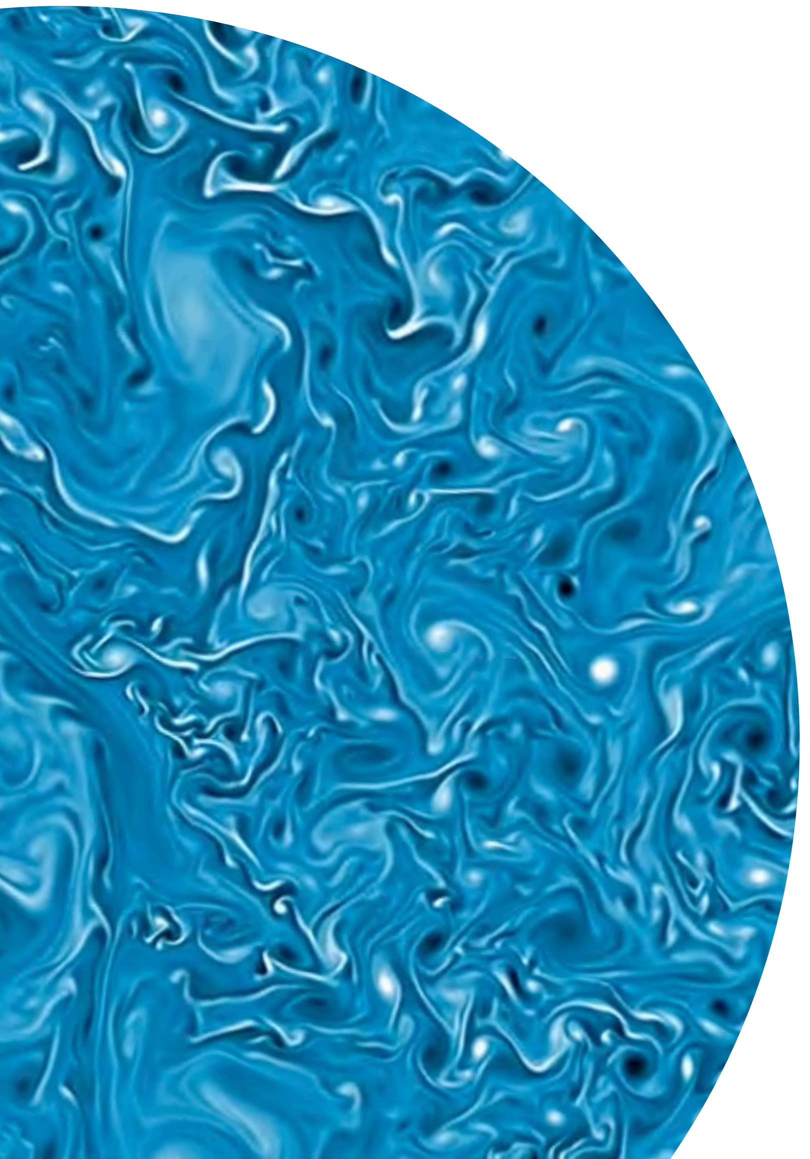
Atelier turbulence OSUG

5 Dec 2024

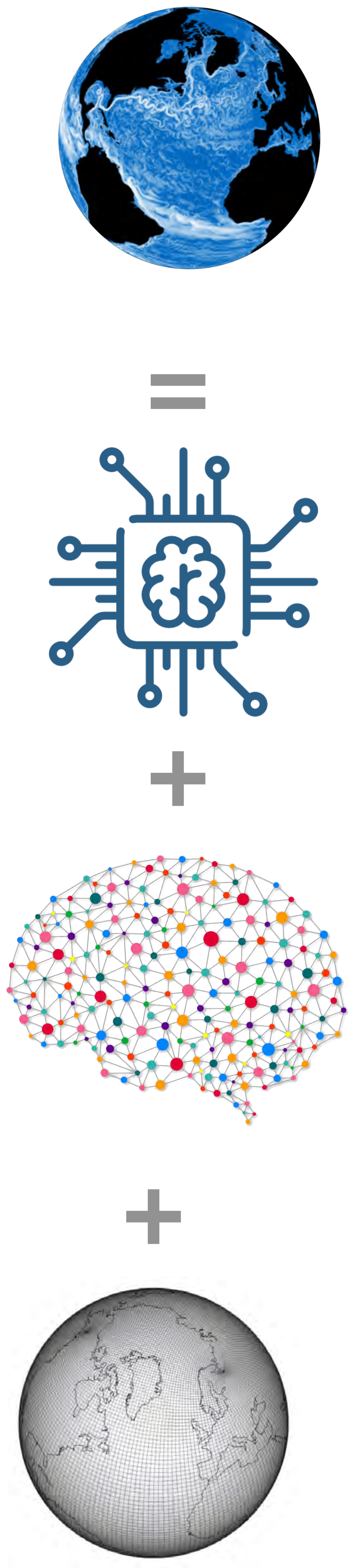
Combining physics and machine learning in hybrid climate models

Julien Le Sommer - computational oceanographer

Institut des Géosciences de l'Environnement, Grenoble



Objectives of this talk

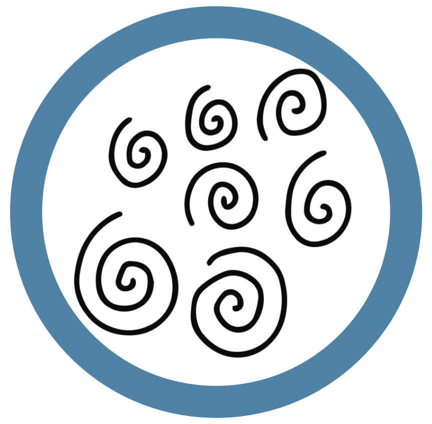


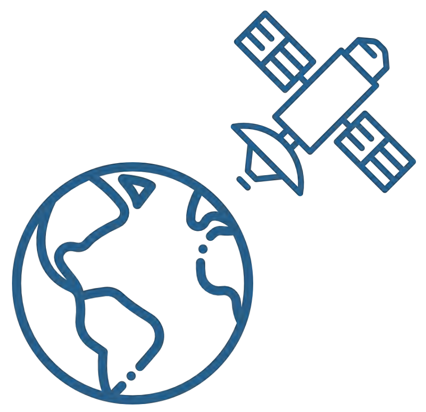
- how ML is leveraged in computational oceanography

- with methods from the emerging field of SciML

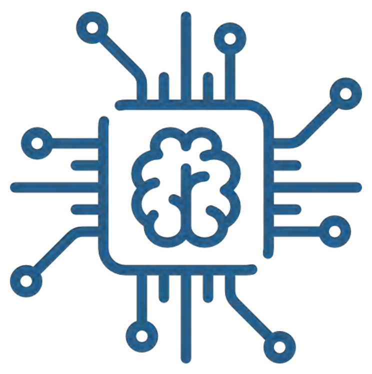
- How this leads to to deep changes in our systems

- and some interesting questions for the future…

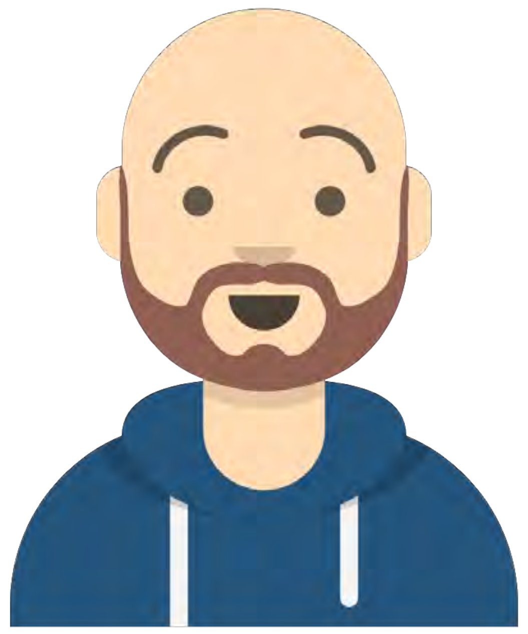




Observations

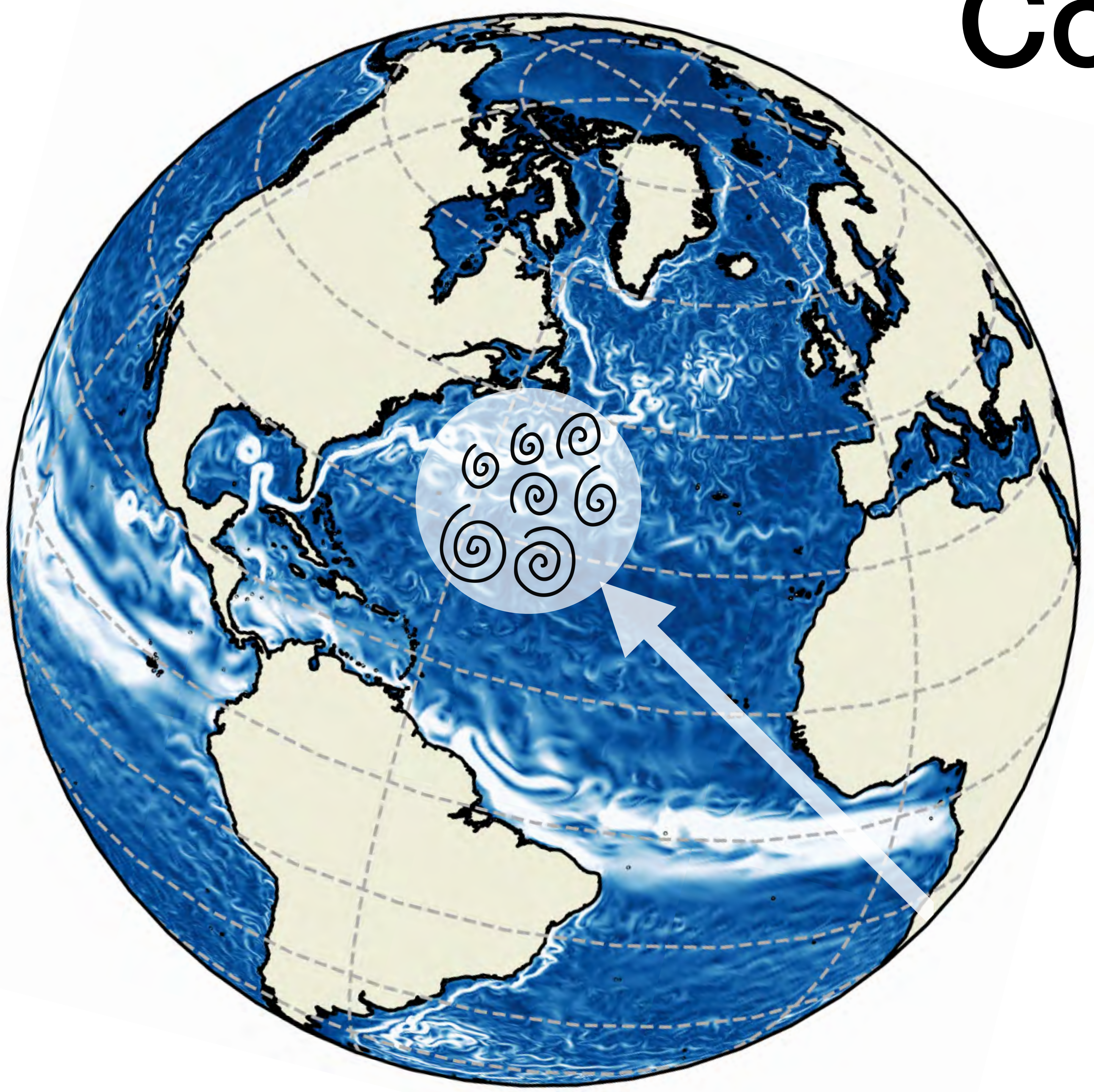


Models / AI



The context of computational oceanography

Computational oceanography



physical oceanography

currents, parameters

understand the functioning

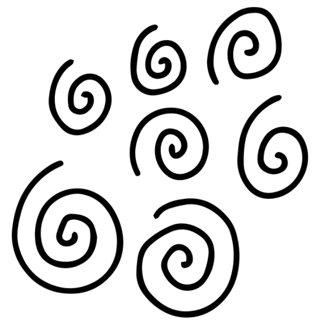
forecast its evolution (timescales)

climate - environnemental changes

human activities

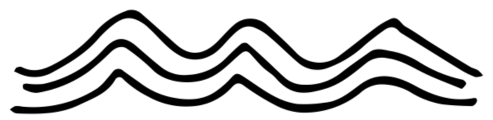
Scale interactions, processes

Macro-turbulence

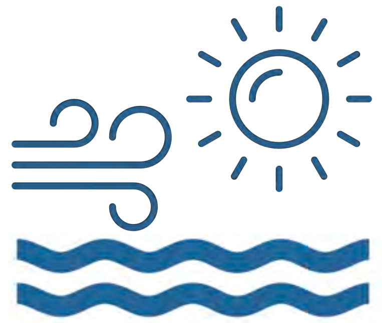


Surface waves

internal waves (tides)

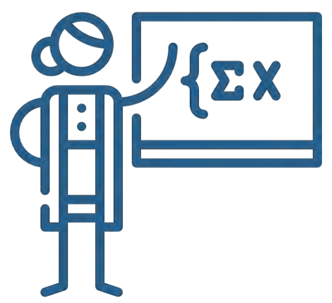


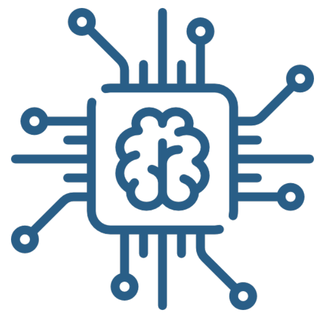
Interactions with components



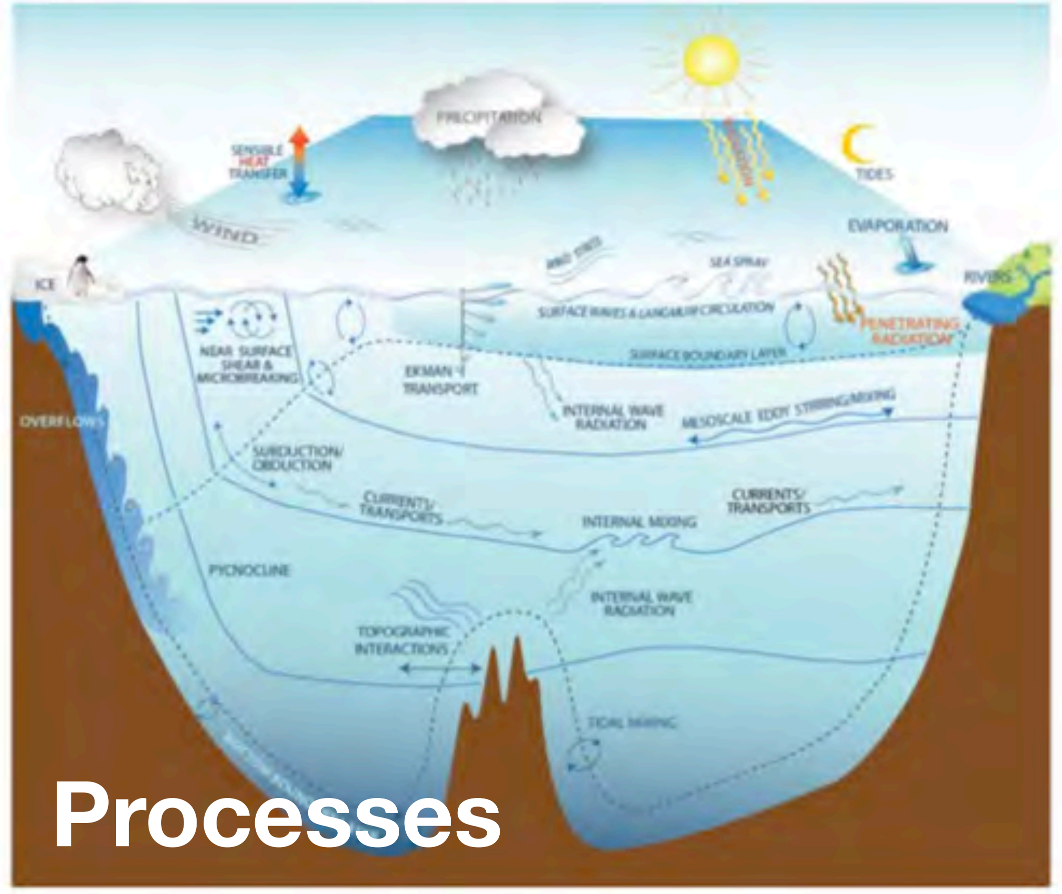
develops and use **numerical tools** and methods

maths, numerics, compute, data

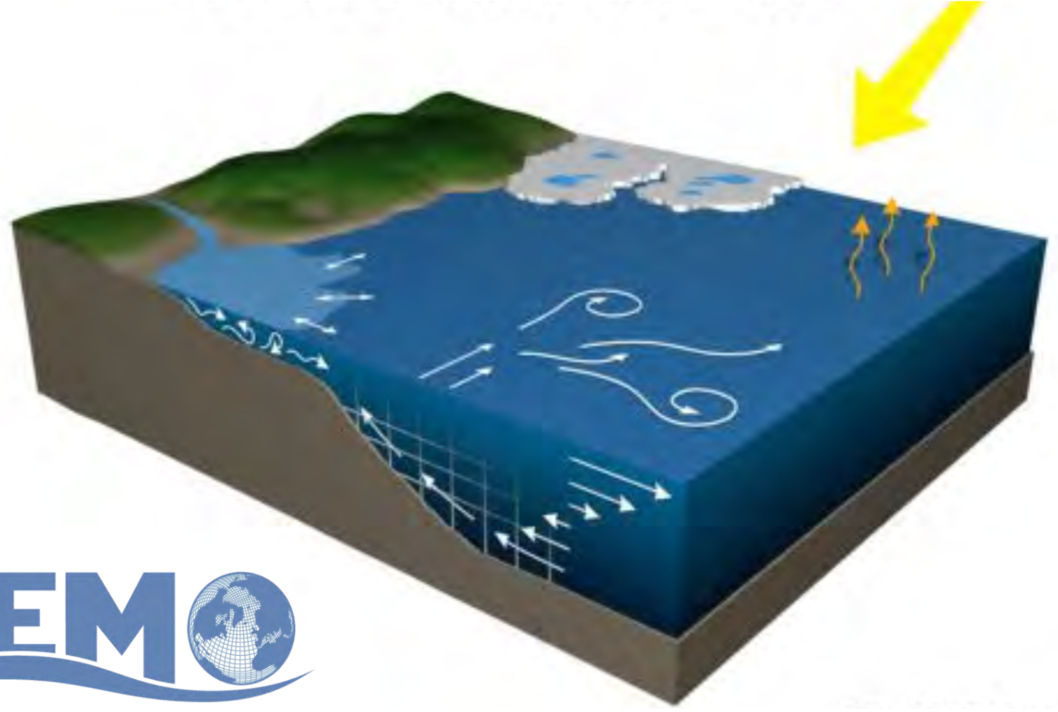


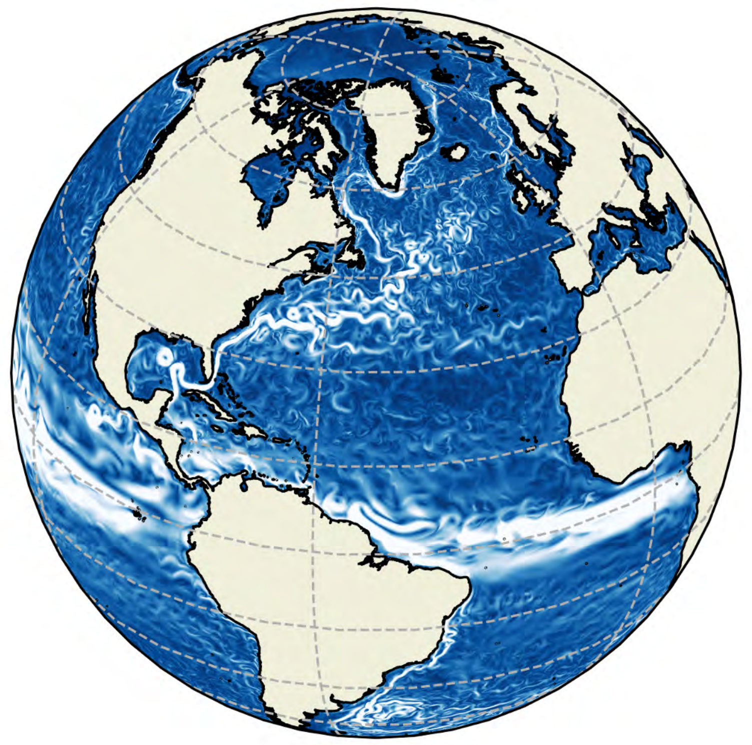


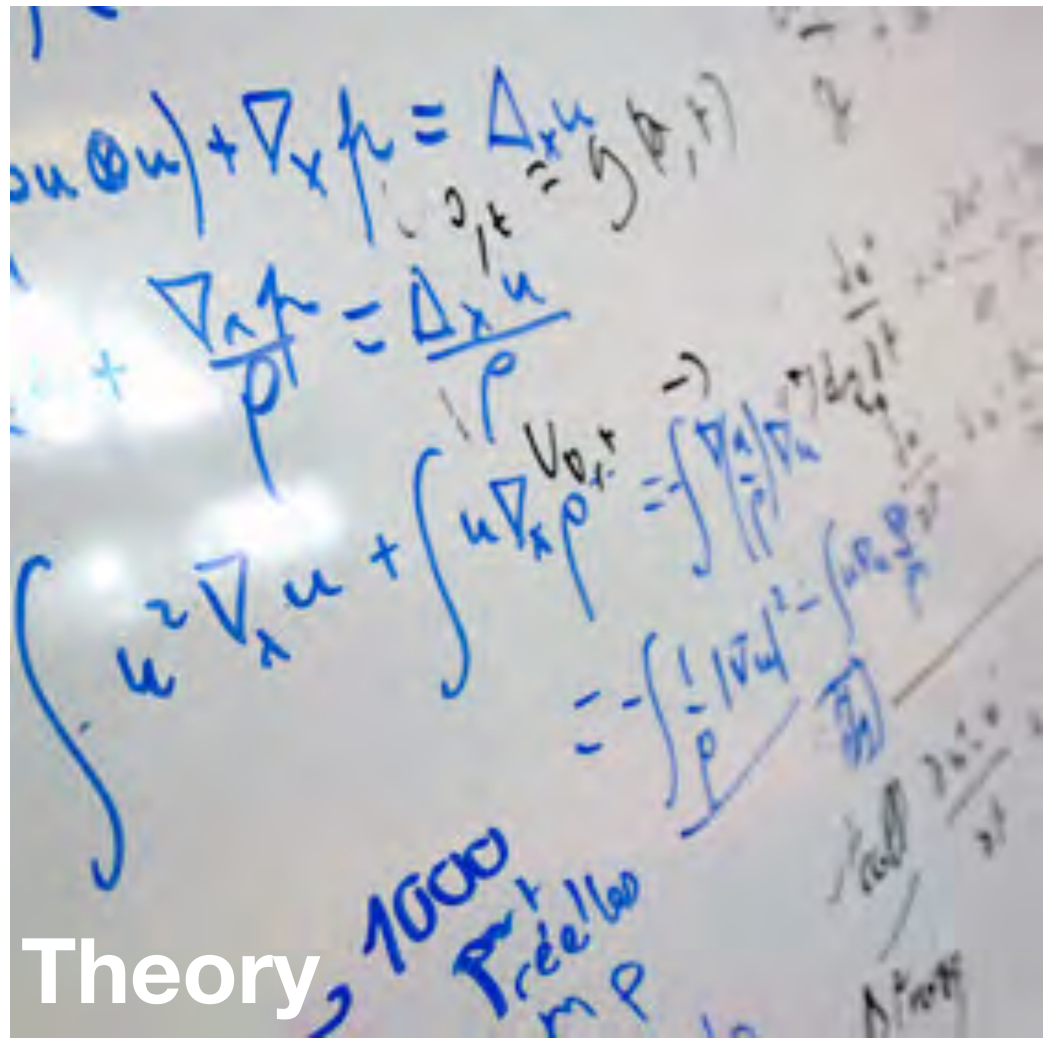
A key tool : ocean models

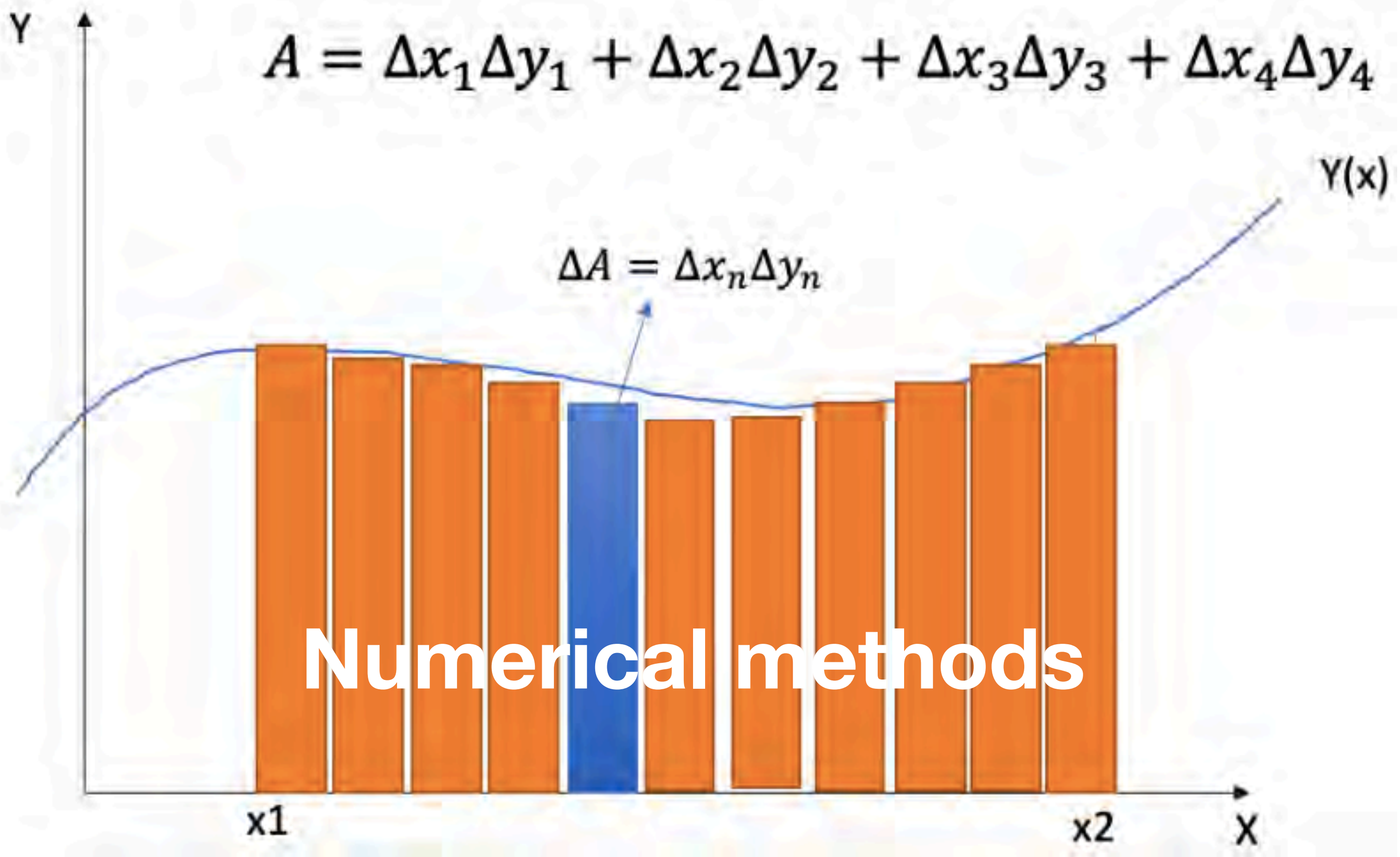


Physics-based models
(ocean circulation models)



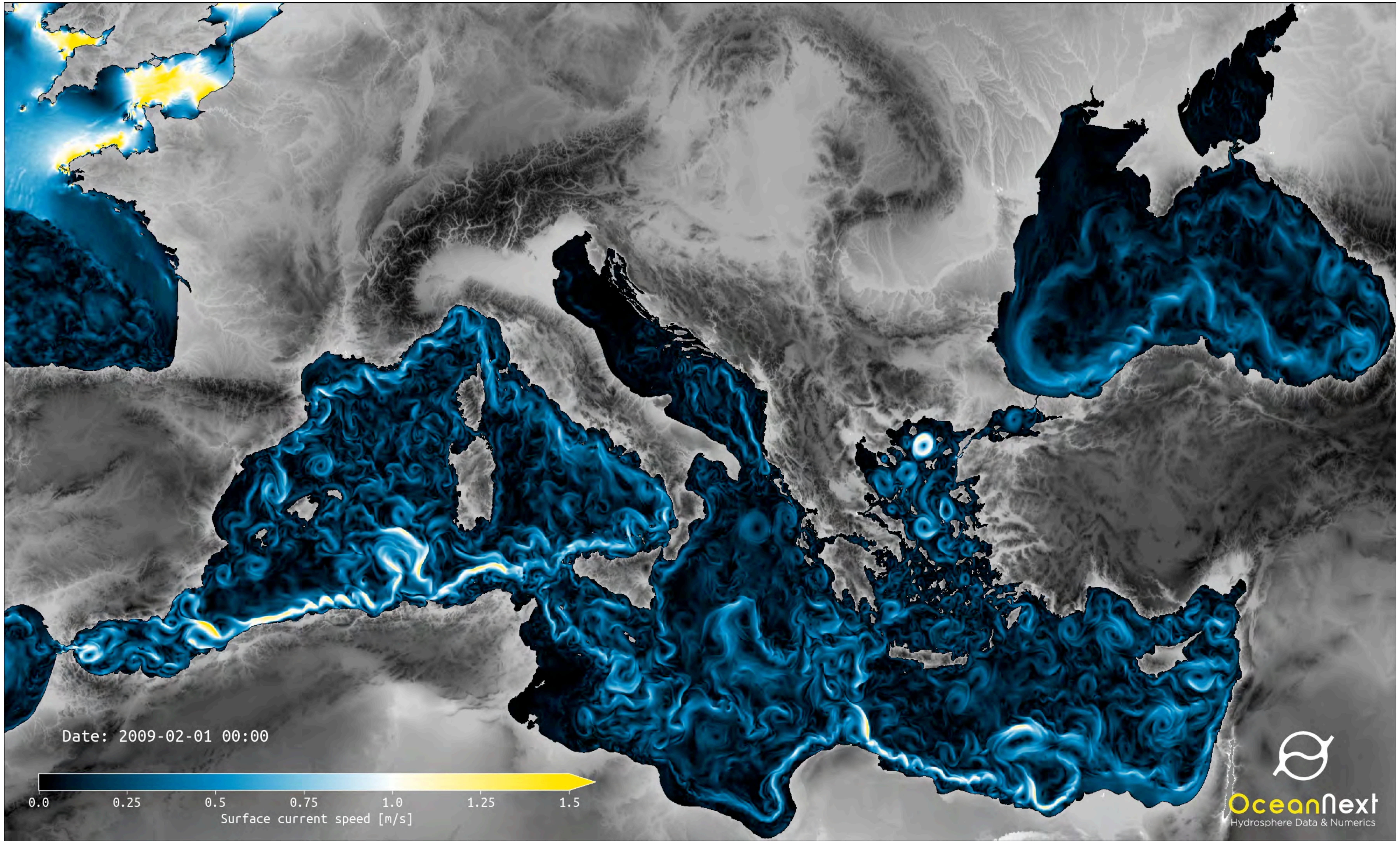






Physical models summarize our understanding of physical systems

A key tool : ocean models



dx ~ 1km

tides, eddies

NEMO

Tier 1 HPC

>200 000 lines

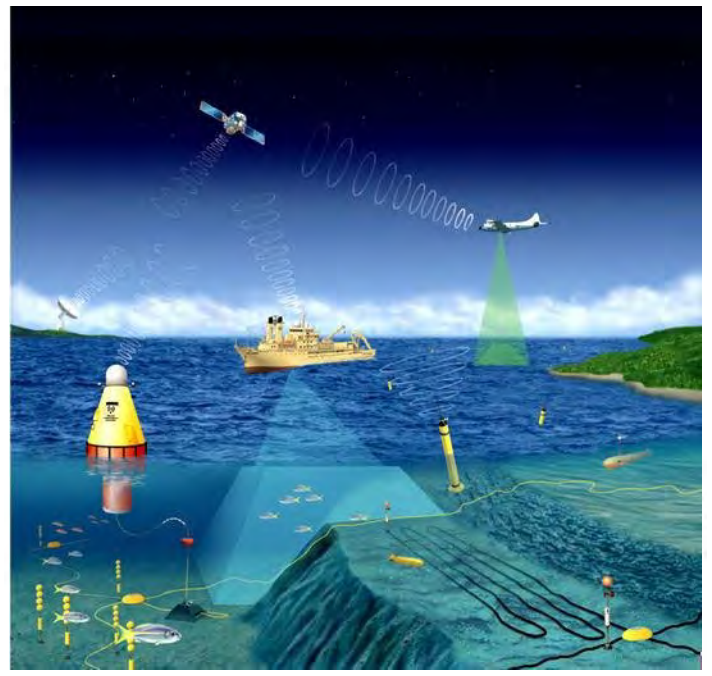
>15 yrs

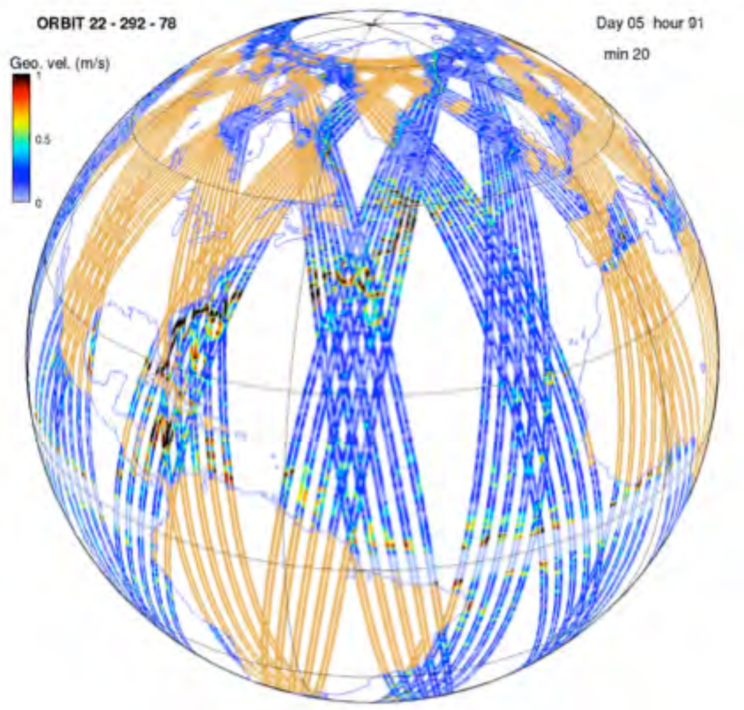
5 institutions

Physical models summarize our understanding of physical systems

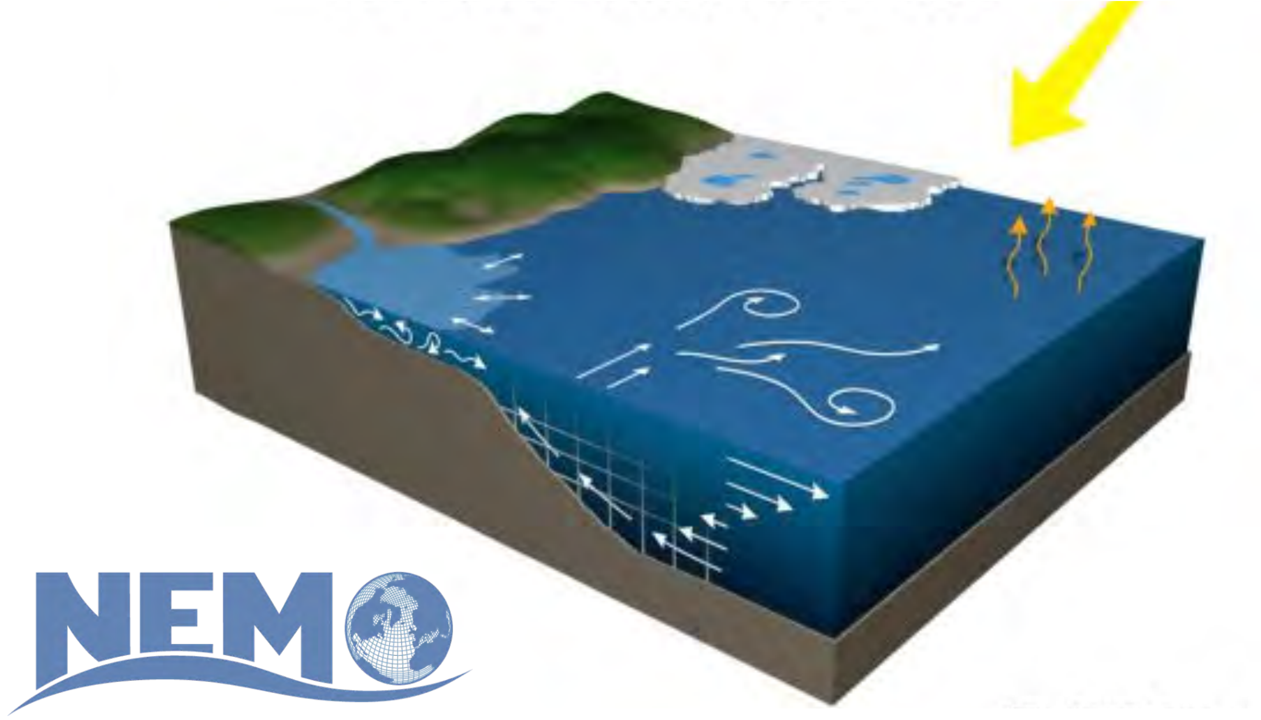
Our toolboxes

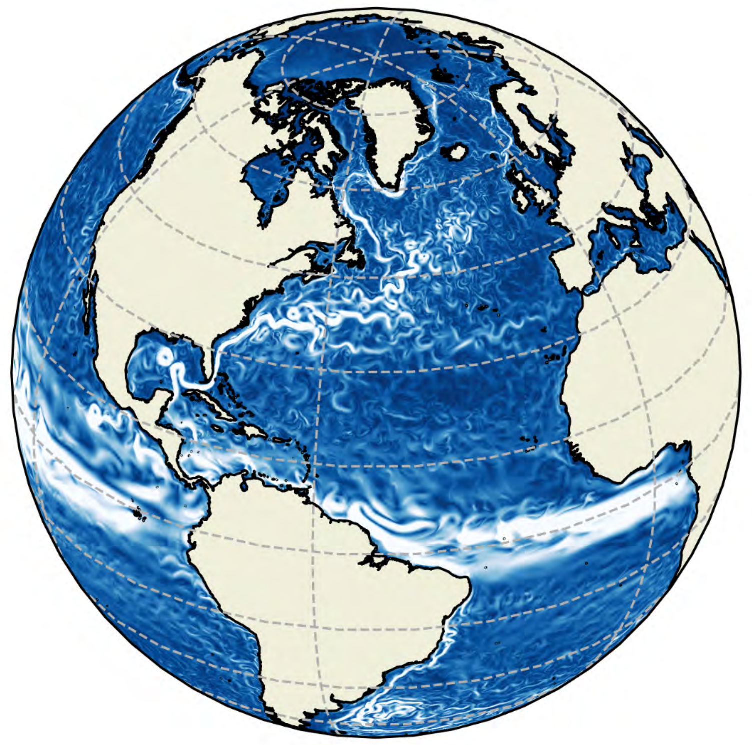
Observations
(in situ/satellite)



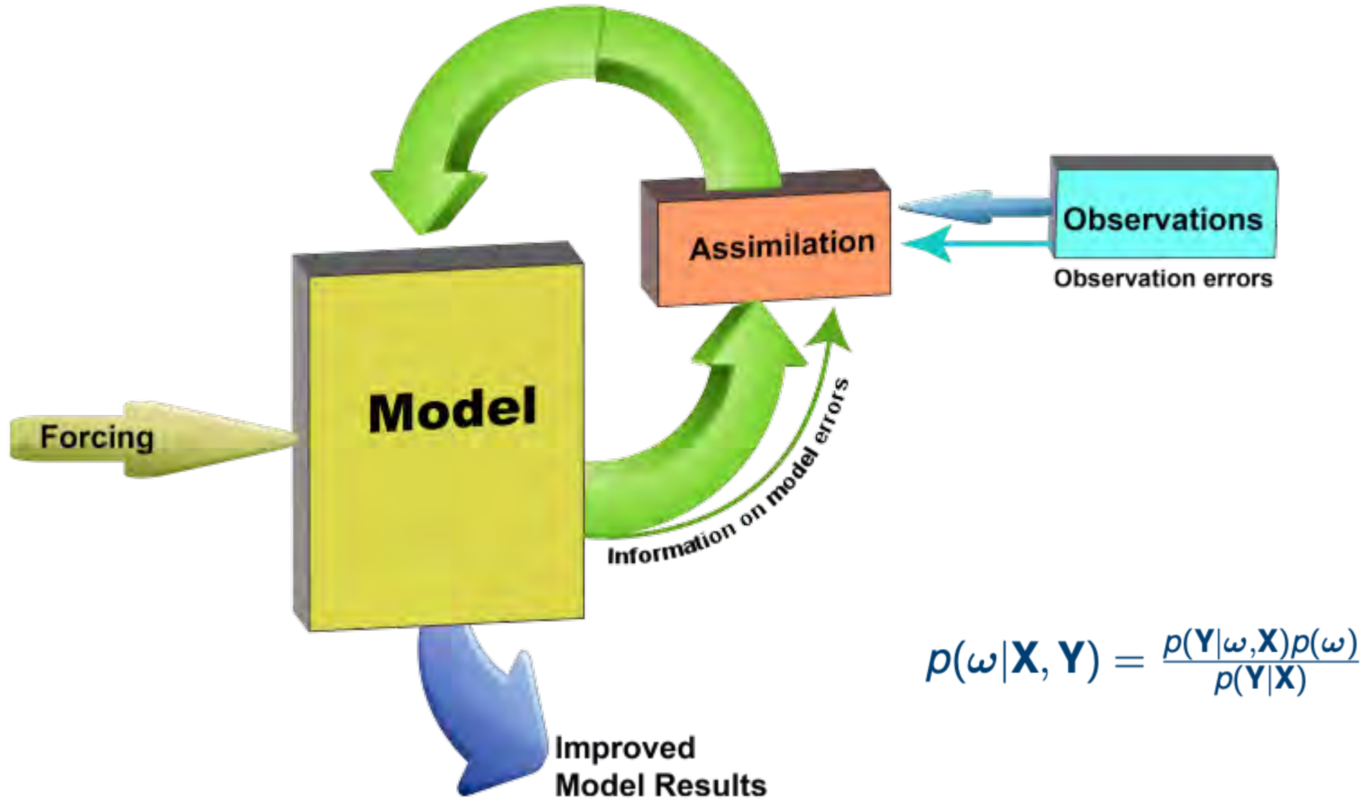


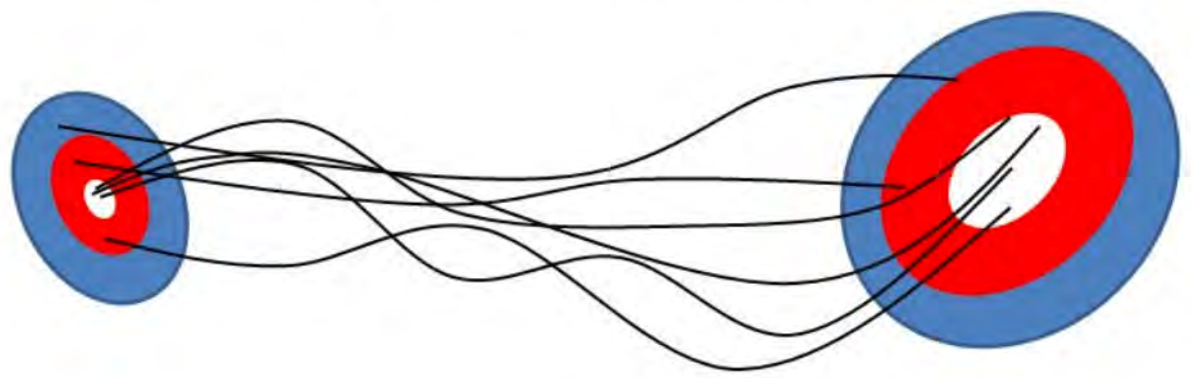
Physical models
(ocean circulation models)





Inverse methods
(data assimilation)



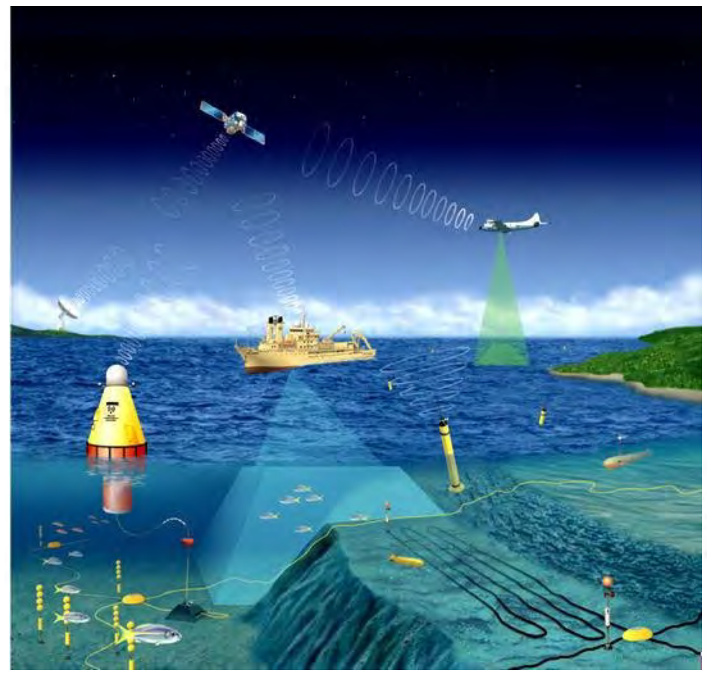


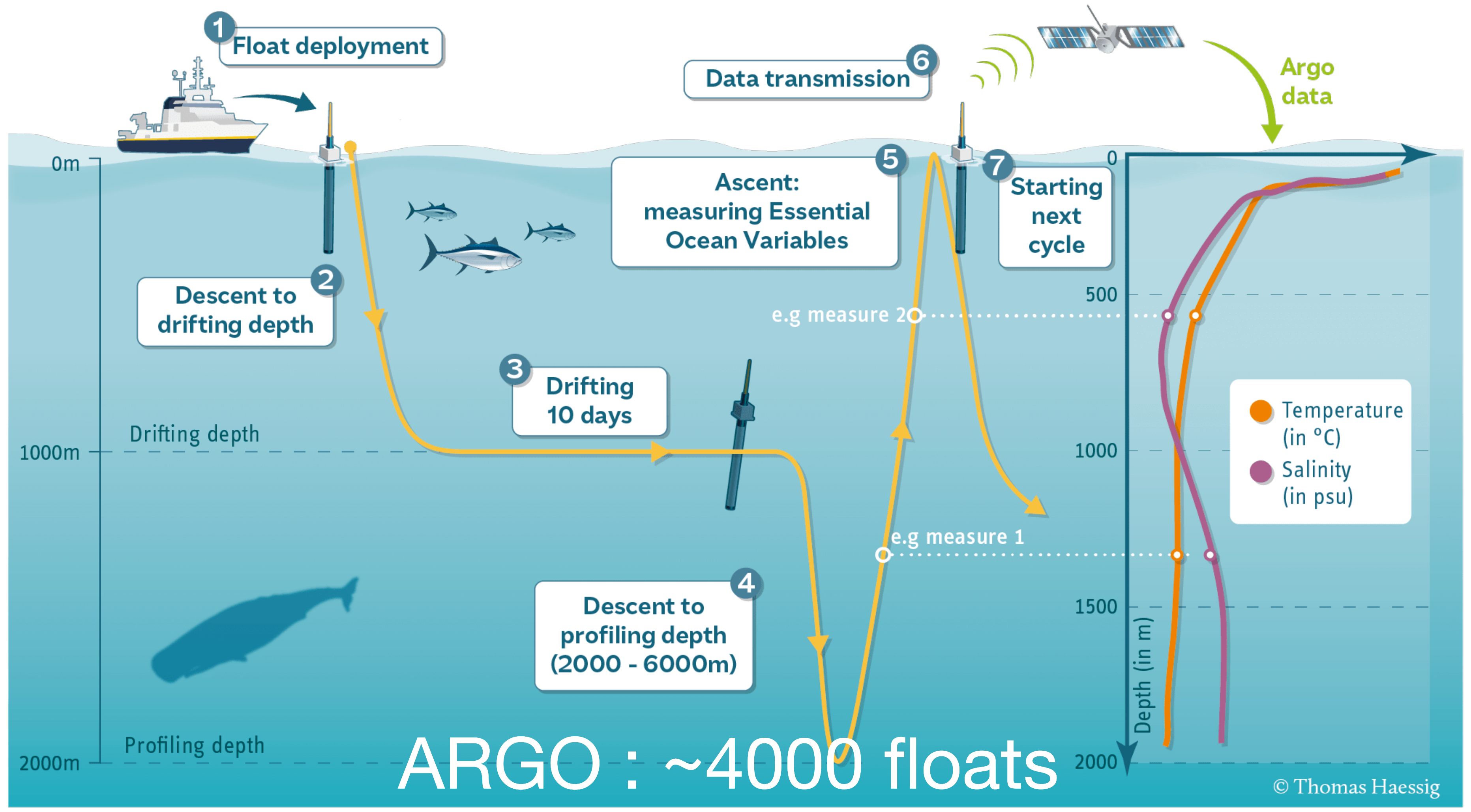
Tools for understanding but also monitoring and forecasting ocean circulation

Observations of the ocean

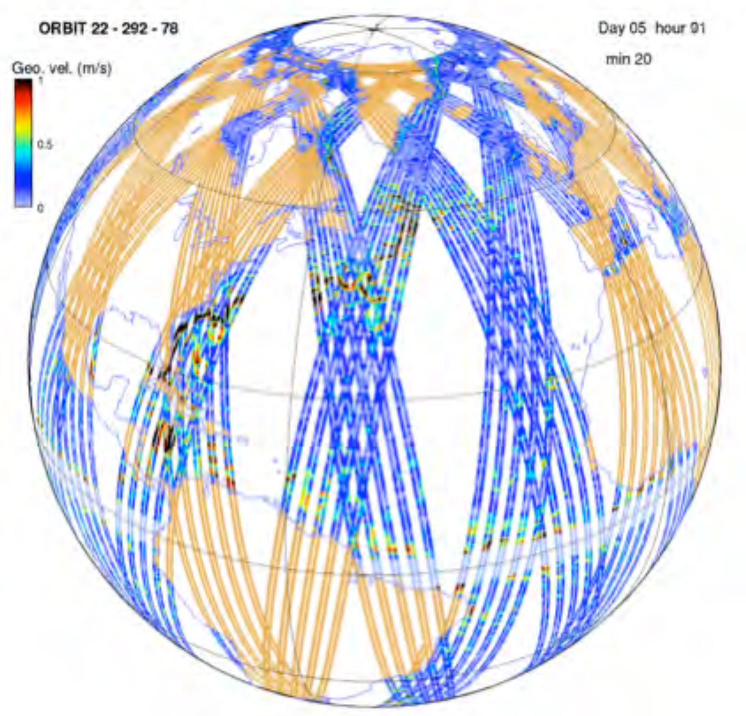
Observations

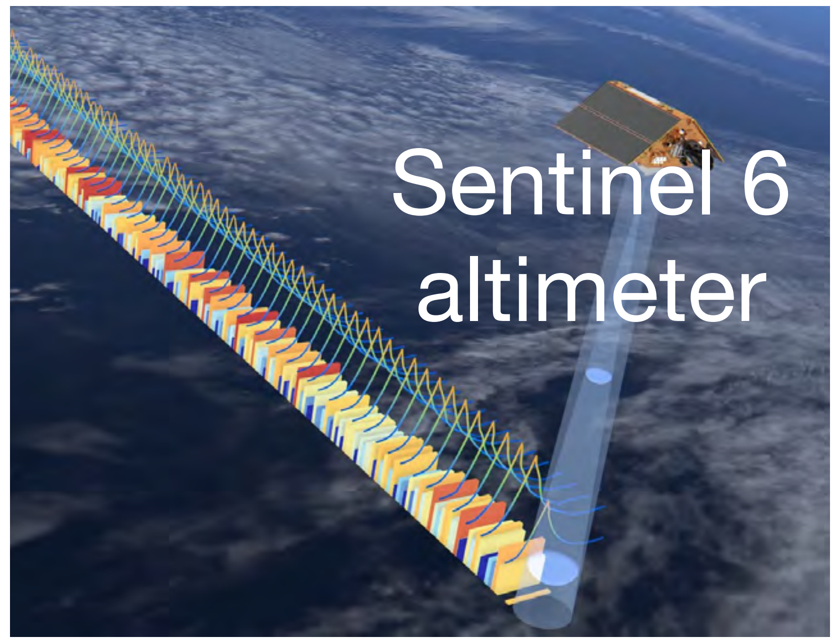
in-situ

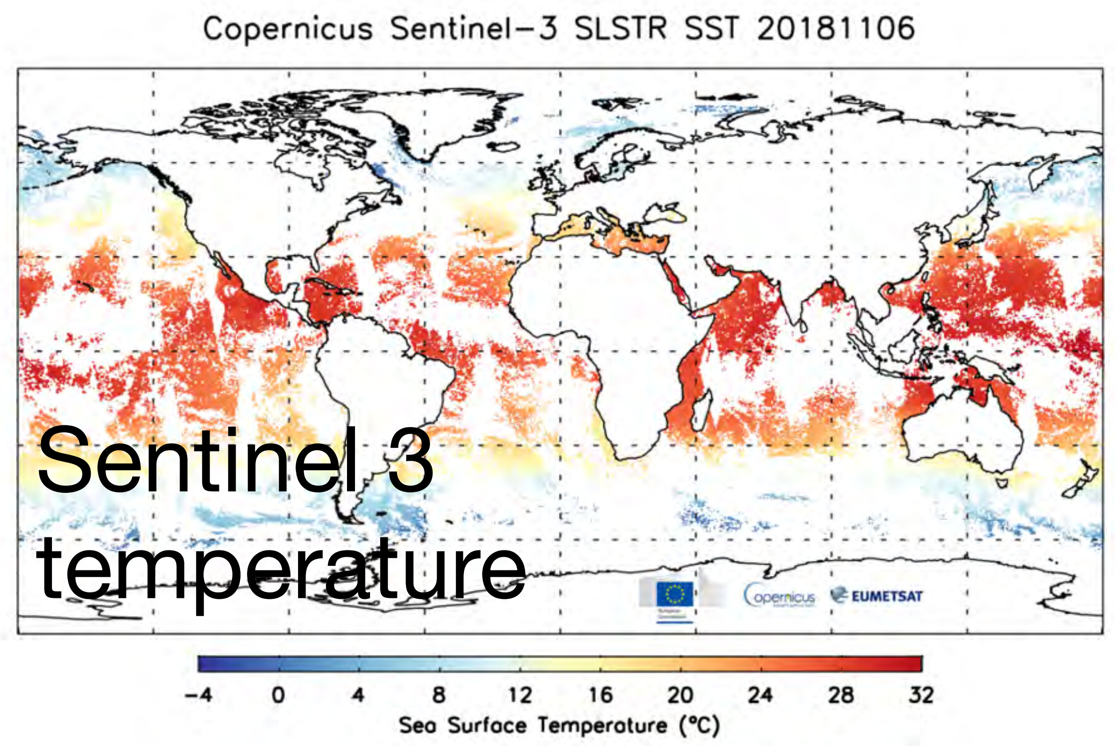


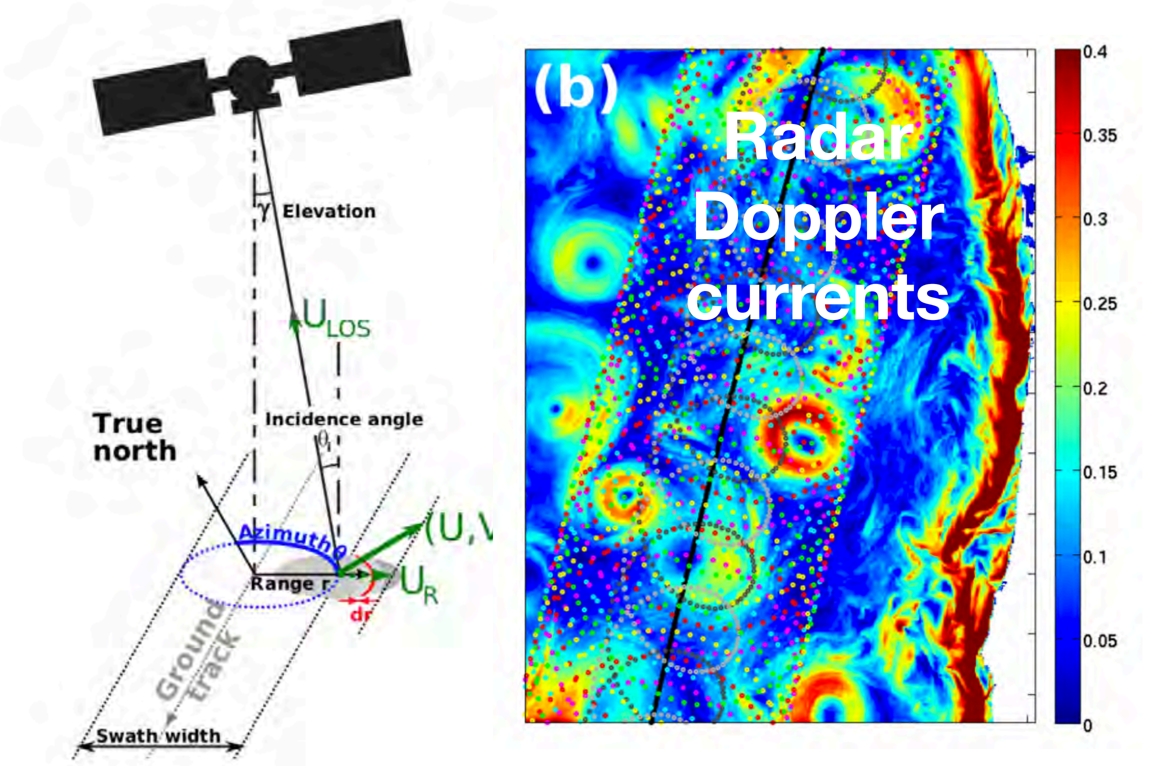


satellite







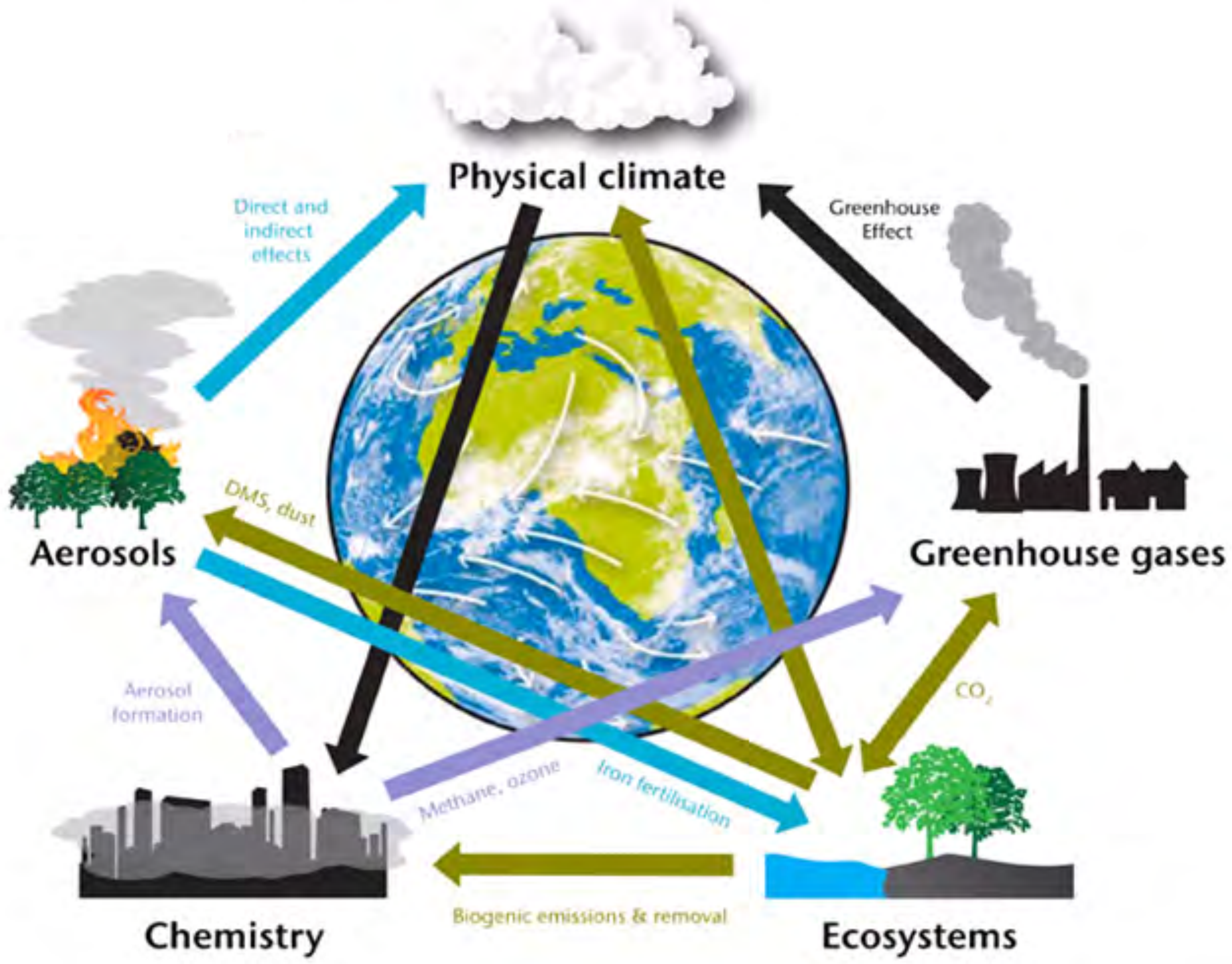


Continuously operated networks

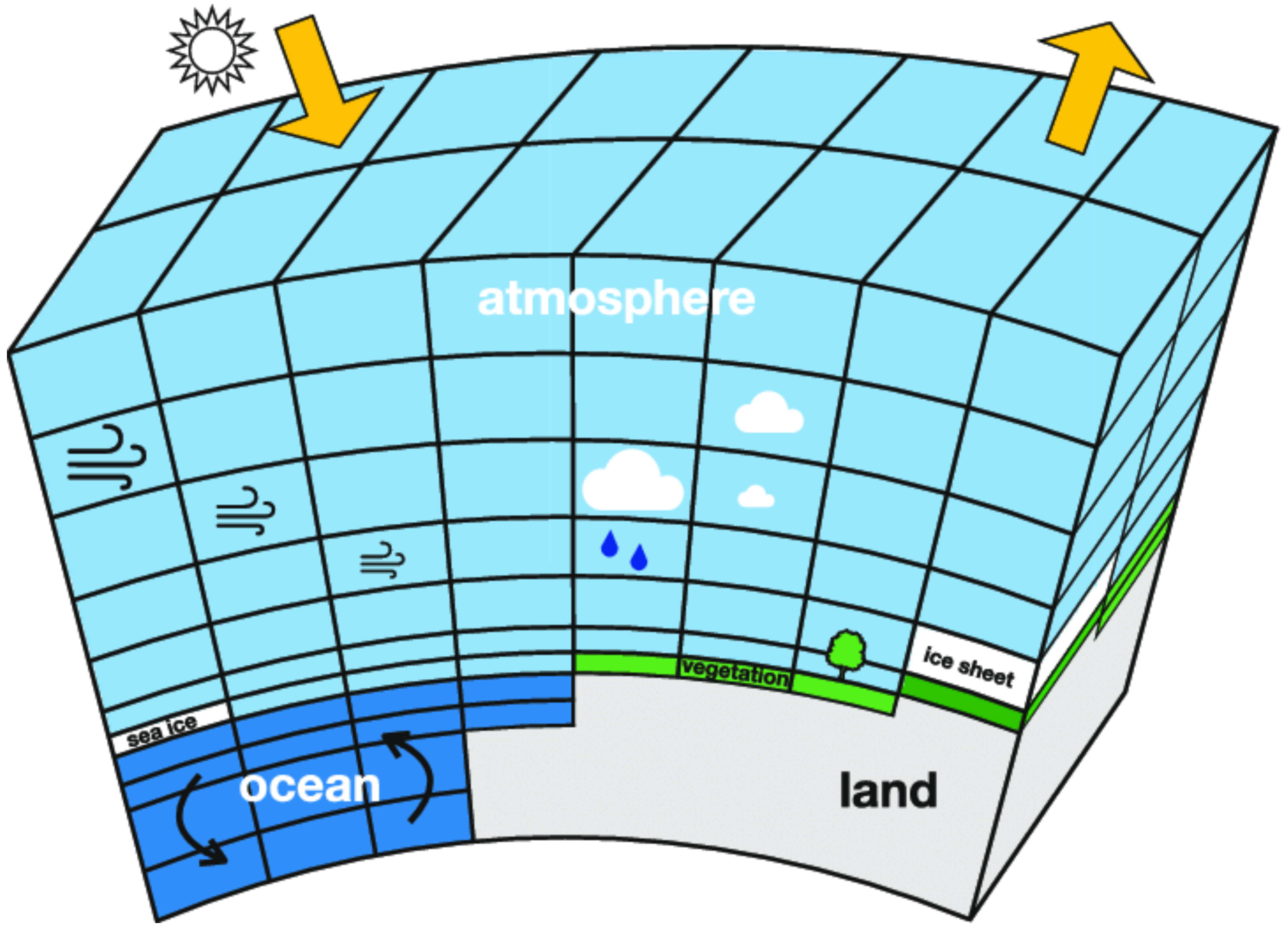
New platforms

How AI (SciML) **is affecting our field ?**

Tools are integrated into systems

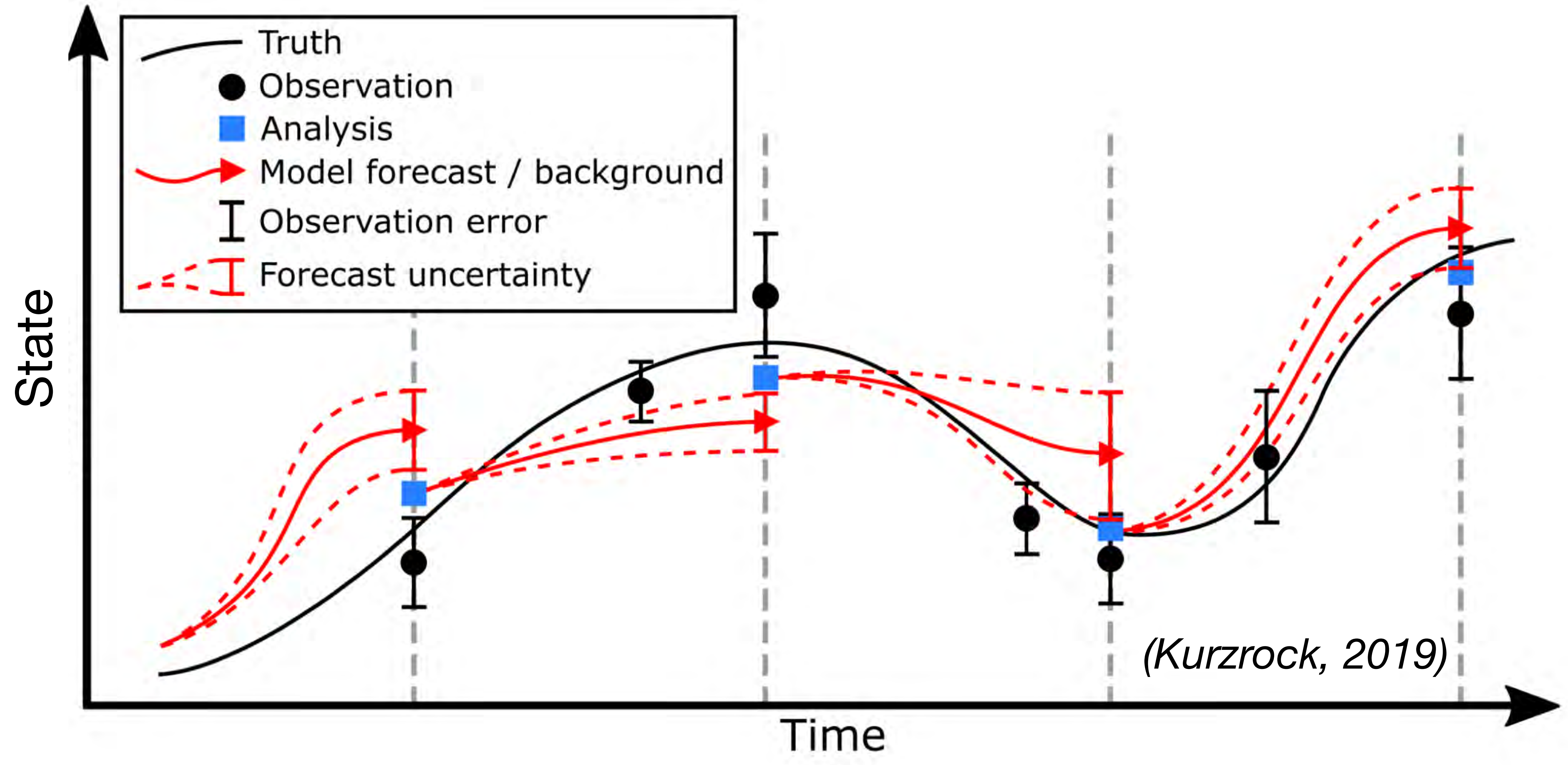


Earth System Models
(IPCC)



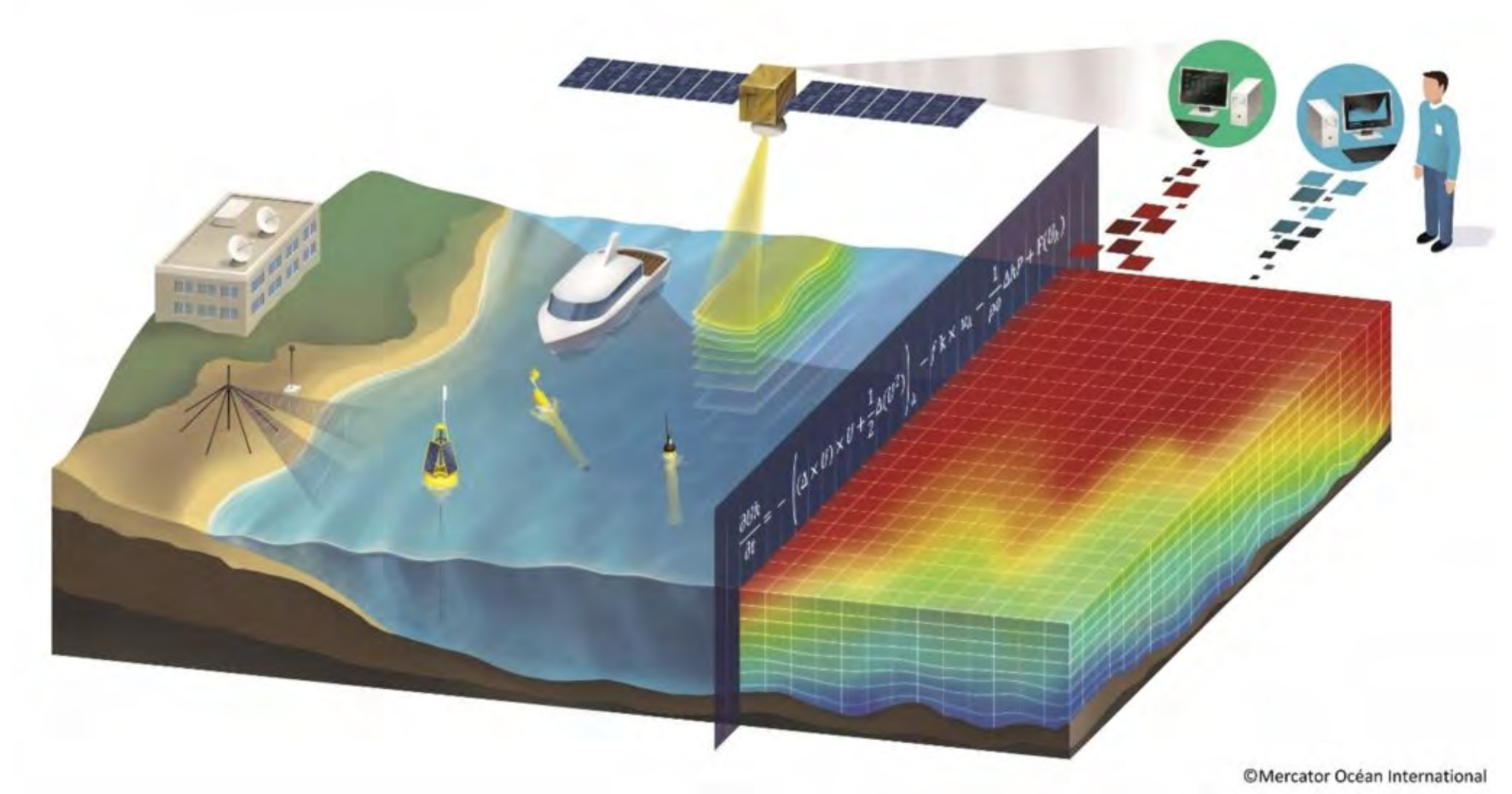
Combining models of each components of the climate system

Tools are integrated into systems



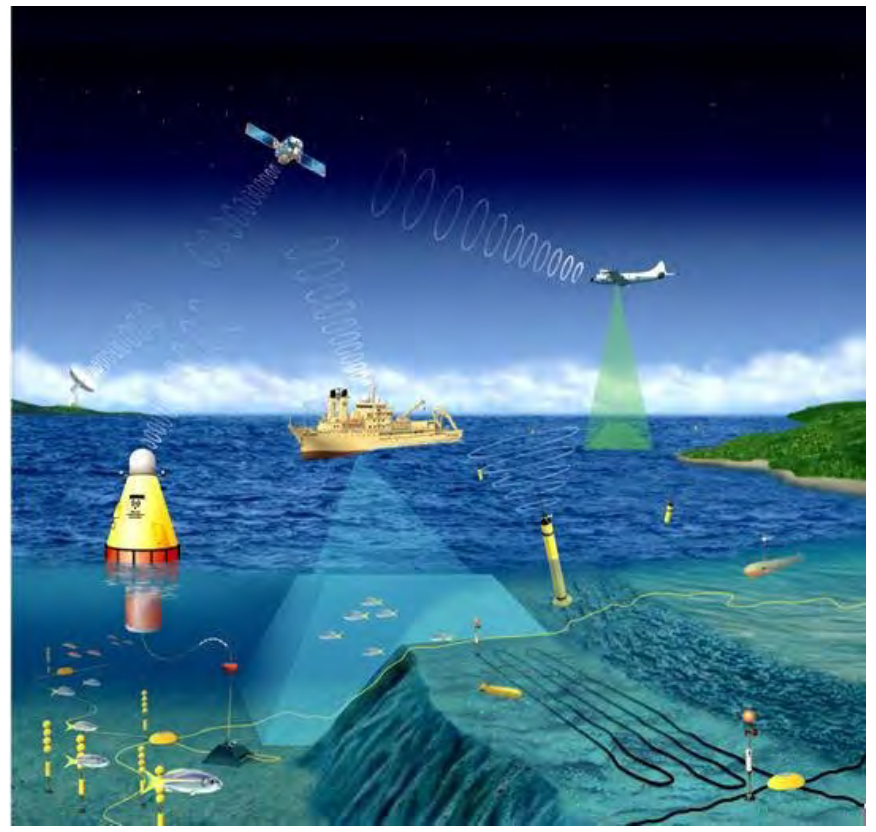
Combining models and observations to produce forecasts

Copernicus Marine Service

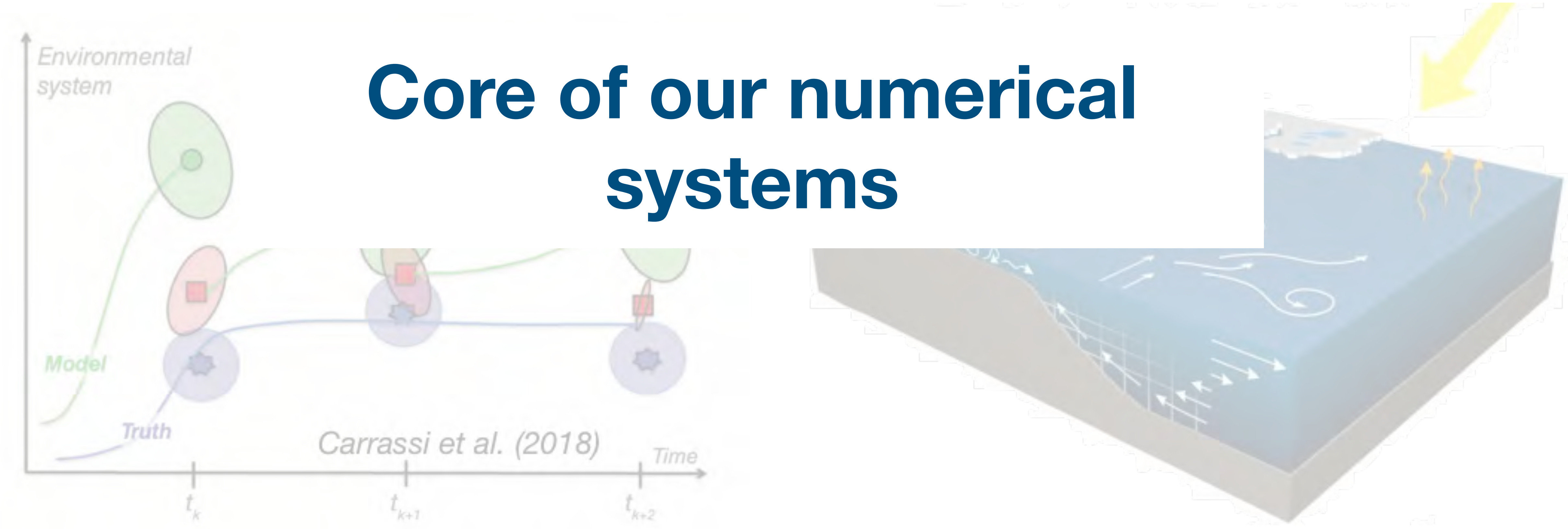


Operational prediction systems
(Copernicus)

How AI is affecting our systems



Core of our numerical systems



Observations → **Data assimilation** → **Physics-based forecast** → **Post-processing, dissemination**

Upstream

Downstream

denoising, inpainting
parameter retrieval
quality control

AI, machine learning & data-driven approaches

data fusion,
tailored services
data mining

AI-based ocean forecasting

XiHe: A Data-Driven Model for Global Ocean Eddy-Resolving Forecasting

Xiang Wang, Renzhi Wang, Ningzi Hu, Pinqiang Wang, Peng Huo, Guihua Wang, Huizan Wang, Senzhang Wang, Junxing Zhu, Jianbo Xu, Jun Yin, Senliang Bao, Ciqiang Luo, Ziqing Zu, Yi Han, Weimin Zhang, Kaijun Ren, Kefeng Deng, Junqiang Song

Abstract—Global ocean forecasting is fundamentally important to support marine activities. The leading operational Global Ocean Forecasting Systems (GOFSs) use physics-driven numerical forecasting models that solve the partial differential equations with expensive computation. Recently, specifically in atmosphere weather forecasting, data-driven models have demonstrated significant potential for speeding up environmental forecasting by orders of magnitude, but there is still no data-driven GOFS that matches the forecasting accuracy of the numerical GOFSs. In this paper, we propose the first data-driven $1/12^\circ$ resolution global ocean eddy-resolving forecasting model named *XiHe*, which is established from the 25-year France Mercator Ocean International's daily GLORYS12 reanalysis data. *XiHe* is a hierarchical transformer-based framework coupled with two special designs. One is the land-ocean mask mechanism for focusing exclusively on the global ocean circulation. The other is the ocean-specific block for effectively capturing both local ocean information and global teleconnection. Extensive experiments are conducted under satellite observations, *in situ* observations, and the IV-TT Class 4 evaluation framework of the world's leading operational GOFSs from January 2019 to December 2020. The results demonstrate that *XiHe* achieves stronger forecast performance in all testing variables than existing leading operational numerical GOFSs including Mercator Ocean Physical SYstem (PSY4), Global Ice Ocean Prediction System (GIOPS), BLUElinK OceanMAPS (BLK), and Forecast Ocean Assimilation Model (FOAM). Particularly, the accuracy of ocean current forecasting of *XiHe* out to 60 days is even better than that of PSY4 in just 10 days. Additionally, *XiHe* is able to forecast the large-scale circulation and the mesoscale eddies. Furthermore, it can make a 10-day forecast in only 0.36 seconds, which accelerates the forecast speed by thousands of times compared to the traditional numerical GOFSs.

Index Terms—Global Ocean Forecasting, Deep Learning, Eddy Resolving, Data-Driven, AI for Science

1 Introduction

Ocean forecasting is critically important for many marine activities. At present, the leading GOFSs (e.g. Mercator Ocean Physical SYstem (PSY4) and Real-Time Ocean Forecast System (RTOFS)) use physics-driven models in fluid mechanics and thermodynamics to predict future ocean motion states and phenomena based on current ocean conditions [1]. The GOFSs adopt numerical methods that rely on supercomputers to solve the partial differential equations of the physical models. Due to their desirable performance, they are operationally run in different countries worldwide. However, numerical forecasting methods are usually computationally expensive and slow. For example, a single forecasting simulation in the numerical GOFSs may take hours on a supercomputer with hundreds of computational nodes [2]. Besides, improving the forecasting accuracy of these methods is exceedingly challenging because they heavily rely on the human cognitive abilities in understanding the physical laws of the ocean environment [3].

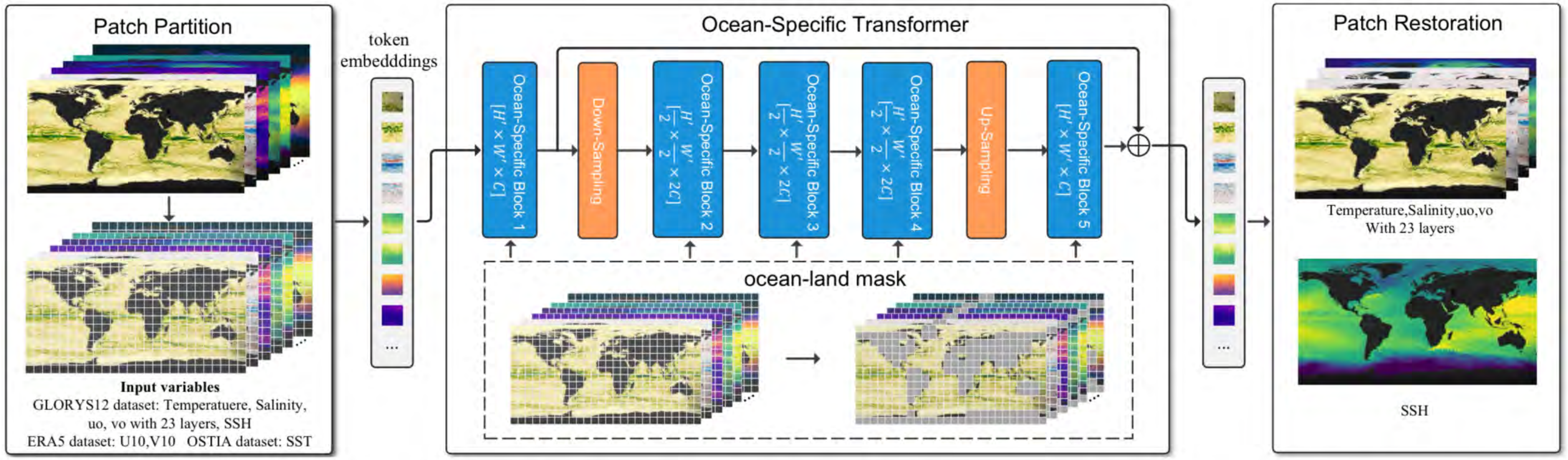
With the recent advances of Artificial Intelligence (AI) techniques, deep learning methods have been widely applied in various prediction/forecasting tasks of different fields and achieved great success. Particularly, some data-driven AI models have shown the potential in atmosphere weather forecasting like *Pangu-Weather* [4] and *GraphCast* [5]. They have achieved comparable or even better prediction results in global medium-range weather forecasting than current leading numerical weather prediction (NWP) methods [4], [5], [6], [7], [8], [9]. One significant advantage of data-driven models is that they can make the forecasting thousands or even tens of thousands of times faster than NWP methods [4]. Furthermore, they can automatically learn the spatial-temporal relationships from massive meteorological data, and effectively capture the rules of weather changing, without introducing the prior knowledge of physics mechanisms.

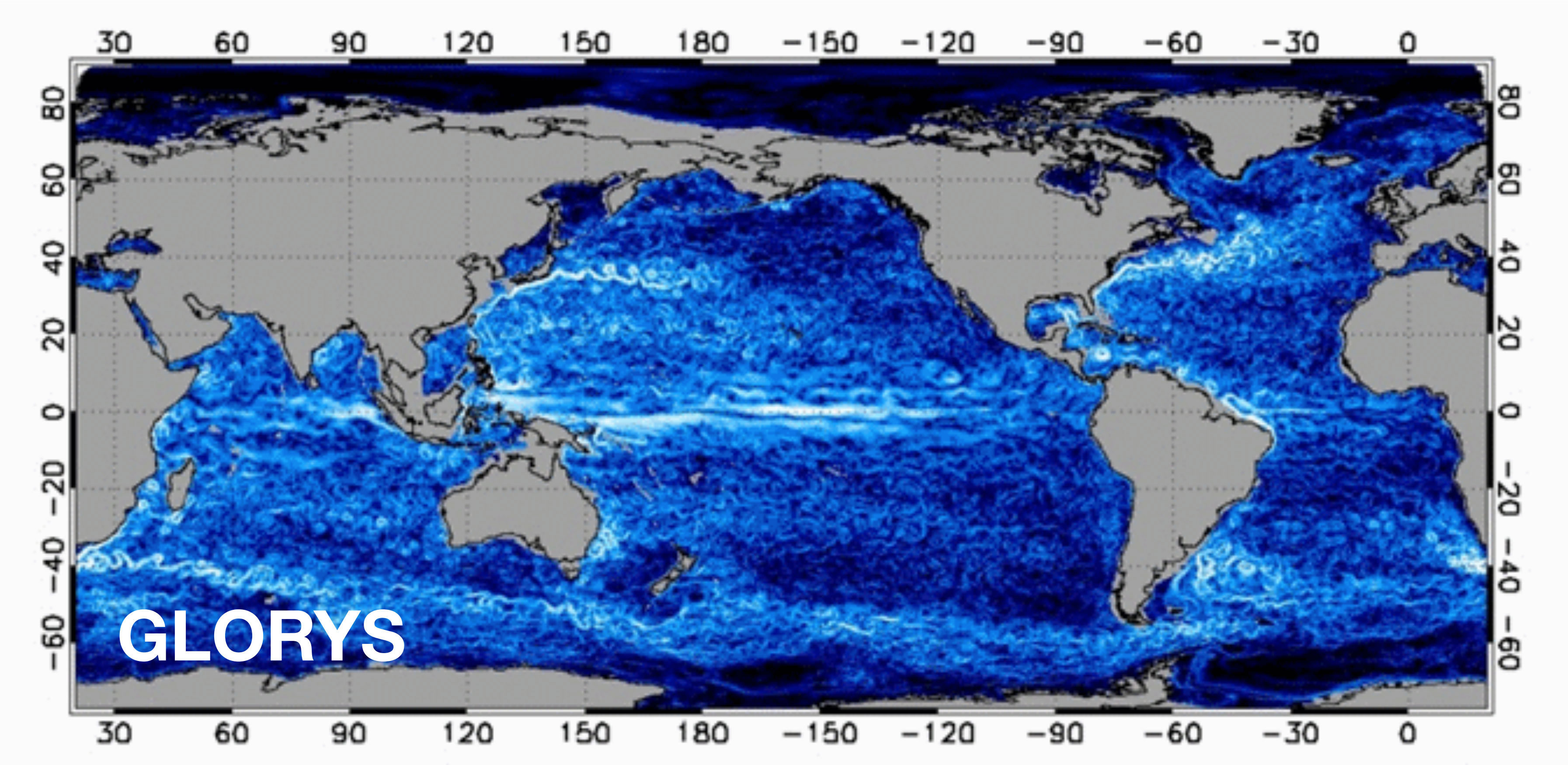
Although data-driven models have achieved promising results in atmosphere weather forecasting, how to build a more accurate and efficient data-driven ocean forecasting model remains an open research issue due to the following

- *Xiang Wang, Pinqiang Wang, Huizan Wang, Junxing Zhu, Jianbo Xu, Senliang Bao, Ciqiang Luo, Yi Han, Weimin Zhang, Kaijun Ren, Kefeng Deng, and Junqiang Song are with the College of Meteorology and Oceanography, National University of Defense Technology, Changsha 410073, China.*
- *Renzhi Wang, Senzhang Wang and Jun Yin are with the School of Computer Science and Engineering, Central South University, Changsha 410083, China.*
- *Ningzi Hu is with the College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao 266580, China.*
- *Peng Huo is with the College of Artificial Intelligence, Tianjin University of Science and Technology, Tianjin 300457, China.*
- *Guihua Wang is with the Department of Atmospheric and Oceanic Sciences, Fudan University, Shanghai 200438, China.*
- *Ziqing Zu, Key Laboratory of Marine Hazards Forecasting, National Marine Environmental Forecasting Center, Ministry of Natural Resources, Beijing 100081, China.*
- *Guihua Wang, Huizan Wang, Senzhang Wang, and Weimin Zhang are the corresponding authors.*

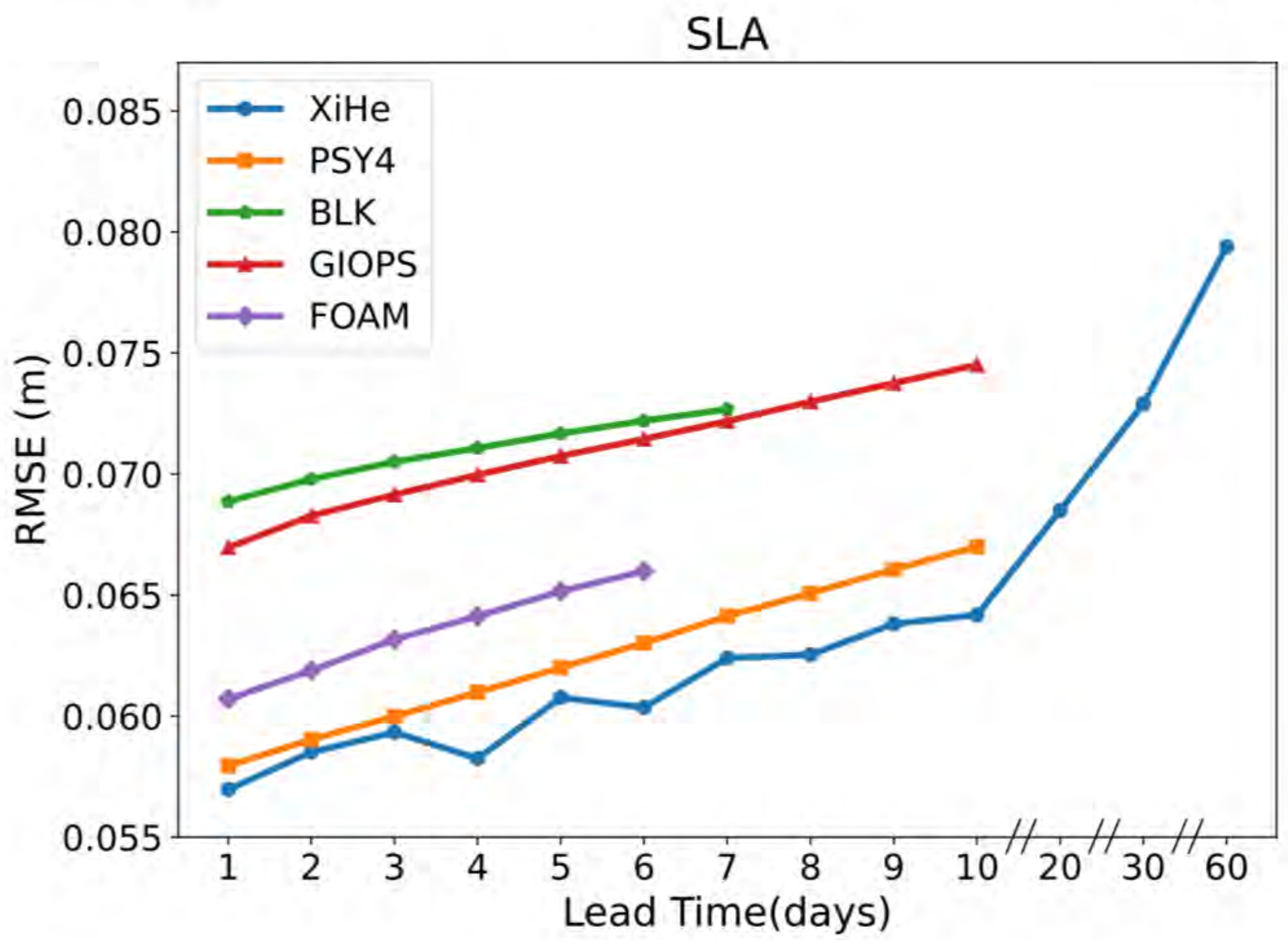
https://arxiv.org/abs/2402.02995

Wang et al. (2024)





Trained from ocean reanalyses



Short term forecast skill

COMING UP

AI-native hybrid climate models

Article

Neural general circulation models for weather and climate

https://doi.org/10.1038/s41586-024-07744-y

Received: 13 November 2023

Accepted: 15 June 2024

Published online: 22 July 2024

Open access

Check for updates

Dmitrii Kochkov[1,6,✉], Janni Yuval[1,6,✉], Ian Langmore[1,6], Peter Norgaard[1,6], Jamie Smith[1,6], Griffin Mooers[1], Milan Klöwer[2], James Lottes[1], Stephan Rasp[1], Peter Düben[3], Sam Hatfield[3], Peter Battaglia[4], Alvaro Sanchez-Gonzalez[4], Matthew Willson[4], Michael P. Brenner[1,5] & Stephan Hoyer[1,6,✉]

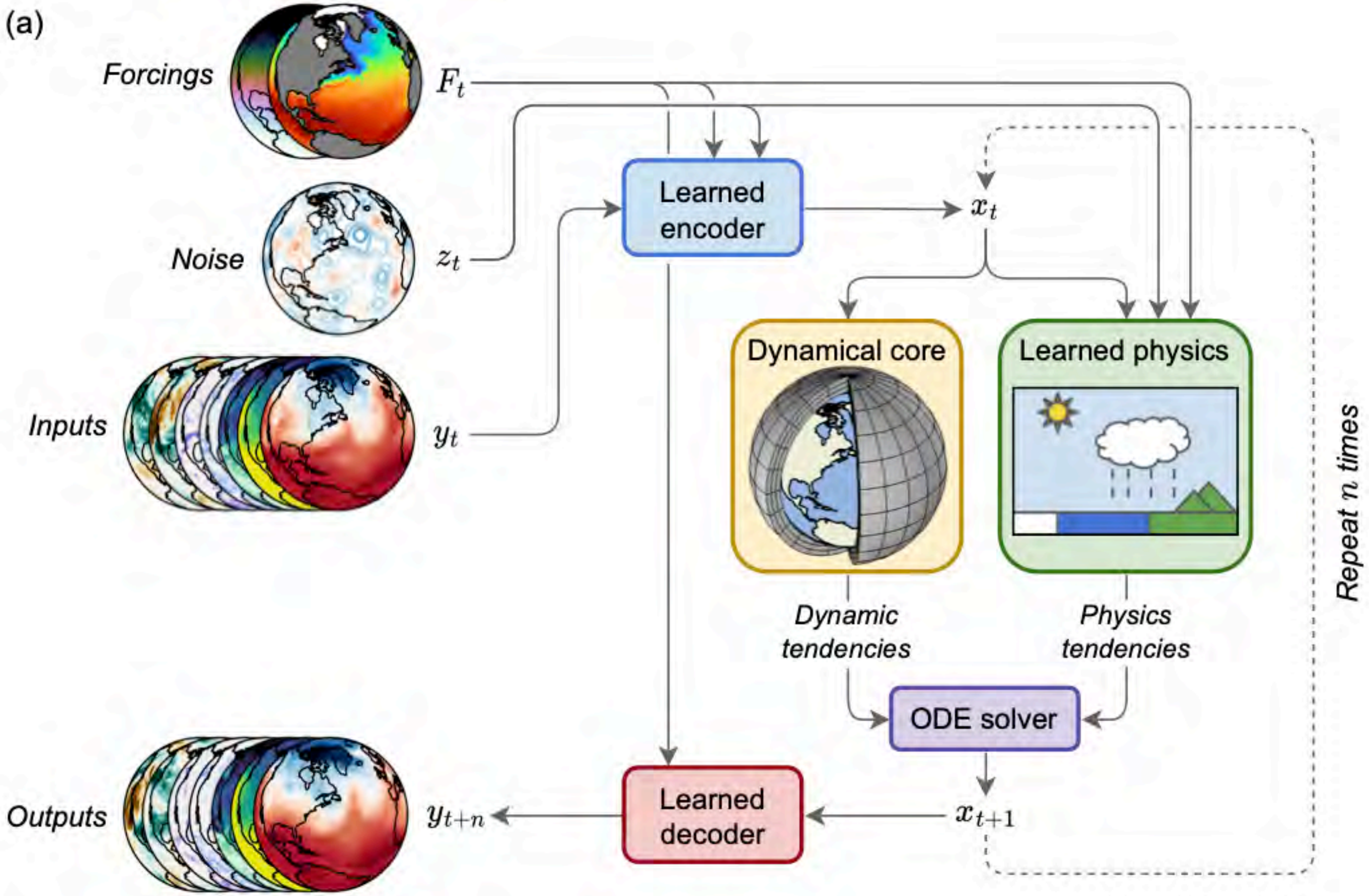
General circulation models (GCMs) are the foundation of weather and climate prediction[1,2]. GCMs are physics-based simulators that combine a numerical solver for large-scale dynamics with tuned representations for small-scale processes such as cloud formation. Recently, machine-learning models trained on reanalysis data have achieved comparable or better skill than GCMs for deterministic weather forecasting[3,4]. However, these models have not demonstrated improved ensemble forecasts, or shown sufficient stability for long-term weather and climate simulations. Here we present a GCM that combines a differentiable solver for atmospheric dynamics with machine-learning components and show that it can generate forecasts of deterministic weather, ensemble weather and climate on par with the best machine-learning and physics-based methods. NeuralGCM is competitive with machine-learning models for one- to ten-day forecasts, and with the European Centre for Medium-Range Weather Forecasts ensemble prediction for one- to fifteen-day forecasts. With prescribed sea surface temperature, NeuralGCM can accurately track climate metrics for multiple decades, and climate forecasts with 140-kilometre resolution show emergent phenomena such as realistic frequency and trajectories of tropical cyclones. For both weather and climate, our approach offers orders of magnitude computational savings over conventional GCMs, although our model does not extrapolate to substantially different future climates. Our results show that end-to-end deep learning is compatible with tasks performed by conventional GCMs and can enhance the large-scale physical simulations that are essential for understanding and predicting the Earth system.

Solving the equations for Earth's atmosphere with general circulation models (GCMs) is the basis of weather and climate prediction[1,2]. Over the past 70 years, GCMs have been steadily improved with better numerical methods and more detailed physical models, while exploiting faster computers to run at higher resolution. Inside GCMs, the unresolved physical processes such as clouds, radiation and precipitation are represented by semi-empirical parameterizations. Tuning GCMs to match historical data remains a manual process[5], and GCMs retain many persistent errors and biases[6–8]. The difficulty of reducing uncertainty in long-term climate projections[9] and estimating distributions of extreme weather events[10] presents major challenges for climate mitigation and adaptation[11].

Recent advances in machine learning have presented an alternative for weather forecasting[3,4,12,13]. These models rely solely on machine-learning techniques, using roughly 40 years of historical data from the European Center for Medium-Range Weather Forecasts (ECMWF) reanalysis v5 (ERA5)[14] for model training and forecast initialization. Machine-learning methods have been remarkably successful, demonstrating state-of-the-art deterministic forecasts for 1- to 10-day weather prediction at a fraction of the computational cost of traditional models[3,4]. Machine-learning atmospheric models also require considerably less code, for example GraphCast[3] has 5,417 lines versus 376,578 lines for the National Oceanic and Atmospheric Administration's FV3 atmospheric model[15] (see Supplementary Information section A for details).

Nevertheless, machine-learning approaches have noteworthy limitations compared with GCMs. Existing machine-learning models have focused on deterministic prediction, and surpass deterministic numerical weather prediction in terms of the aggregate metrics for which they are trained[3,4]. However, they do not produce calibrated uncertainty estimates[4], which is essential for useful weather forecasts[1]. Deterministic machine-learning models using a mean-squared-error loss are rewarded for averaging over uncertainty, producing unrealistically blurry predictions when optimized for multi-day forecasts[3,13]. Unlike physical models, machine-learning models misrepresent derived (diagnostic) variables such as geostrophic wind[16]. Furthermore,

[1]Google Research, Mountain View, CA, USA. [2]Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. [3]European Centre for Medium-Range Weather Forecasts, Reading, UK. [4]Google DeepMind, London, UK. [5]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. [6]These authors contributed equally: Dmitrii Kochkov, Janni Yuval, Ian Langmore, Peter Norgaard, Jamie Smith, Stephan Hoyer. ✉e-mail: dkochkov@google.com; janniyuval@google.com; shoyer@google.com



https://doi.org/10.1038/s41586-024-07744-y

Kochkov et al. (2024)

https://github.com/google-research/dinosaur

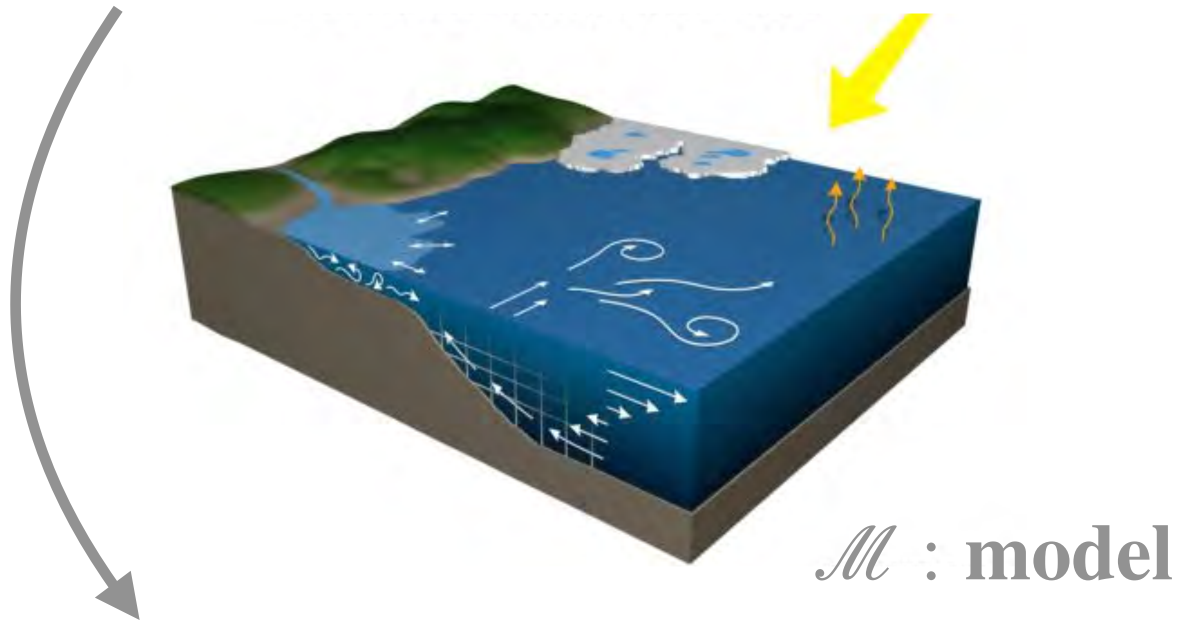
https://github.com/google-research/neuralgcm

Hybrid models combining physics and ML

Augmenting ocean models with ML components

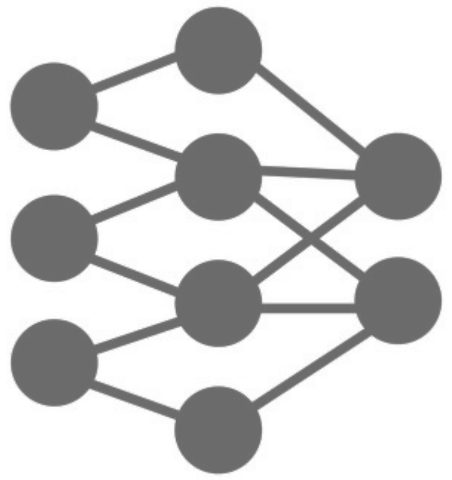
Input

step n



$\mathscr{M}$: **model**

+



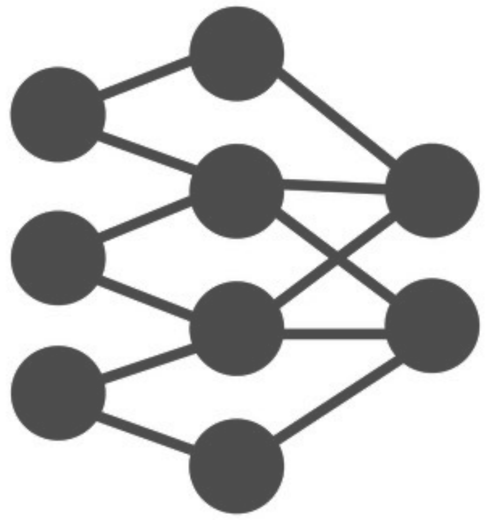
θ : **parameters**

step n+1

Output

The model is augmented with a **trainable** component

with



θ : **parameters**

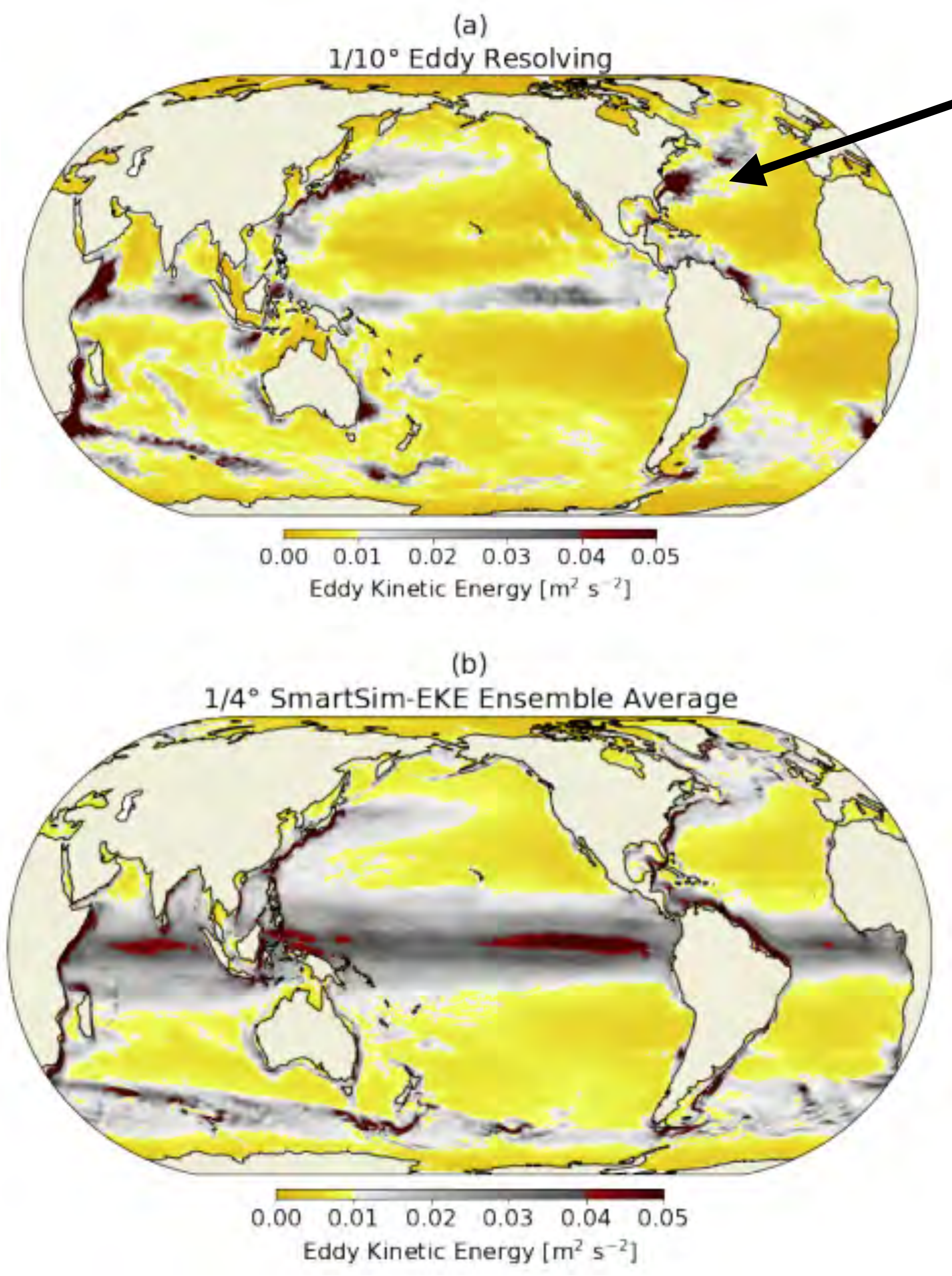
trained to minimise :

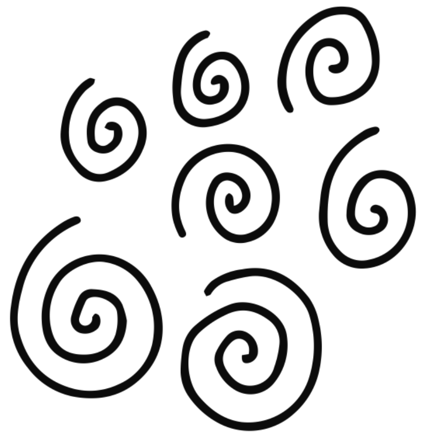
$$\mathscr{L}(\theta) = \textbf{training objective}$$

- improving physical **consistency**
- correcting model **errors** (vs obs.)
- **accelerating** execution (x10-100)

NB : does not have to be deterministic

ML for ocean models subgrid physics (1/2)



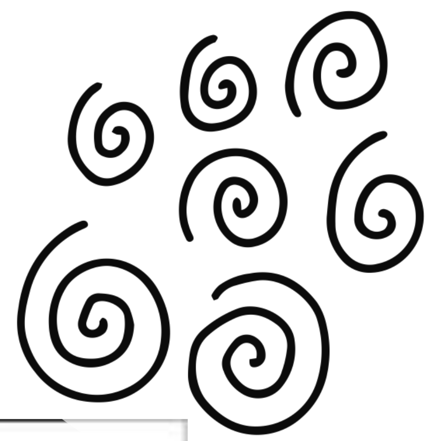


oceanic macro-scale turbulence

- missing terms from resolved quantities
- closures for turbulent processes
- leveraging hi-res/process model data
- encoded as closed forms or ML models
- a very active field (5-10 papers / months)

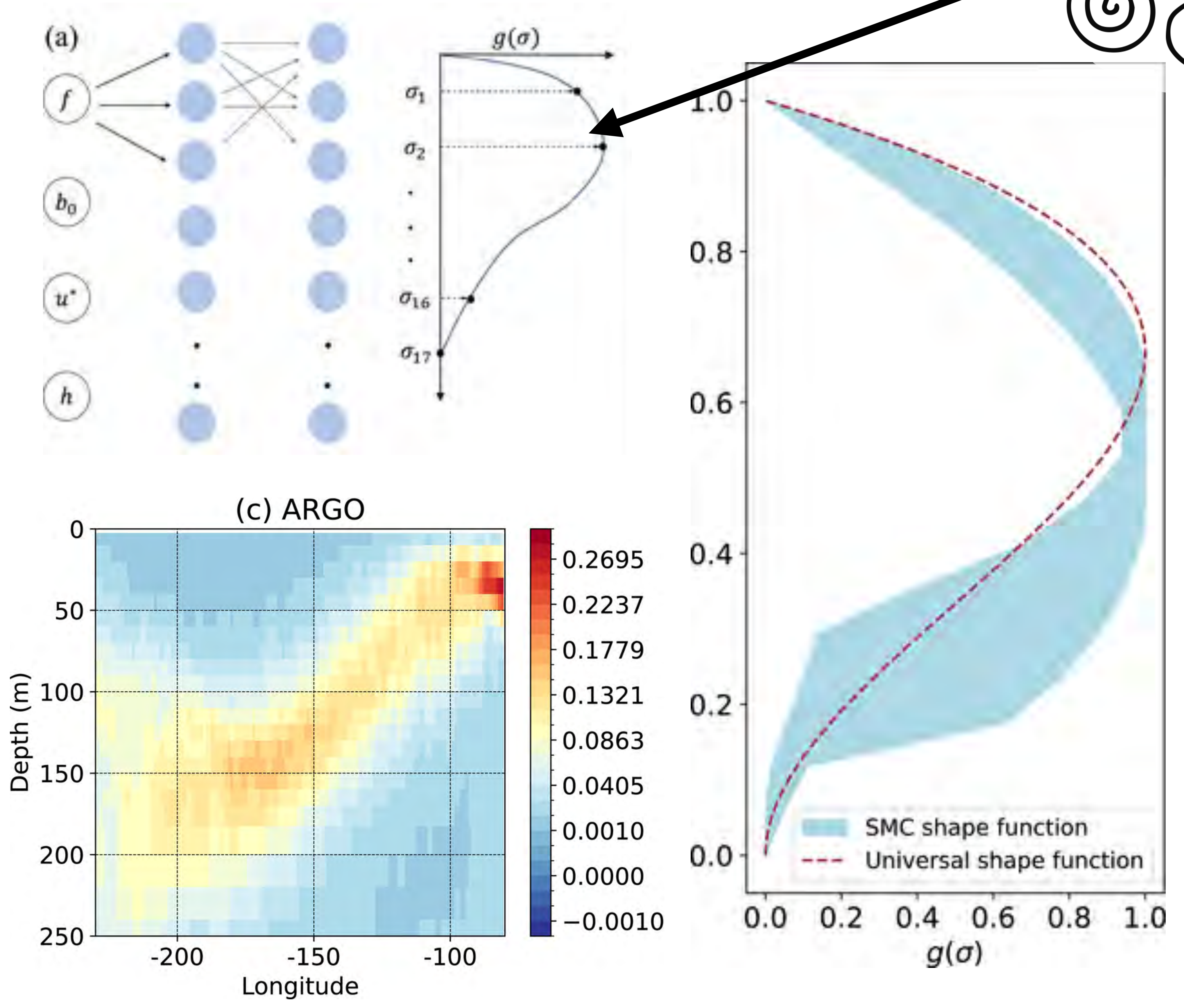
Partee et al. 2022
https://doi.org/10.1016/j.jocs.2022.101707

ML for ocean models subgrid physics (1/2)



oceanic micro-scale turbulence

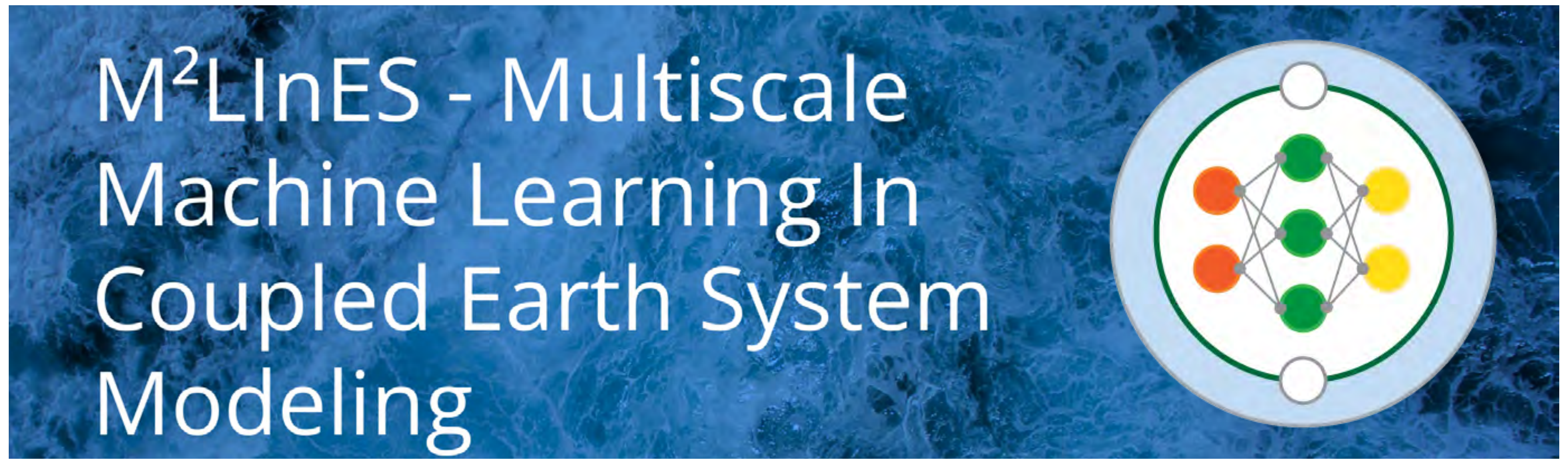
- missing terms from resolved quantities
- closures for turbulent processes
- leveraging hi-res/process model data
- encoded as closed forms or ML models
- a very active field (5-10 papers / months)



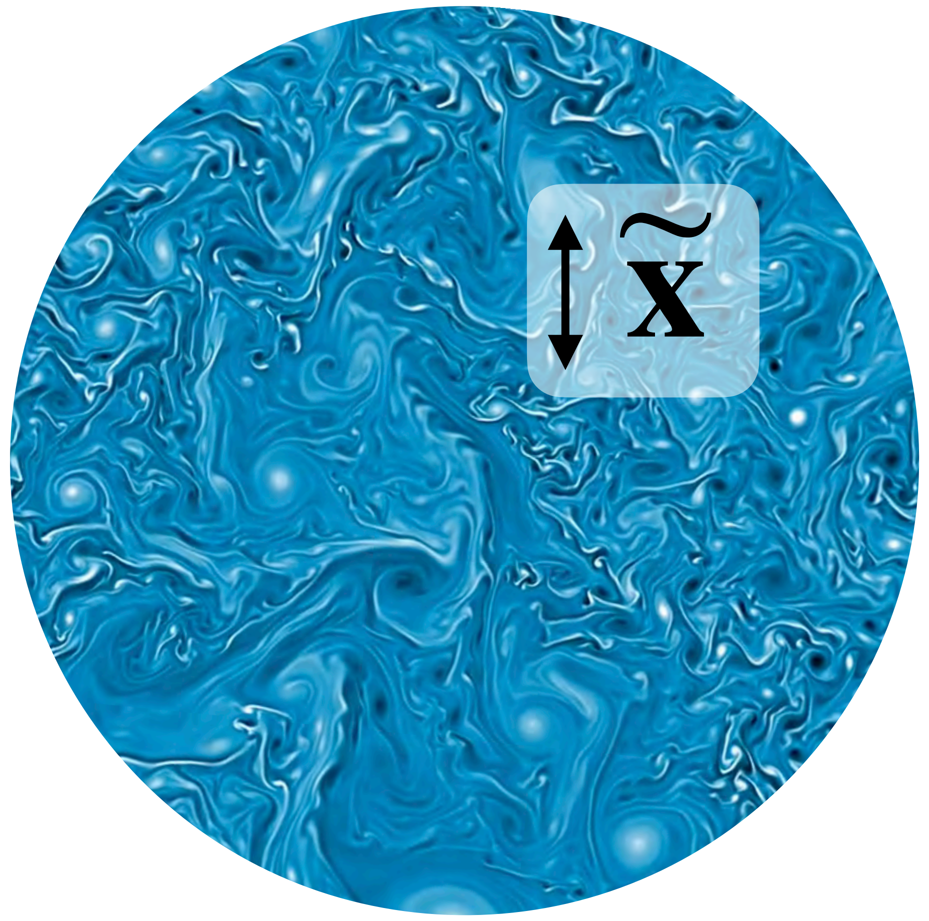
Sane et al. 2023
https://doi.org/10.1029/2023MS003890

See for instance :
M2LInES consortium

https://m2lines.github.io



ML for ocean models subgrid physics (2/2)

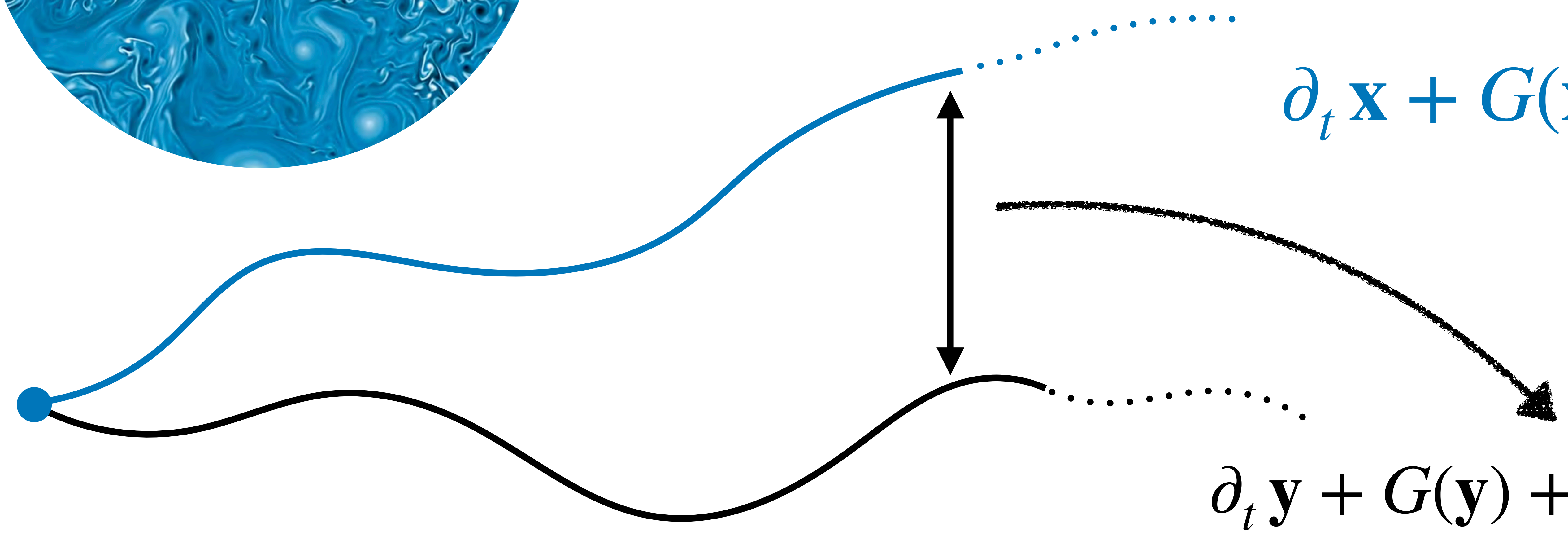


Dynamical system

$$\partial_t \mathbf{x} + \mathscr{L}\,\mathbf{x} + \mathscr{N}(\mathbf{x}) = 0$$

Resolved equations

$$\partial_t \widetilde{\mathbf{x}} + \mathscr{L}\,\widetilde{\mathbf{x}} + \mathscr{N}(\widetilde{\mathbf{x}}) = \mathscr{N}(\widetilde{\mathbf{x}}) - \widetilde{\mathscr{N}(\mathbf{x})}$$



$$\partial_t \mathbf{x} + G(\mathbf{x}) = 0$$

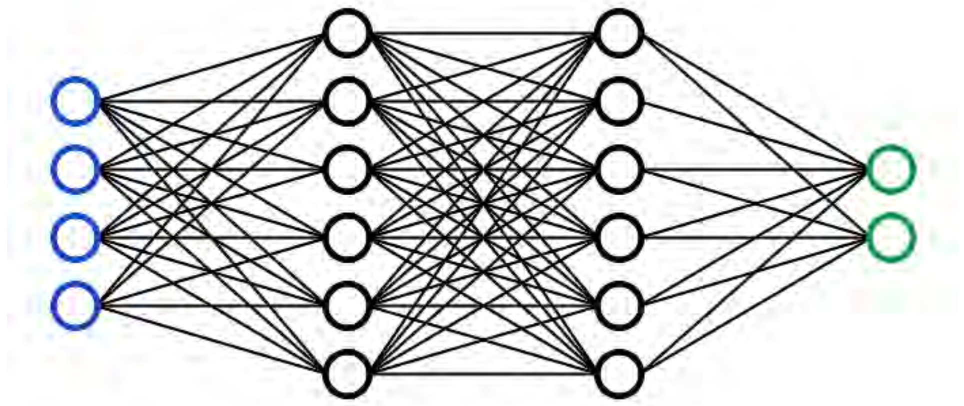
Subgrid closure

$$\mathscr{M}(\widetilde{\mathbf{x}}) \simeq \mathscr{N}(\widetilde{\mathbf{x}}) - \widetilde{\mathscr{N}(\mathbf{x})}$$

$$\partial_t \mathbf{y} + G(\mathbf{y}) + \mathscr{M}_{NN}(\mathbf{y}) = 0$$

Learning the mapping

$$\widetilde{\mathbf{x}}(t) \rightarrow \mathscr{M}(\widetilde{\mathbf{x}}(t))$$



θ : **parameters**

Frezat et al. (2021)

Physical consistency

Symmetries, invariances
loss function / architecture

Frezat et al. (2022)

End-to-end training

Differentiable programming,
different loss function
w/ same architecture

Frezat et al. (2024)

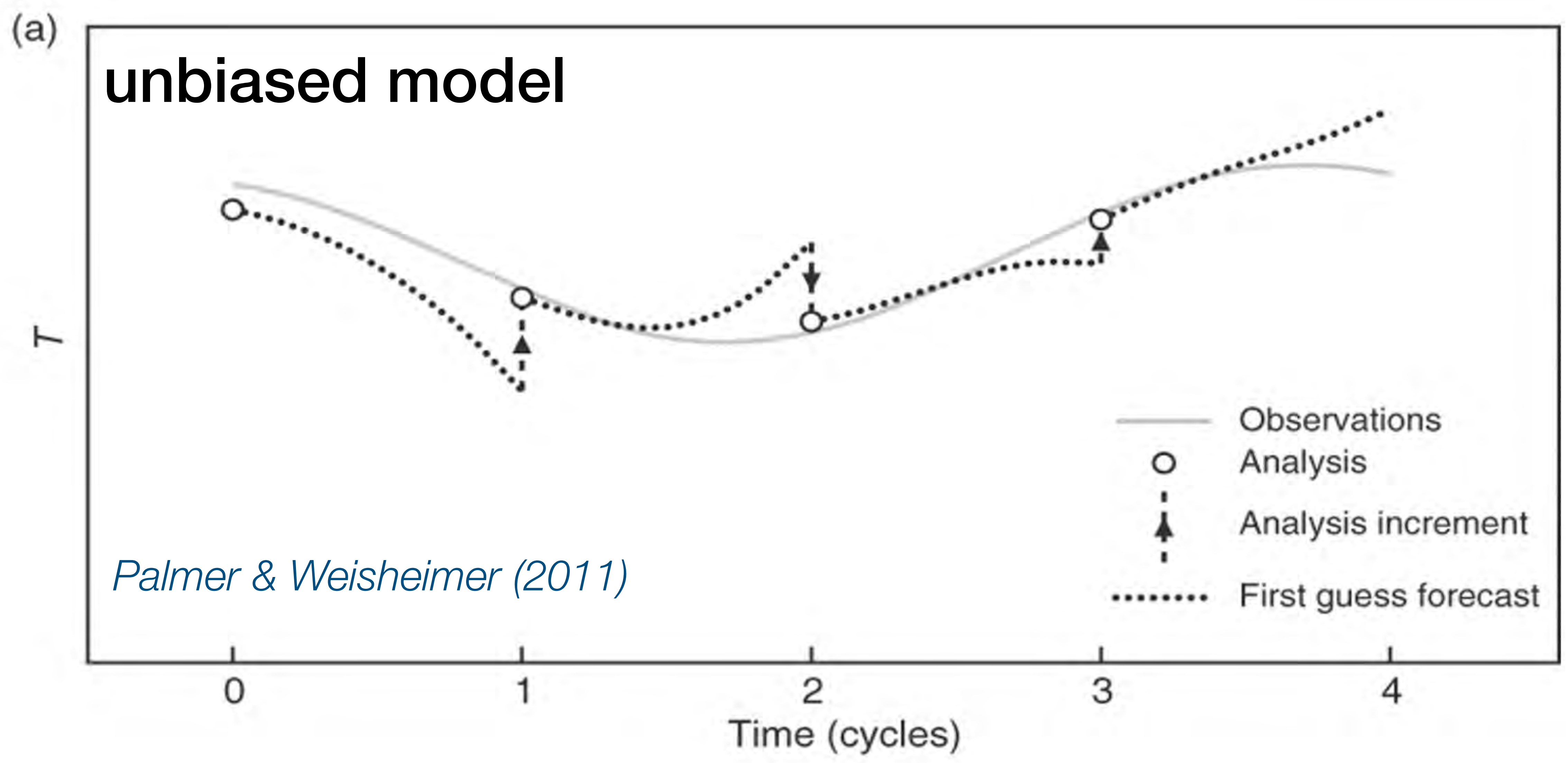
Gradient-free training

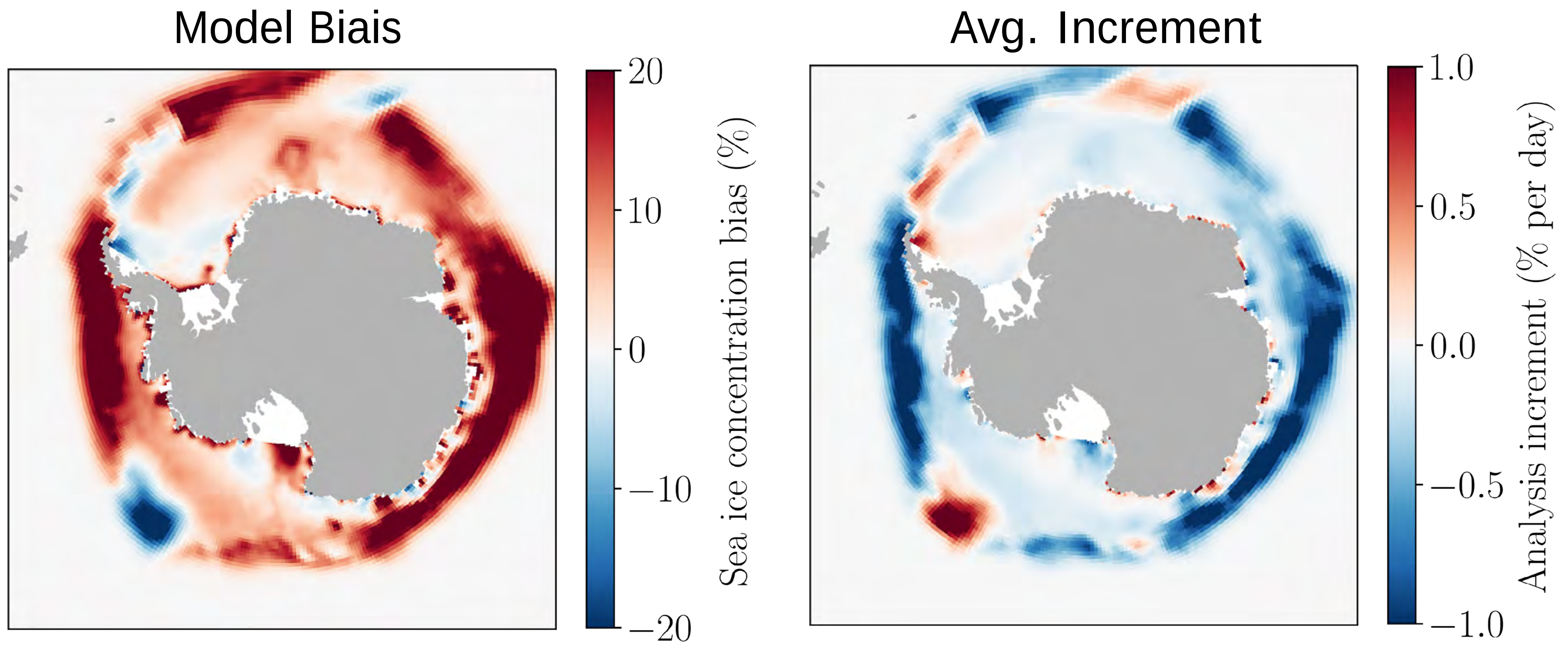
training model emulator
for approx. gradient
wrt NN. parameters

Performance, stability

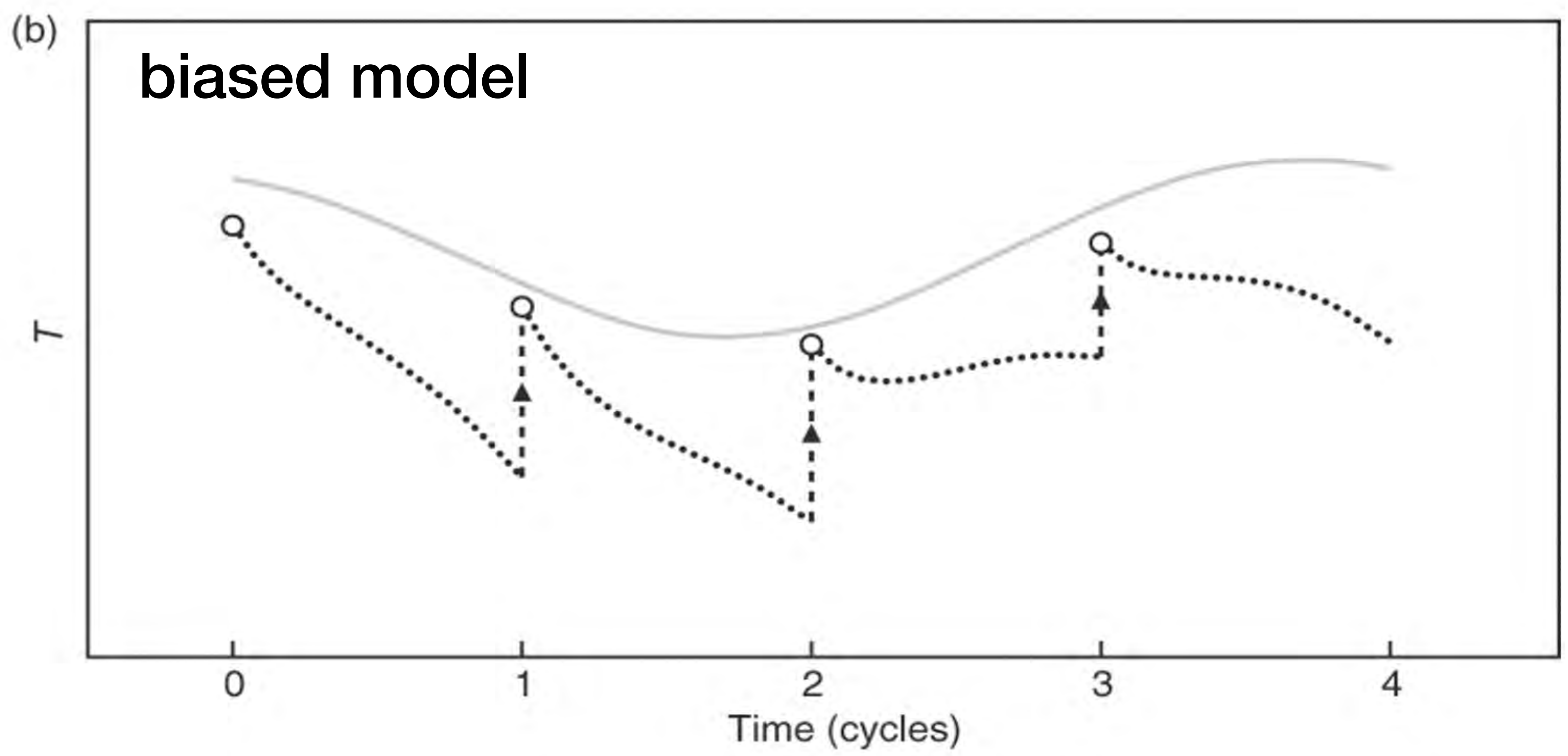
Generalisation, interpretability

Learning model error from observations (1/2)



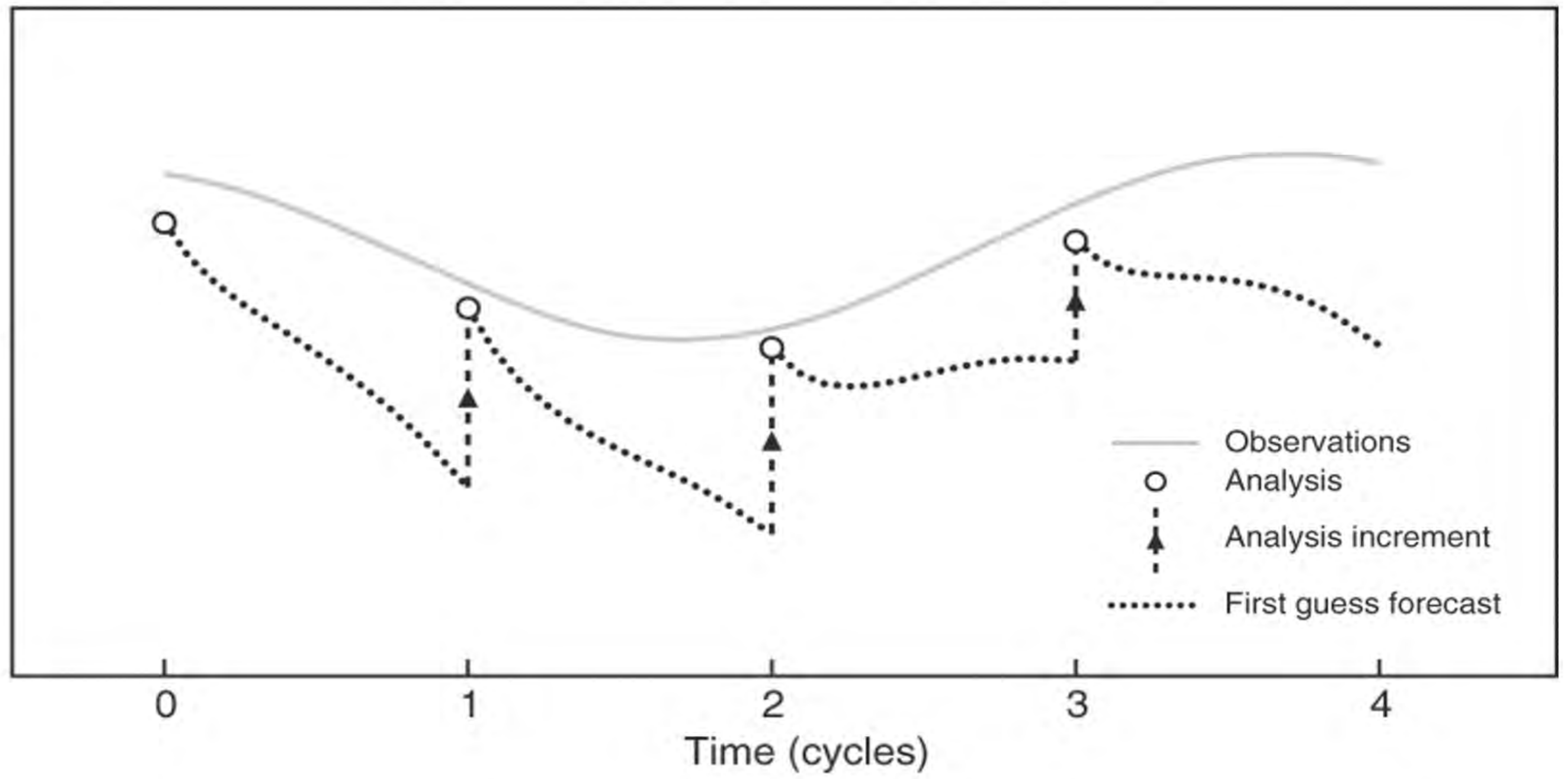


Gregory et al. (2023)

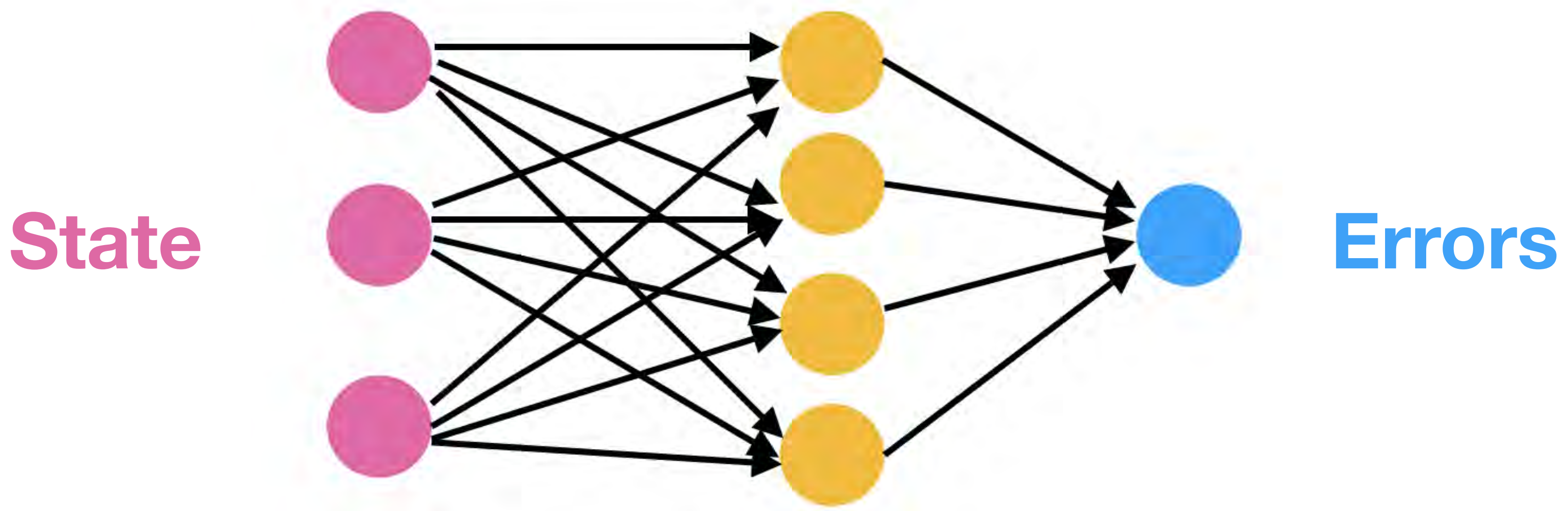


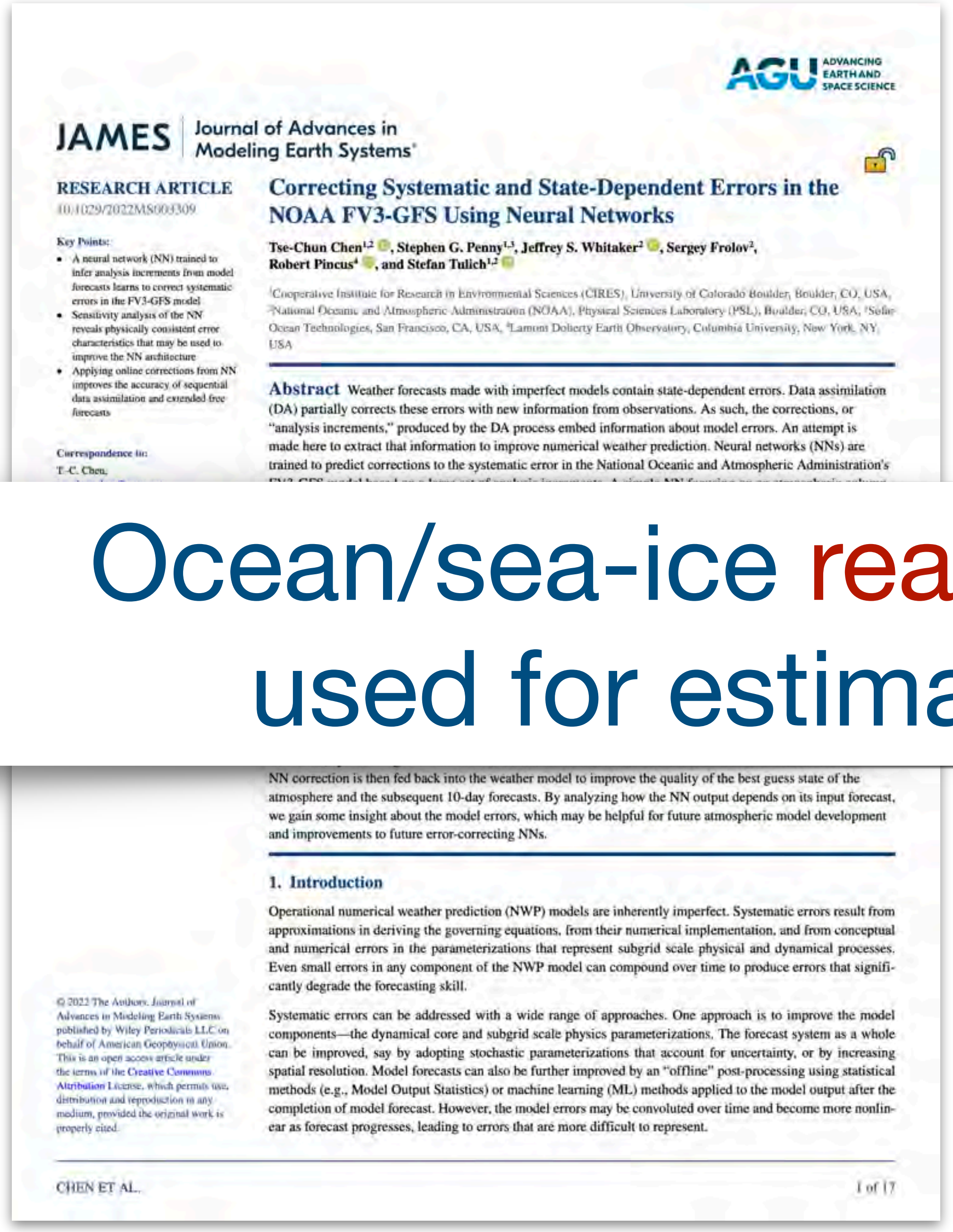
- w/ unbiased observations, analysis increments compensate for model biais
- estimating state-dependent bias corrections (Leith, 1978; Saha, 1992; DelSole and Hou, 1999)
- state-dependent biais corrections provide a representation of model errors

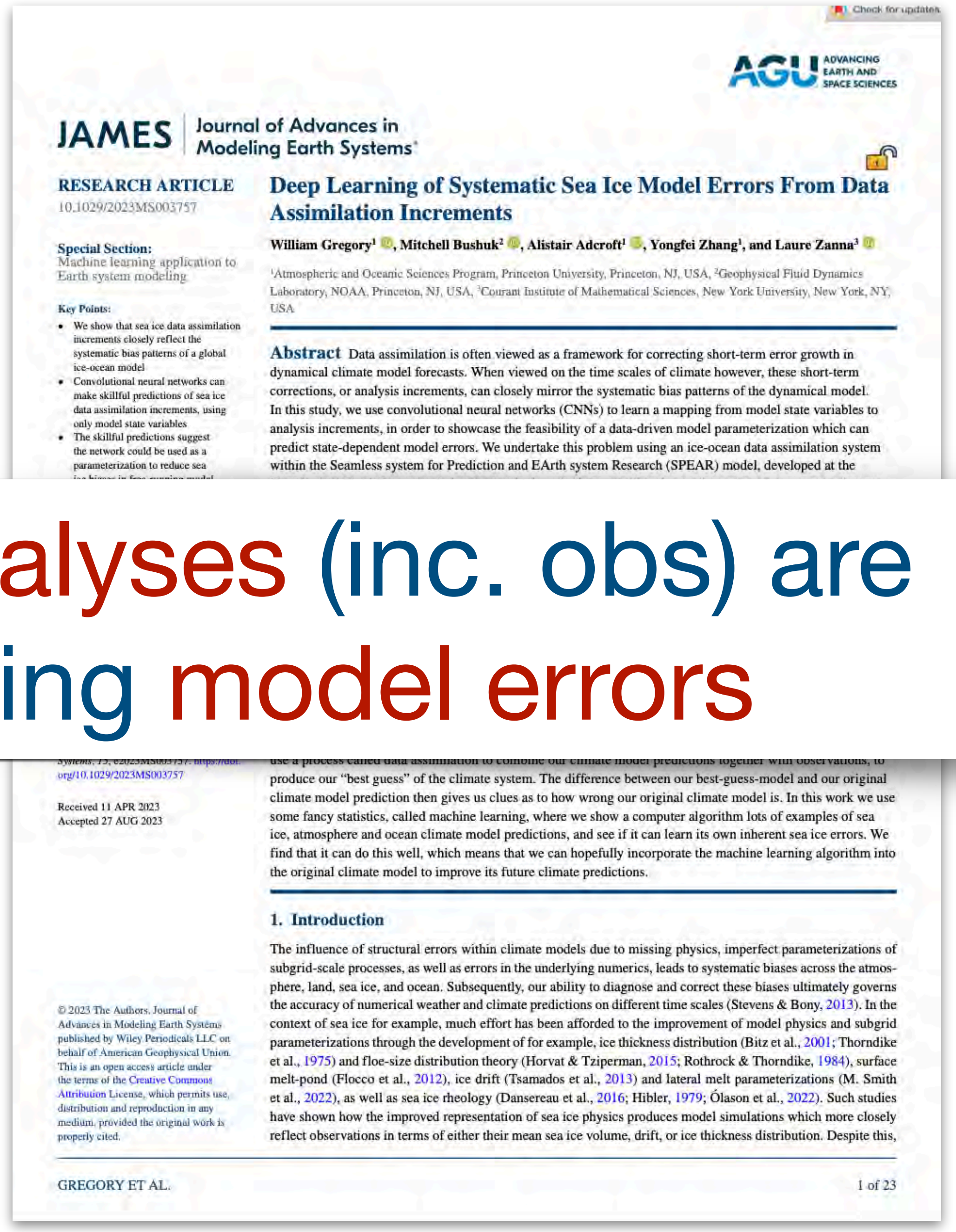
Learning model error from observations (2/2)



Offline







Ocean/sea-ice reanalyses (inc. obs) are used for estimating model errors

https://doi.org/10.1029/2022MS003309

https://doi.org/10.1029/2023MS003757

- NN for learning state-dependant biais corrections from analysis increments
- w/ applications in GCMs (atmosphere and ocean/sea-ice)
- showing success in improving the modeled climate state & forecast skill

Bonavita and Laloyaux, 2020; Watt-Meyer et al., 2021; Chen et al., 2022; Gregory et al. 202
Chapman and Berner 2023

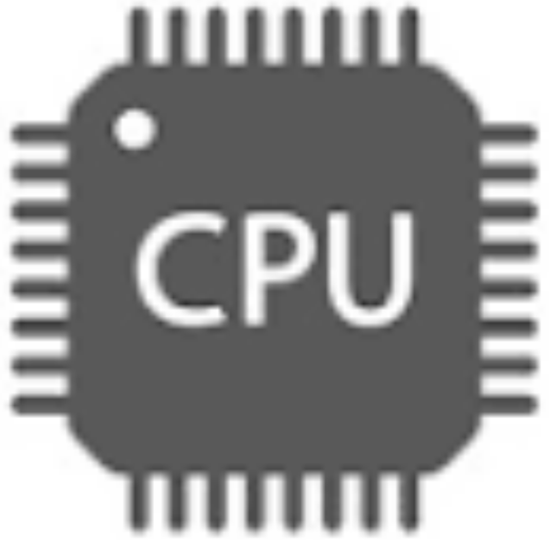
The plumbing challenges of hybrid modelling

Interfacing ocean models with DL frameworks (1/3)

stable, robust, low abstraction languages

+

supercomputers

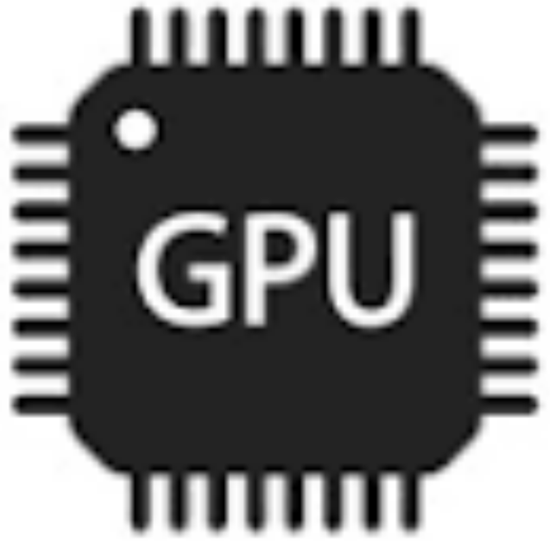


runs only on CPUs

high abstraction, fast evolving languages

+

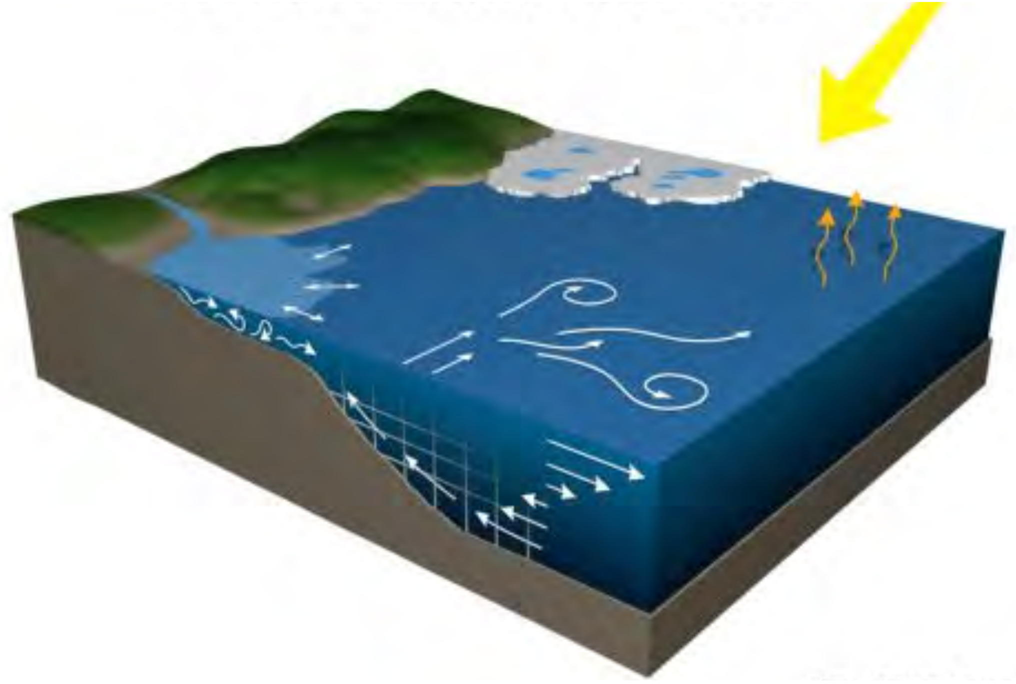
cloud ready

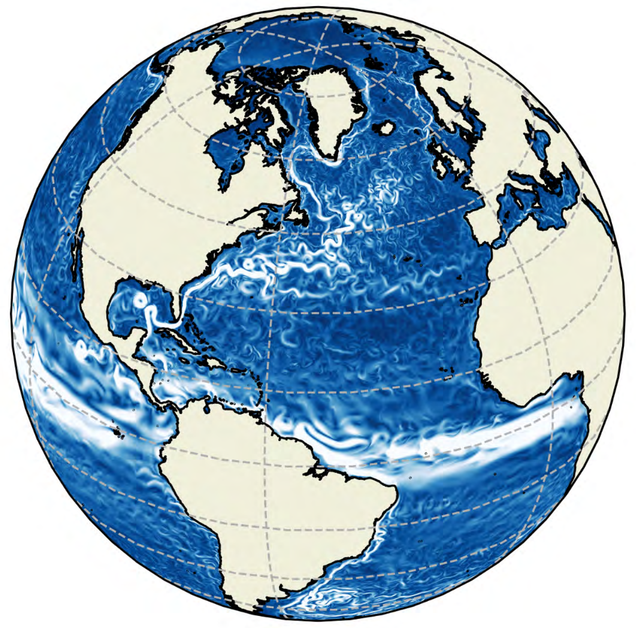


natively runs on GPUs

Interfacing ocean models with DL frameworks (2/3)

Ocean circulation models





Input

step n

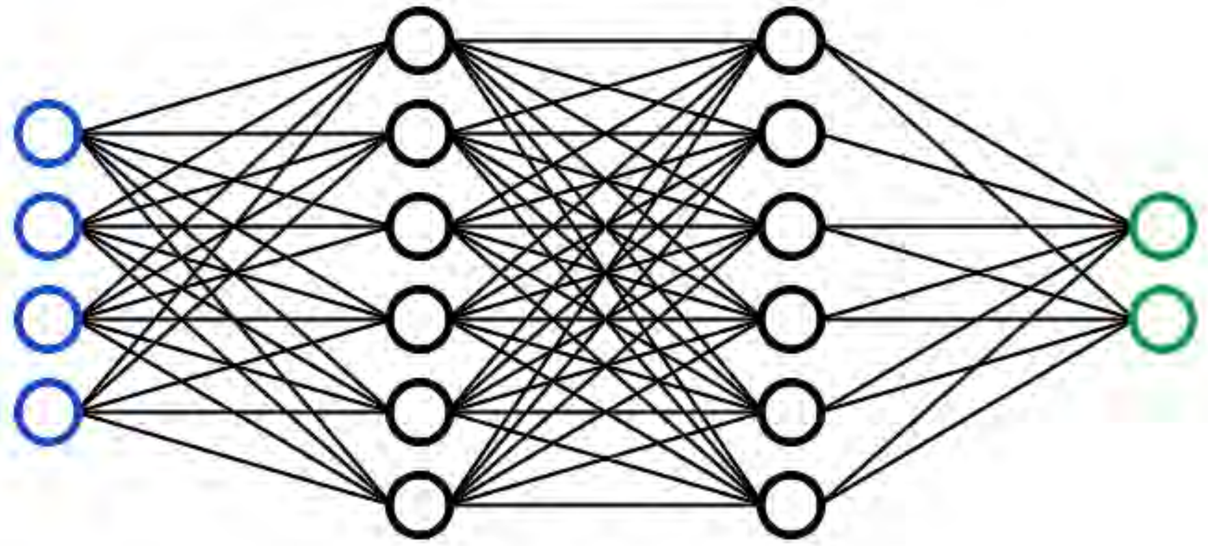
+

step n+1

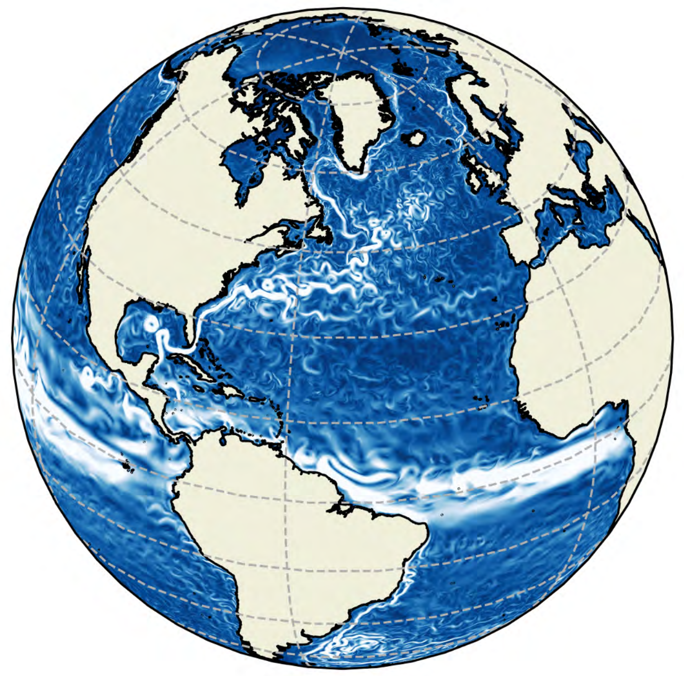
Output

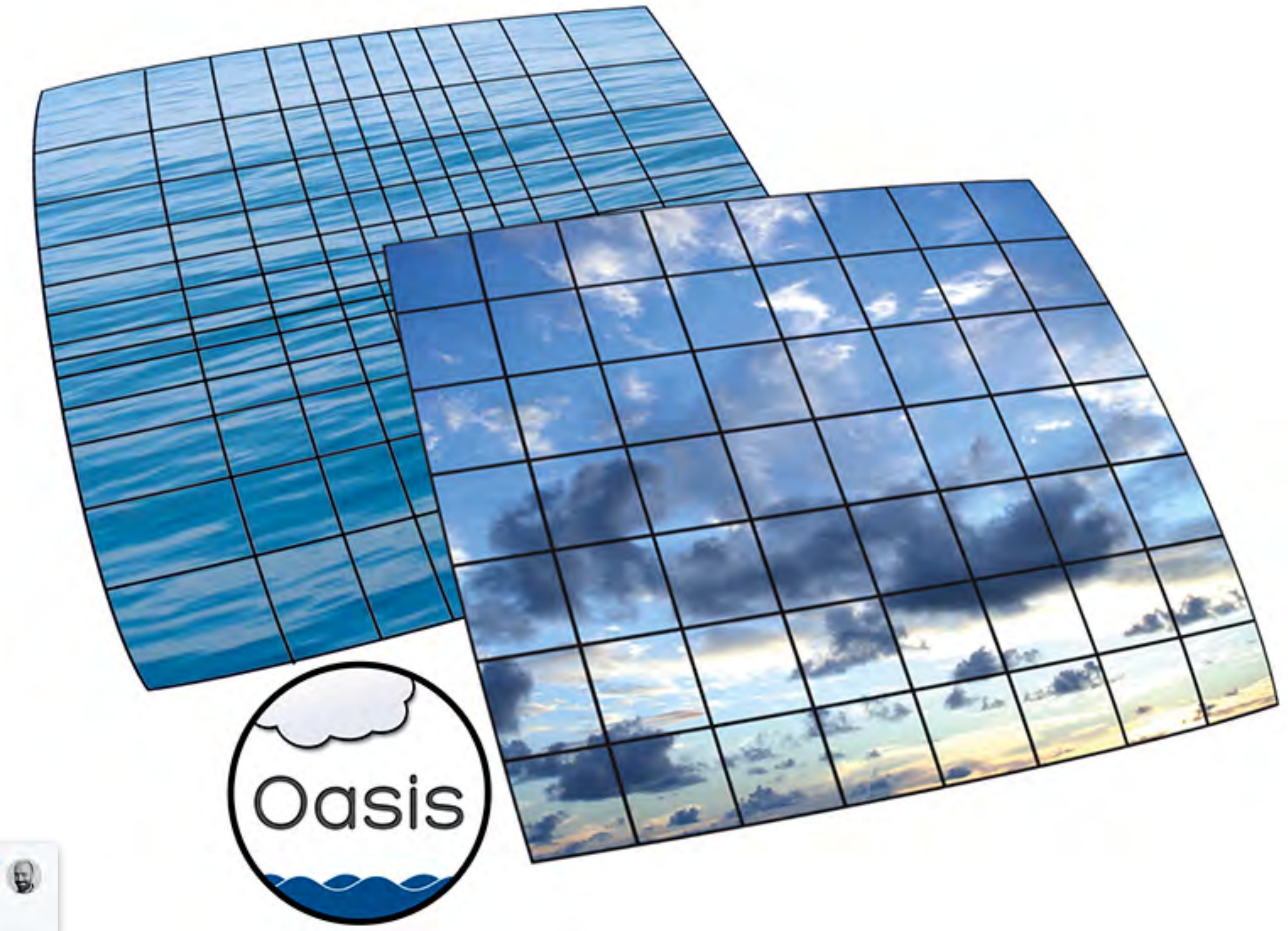
Trainable components

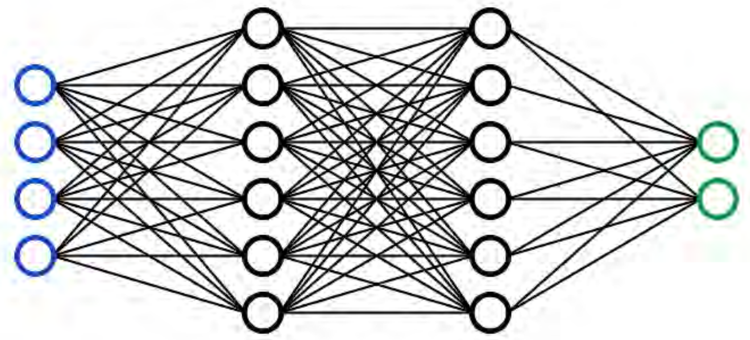
(closures, error corrections)



Interfacing ocean models with DL frameworks (3/3)

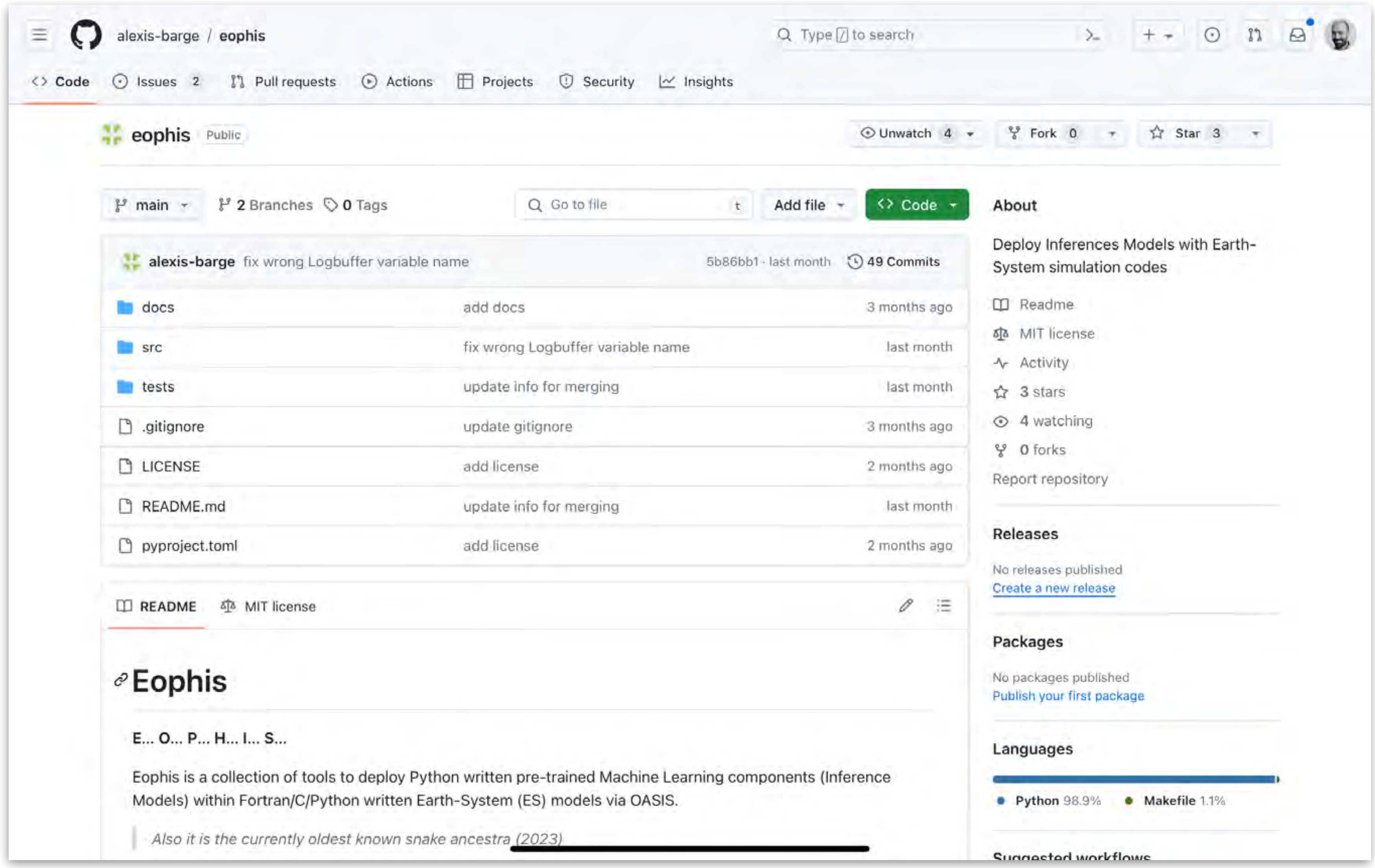






A. Barge

Work by Alexis Barge at IGE

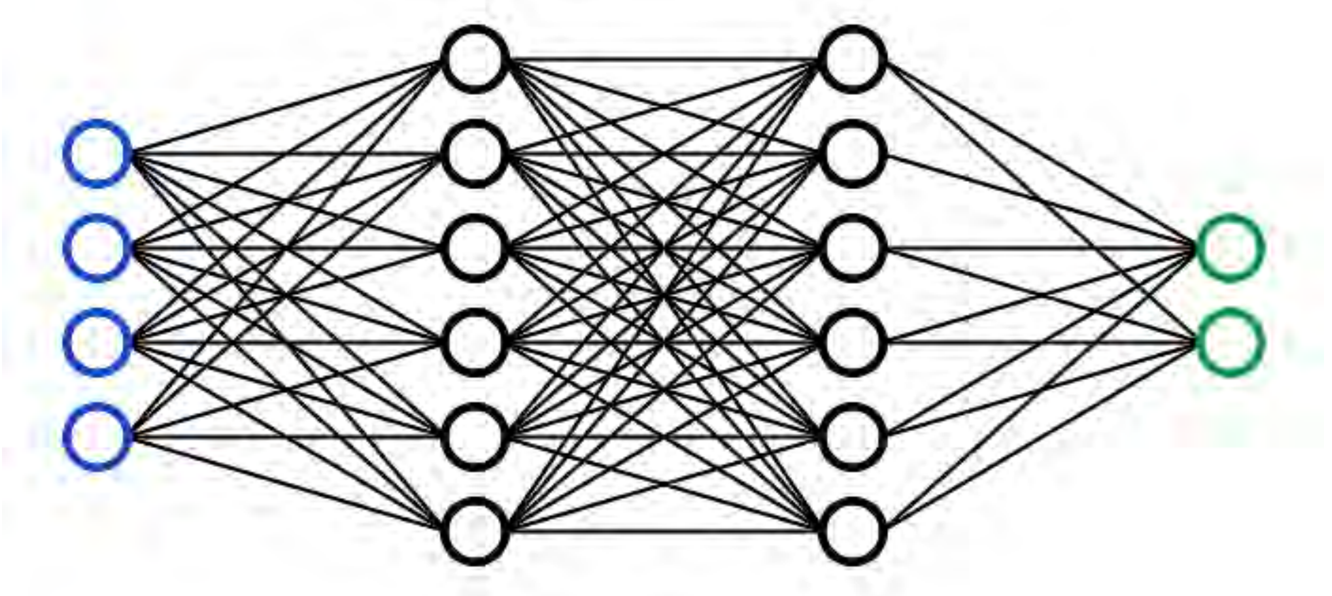


https://github.com/meom-group/eophis

- OASIS : exchange of 3D data between different codes
- Eophis : simplified deployment of ML models w/ OASIS
- Requires some change to the NEMO code
- Key : portability, domain decomposition

The challenge of online training strategies (1/2)

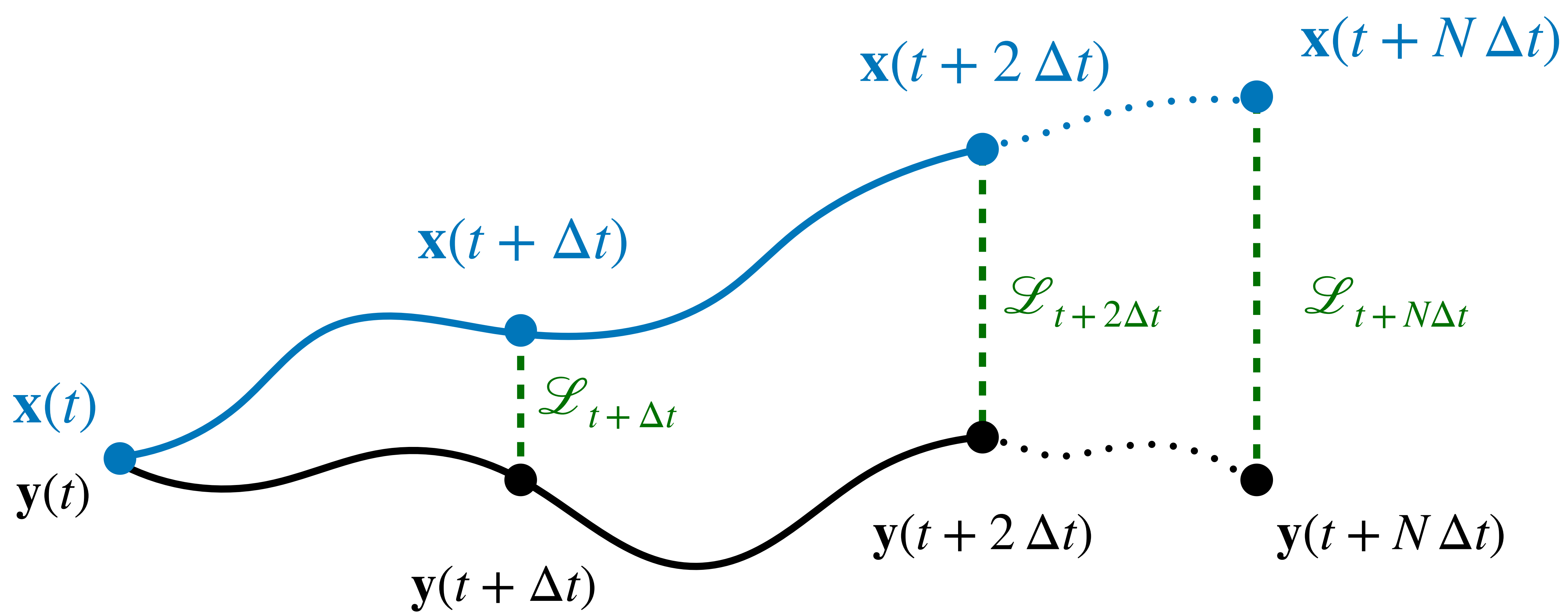
offline learning



mapping $\overline{\mathbf{x}} \rightarrow \overline{\mathcal{N}(\mathbf{x})}$

from pre-existing data

online learning



$$\partial_t \mathbf{y} + G(\mathbf{y}) + \mathscr{M}_{NN}(\mathbf{y}) = f$$

along a trajectory

(a.k.a : a posteriori, solver-in-the-loop, end-to-end, auto-regressive roll-outs)

Online training improves performance, stability, generalisation *Frezat et al. 2022; List et al. 2024*

The challenge of online training strategies (2/2)

target

prediction

$$\arg\min_{\theta} \mathscr{L}(\mathbf{z}, \mathscr{M}(\mathbf{y} \mid \theta))$$

$$\frac{\partial \mathscr{L}}{\partial \theta}(\mathbf{z}, \mathscr{M}(\mathbf{y} \mid \theta)) = \frac{\partial \mathscr{M}}{\partial \theta}(\mathbf{y} \mid \theta)\frac{\partial \mathscr{L}}{\partial \mathscr{M}}$$

gradient of the loss

For time evolving problems, with

$$\mathbf{y}(t + \Delta t) = E_m \circ \cdots \circ E_1(\mathbf{y}(t))$$

$\mathscr{M} \equiv E$ Auto-regressive operator (time)

The gradient of the loss involves

tricky without Automatic Differenciation (AD) !

$$\frac{\partial \mathscr{M}}{\partial \theta} \equiv \frac{\partial E}{\partial \theta} = \frac{\partial (E_m \circ \cdots \circ E_1)}{\partial \theta} = \frac{\partial E_m}{\partial E_{m-1}} \cdots \frac{\partial E_2}{\partial E_1}\frac{\partial E_1}{\partial \theta}$$

The challenge of online training strategies (2/2)

target   prediction

$$\arg\min_{\theta} \mathscr{L}(\mathbf{z}, \mathscr{M}(\mathbf{y} \mid \theta))$$

AD is readily available in some language

For time evolving problems, with

$$\mathbf{y}(t + \Delta t) = E_m \circ \cdots \circ E_1(\mathbf{y}(t))$$

$\mathscr{M} \equiv E$ temporal evolution operator

The gradient of the loss involves

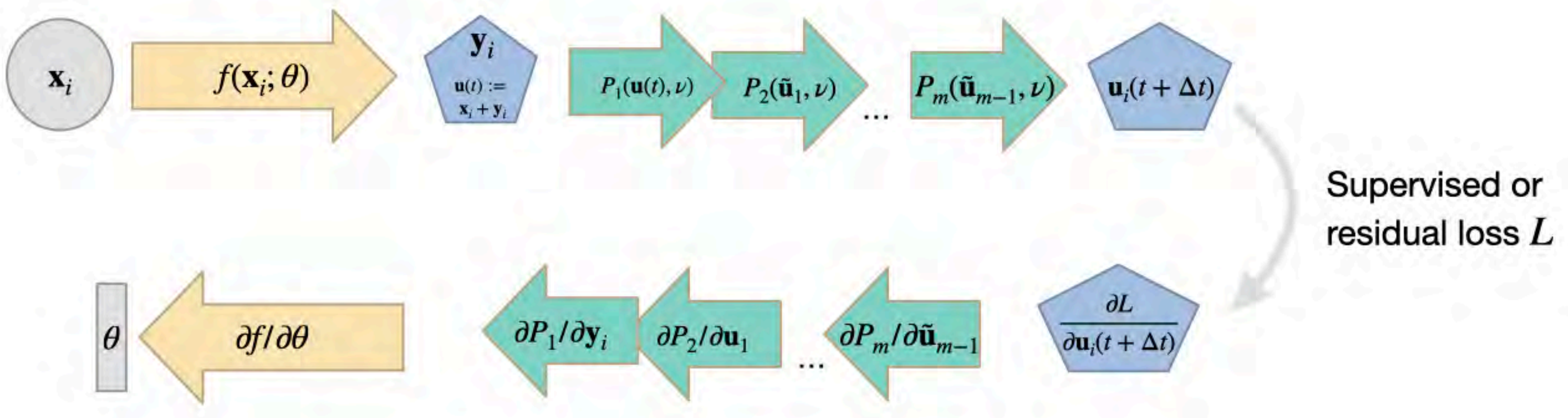
$$\frac{\partial \mathscr{M}}{\partial \theta} \equiv \frac{\partial E}{\partial \theta} = \frac{\partial (E_m \circ \cdots \circ E_1)}{\partial \theta} = \frac{\partial E_m}{\partial E_{m-1}} \cdots \frac{\partial E_2}{\partial E_1} \frac{\partial E_1}{\partial \theta}$$

tricky without Automatic Differenciation (AD) !

The challenge of online training strategies (2/2)

target prediction

$$\underset{\theta}{\arg\min}\, \mathscr{L}(\mathbf{z}, \mathscr{M}(\mathbf{y} \mid \theta))$$



See eg Thuerey et al. 2021

https://arxiv.org/abs/2109.05237

AD is readily available in some language

But AD used yet in climate models…

Differentiable programming

- programs composed of differentiable building blocks
- building blocks : trainable and procedural code components
- trainable end-to-end with gradient based optimisation

AI-native hybrid geoscientific models

Article

Neural general circulation models for weather and climate

https://doi.org/10.1038/s41586-024-07744-y

Received: 13 November 2023

Accepted: 15 June 2024

Published online: 22 July 2024

Open access

Check for updates

Dmitrii Kochkov[1,6], Janni Yuval[1,6], Ian Langmore[1,6], Peter Norgaard[1,6], Jamie Smith[1,6], Griffin Mooers[1], Milan Klöwer[2], James Lottes[1], Stephan Rasp[1], Peter Düben[3], Sam Hatfield[3], Peter Battaglia[4], Alvaro Sanchez-Gonzalez[4], Matthew Willson[4], Michael P. Brenner[1,5] & Stephan Hoyer[1,6]

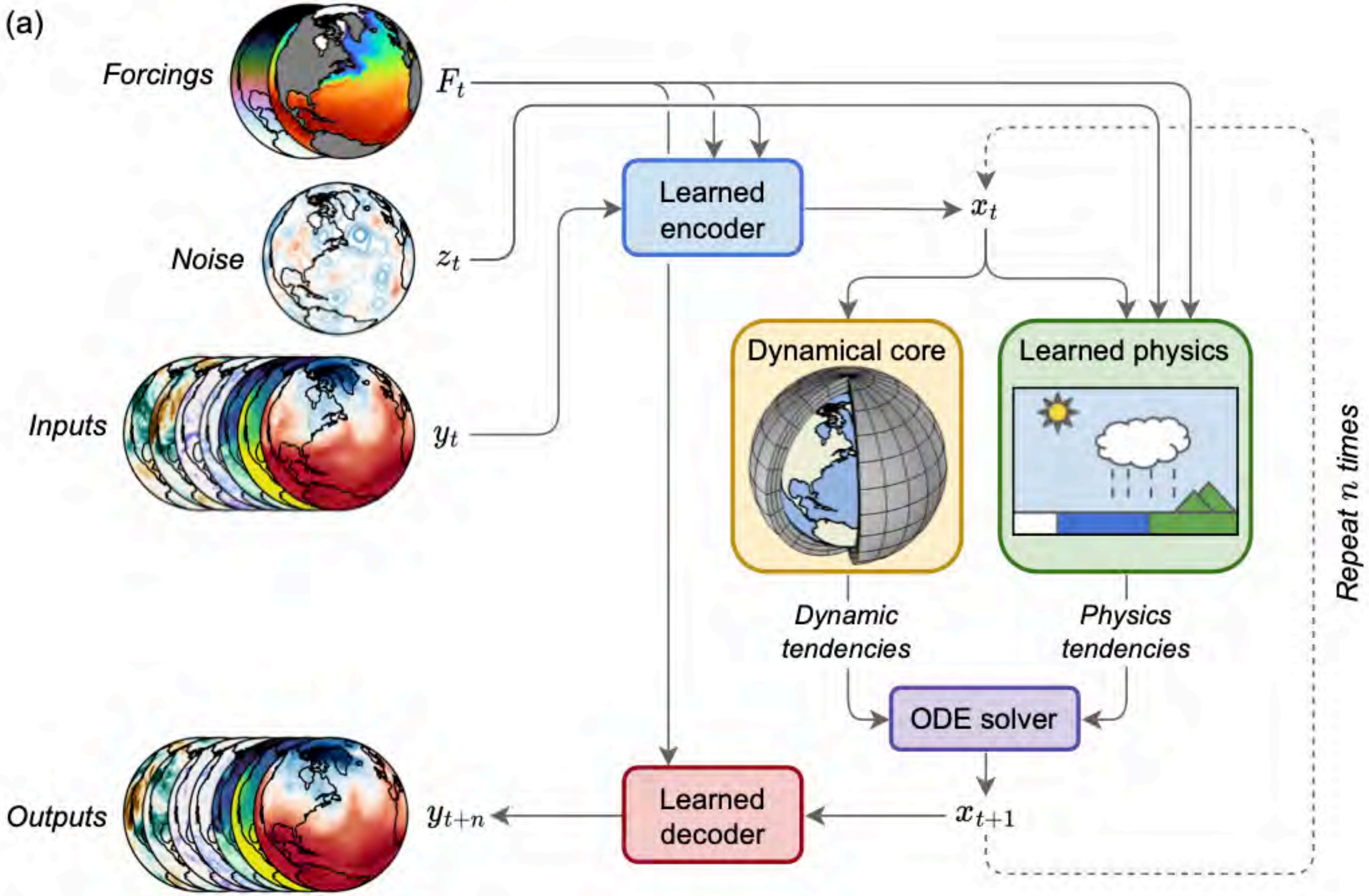
General circulation models (GCMs) are the foundation of weather and climate prediction[1,2]. GCMs are physics-based simulators that combine a numerical solver for large-scale dynamics with tuned representations for small-scale processes such as cloud formation. Recently, machine-learning models trained on reanalysis data have achieved comparable or better skill than GCMs for deterministic weather forecasting[3,4]. However, these models have not demonstrated improved ensemble forecasts, or shown sufficient stability for long-term weather and climate simulations. Here we present a GCM that combines a differentiable solver for atmospheric dynamics with machine-learning components and show that it can generate forecasts of deterministic weather, ensemble weather and climate on par with the best machine-learning and physics-based methods. NeuralGCM is competitive with machine-learning models for one- to ten-day forecasts, and with the European Centre for Medium-Range Weather Forecasts ensemble prediction for one- to fifteen-day forecasts. With prescribed sea surface temperature, NeuralGCM can accurately track climate metrics for multiple decades, and climate forecasts with 140-kilometre resolution show emergent phenomena such as realistic frequency and trajectories of tropical cyclones. For both weather and climate, our approach offers orders of magnitude computational savings over conventional GCMs, although our model does not extrapolate to substantially different future climates. Our results show that end-to-end deep learning is compatible with tasks performed by conventional GCMs and can enhance the large-scale physical simulations that are essential for understanding and predicting the Earth system.

Solving the equations for Earth's atmosphere with general circulation models (GCMs) is the basis of weather and climate prediction[1,2]. Over the past 70 years, GCMs have been steadily improved with better numerical methods and more detailed physical models, while exploiting faster computers to run at higher resolution. Inside GCMs, the unresolved physical processes such as clouds, radiation and precipitation are represented by semi-empirical parameterizations. Tuning GCMs to match historical data remains a manual process[5], and GCMs retain many persistent errors and biases[6–8]. The difficulty of reducing uncertainty in long-term climate projections[9] and estimating distributions of extreme weather events[10] presents major challenges for climate mitigation and adaptation[11].

Recent advances in machine learning have presented an alternative for weather forecasting[3,4,12,13]. These models rely solely on machine-learning techniques, using roughly 40 years of historical data from the European Center for Medium-Range Weather Forecasts (ECMWF) reanalysis v5 (ERA5)[14] for model training and forecast initialization. Machine-learning methods have been remarkably successful, demonstrating state-of-the-art deterministic forecasts for 1- to 10-day weather prediction at a fraction of the computational cost of traditional models[3,4]. Machine-learning atmospheric models also require considerably less code, for example GraphCast[3] has 5,417 lines versus 376,578 lines for the National Oceanic and Atmospheric Administration's FV3 atmospheric model[15] (see Supplementary Information section A for details).

Nevertheless, machine-learning approaches have noteworthy limitations compared with GCMs. Existing machine-learning models have focused on deterministic prediction, and surpass deterministic numerical weather prediction in terms of the aggregate metrics for which they are trained[3,4]. However, they do not produce calibrated uncertainty estimates[4], which is essential for useful weather forecasts[1]. Deterministic machine-learning models using a mean-squared-error loss are rewarded for averaging over uncertainty, producing unrealistically blurry predictions when optimized for multi-day forecasts[3,13]. Unlike physical models, machine-learning models misrepresent derived (diagnostic) variables such as geostrophic wind[16]. Furthermore,

[1]Google Research, Mountain View, CA, USA. [2]Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. [3]European Centre for Medium-Range Weather Forecasts, Reading, UK. [4]Google DeepMind, London, UK. [5]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. [6]These authors contributed equally: Dmitrii Kochkov, Janni Yuval, Ian Langmore, Peter Norgaard, Jamie Smith, Stephan Hoyer. e-mail: dkochkov@google.com; janniyuval@google.com; shoyer@google.com



https://doi.org/10.1038/s41586-024-07744-y

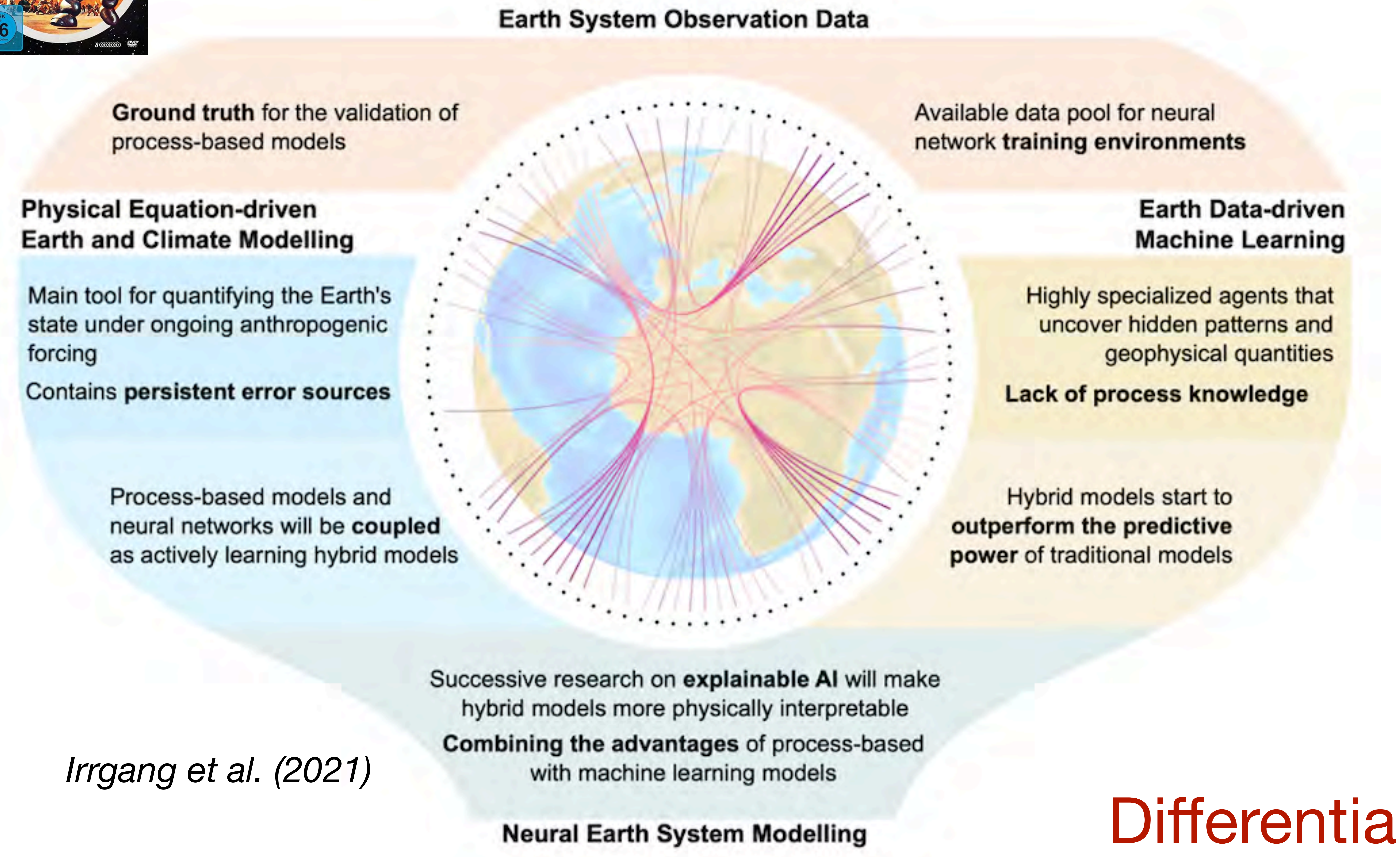
Kochkov et al. (2024)

https://github.com/google-research/dinosaur

https://github.com/google-research/neuralgcm

Towards AI-native hybrid geoscientific models ?

AI-native hybrid geoscientific models ?



Irrgang et al. (2021)

Betting harnessing **observations** & hi-fidelity **simulations**

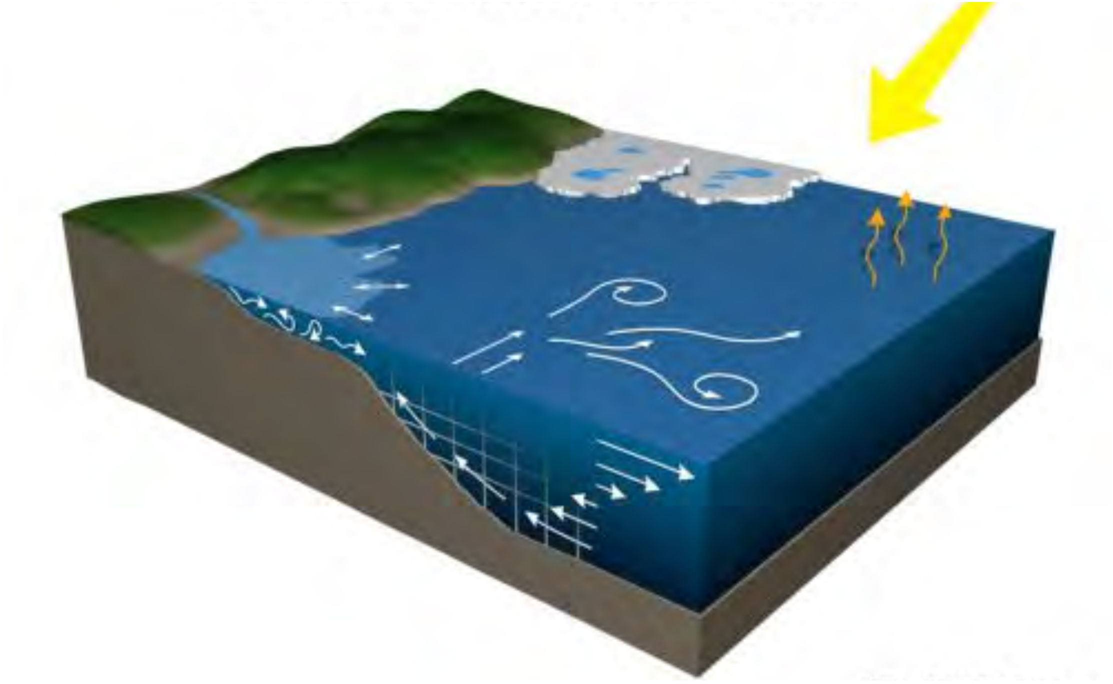
… for optimising

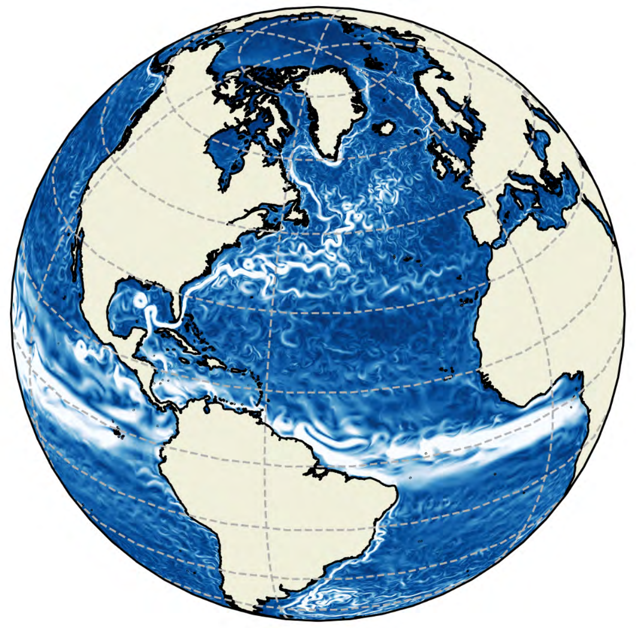
- model **parameters**
- **numerical** schemes
- subgrid **closures**
- …

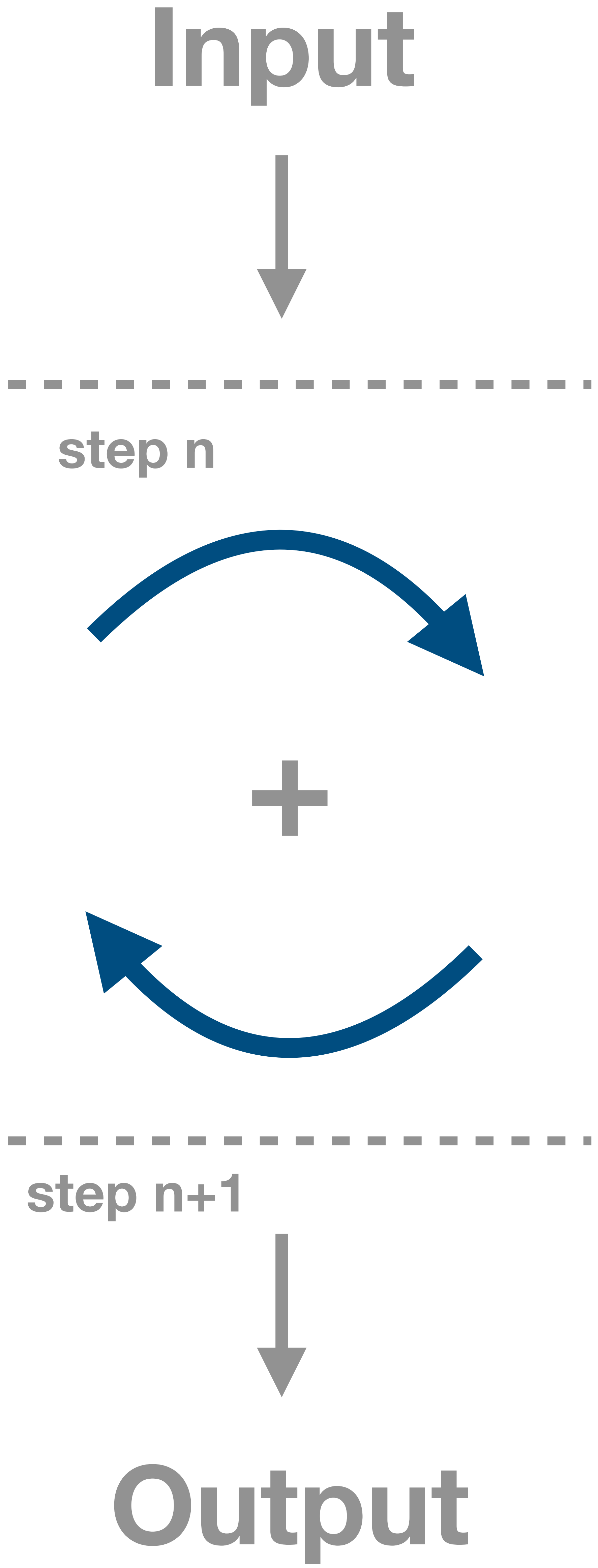
Differentiable programming in earth system models ?

Current generation hybrid models

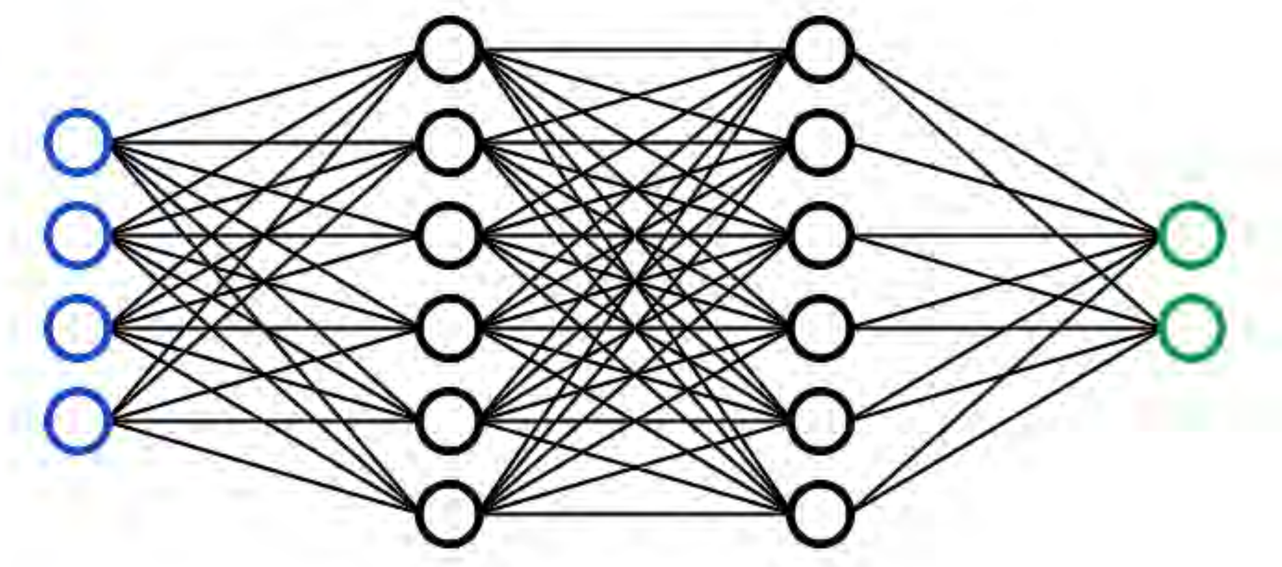
Ocean circulation models







Trainable components
(closures, error corrections)

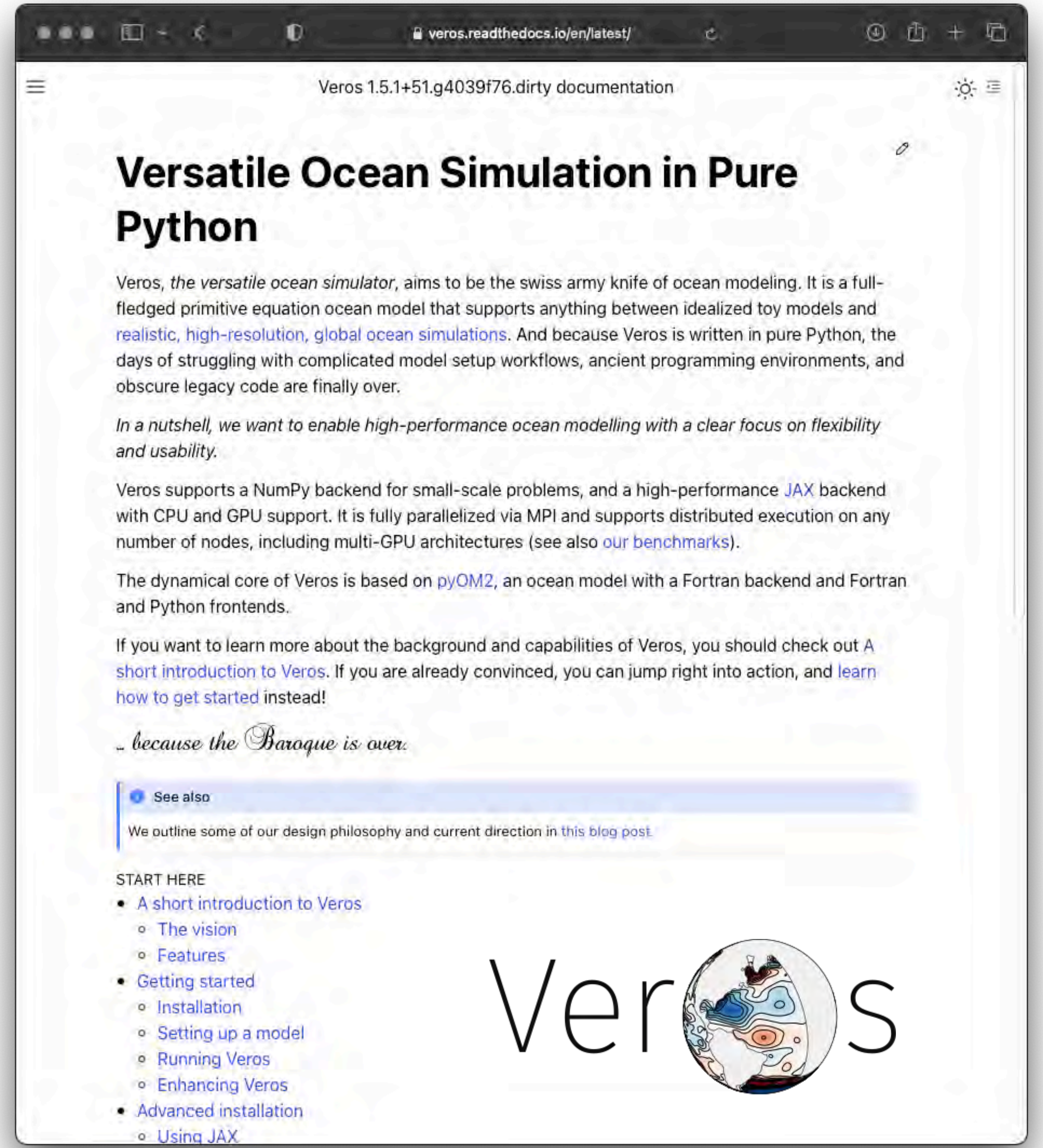


Current generation hybrid models

Existing systems
Less flexible software design
(APIs, DevOps, CI, …)

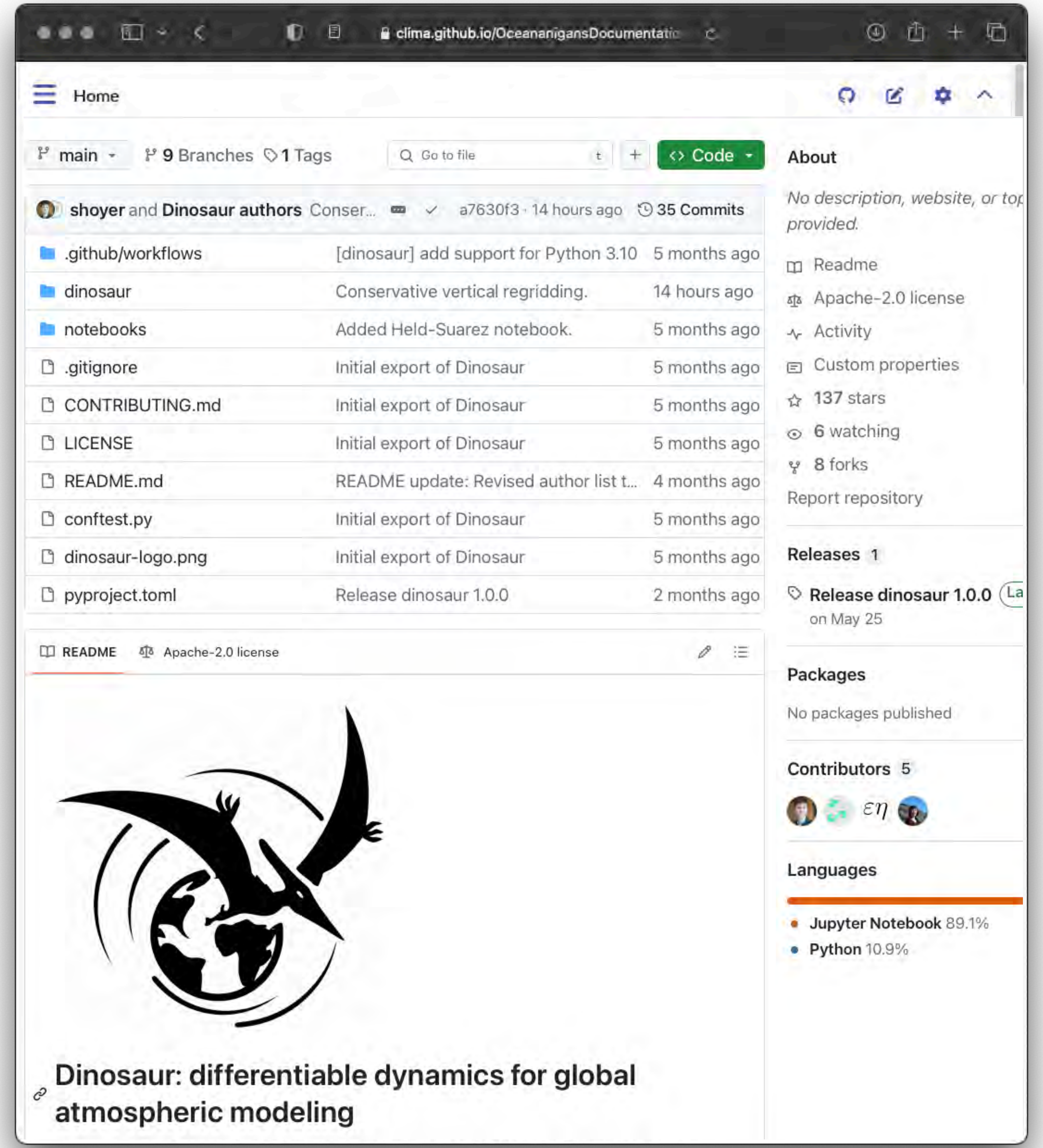
Modern code practices :
Robust and versatile APIs and MLOPs

A new generation of geoscientific models



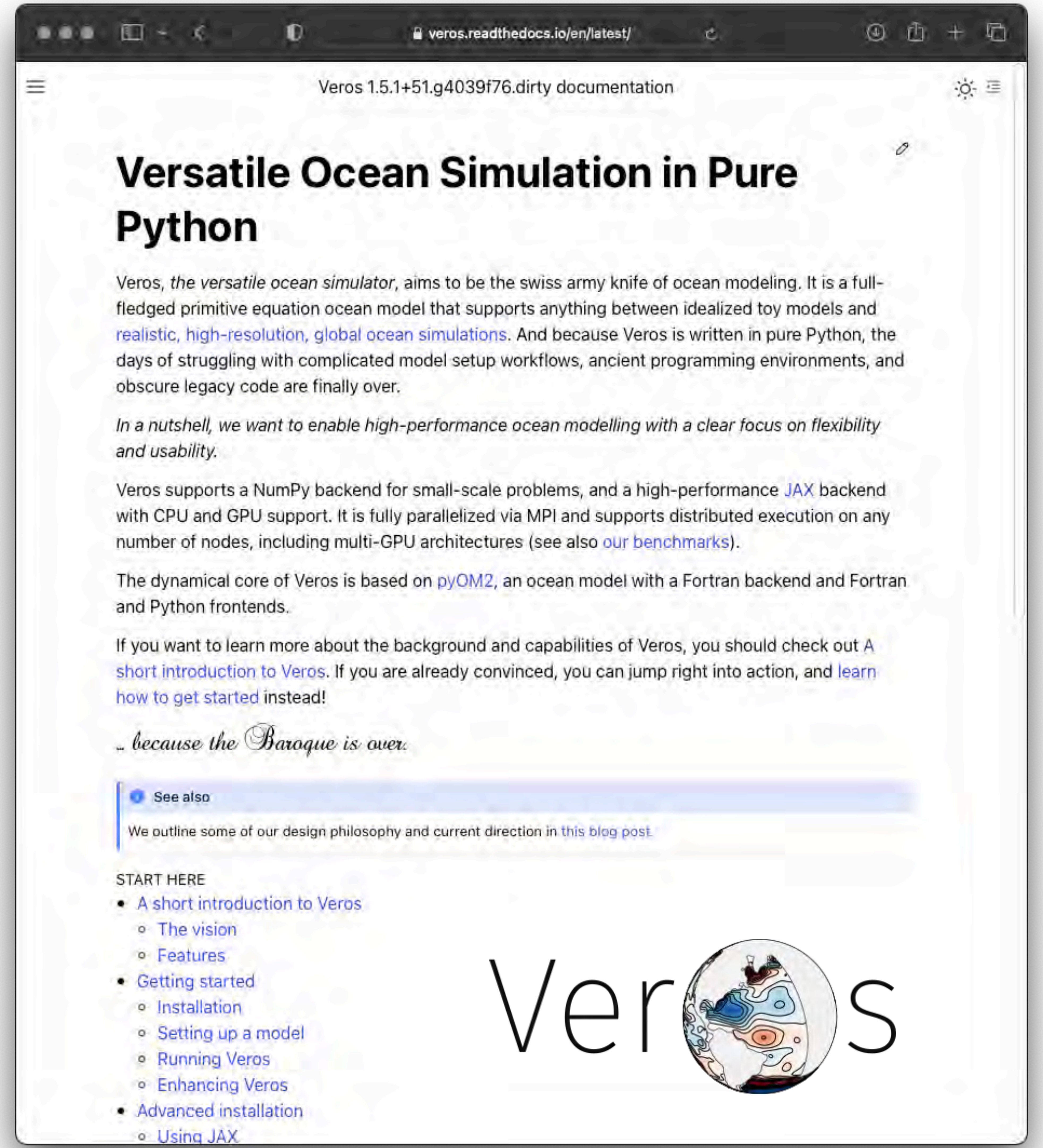
Atmos

Ocean



Modern code and compute : simple to write, scales, runs on any hardware

A new generation of geoscientific models



veros.readthedocs.io/en/latest/

Veros 1.5.1+51.g4039f76.dirty documentation

Versatile Ocean Simulation in Pure Python

Veros, *the versatile ocean simulator*, aims to be the swiss army knife of ocean modeling. It is a full-fledged primitive equation ocean model that supports anything between idealized toy models and realistic, high-resolution, global ocean simulations. And because Veros is written in pure Python, the days of struggling with complicated model setup workflows, ancient programming environments, and obscure legacy code are finally over.

In a nutshell, we want to enable high-performance ocean modelling with a clear focus on flexibility and usability.

Veros supports a NumPy backend for small-scale problems, and a high-performance JAX backend with CPU and GPU support. It is fully parallelized via MPI and supports distributed execution on any number of nodes, including multi-GPU architectures (see also our benchmarks).

The dynamical core of Veros is based on pyOM2, an ocean model with a Fortran backend and Fortran and Python frontends.

If you want to learn more about the background and capabilities of Veros, you should check out A short introduction to Veros. If you are already convinced, you can jump right into action, and learn how to get started instead!

… because the Baroque is over.

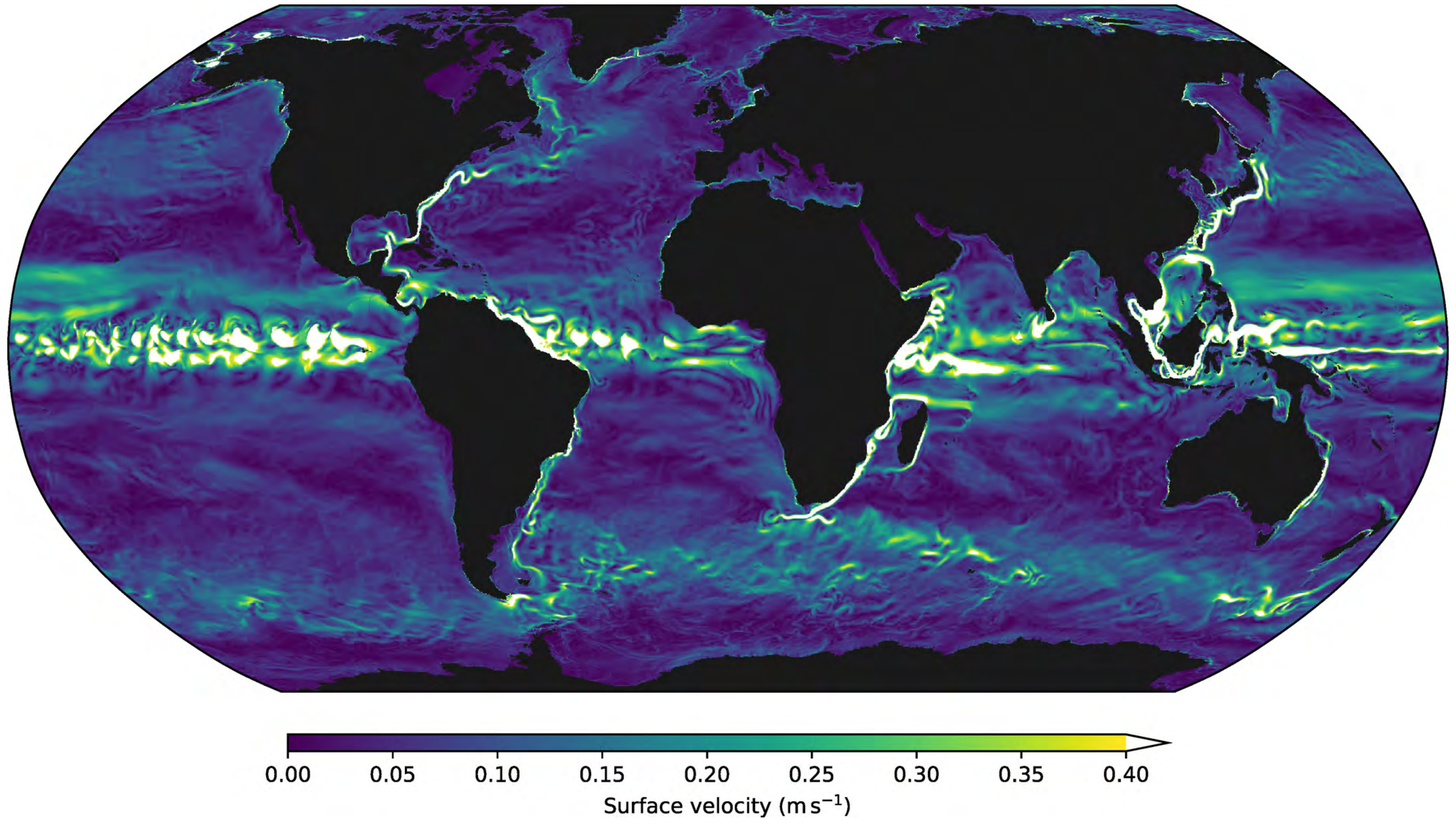
See also

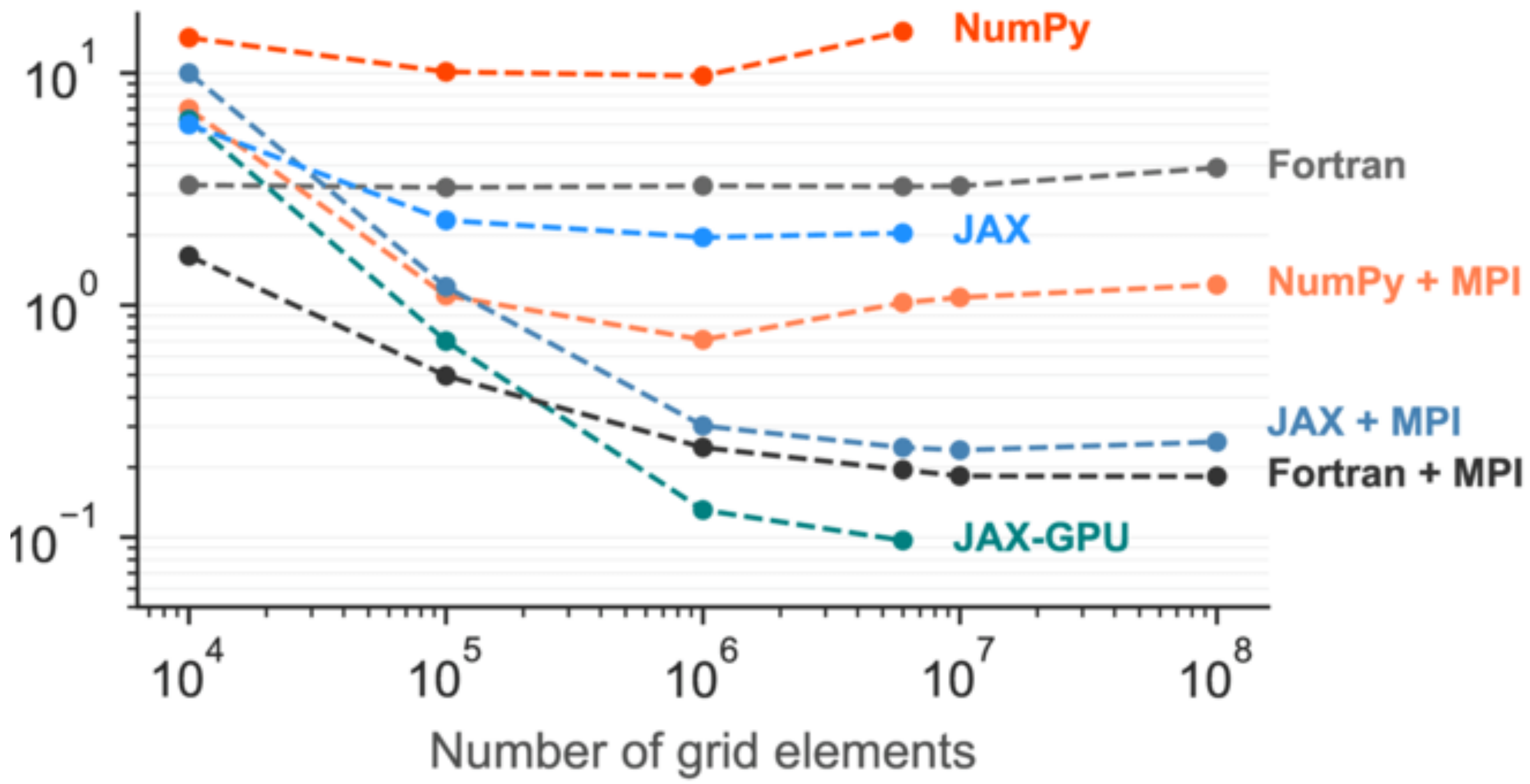
We outline some of our design philosophy and current direction in this blog post.

START HERE

- A short introduction to Veros
  - The vision
  - Features
- Getting started
  - Installation
  - Setting up a model
  - Running Veros
  - Enhancing Veros
- Advanced installation
  - Using JAX

Veros

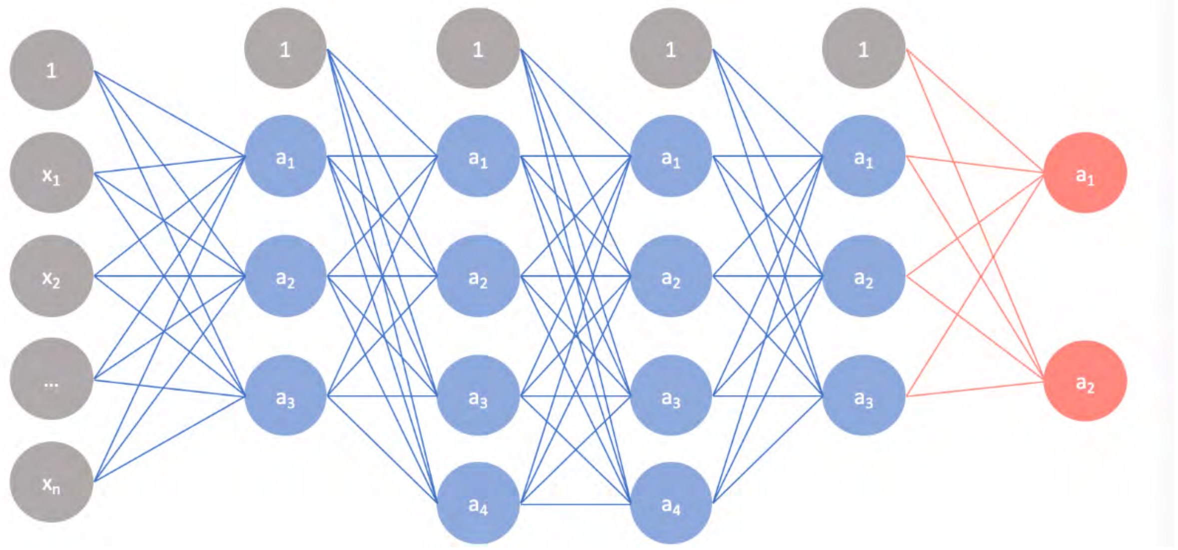




Modern code and compute : simple to write, scales, runs on any hardware

Allowing a seamless integration with AI

AI



Physics

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

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

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

PyTorch

high abstraction, fast evolving languages

PyTorch

high abstraction, fast evolving languages

AI-native hybrid geoscientific models

Neural general circulation models for weather and climate

https://doi.org/10.1038/s41586-024-07744-y
Received: 13 November 2023
Accepted: 15 June 2024
Published online: 22 July 2024
Open access
Check for updates

Dmitrii Kochkov[1,6✉], Janni Yuval[1,6✉], Ian Langmore[1,6], Peter Norgaard[1,6], Jamie Smith[1,6], Griffin Mooers[1], Milan Klöwer[2], James Lottes[1], Stephan Rasp[1], Peter Düben[3], Sam Hatfield[3], Peter Battaglia[4], Alvaro Sanchez-Gonzalez[4], Matthew Willson[4], Michael P. Brenner[1,5] & Stephan Hoyer[1,6✉]

General circulation models (GCMs) are the foundation of weather and climate prediction[1,2]. GCMs are physics-based simulators that combine a numerical solver for large-scale dynamics with tuned representations for small-scale processes such as cloud formation. Recently, machine-learning models trained on reanalysis data have achieved comparable or better skill than GCMs for deterministic weather forecasting[3,4]. However, these models have not demonstrated improved ensemble forecasts, or shown sufficient stability for long-term weather and climate simulations. Here we present a GCM that combines a differentiable solver for atmospheric dynamics with machine-learning components and show that it can generate forecasts of deterministic weather, ensemble weather and climate on par with the best machine-learning and physics-based methods. NeuralGCM is competitive with machine-learning models for one- to ten-day forecasts, and with the European Centre for Medium-Range Weather Forecasts ensemble prediction for one- to fifteen-day forecasts. With prescribed sea surface temperature, NeuralGCM can accurately track climate metrics for multiple decades, and climate forecasts with 140-kilometre resolution show emergent phenomena such as realistic frequency and trajectories of tropical cyclones. For both weather and climate, our approach offers orders of magnitude computational savings over conventional GCMs, although our model does not extrapolate to substantially different future climates. Our results show that end-to-end deep learning is compatible with tasks performed by conventional GCMs and can enhance the large-scale physical simulations that are essential for understanding and predicting the Earth system.

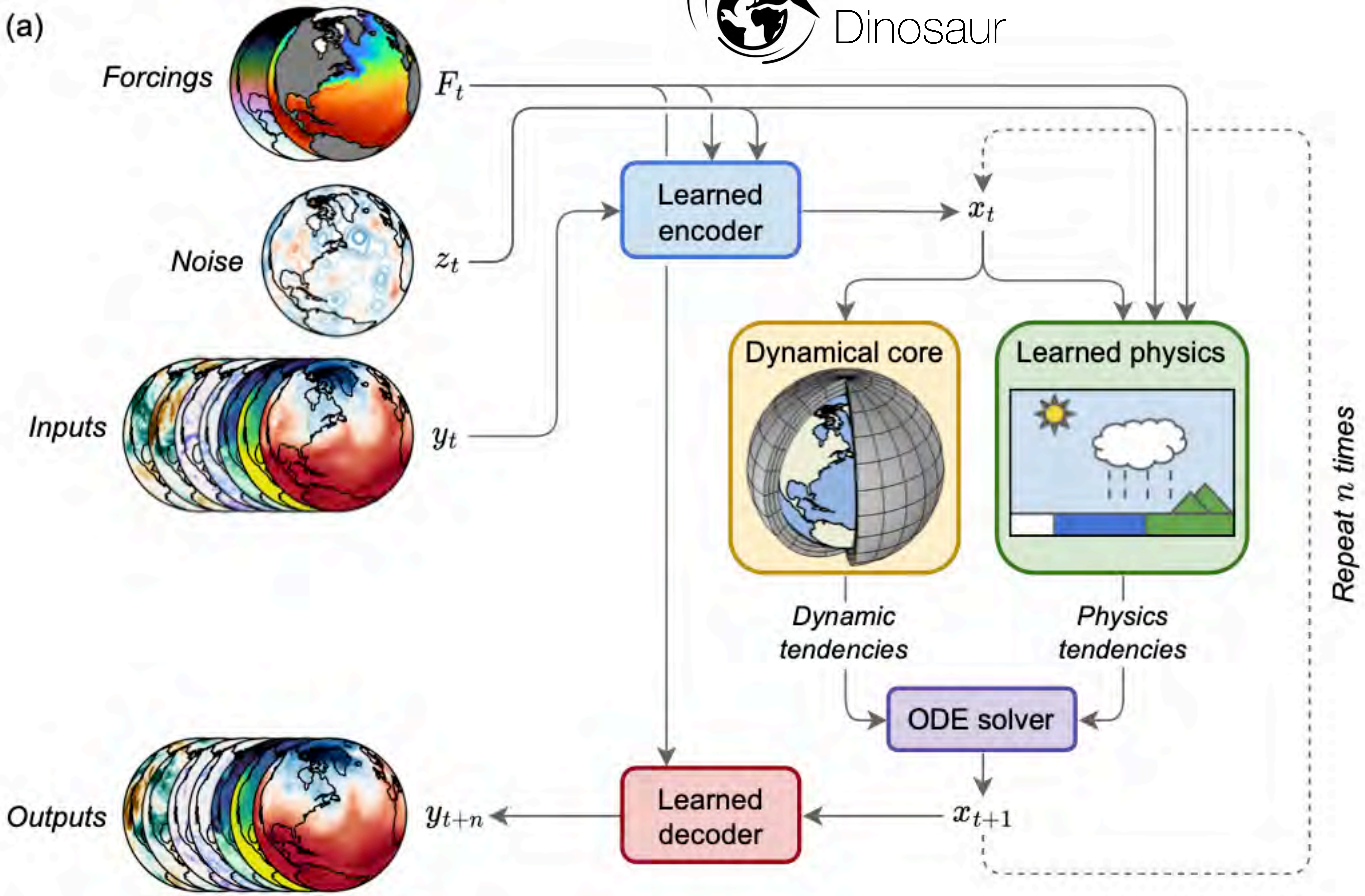
Solving the equations for Earth's atmosphere with general circulation models (GCMs) is the basis of weather and climate prediction[1,2]. Over the past 70 years, GCMs have been steadily improved with better numerical methods and more detailed physical models, while exploiting faster computers to run at higher resolution. Inside GCMs, the unresolved physical processes such as clouds, radiation and precipitation are represented by semi-empirical parameterizations. Tuning GCMs to match historical data remains a manual process[5], and GCMs retain many persistent errors and biases[6–8]. The difficulty of reducing uncertainty in long-term climate projections[9] and estimating distributions of extreme weather events[10] presents major challenges for climate mitigation and adaptation[11].

Recent advances in machine learning have presented an alternative for weather forecasting[3,4,12,13]. These models rely solely on machine-learning techniques, using roughly 40 years of historical data from the European Center for Medium-Range Weather Forecasts (ECMWF) reanalysis v5 (ERA5)[14] for model training and forecast initialization. Machine-learning methods have been remarkably successful, demonstrating state-of-the-art deterministic forecasts for 1- to 10-day weather prediction at a fraction of the computational cost of traditional models[3,4]. Machine-learning atmospheric models also require considerably less code, for example GraphCast[3] has 5,417 lines versus 376,578 lines for the National Oceanic and Atmospheric Administration's FV3 atmospheric model[15] (see Supplementary Information section A for details).

Nevertheless, machine-learning approaches have noteworthy limitations compared with GCMs. Existing machine-learning models have focused on deterministic prediction, and surpass deterministic numerical weather prediction in terms of the aggregate metrics for which they are trained[3,4]. However, they do not produce calibrated uncertainty estimates[4], which is essential for useful weather forecasts[1]. Deterministic machine-learning models using a mean-squared-error loss are rewarded for averaging over uncertainty, producing unrealistically blurry predictions when optimized for multi-day forecasts[3,13]. Unlike physical models, machine-learning models misrepresent derived (diagnostic) variables such as geostrophic wind[16]. Furthermore,

[1]Google Research, Mountain View, CA, USA. [2]Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. [3]European Centre for Medium-Range Weather Forecasts, Reading, UK. [4]Google DeepMind, London, UK. [5]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. [6]These authors contributed equally: Dmitrii Kochkov, Janni Yuval, Ian Langmore, Peter Norgaard, Jamie Smith, Stephan Hoyer. ✉e-mail: dkochkov@google.com; janniyuval@google.com; shoyer@google.com

Dinosaur



https://github.com/google-research/dinosaur

https://github.com/google-research/neuralgcm

https://doi.org/10.1038/s41586-024-07744-y

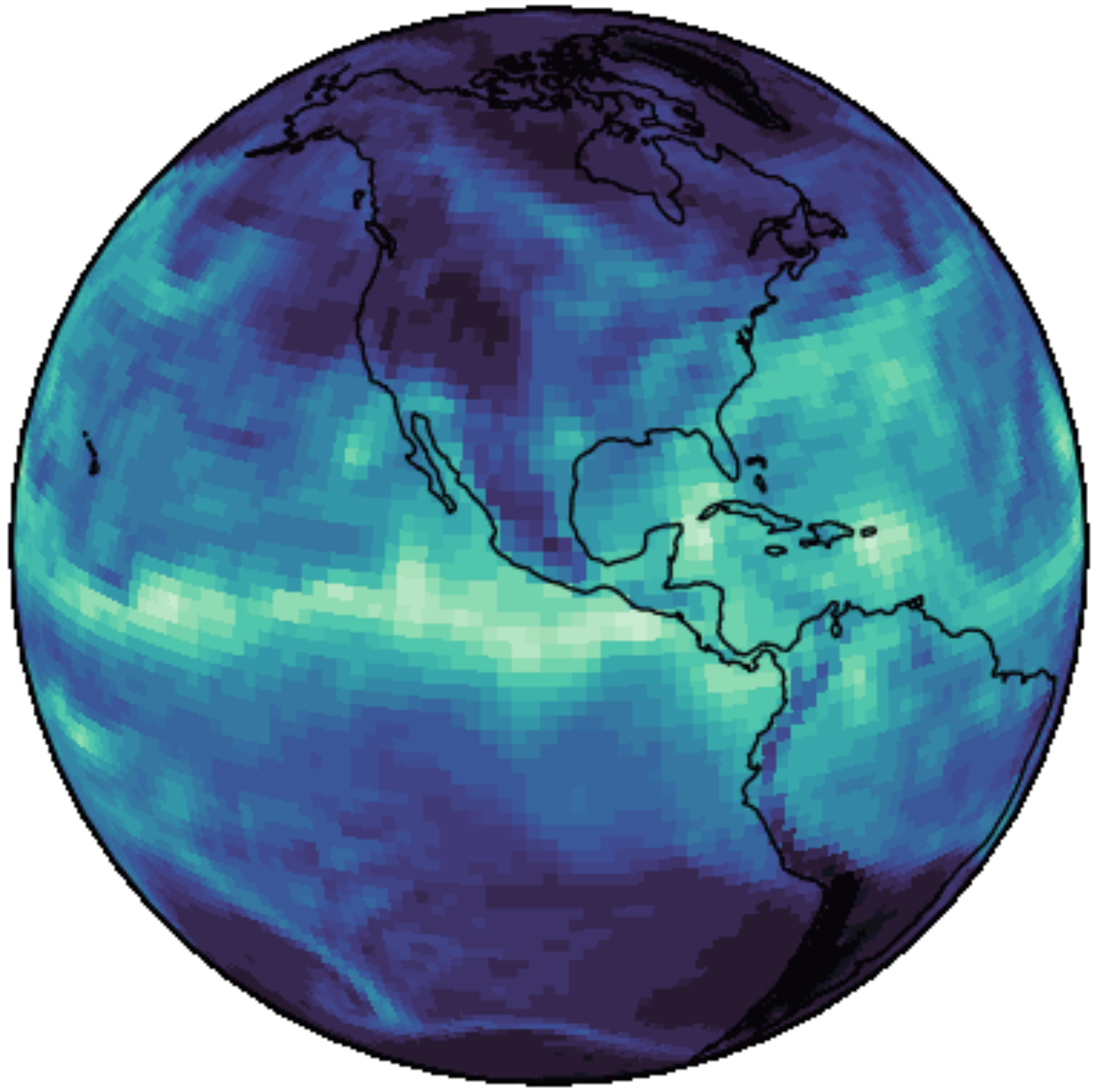
Kochkov et al. (2024)

AI-native hybrid geoscientific models

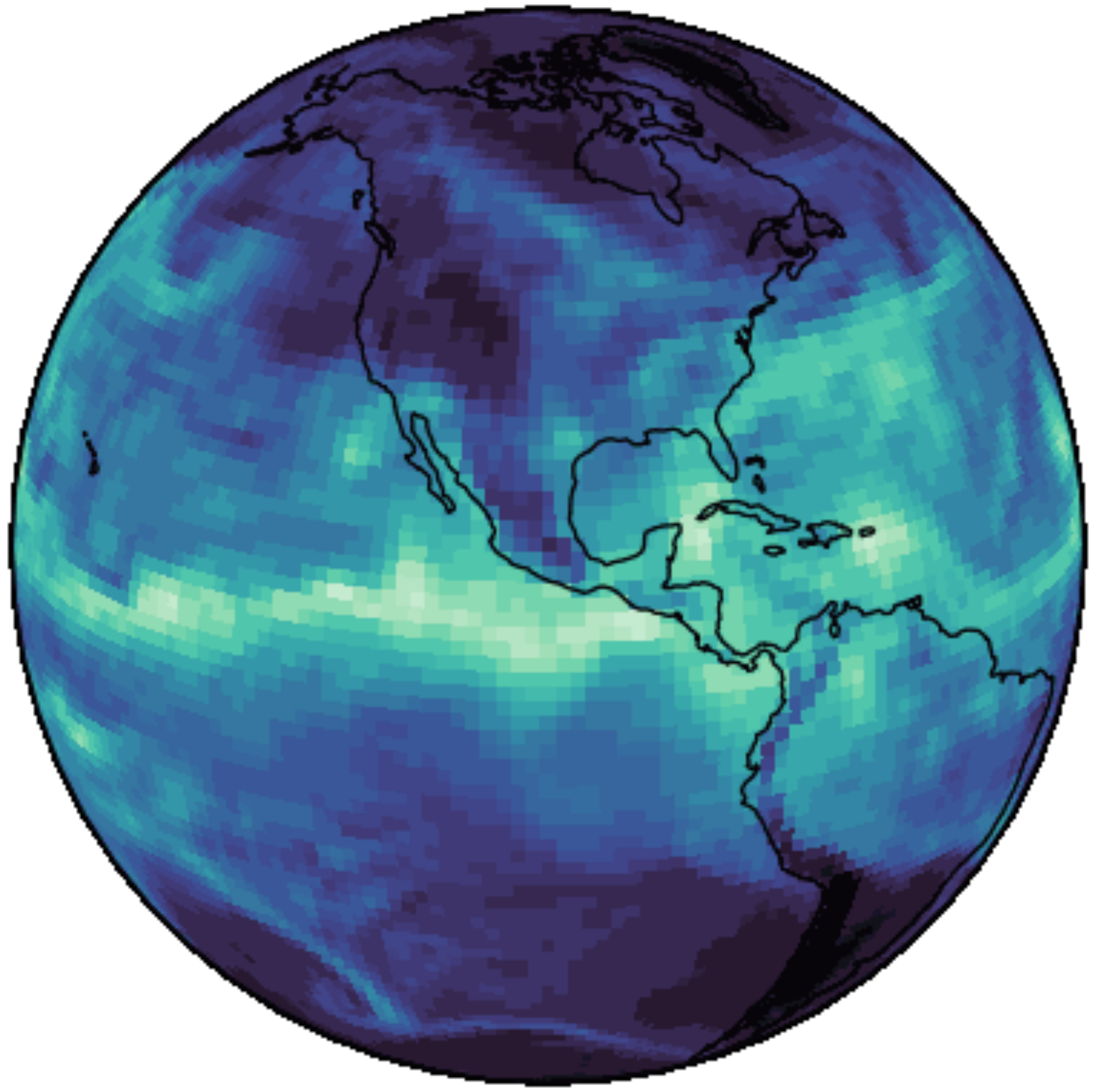
Kochkov et al. (2024)

https://arxiv.org/abs/2311.07222

Total column water, 0-15 days



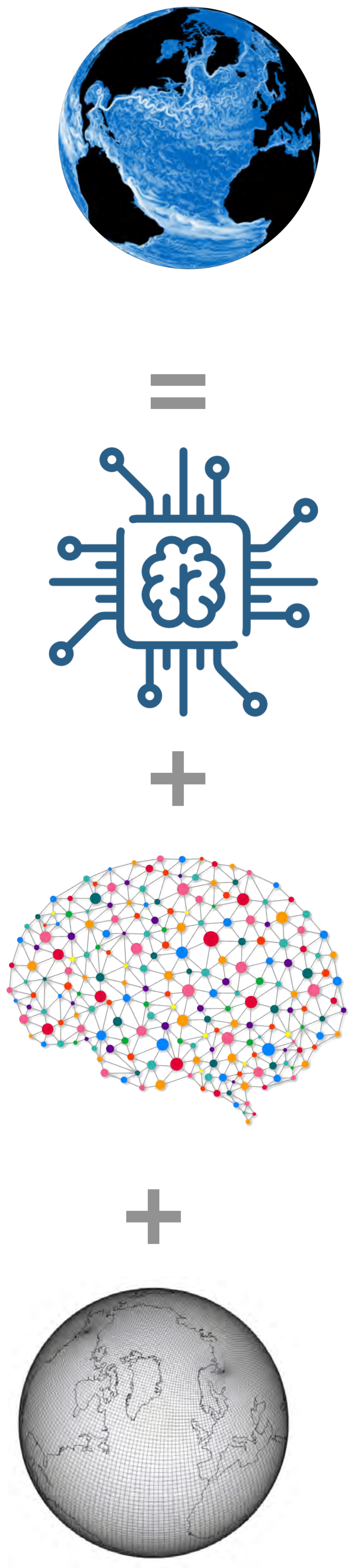
ERA5



NeuralGCM

Hybrid w/ online : non-blurry forecast + stable simulators (runs ~10 years)

Summary



- Illustrated why we are **augmenting models** with ML
- Described how this can be **done in practice** today
- Advocated that a **deep recast** of our models is needed
- Described upcoming **AI-native hybrid** models
- Exciting time for **cross-disciplinary** investigations !

IGE
MIAI Grenoble Alpes
cnrs
UGA Université Grenoble Alpes

References

- Frezat et al. (2021). Physical invariance in neural networks for subgrid-scale scalar flux modeling. https://doi.org/10.1103/PhysRevFluids.6.024607

- Frezat et al. (2022) A posteriori learning for quasi-geostrophic turbulence parametrization, JAMES. 14. https://doi.org/10.1029/2022MS003124

- Frezat et al. (2024) Gradient-free online learning of subgrid-scale dynamics with neural emulators, sub., https://doi.org/10.48550/arXiv.2310.19385

- Yan (2024). Adjoint-based online learning of two-layer quasi-geostrophic baroclinic turbulence. arXiv. https://doi.org/10.48550/arXiv.2411.14106