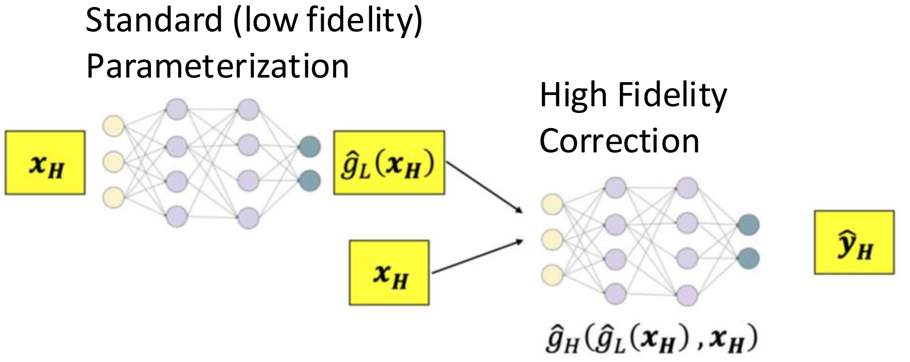
Rasp, Pritchard, Gentine., 2018 PNAS; Behrens, Eyring, Gentine submitted

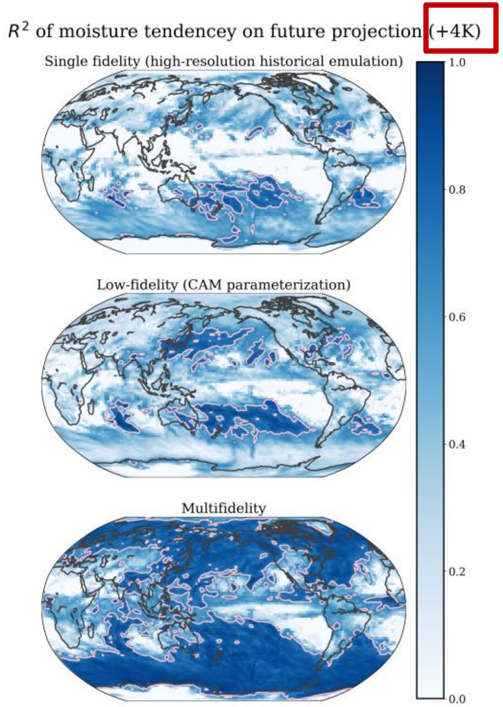
Main Challenge for ML: Generalization



- Challenge: out-of-distribution prediction: climate change
- Solution: **Hybrid Multifidelity approach**: merges physics with ML to extrapolate
  (online with MeteoFrance to improve subseasonal/seasonal forecast)

→ Best of both worlds



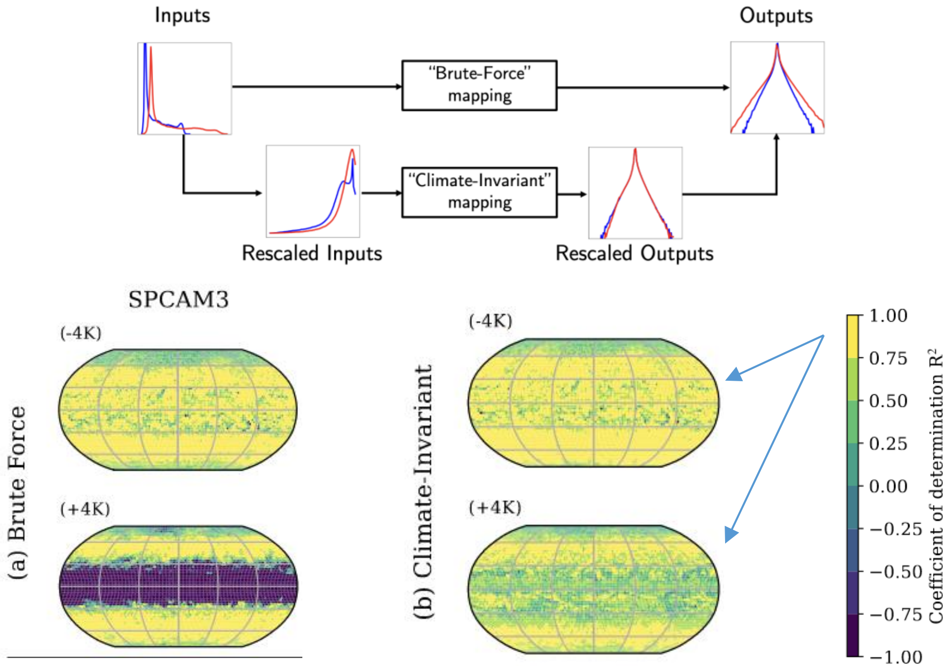


Other strategies:
- embedding physical invariances
- merging causality & ML

Bhouri, Pritchard, Gentine, in review @ Science Advances; Beucler, Pritchard,..., Gentine., 2023 in revision @ Sci Advances, Connolly, Yu, Gentine, 2023 submitted; Iglesiaz-Suarez, Eyring and Gentine., submitted

Challenge 1: Generalization

- Challenge: out-of-sample, out-of-distribution prediction/sampling bias
- Solution: **embed physical invariances/equivariance** along Lie groups in ML (e.g., dimensionless numbers or rotational equivariance to collapse distributions)

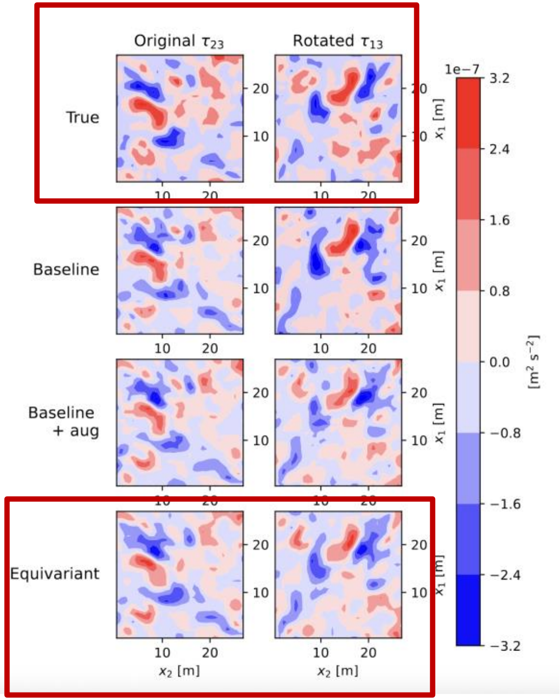


Beucler, Pritchard,…, Gentine., 2023 in revision @ Sci Advances, Connolly, Yu, Gentine, 2023 submitted

Challenge 1: Generalization

- Challenge: out-of-sample, out-of-distribution prediction/sampling bias
- Solution: **embed physical invariances/equivariance** along Lie groups in ML (e.g., dimensionless numbers or rotational equivariance to collapse distributions)

$$f(\rho_{\text{in}}(g)x^*) = f(x^*)$$

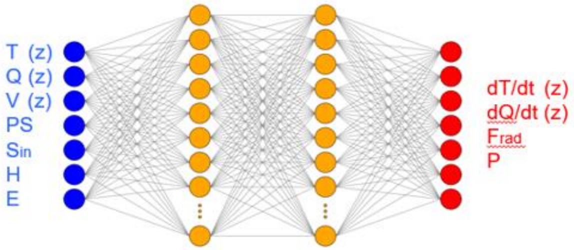


Connolly, Yu, Gentine, submitted

Challenge 2: Trust + interpretability

- Challenge: cross-correlations can fool ML
- Solution: merge **causal discovery with ML**
  → more interpretable, more trustworthy

NN (Rasp et al., 2018)



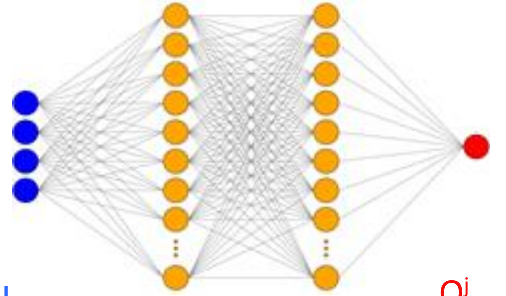
$I_{N=94}$

Hidden layers
Number = 9
…

$O_{N=65}$

Drop connections based on causality: a very sparse NN

Causal-NNs



$I_{N=causal}$

Hidden layers
Number = 9

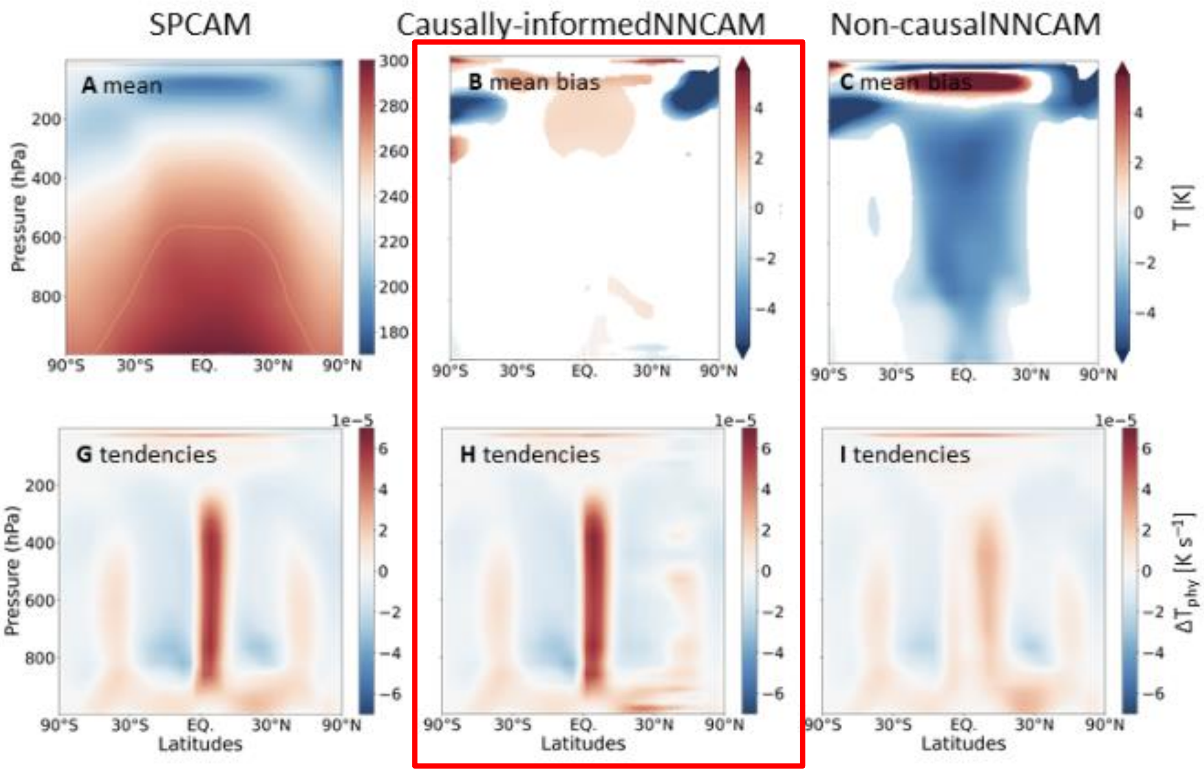
$O^j_{(N=65)}$

$$I^i_{causal} \rightarrow O^j \in P(O^j)$$

Iglesiaz-Suarez, Eyring and Gentine., submitted

Causality: going beyond correlations

Online global simulations



Iglesiaz-Suarez et al., submitted

Journey through the use of ML in climate science

Journey through ML use: a roller coaster 🎢

Initial work:

- Correct emulation of convection 🤘 🍾 🎉
- But cannot generalize, may lack stability, misses physical invariances 🙈 😬 😫

Recent work focusing on:

- Embedding physical invariances improves model stability, including causality and generalization 👍

More fundamentally, what did we (really) learn? 🤷‍♂️

- Using machine learning for new **discoveries** 🤩

Rasp, Pritchard, Gentine., 2018 PNAS

1. Introduction: climate uncertainties and the role of subgrid processes

2. AI: from climate process emulation
to new discoveries

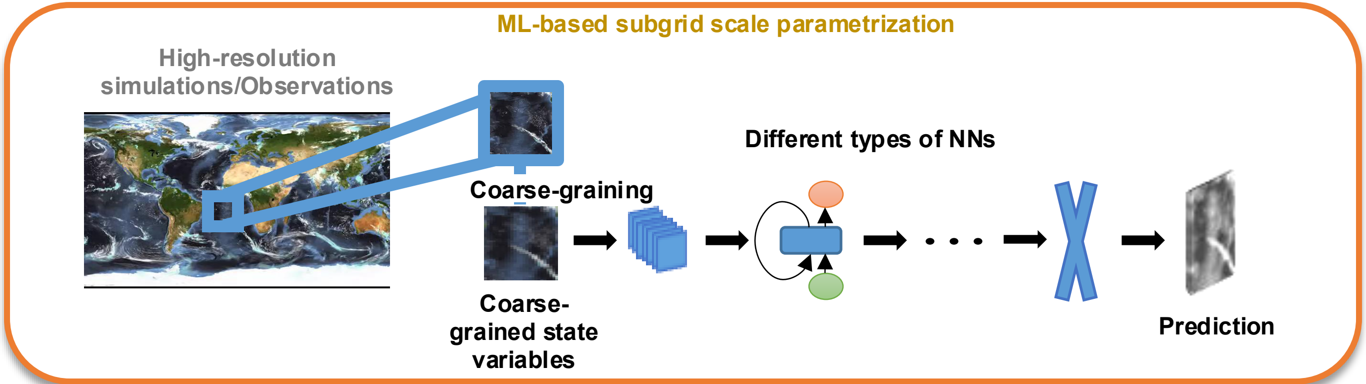
3. Solutions and challenges for next-generation climate models

4. A vision for next-generation climate models

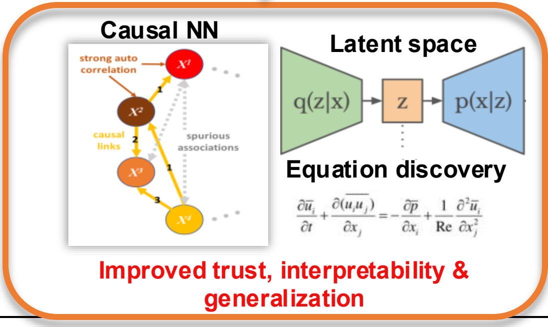
22

From emulation to understanding

Data analysis is extremely challenging in high-res simulations and remote sensing (size)



Improved Earth system understanding/ discoveries



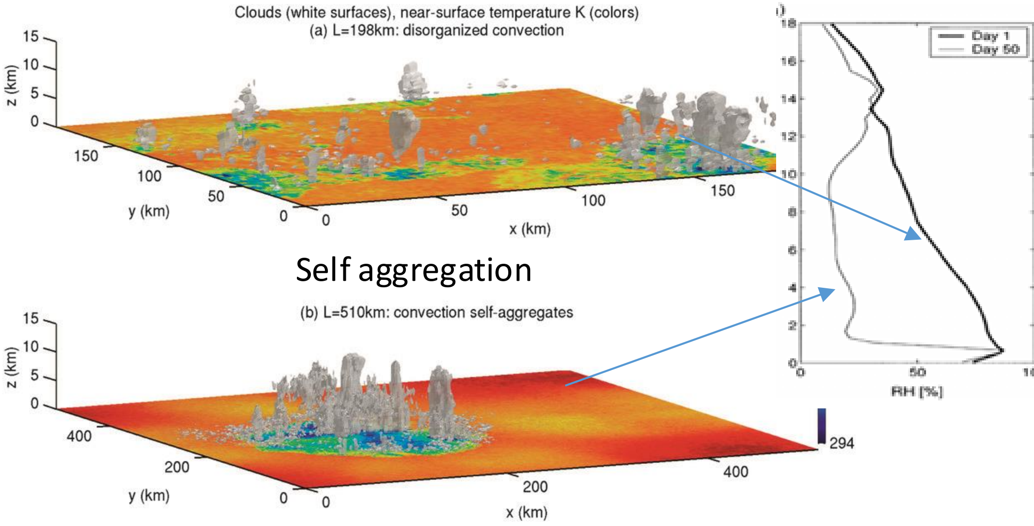
○ Convection can **aggregate**

○ Aggregation has a large impact on:

- Humidity
- Radiative cooling
  → Enhances radiative cooling
  → Potentially impacts climate sensitivity
- Precipitation
  → Increases accumulated precipitation

Random convection

Self aggregation



Muller and Held 2012, JAS

Can we learn the implicit role of subgrid (micro) scale on precipitation?

- Science questions:
  - microscale
  - Does P = F (X coarse-scale, **subgrid scale stuff**) improve prediction?
  - Can we explain (some of) the **stochasticity**? Still unclear how to model it.

STOCHASTIC PARAMETERIZATION

Toward a New View of Weather and Climate Models

Judith Berner, Ulrich Achatz, Lauriane Batté, Lisa Bengtsson, Alvaro de la Cámara, Hannah M. Christensen, Matteo Colangeli, Danielle R. B. Coleman, Daan Crommelin, Stamen I. Dolaptchiev, Christian L. E. Franzke, Petra Friederichs, Peter Imkeller, Heikki Järvinen, Stephan Juricke, Vassili Kitsios, François Lott, Valerio Lucarini, Salil Mahajan, Timothy N. Palmer, Cécile Penland, Mirjana Sakradzija, Jin-Song von Storch, Antje Weisheimer, Michael Weniger, Paul D. Williams, and Jun-Ichi Yano

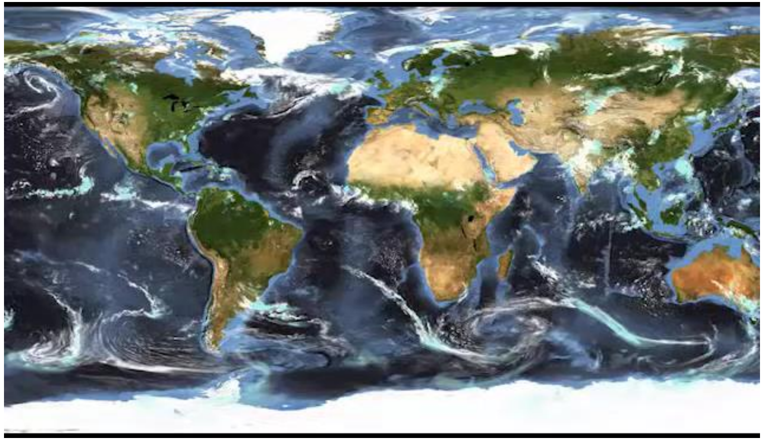
Stochastic parameterizations—empirically derived or based on rigorous mathematical and statistical concepts—have great potential to increase the predictive capability of next-generation weather and climate models.

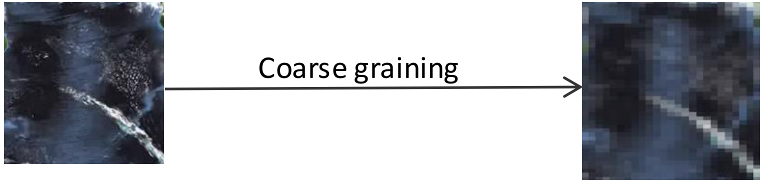
Using high-res models to learn the role of subgrid microscale on P

- Data:

DYAMOND Storm-resolving model experiment
Tropical band (20S-20N), ~2.5km resolution,
10 days of simulations
10^8 data points

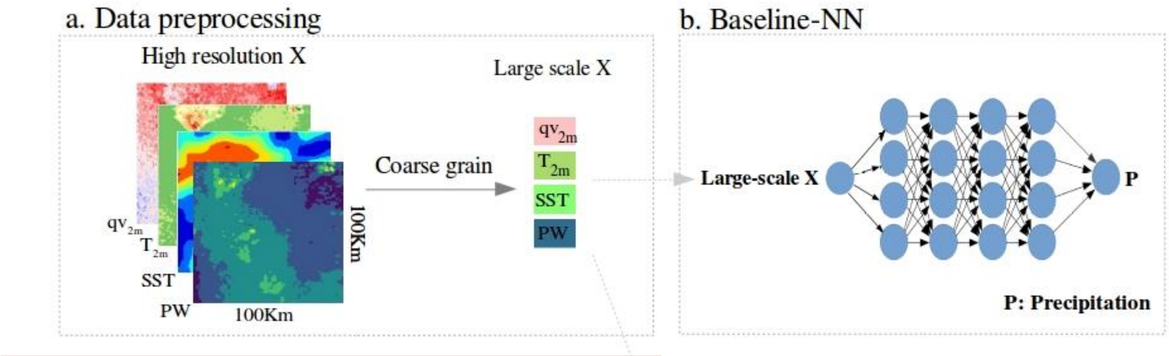
→ Predicting precipitation





Shamekh, Lamb, Huang and Gentine, PNAS 2023

- Strategy: Learning precipitation and organization (implicitly) in tandem



Learnt organization dramatically improves precipitation predictions

- Precipitation and its stochasticity are very well predicted with org

What did we learn?

✅ Organization regulates precipitation extremes and their prediction

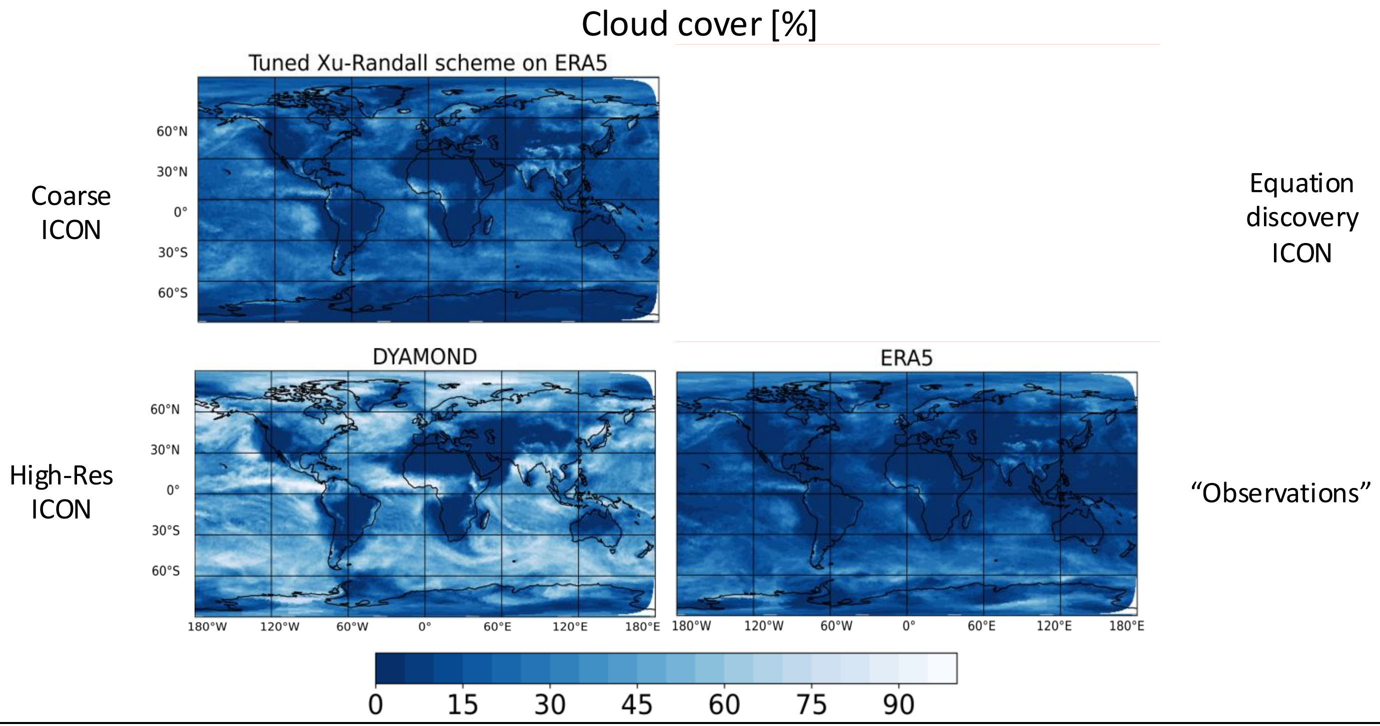
✅ Precipitation stochasticity is mainly due to convective aggregation

New discoveries 2: **Symbolic regression – equation discovery**

- Example for ICON (offline) **cloud cover**
- Projection on library of functions

$$C = P_3(RH, T) + (c_1 \partial_z RH + c_2)(\partial_z RH)^2 - \frac{1}{c_3 q_c + c_4 q_i + \varepsilon}$$

More interpretable, more trustworthy. Improve upon high-res ICON model.



Grundner, Beucler, Gentine , Eyring, *JAMES* 2023

Machine learning improves climate model processes and understanding

Machine learning can improve the representation of subgrid processes in climate models and **leap across scales**

ML closures are working now in full climate models (CAM/ICON) with leap in accuracy and major reduction in biases (implementation at a pace faster than climate change).
→ Not a hypothesis anymore.

Still some challenges – several of them discussed in this talk.

Machine learning can be used for **new discoveries** on big data.

Outline

1. Introduction: climate uncertainties and the role of subgrid processes

2. AI: from climate process emulation
to new discoveries

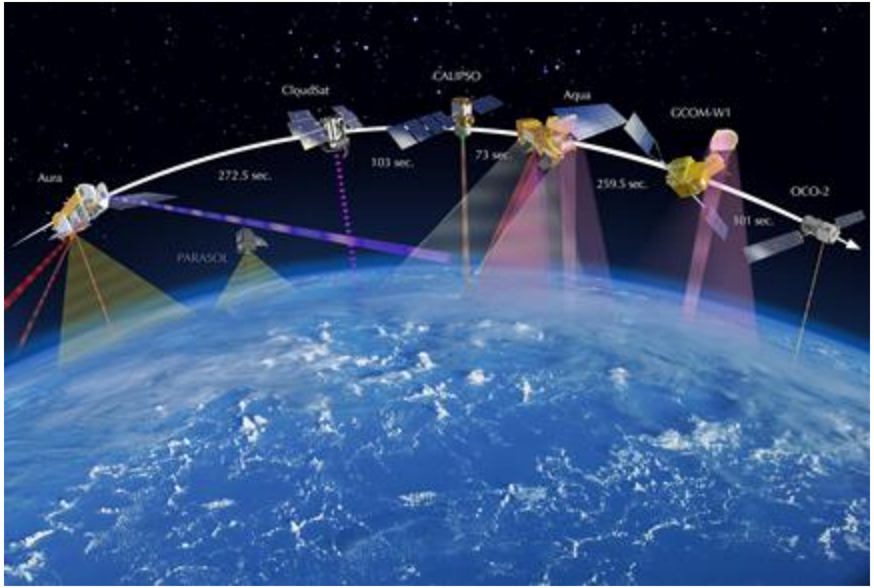
3. Solutions and challenges for next-generation climate models

4. A vision for next-generation climate models

31

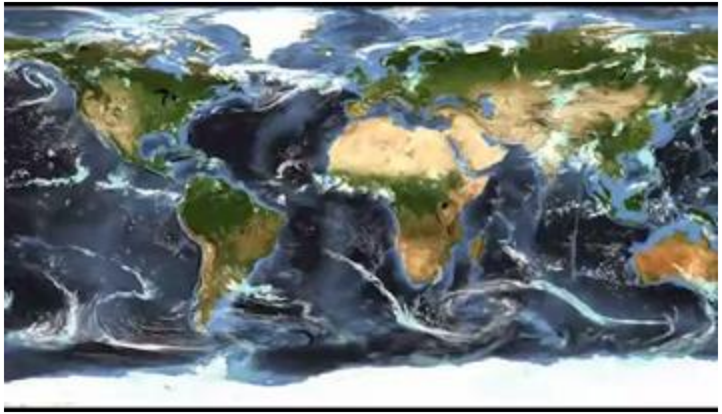
How can we improve climate projections?

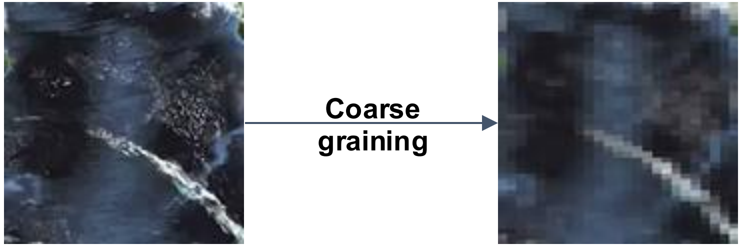
Solution 1. Using AI to harvest the data revolution: high-resolution simulations *or* Earth's observations to better emulate processes



Challenge: Climate models cannot easily integrate AI (old languages like Fortran) and assimilate data

→ Difficult to harvest the data revolution

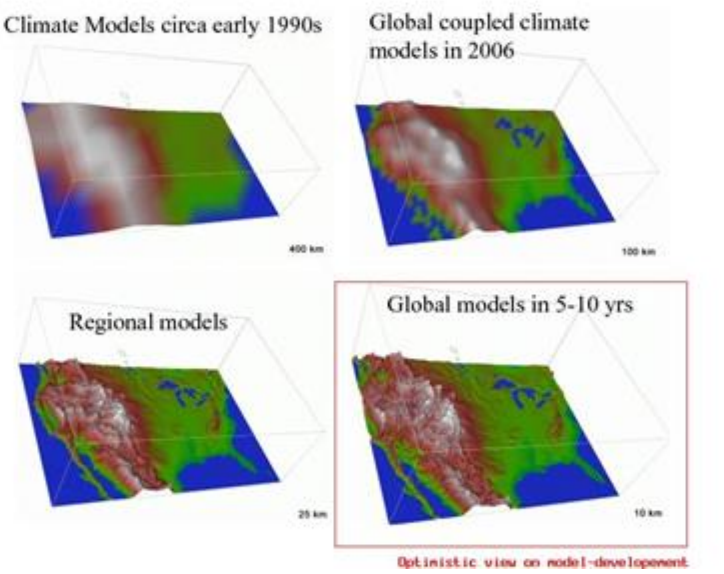




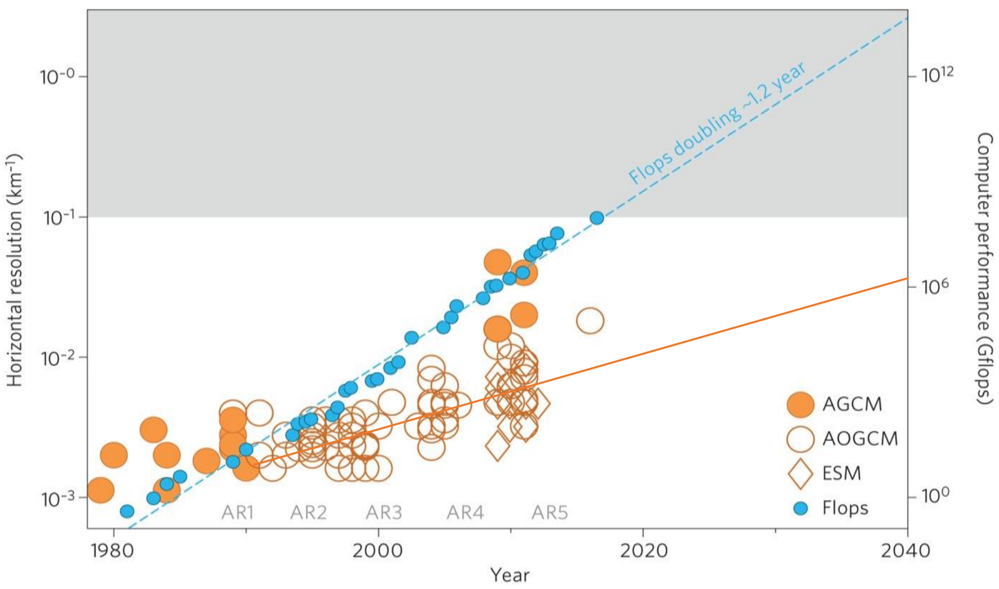
Gentine et al., 2018 *GRL*; Rasp et al., 2018 *PNAS*, Shamekh et al 2023, *PNAS*

How can we improve climate projections?

Solution 2. Increasing compute to better resolve processes: Exponential cost with resolution



Challenge: Climate models do not leverage modern hardware GPUs or TPUs and are based on old hardware infrastructure (CPUs). Legacy of low-level Fortran code, not agile to different hardware, plateauing performance



Schneider et al. 2019 Nat Climate Change

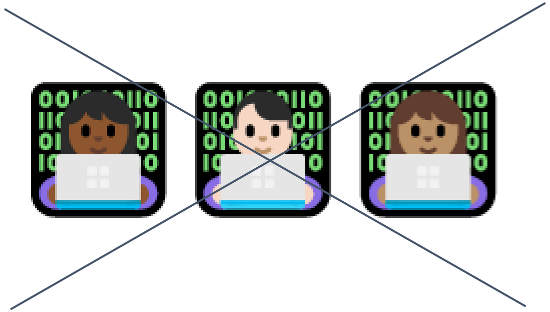
How can we improve climate projections?

Solution 3. Accelerate progress on climate model developments and theories

Challenge: Climate models are based on old programming languages and are very difficult to use → **not inclusive:** huge barrier to progress

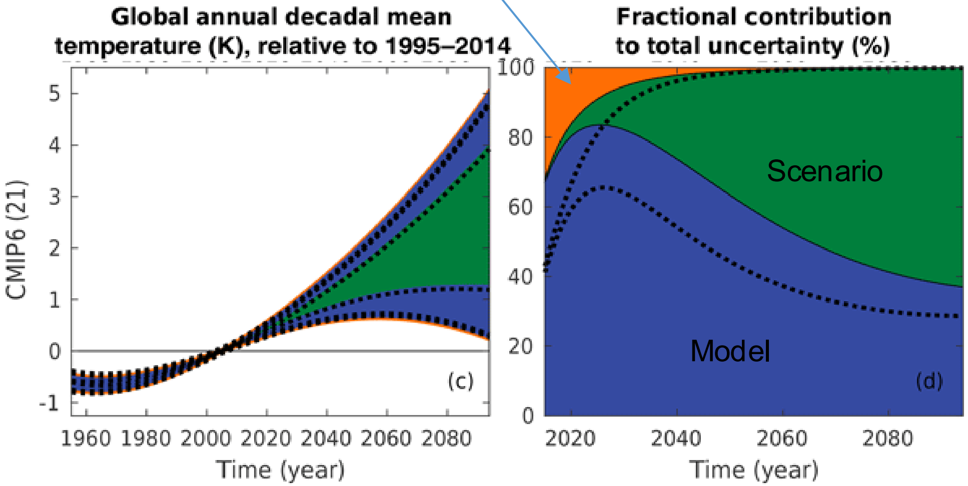
$$\frac{\partial \vec{v}_h}{\partial t} + (\vec{v} \cdot \nabla)\vec{v}_h + f\hat{k} \times \vec{v}_h + \frac{1}{\rho_0}\nabla_h p' = \vec{\mathcal{F}}$$

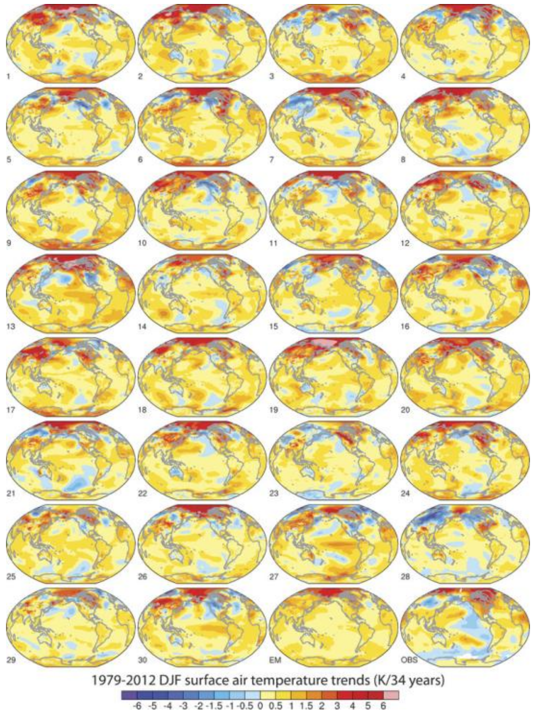
$$\nabla_h \cdot \vec{v}_h + \frac{\partial w}{\partial z} = 0$$



Solution 4. Accurate quantification of internal variability (and extremes)

Challenge: Computationally expensive – tradeoff between resolution and ensemble members





Outline

1. Introduction: climate uncertainties and the role of subgrid processes

2. AI: from climate process emulation
to new discoveries

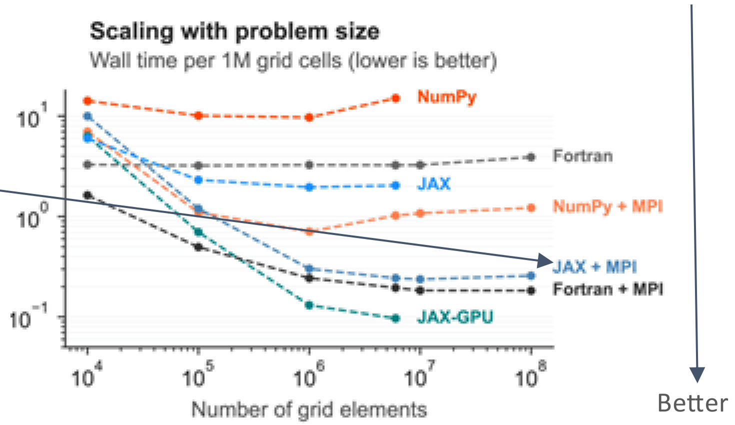
3. Solutions and challenges for next-generation climate models

4. A vision for next-generation climate models

A vision for next-generation climate modeling

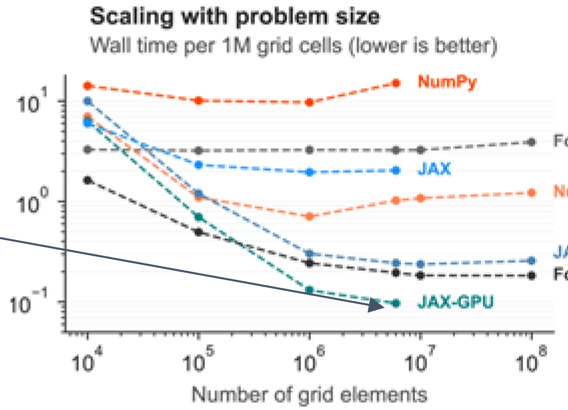
THREE PILLARS

1. Modern code & compute:
Python-based: more inclusive
+ JAX /GT4Py for fast compute (GPUs)



THREE PILLARS

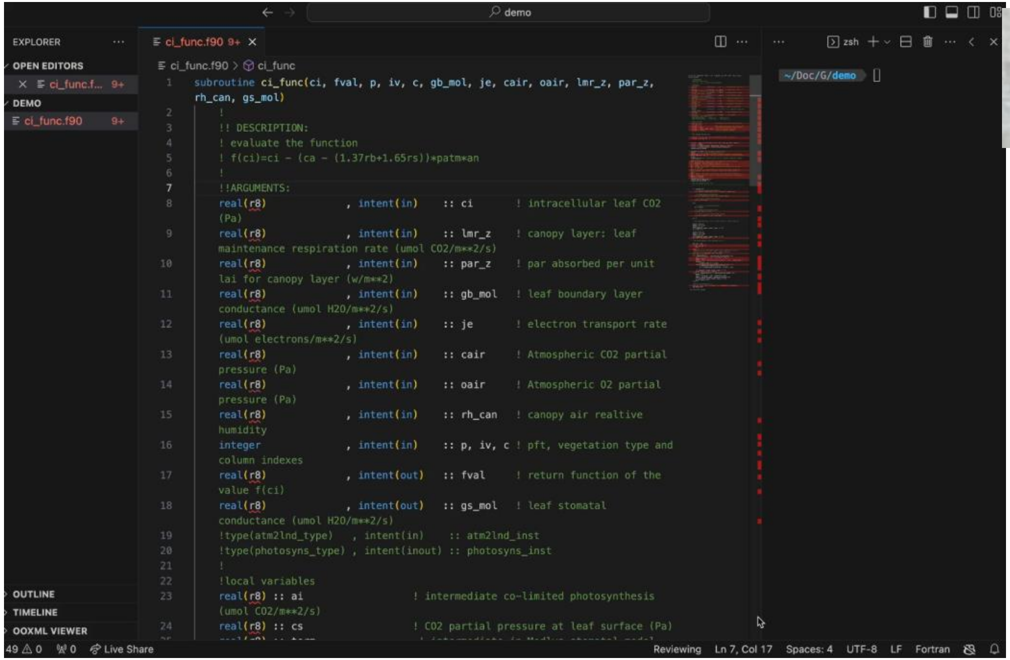
1. Modern code & compute:
Python-based: more inclusive
+ JAX /GT4Py for fast compute (GPUs)

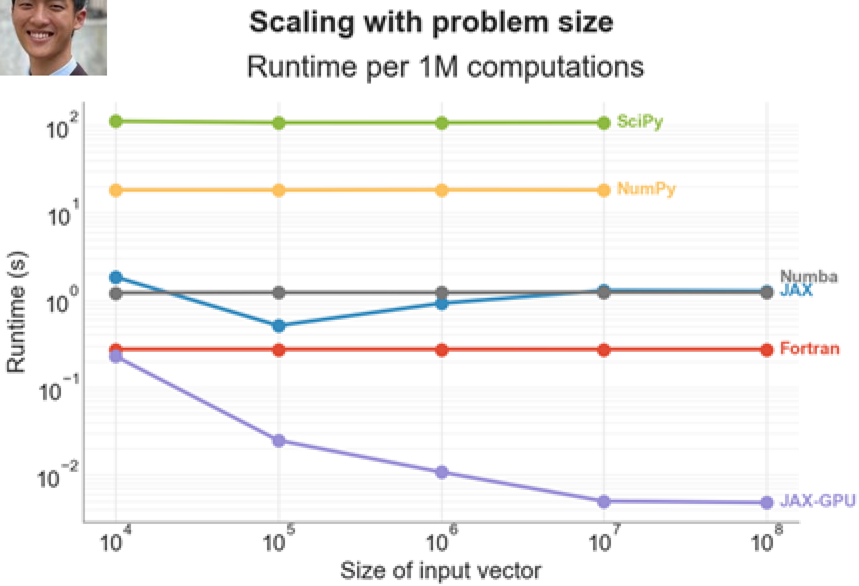


Better

How can we convert codes (faster)?

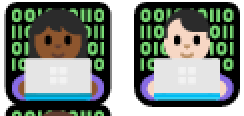
Generative AI with human-in-the-loop can help translate code faster: test on Community Land Model





A vision for next-generation climate modeling

THREE PILLARS

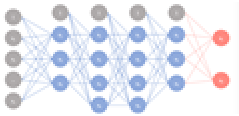


1. Modern code & compute:
Python-based: more inclusive
+ JAX /GT4Py for fast compute (GPUs)

GridTools 4python

XCLAIM

AI



Physics

$$\frac{\mathrm{D}u}{\mathrm{D}t} = \frac{uv\tan\phi}{r} - \frac{uw}{r} + fv - f'w - \frac{c_p\theta}{r\cos\phi}\frac{\partial\Pi}{\partial\lambda} + \mathrm{D}(u)$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} = -\frac{u^2\tan\phi}{r} - \frac{vw}{r} - uf - \frac{c_p\theta}{r}\frac{\partial\Pi}{\partial\phi} + \mathrm{D}(v),$$

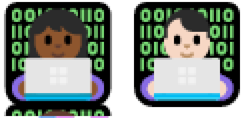
$$\delta\frac{\mathrm{D}w}{\mathrm{D}t} = \frac{u^2+v^2}{r} + uf' - g(r) - c_p\theta\frac{\partial\Pi}{\partial r},$$

2. Seamless AI-integration:
hybrid Earth system model: unify physics and AI

Häfner et al. *JAMES* 2022; Automatic differentiation: Chen, Appling, Gentine et al *Nature Reviews Earth & Environment* 2022

A vision for next-generation climate modeling

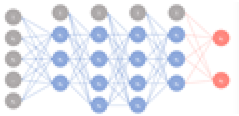
THREE PILLARS



1. Modern code & compute:
Python-based: more inclusive
+ JAX /GT4Py for fast compute (GPUs)

JAX GridTools 4python XCLAIM

AI



Physics

$$\frac{Du}{Dt} = \frac{uv\tan\phi}{r} - \frac{uw}{r} + fv - f'w - \frac{c_p\theta}{r\cos\phi}\frac{\partial\Pi}{\partial\lambda} + D(u)$$

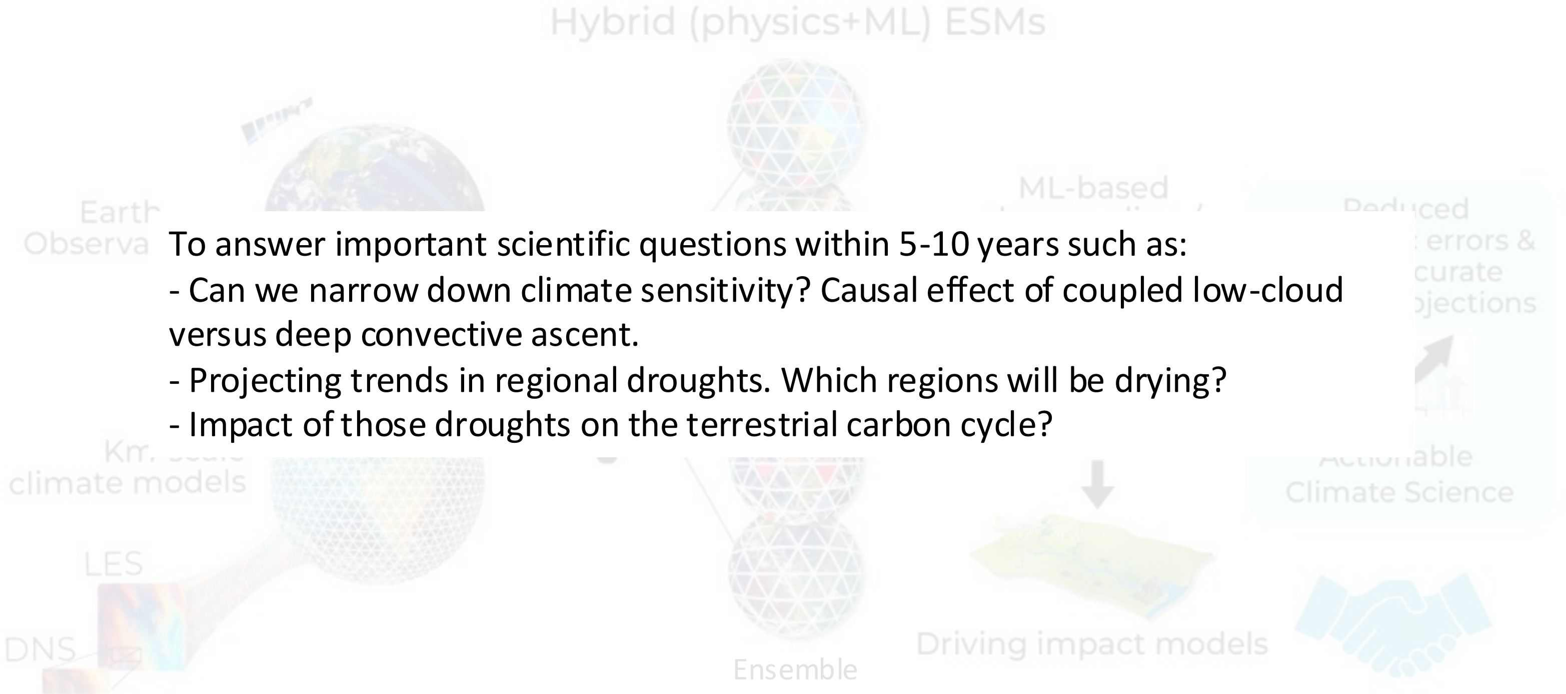
$$\frac{Dv}{Dt} = -\frac{u^2\tan\phi}{r} - \frac{vw}{r} - uf - \frac{c_p\theta}{r}\frac{\partial\Pi}{\partial\phi} + D(v),$$

$$\delta\frac{Dw}{Dt} = \frac{u^2+v^2}{r} + uf' - g(r) - c_p\theta\frac{\partial\Pi}{\partial r},$$

2. Seamless AI-integration:
hybrid Earth system model: unify physics and AI

3. Harnessing global observations:
with new data assimilation
- targeting statistics (e.g. precipitation distribution)
- even high-res models are imperfect

Li, Gentine, Bhouri, Bocquet, Farchi, Zheng, 2024a and b in preparation

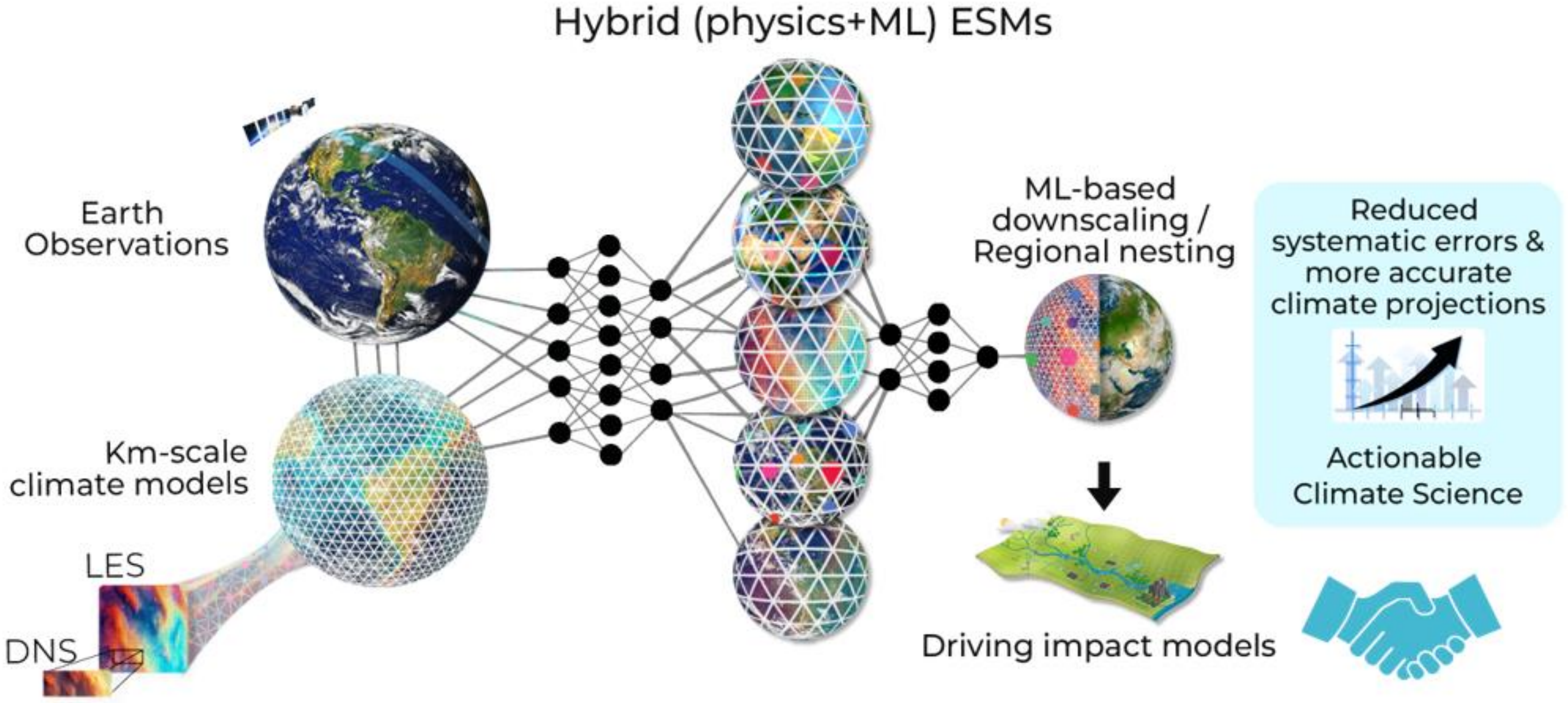


To answer important scientific questions within 5-10 years such as:
- Can we narrow down climate sensitivity? Causal effect of coupled low-cloud versus deep convective ascent.
- Projecting trends in regional droughts. Which regions will be drying?
- Impact of those droughts on the terrestrial carbon cycle?

With focus on land and atmosphere in my group

Eyring, Gentine, Camps-Valls, Lawrence, *Nature Geo*, in review; Schiro .. Gentine et al. 2022 *Nature Communications*

Questions and Answers



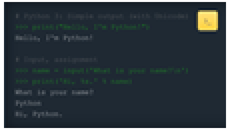
Eyring, Gentine, Lawrence, Camps-Valls, Reichstein; Nature Geoscience. 2024 accepted

Supplementary slides

A vision for next generation climate model to address societal needs

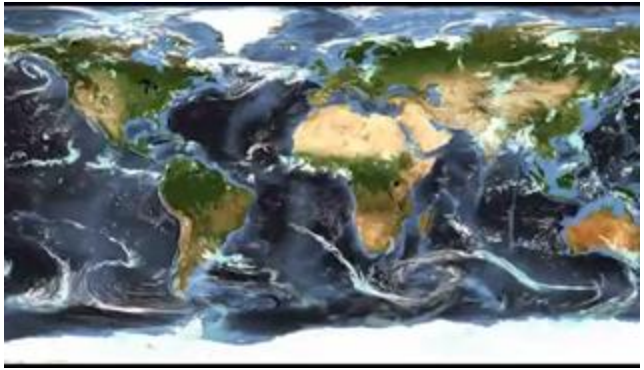
Goal: build the <u>first</u> full climate model based on:

- modern language (Python)



- scalable + flexible hardware (GPUs)

- seamless integration of AI to harvest global data

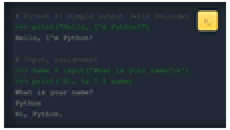


Summary

A vision for next-generation climate model to address societal needs

Goal: build the first full climate model (starting with atmosphere and land) based on:

- modern language (Python)



- scalable + flexible hardware (GPUs)

- seamless integration of AI to harvest global data

to provide more accurate climate projections, support global climate adaptation and make new discoveries

A vision for next-generation climate modeling

Can we replicate the machine learning revolution in physical/climate modeling?

Backbones of the ML revolution:

- Data
- Efficient hardware use
- Algorithms for optimization (backpropagation, automatic differentiation)
- Under-appreciated: high-level programming with expressive, flexible and optimized low-level operations (few lines of codes)

```
model_vanilla = Sequential()

model_vanilla.add(Dense(n_neuron,  activation=activation,input_shape=(X_train.shape[1],))) #  the 1st hidden layer
model_vanilla.add(Dense(n_neuron,  activation=activation)) # the 2nd hidden layer
model_vanilla.add(Dense(n_neuron,  activation=activation)) # the 3rd hidden layer
model_vanilla.add(Dense(y_train.shape[1],  activation='linear')) # the output layer

model_vanilla.compile(loss='mse',optimizer=tf.keras.optimizers.Adam(learning_rate=learning_rate))
```

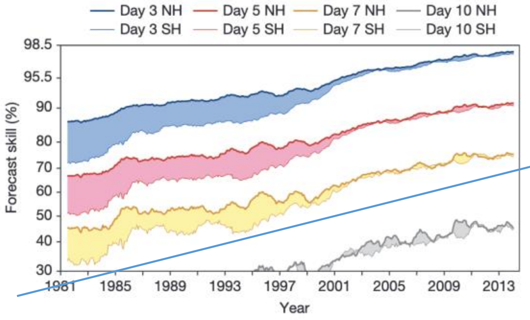
```
early_stop = keras.callbacks.EarlyStopping(monitor='val_loss', patience=20)

history = model_vanilla.fit(X_train, y_train,
                            batch_size      = minibatch_size,
                            epochs          = num_epochs,
                            validation_split= 0.2,
                            verbose         = 1,
                            callbacks       = [early_stop])
```

Eyring, Gentine, Camps-Valls, Lawrence, *Nature Geo*, in review; Automatic differentiation: Chen, Appling, Gentine et al *Nature Reviews Earth & Environment* 2022

Bringing observational data into climate model development core

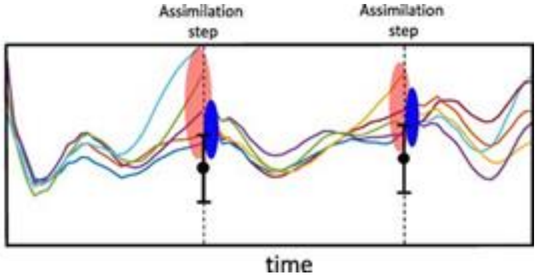
- No perfect model, even at high-resolution (e.g high-res ICON ~10W/m2 bias or cloud cover)

- Data assimilation led to weather quiet revolution due to increasing data (~ like AI)



More data

Traditional data assimilation



- Traditional data assimilation focuses on trajectory correction - **initial conditions**

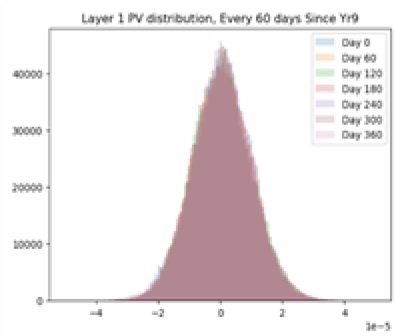
→ Not applicable to climate: need excellent future model

Bauer et al. Science 2015

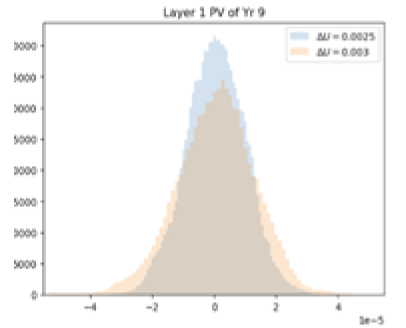
- Key differences with traditional techniques:
  1. Does not correct trajectories but **corrects the model parameters/structure** (predicting into the future)

  2. Does not consider individual realizations but **statistics**

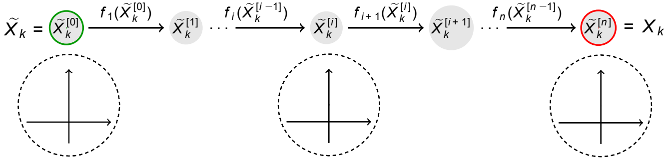
Potential Vorticity



Statistically steady state



Parameter dependence of distribution

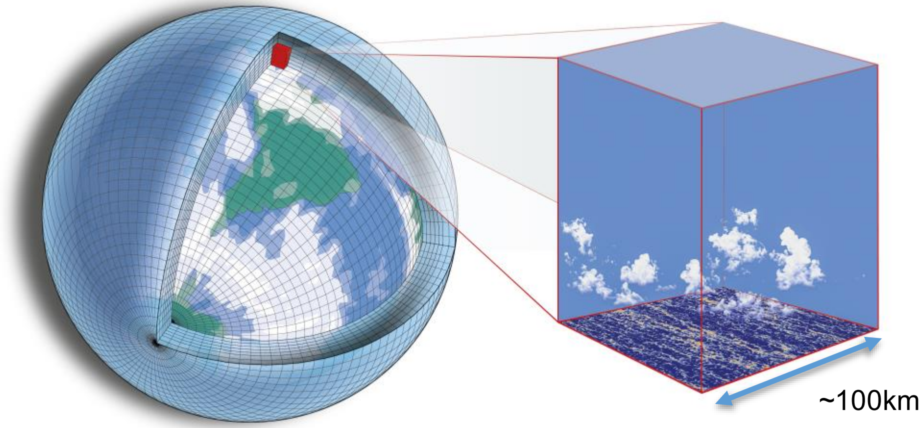


3. Merges Machine Learning with Data Assimilation techniques:
   harvest data at scale + uncertainty quantification + indirect observations

Li, Gentine, Bhouri, Bocquet, Farchi, Zheng, 2024a and b in preparation

Climate models use parameterizations of subgrid processes

Parameterization: represents (physically or statistically) a physical process that cannot be resolved (e.g. clouds). Typically physically based.

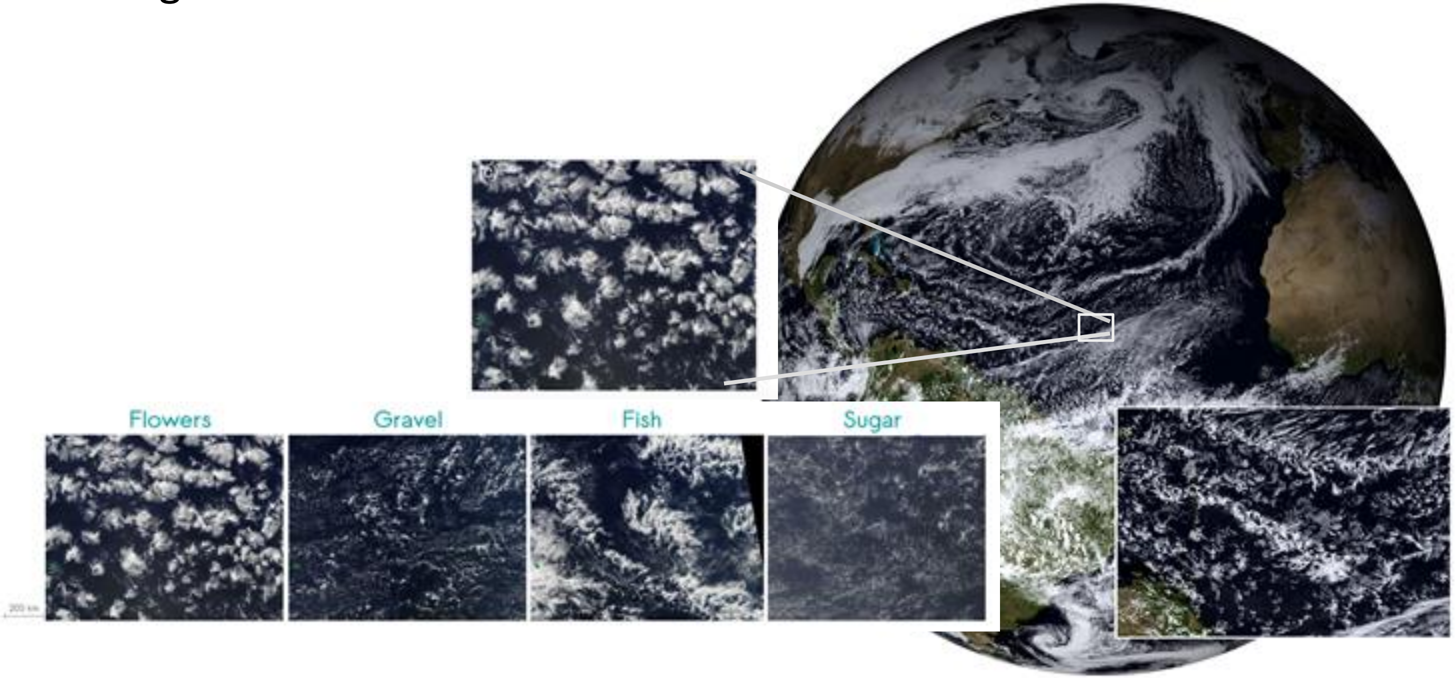


However: it has failed for ~40 years

This largely **explains inter-model spread in climate projection**

Schneider et al., 2019 GRL; Randall et al 2003, BAMES

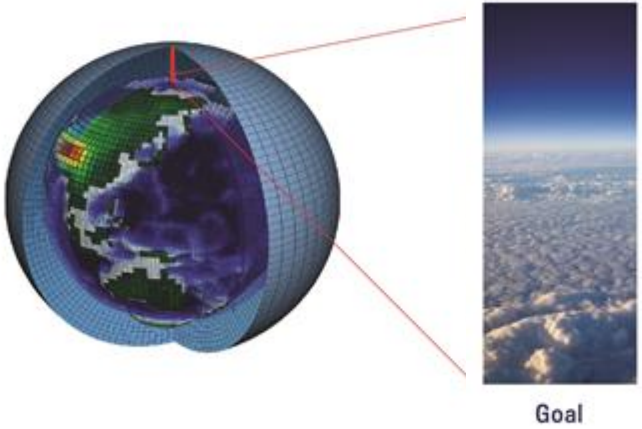
- Clouds organize in observations



More reliable predictions is one problem …

effective climate action *is another problem altogether*

Climate Prediction Dat



Knowledge Transfer

+

Diversity, Equity, Inclusion

Effective Climate Action

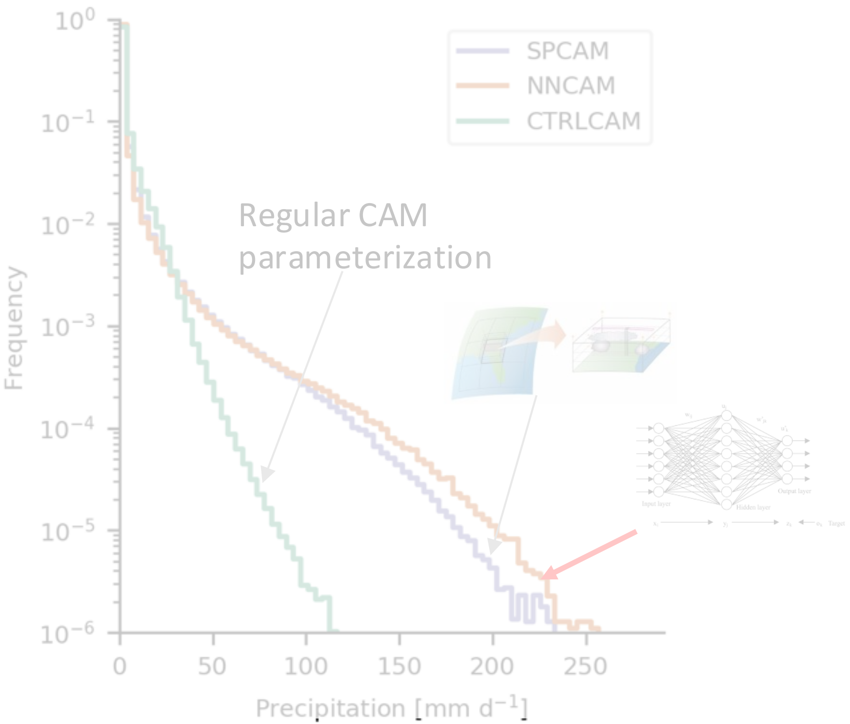
Modern cloud infrastructure to empower climate action with an ecosystem of local partners, especially in Global South (with Climatematch Academy)

Transparent, inclusive and ethical: same climate data accessible by anyone across the globe

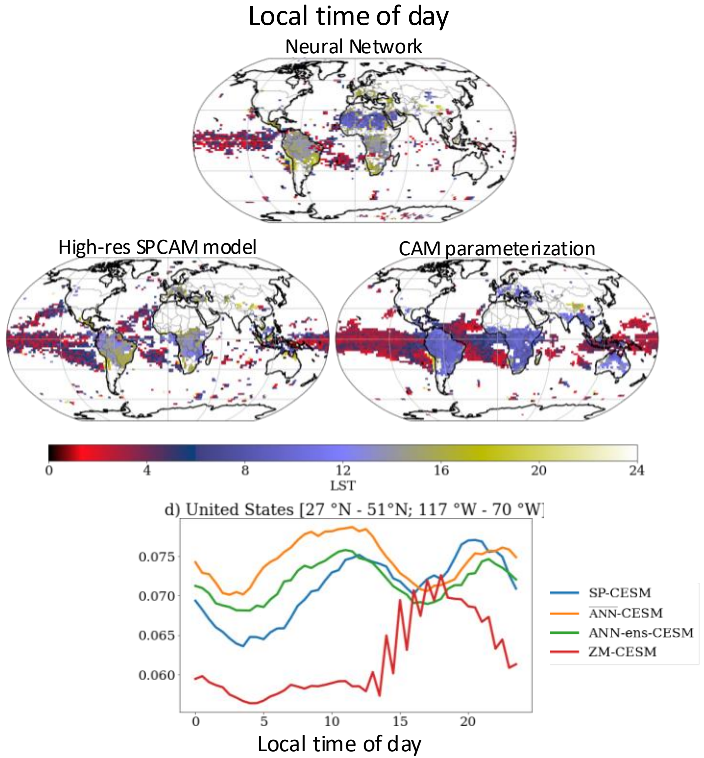
Inclusion of ML closure leads to **Step-change** in climate model quality

Online global simulations 🌎

Precipitation distribution (CAM)



Precipitation diurnal cycle (CAM+ICON)



Rasp, Pritchard, Gentine., 2018 PNAS; Behrens, Eyring, Gentine submitted

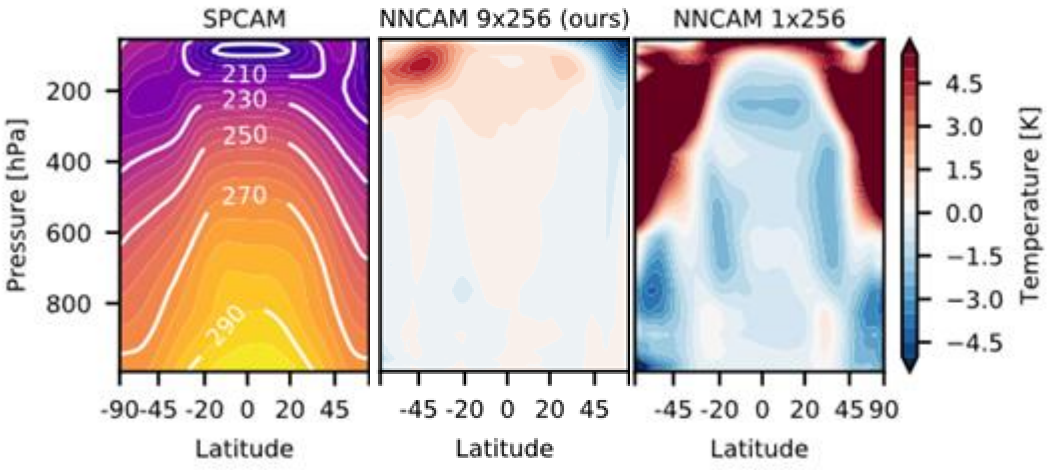
Issues: Convection – DL instabilities & spurious correlations

SPCAM → **Artificial Neural Network (NN)** → **NNCAM**

learning subgrid processes as represented by the SP component

(online)

Zonal-mean temperature

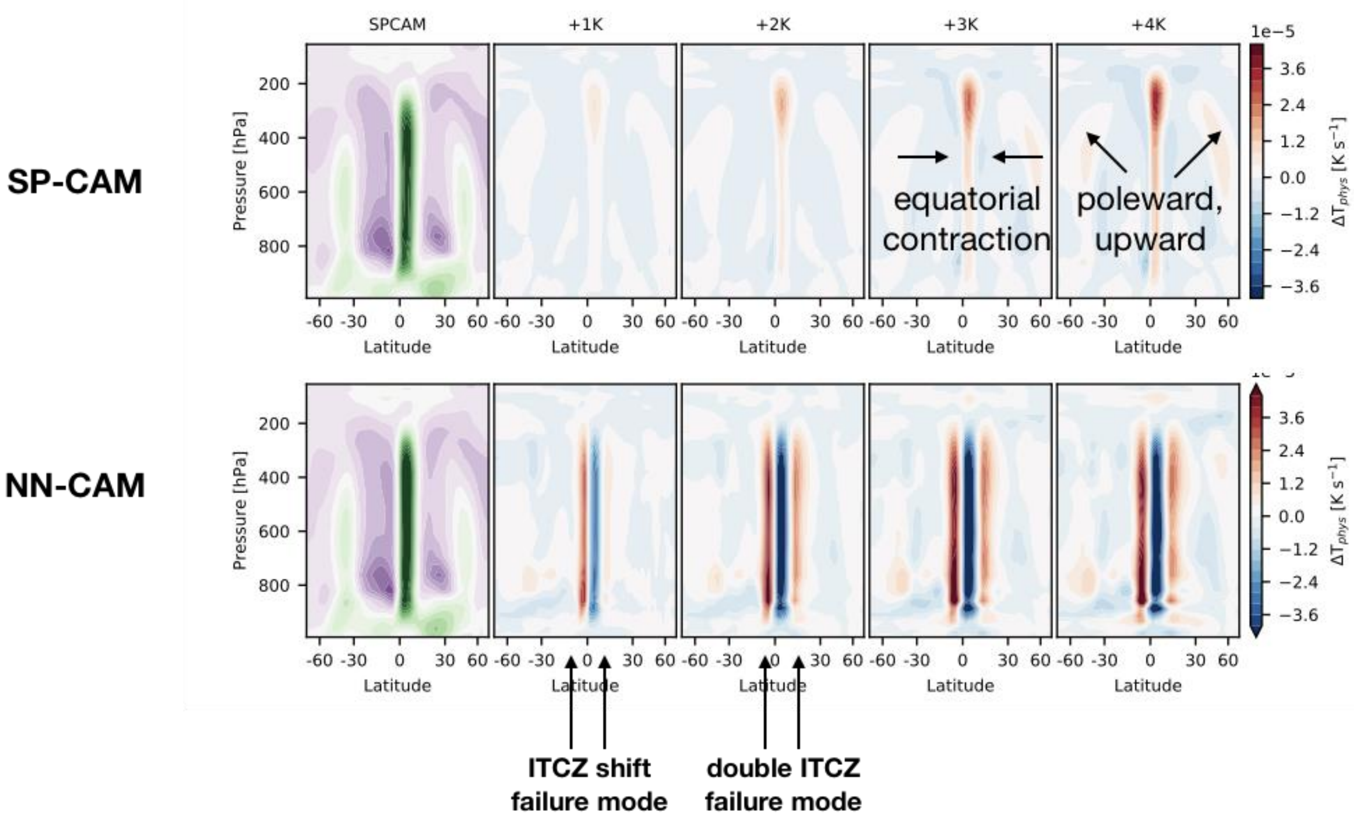


NN-parametrizations can capture many aspects of CRMs

Issues

- Spurious correlations between stratosphere and boundary layer
- Instabilities in the coupled runs (NNCAM) under a number of setups
- Generalization: Limitations with out-of-sample temperatures
- Does not conserve energy

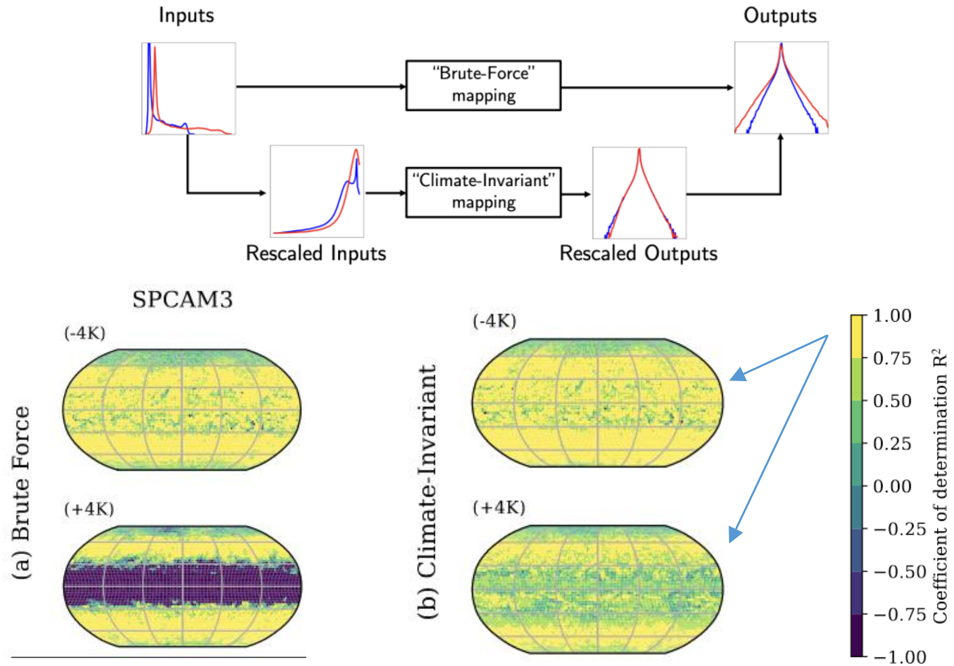
Rasp et al. (2018) *PNAS;* Beucler et al. et al. (2022). *Phys Rev Letters*



SP-CAM
NN-CAM
SPCAM
+1K
+2K
+3K
+4K
Pressure [hPa]
Latitude
1e−5
ΔT_{phys} [K s^{-1}]
equatorial contraction
poleward, upward
ITCZ shift failure mode
double ITCZ failure mode

Challenge 1: Generalization

- Challenge: out-of-sample, out-of-distribution prediction/sampling bias
- Solution: **embed physical invariances/equivariance** along Lie groups in ML (e.g., dimensionless numbers or rotational equivariance to collapse distributions)

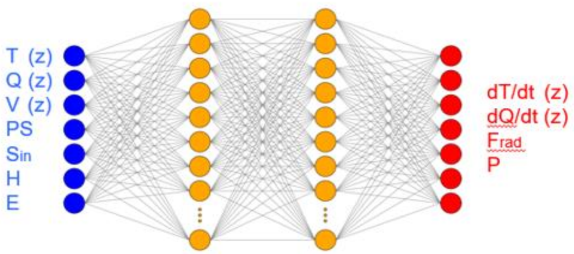


Beucler, Pritchard,…, Gentine., 2023 in revision @ Sci Advances, Connolly, Yu, Gentine, 2023 submitted

Challenge 2: Trust + interpretability

- Challenge: cross-correlations can fool ML
- Solution: merge **causal discovery with ML**
  → more interpretable, more trustworthy

NN (Rasp et al., 2018)



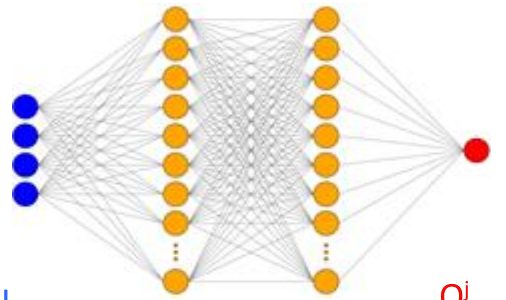
$I_{N=94}$

Hidden layers
Number = 9
…

$O_{N=65}$

Drop connections based on causality: a very sparse NN

Causal-NNs



$I_{N=causal}$

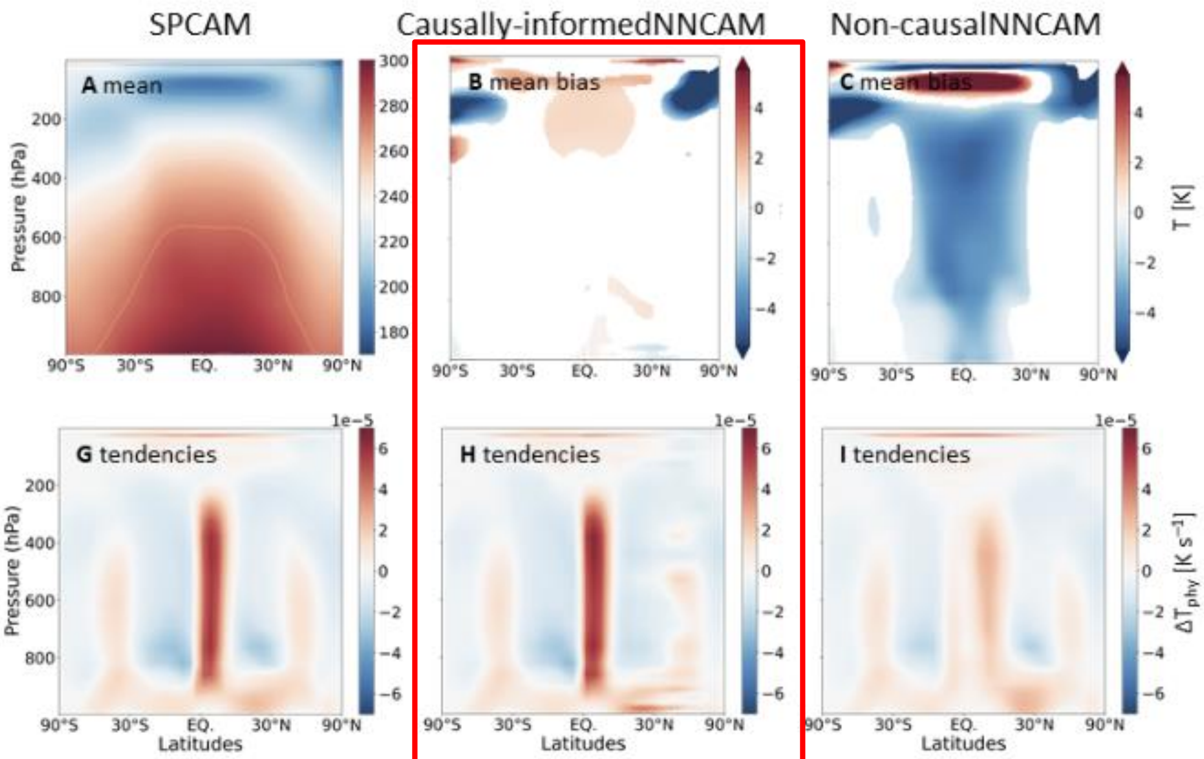
Hidden layers
Number = 9

$O^j_{(N=65)}$

$\boxed{I^i_{causal} \rightarrow O^j \in P(O^j)}$

Causality: going beyond correlations

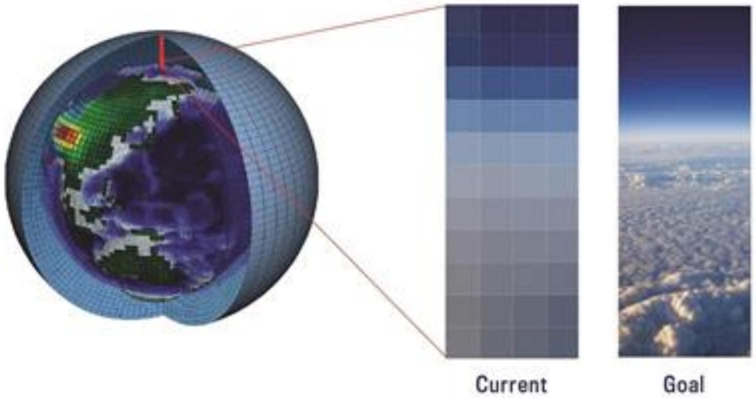
Online global simulations



Iglesiaz-Suarez et al., submitted

Why ETH

- ICON ongoing work
- Aligned vision on high-level code and high compute (MeteoSwiss-ETH)
- Continued collaboration on land (MPI)
- Strong climate group (including impact), who would be using new model outputs
- Strong CS group, plus interested in climate and climate modeling
- Key partnership with MeteoSwiss, pioneer in the use of GPUs



A realistic Earth's Twin

Other strategy

Use physical insights to improve (phenomenological) subgrid models

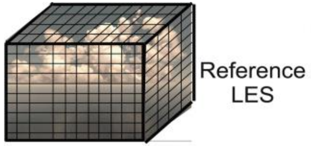
\+ **systematic model tuning**

(with statistical emulation for speedup)

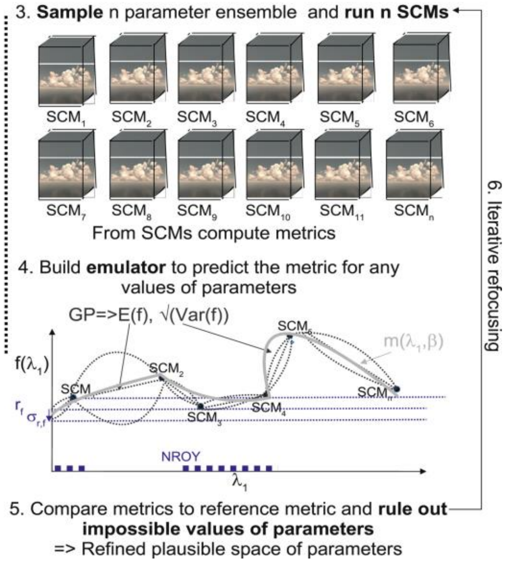
e.g., IPSL (France), NASA GISS (US), NCAR (US), Clima (US)

Goal: Estimate parameterization parameters *Θ*

e.g. clouds



High-resolution
Simulation as truth



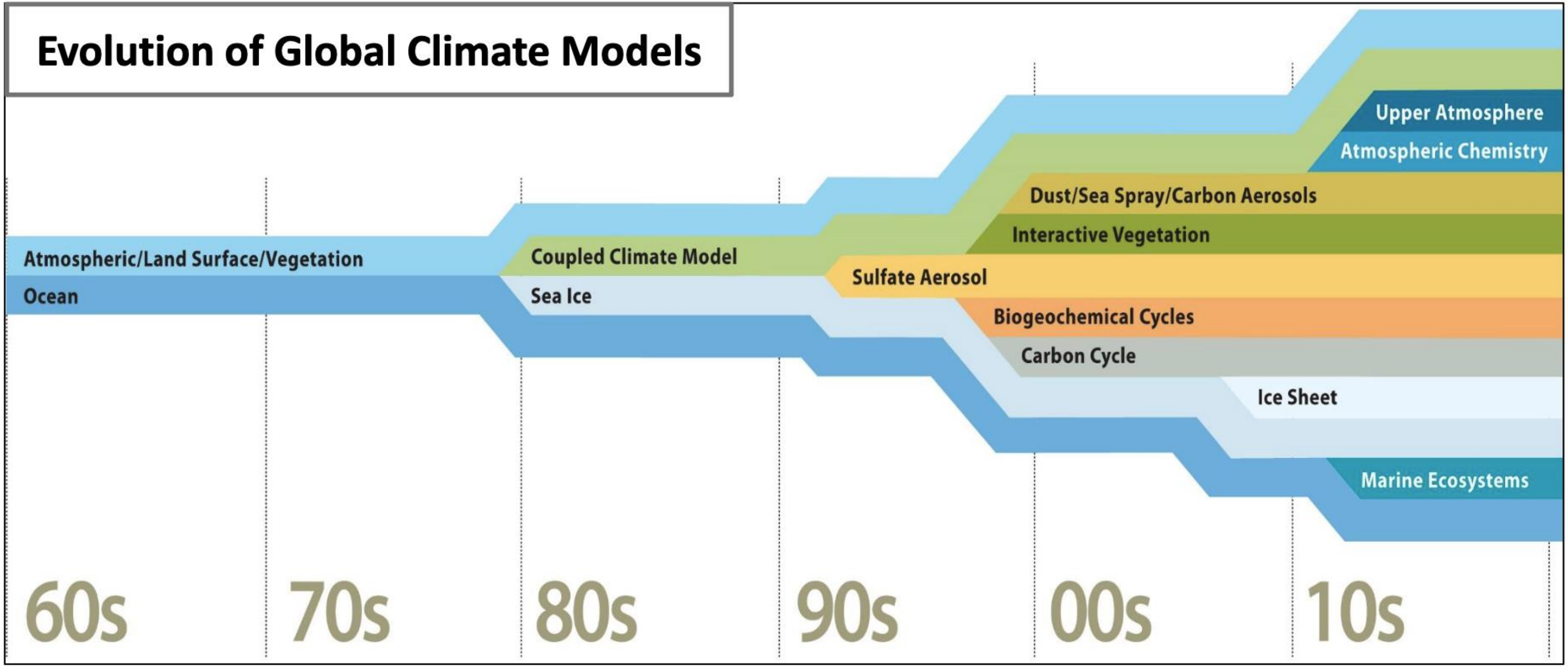
Sample: Coarse-resolution model MCMC on *Θ*

Emulator on *Θ*

Calibrate *Θ*

Couvreux et al. 2020 *JAMES*; Cleary et al 2021 *J Comp Phys*

From climate models to Earth system models



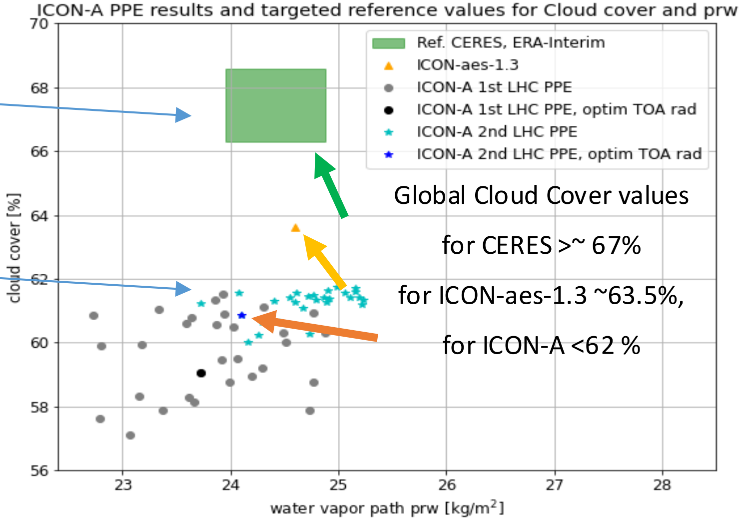
Simple climate models have evolved into complex Earth System models
to **answer many questions** (not just climate projections)

Caveat: parameterizations can have large structural errors

- Parameter tuning might be **impossible**

observations

model MCMC



Global Cloud Cover values

for CERES >~ 67%

for ICON-aes-1.3 ~63.5%,

for ICON-A <62 %

- **Structural errors dominate many processes**

→ traditional data assimilation may not be feasible for climate

Bonnet, Eyring, Gentine et al., submitted

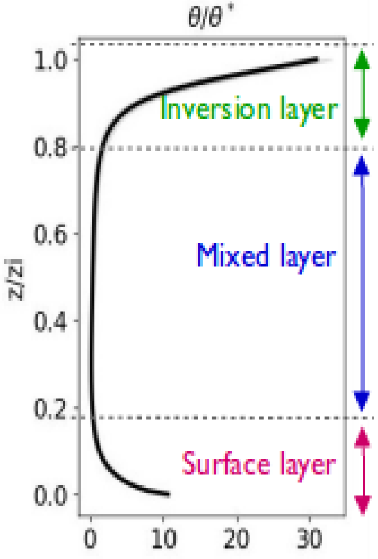
Caveat: parameterizations can have large structural errors

- Example of dry turbulence with state-of-the-art model:

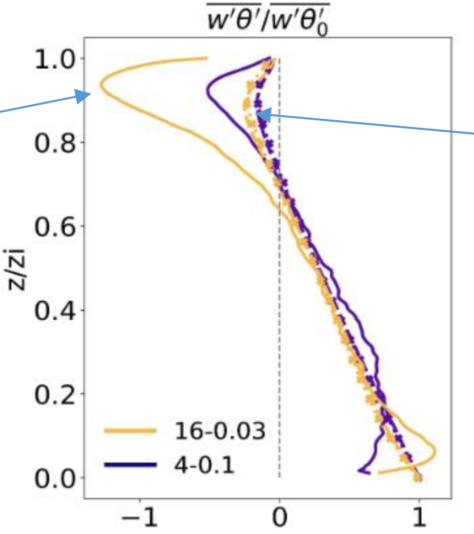
Eddy-diffusion mass-flux model

$$\overline{w'\theta'} = -K\frac{\partial\overline{\theta}}{\partial z} + M_u(\theta_u - \overline{\theta})$$

(ECMWF, IPSL, Clima…)



Best fit coarsem odel



Vertical turbulent heat flux

High-resolution truth

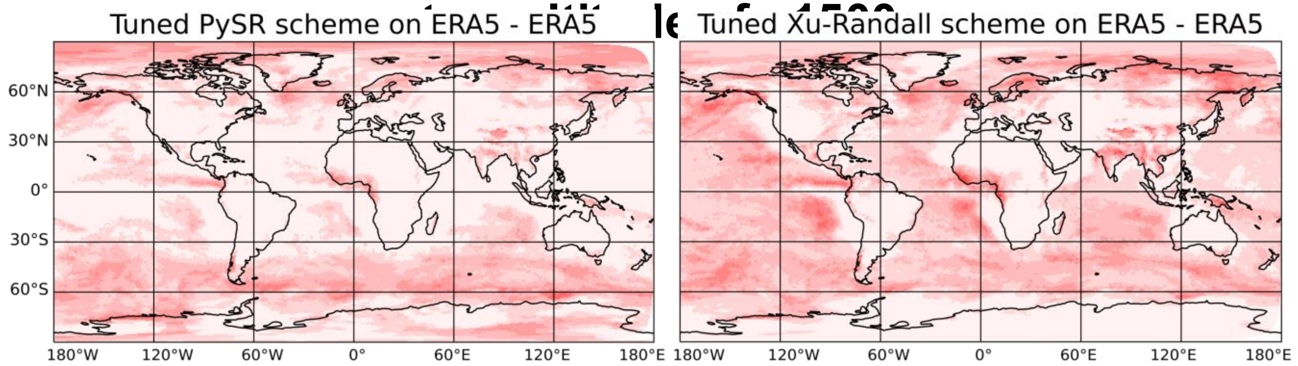
- Even the best set of parameters still has substantial deficiencies

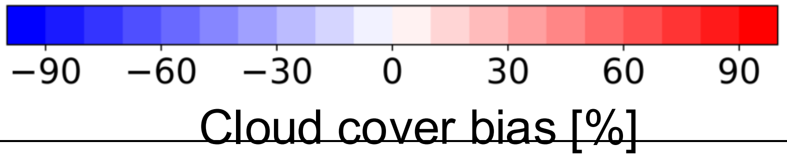
Shamekh and Gentine, submitted

Bias of averaged cloud cover (area fraction) of 3-hourly data from 20160811 to 20160820

$$C_{\text{Xu–Randall}} \stackrel{\text{def}}{=} \min\{\text{RH}^{\beta}(1-\exp(-\alpha(q_c+q_i))),1\}$$

(multiplied by 100)





The data-driven analytical PySR equation

$$f(RH, T, \partial_z \mathrm{RH}, q_c, q_i) = I_1(RH, T) + I_2(\partial_z \mathrm{RH}) + I_3(q_c, q_i),$$

$$I_1(\mathrm{RH}, T) \stackrel{\mathrm{def}}{=} a_1 + a_2(\mathrm{RH} - \overline{\mathrm{RH}}) + a_3(T - \overline{T}) + \frac{a_4}{2}(\mathrm{RH} - \overline{\mathrm{RH}})^2 + \frac{a_5}{2}(T - \overline{T})^2(\mathrm{RH} - \overline{\mathrm{RH}})$$

$$I_2(\partial_z \mathrm{RH}) \stackrel{\mathrm{def}}{=} a_6^3 \left(\partial_z \mathrm{RH} + \frac{3a_7}{2}\right)(\partial_z \mathrm{RH})^2$$

$$I_3(q_c, q_i) \stackrel{\mathrm{def}}{=} \frac{-1}{q_c/a_8 + q_i/a_9 + \epsilon}.$$

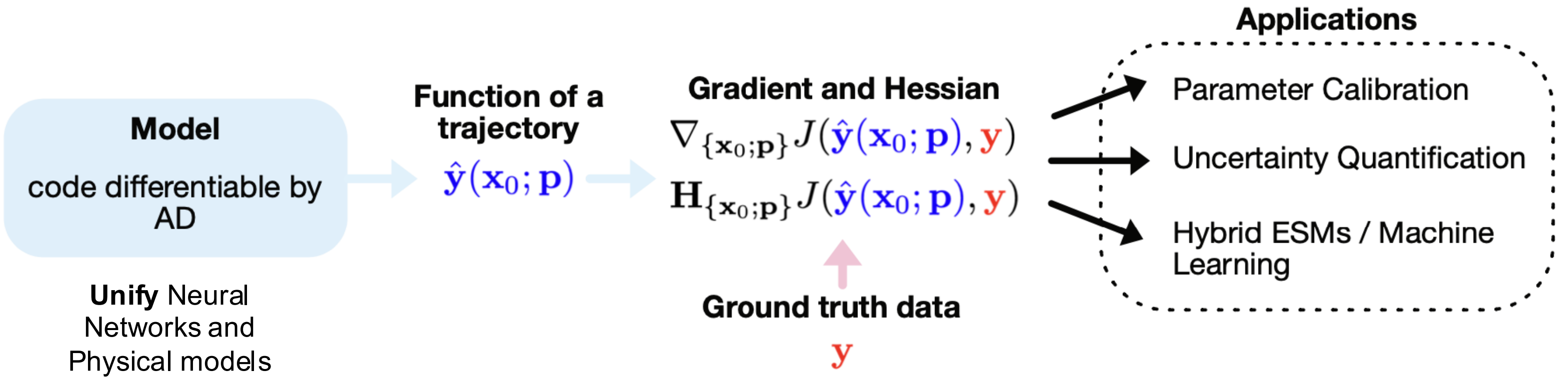
$$\{a_1, \ldots, a_9, \epsilon\} = \{0.4435, 1.1593, -0.0145\,\mathrm{K}^{-1}, 4.06, 1.3176 \cdot 10^{-3}\,\mathrm{K}^{-2}, 584.8036\,\mathrm{m}, 2\,\mathrm{km}^{-1}, 1.1573\,\mathrm{mg/kg}, 0.3073\,\mathrm{mg/kg}, 1.06\}$$

$$C(X) = \max\{\min\{100\, f(X), 100\}, 0\}$$

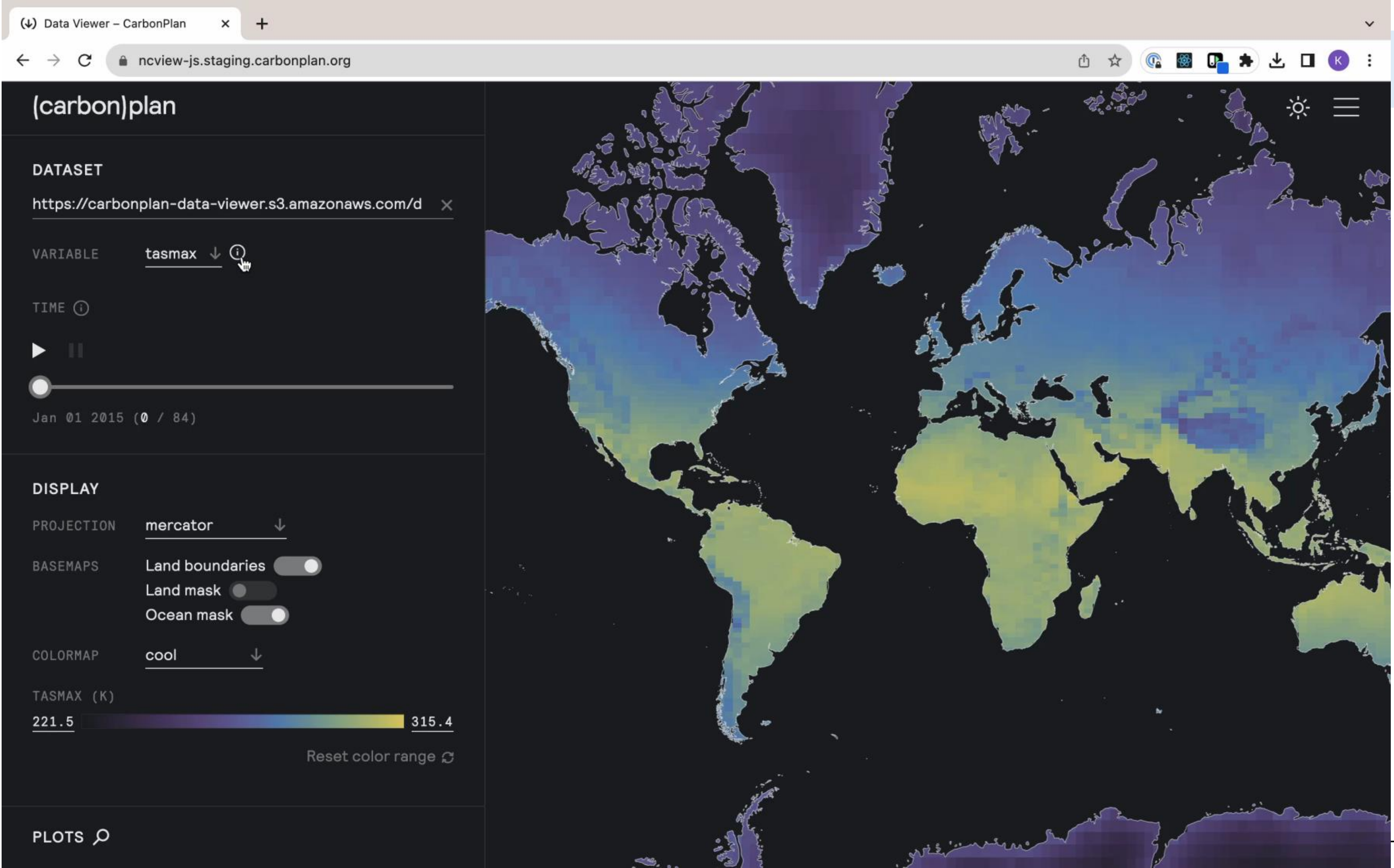
Abbreviated form:

$$C = P_3(RH, T) + (c_1 \partial_z RH + c_2)(\partial_z RH)^2 - \frac{1}{c_3 q_c + c_4 q_i + \varepsilon}$$

- From weather to climate



Bhouri et al., 2022 *arxiv*; Gelbrecht et al. 2022 *arxiv; Shen, Gentine et al, 2023 Nat Comm*


Data Viewer – CarbonPlan
ncview-js.staging.carbonplan.org
(carbon)plan
DATASET
https://carbonplan-data-viewer.s3.amazonaws.com/d
VARIABLE
tasmax
TIME
Jan 01 2015 (0 / 84)
DISPLAY
PROJECTION
mercator
BASEMAPS
Land boundaries
Land mask
Ocean mask
COLORMAP
cool
TASMAX (K)
221.5
315.4
Reset color range
PLOTS

Intermission 1:
Advances in Machine Learning: *backpropagation*

- *Backpropagation*
- Chain's rule

$$x \xrightarrow{w_1} z_1 \xrightarrow{w_2} \hat{y} \longrightarrow J(\boldsymbol{W})$$

$$\frac{\partial J(\boldsymbol{W})}{\partial w_1} = \frac{\partial J(\boldsymbol{W})}{\partial \hat{y}} * \frac{\partial \hat{y}}{\partial z_1} * \frac{\partial z_1}{\partial w_1}$$

- Efficient computation of gradients is key (analytical or automatic differentiation)

Intermission 2:
Advances in Machine Learning: *automatic differentiation*

- What is **automatic differentiation**?

- Propagate ε differential throughout code $f(x+\varepsilon)$
  with property: $\varepsilon^2=0$

- Taylor series:

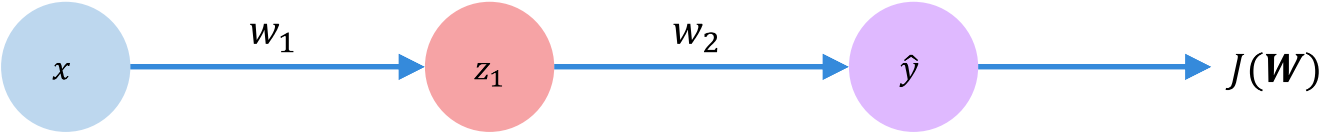
$$f(x+\varepsilon) = f(x) + \varepsilon f'(x) + \varepsilon^2/2\, f''(x) + O(\varepsilon^2)$$

0

Efficient computation of gradients ~ almost free

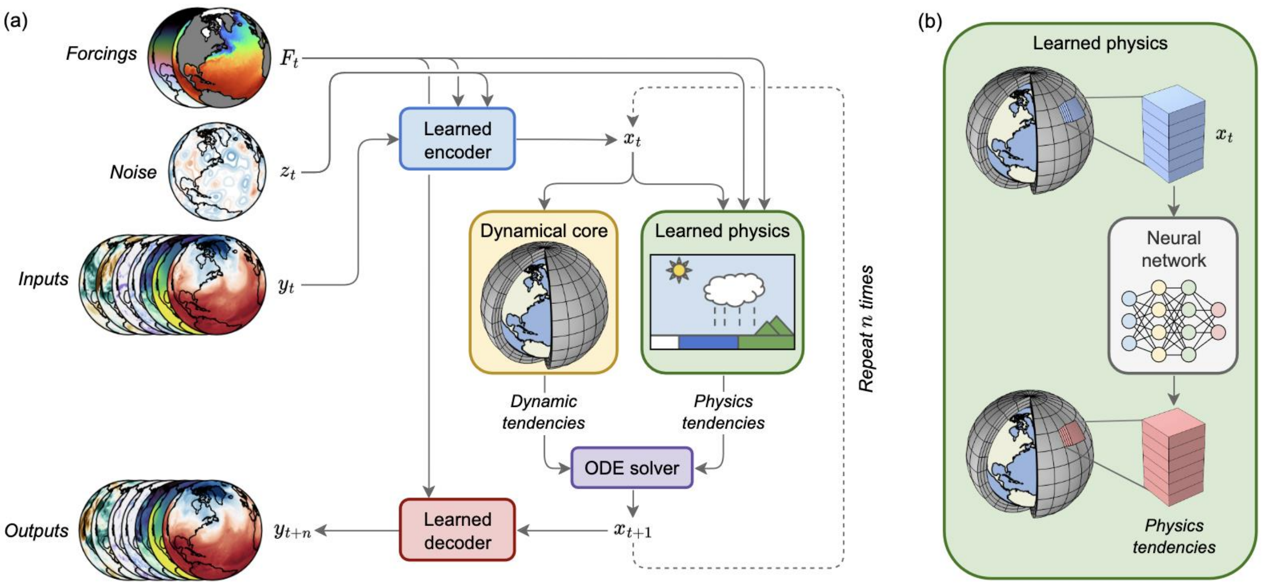
- Used in many modern ML toolboxes:
  Pytorch, JAX or modern computing languages: Julia

- Dependence of neural network weights, physical parameters or initial conditions is the same!



NN weight, initial condition (x0), physical parameter θ

Differentiable model

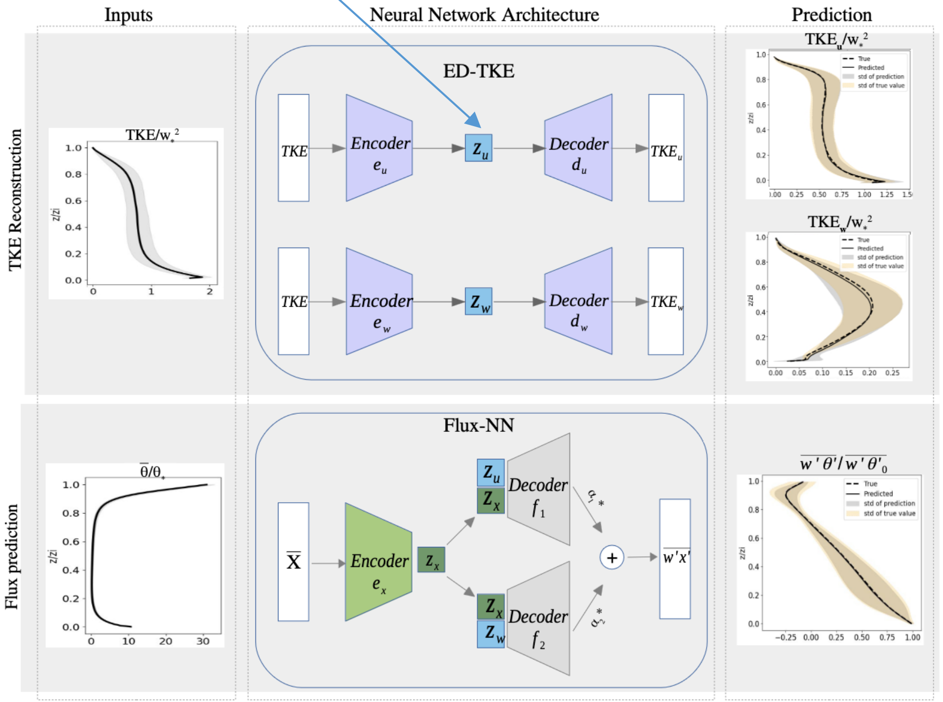


Neural General Circulation Models

Dmitrii Kochkov[1*†], Janni Yuval[1*†], Ian Langmore[1†], Peter Norgaard[1†], Jamie Smith[1†], Griffin Mooers[1], James Lottes[1], Stephan Rasp[1], Peter Düben[3], Milan Klöwer[4], Sam Hatfield[3], Peter Battaglia[2], Alvaro Sanchez-Gonzalez[2], Matthew Willson[2], Michael P. Brenner[1,5], Stephan Hoyer[1*†]

[1]Google Research, Mountain View, CA.
[2]Google DeepMind, London, UK.
[3]European Centre for Medium-Range Weather Forecasts, Reading, UK.
[4]Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology.
[5]School of Engineering and Applied Sciences, Harvard University.

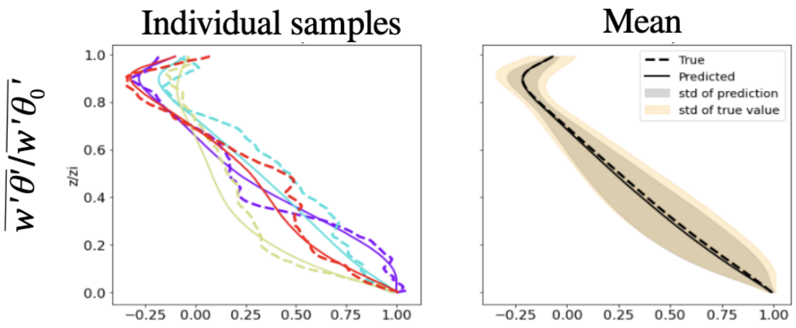
- Example to dry atmospheric turbulence using ML: latent representation of turbulence



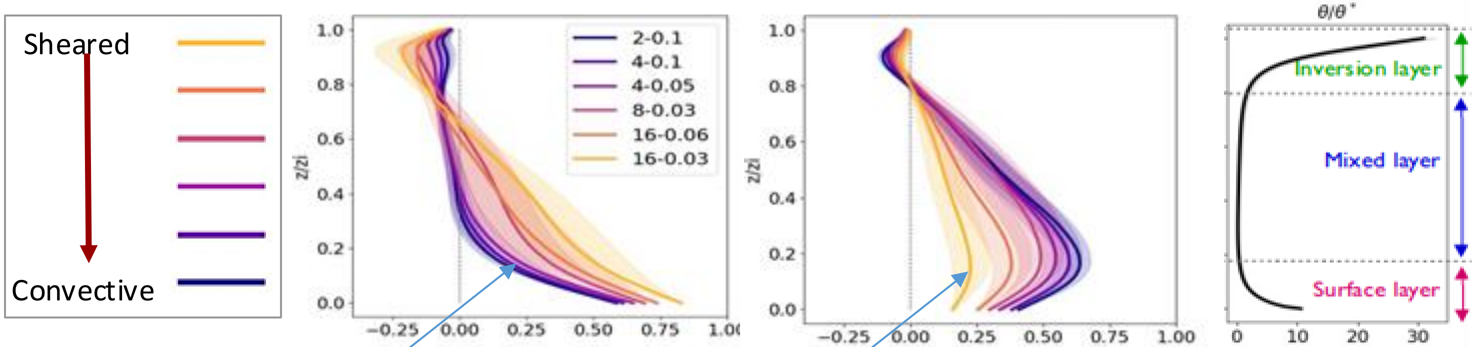
Shamekh and Gentine 2023 submitted

Latent model for dry atmospheric turbulence

- Great at **prediction**



- Can be used for **understanding** (2D only)



Shear:Diffusive

Convective: Non-diffusive, not a mass flux (downward motion is important)

Shamekh and Gentine 2023 submitted

Supplementary: ML-accelerated dynamical core

There are now methods to also accelerate PDE resolution (here JAX-based again)

Check for updates

Machine learning–accelerated computational fluid dynamics

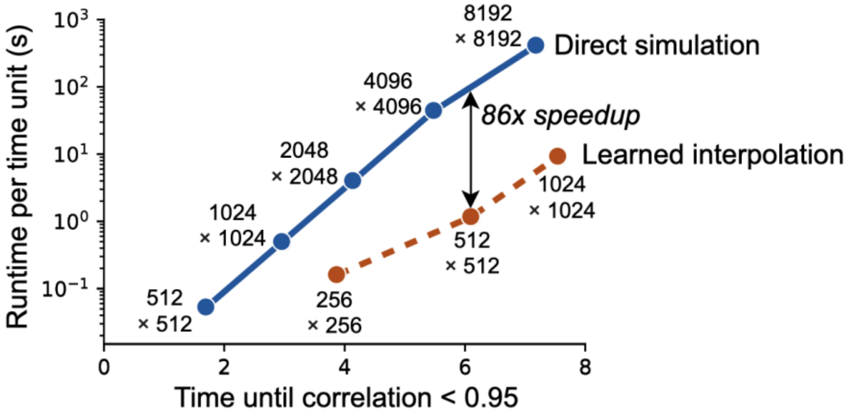
Dmitrii Kochkov[a,1,2], Jamie A. Smith[a,1,2], Ayya Alieva[a], Qing Wang[a], Michael P. Brenner[a,b,2], and Stephan Hoyer[a,2]

[a]Google Research, Mountain View, CA 94043; and [b]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138

Edited by Andrea L. Bertozzi, University of California, Los Angeles, CA, and approved March 25, 2021 (received for review January 29, 2021)

Numerical simulation of fluids plays an essential role in modeling many physical phenomena, such as weather, climate, aerodynamics, and plasma physics. Fluids are well described by the Navier–Stokes equations, but solving these equations at scale remains daunting, limited by the computational cost of resolving the smallest spatiotemporal features. This leads to unfavorable trade-offs between accuracy and tractability. Here we use end-to-end deep learning to improve approximations inside computational fluid dynamics for modeling two-dimensional turbulent flows. For both direct numerical simulation of turbulence and large-eddy simulation, our results are as accurate as baseline solvers with 8 to 10× finer resolution in each spatial dimension, resulting in 40- to 80-fold computational speedups. Our method remains stable during long simulations and generalizes to forcing functions and Reynolds numbers outside of the flows where it is trained, in contrast to black-box machine-learning approaches. Our approach exemplifies how scientific computing can leverage machine learning and hardware accelerators to improve simulations without sacrificing accuracy or generalization.

machine learning | turbulence | computational physics | nonlinear partial differential equations



- **Data assimilation**
- Objective: find **state** that minimizes a cost function *J(**x**)*
- *3D Var*

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + (\mathbf{y} - H[\mathbf{x}])^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - H[\mathbf{x}])$$

$$J(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}_b\|_{\mathbf{B}}^2 + \|\mathbf{y} - H[\mathbf{x}]\|_{\mathbf{R}}^2$$

4D Var

$$J(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}_b\|_{\mathbf{B}}^2 + \sum_i \|\mathbf{y_i} - H[\mathbf{x_i}]\|_{\mathbf{R_i}}^2$$

Talagrand and Courtier *QJRMS* 1987

- **Data assimilation**
- Extension to update *parameters*
- Objective: find **state + parameters** that minimize a cost function *J(**x**)*

$$J(\mathbf{x};\mathbf{p}) = \|\mathbf{x}-\mathbf{x}_b\|^2_{\mathbf{Bx}} + \|\mathbf{p}-\mathbf{p}_b\|^2_{\mathbf{Bp}} + \|\mathbf{y}-H[\mathbf{x}]\|^2_{\mathbf{R}}$$

- Caveat: underlying dynamics is *modeled*: $\frac{d\mathbf{x}}{dt} = g(\mathbf{x},\mathbf{p},t)$

But only an *ansatz* of the world

For weather: even a deficient model can work well because of frequent assimilation
For climate: cannot have a deficient model+DA ➔ diverging

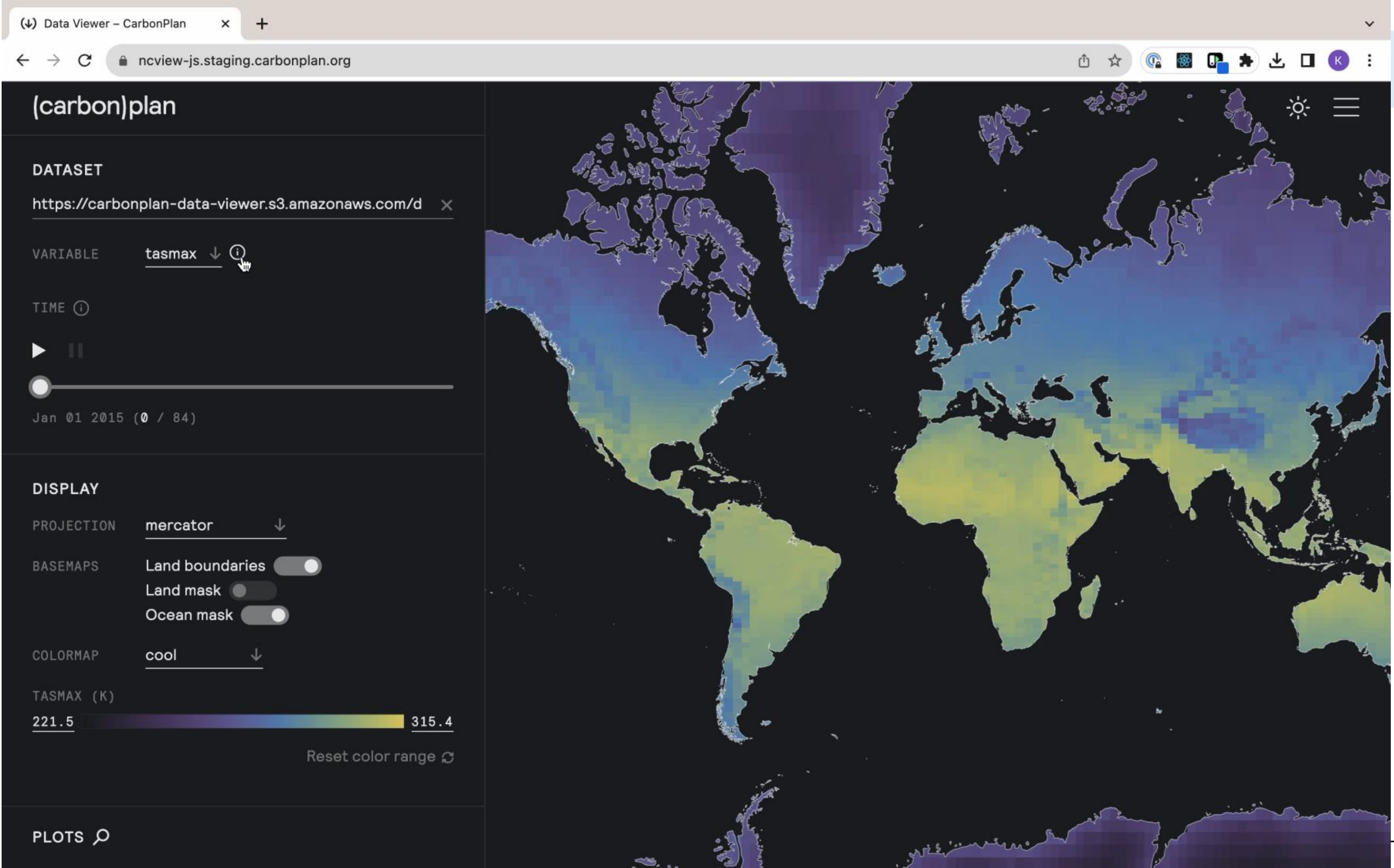
Evensen et al. IEEE 2009

- **Machine learning**
- Objective: find (neural network) **parameters p** that minimize a cost function *J(**p**;**x**)*

$$J(\mathbf{p};\mathbf{x}) = \|\mathbf{y} - f(\mathbf{p};\mathbf{x})]\|_{\mathbf{I}}^{2}$$

- Parameters are estimated over a set $\{\mathbf{x_i}\}$
- Clear similarities with Data Assimilation (DA)
- Caveats: no observational errors + assumes *direct* observations

Farchi et al., 2021 *J Comp Sci*; Cheng et al. 2023 *IEE/CAA*


Data Viewer – CarbonPlan
ncview-js.staging.carbonplan.org
(carbon)plan
DATASET
https://carbonplan-data-viewer.s3.amazonaws.com/d
VARIABLE
tasmax
TIME
Jan 01 2015 (0 / 84)
DISPLAY
PROJECTION
mercator
BASEMAPS
Land boundaries
Land mask
Ocean mask
COLORMAP
cool
TASMAX (K)
221.5
315.4
Reset color range
PLOTS