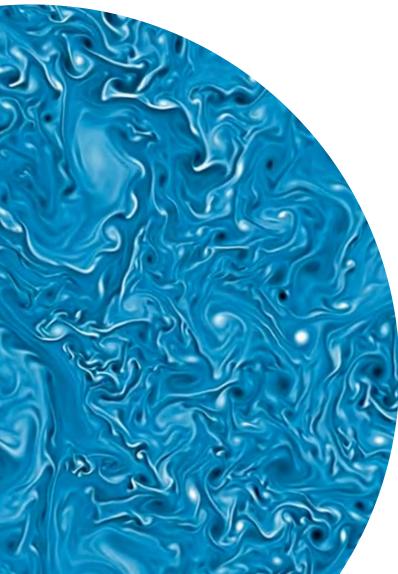
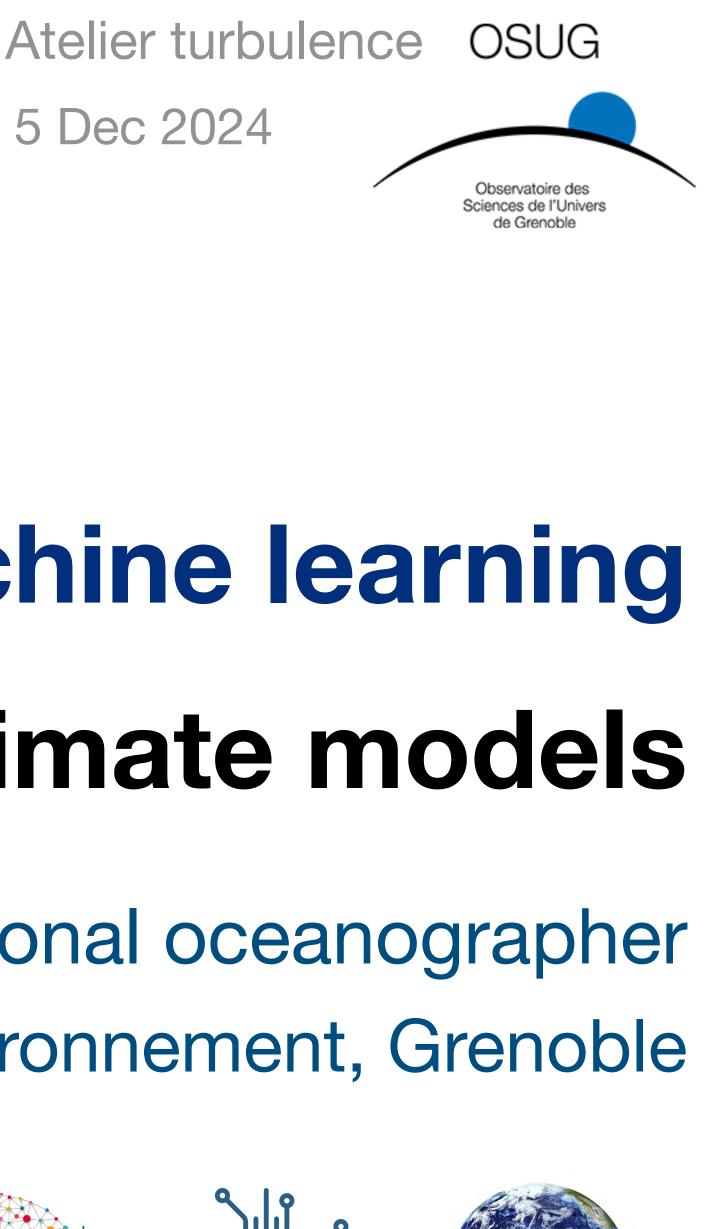


Combining physics and machine learning in hybrid climate models







Julien Le Sommer - computational oceanographer Institut des Géosciences de l'Environnement, Grenoble

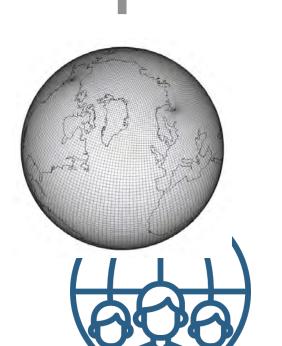
Objectives of this talk



cal

- with methods from the emerging field of SciML
- How this leads to to deep changes in our systems

products and observations

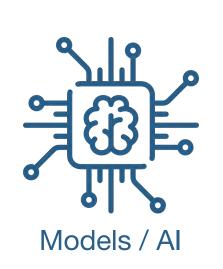






- how ML is leveraged in computational oceanography

- and some interesting questions for the future...





CNIS

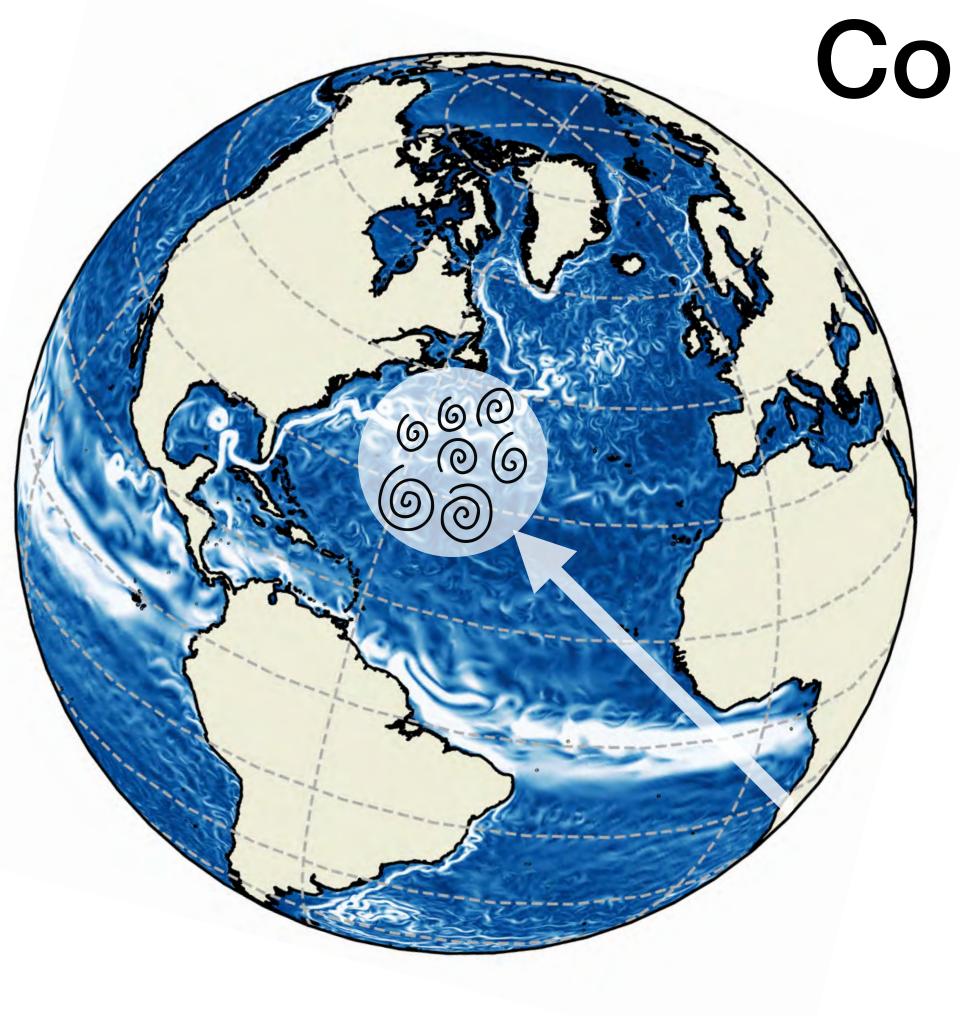






The context of computational oceanography





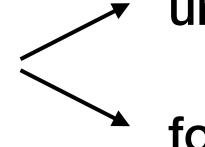
physical oceanography currents, parameters

Macro-turbulence



develops and use numerical tools and metions Next generation maths, numerics, compute, data

Computational oceanography



understand the functioning

forecast its evolution (timescales)

climate - environnemental changes human activities

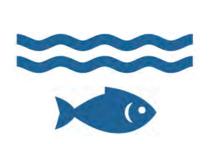


Scale interactions, processes

Surface waves

Interactions with components













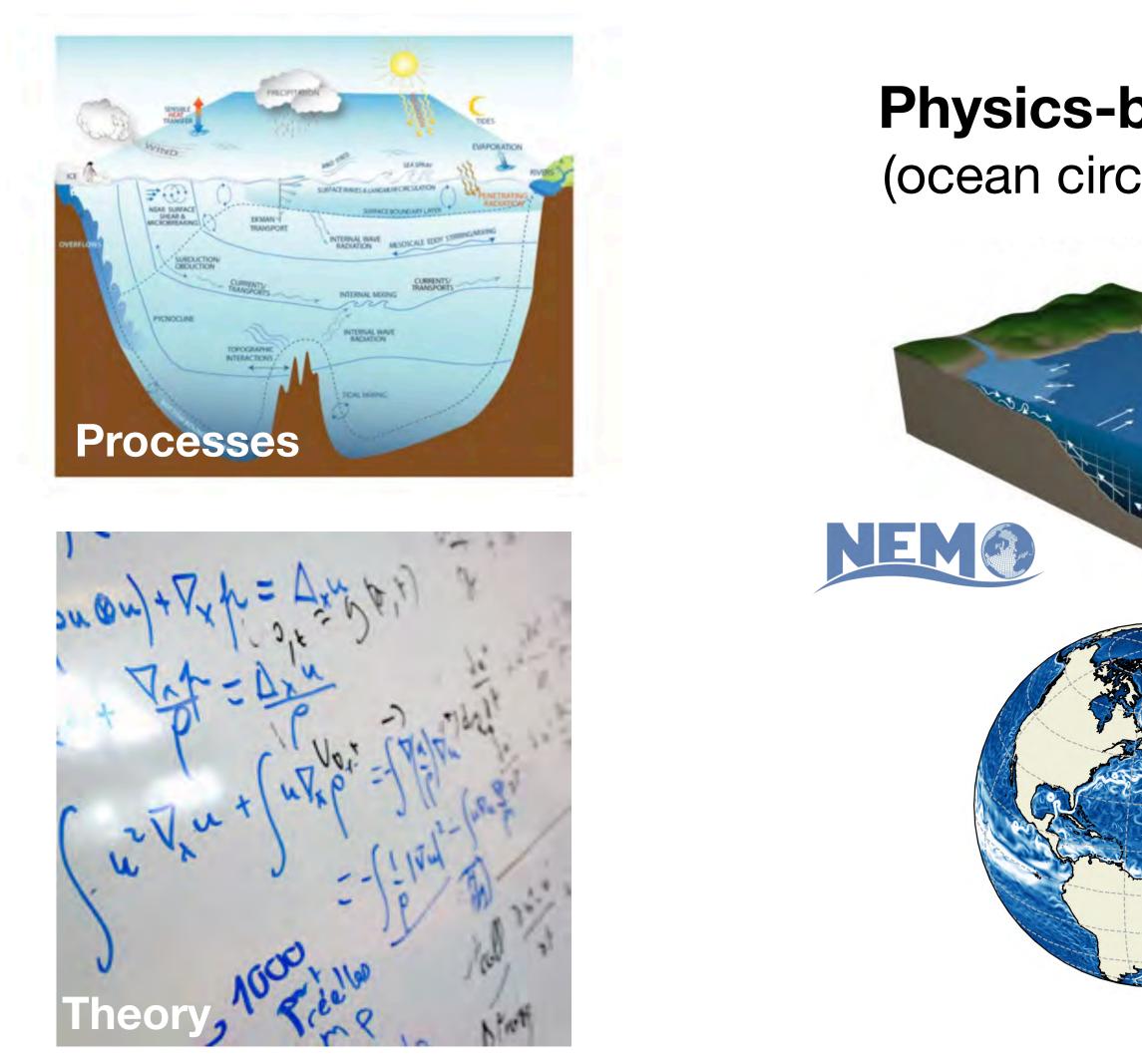


internal waves (tides)

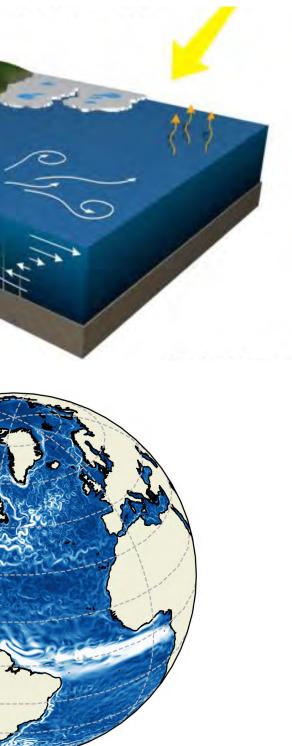


2.

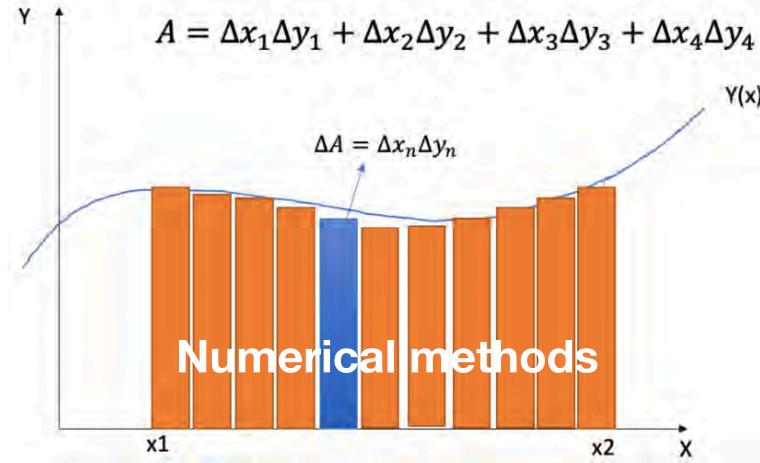
A key tool : ocean models



Physics-based models (ocean circulation models)



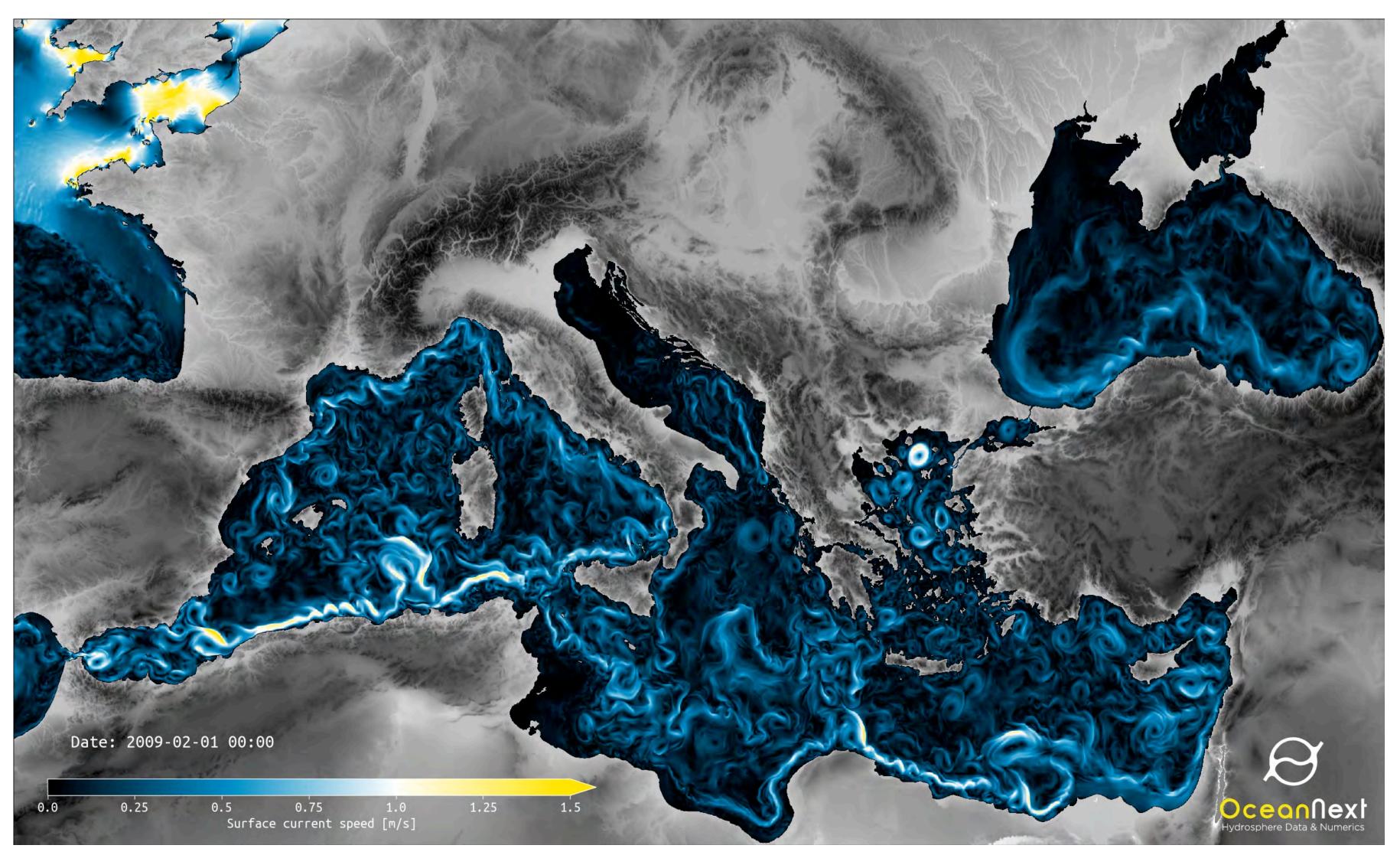




Physical models summarize our understanding of physical systems



A key tool : ocean models



Physical models summarize our understanding of physical systems

dx ~ 1km tides, eddies



Tier 1 HPC >200 000 lines >15 yrs 5 institutions

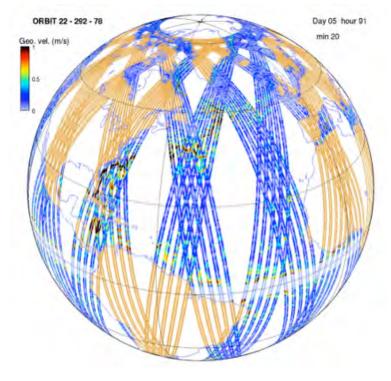




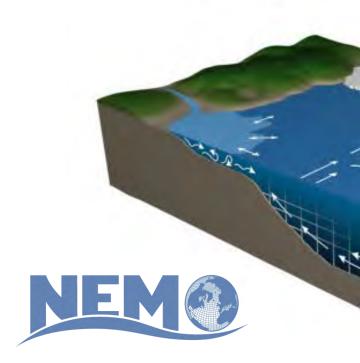
Our toolboxes

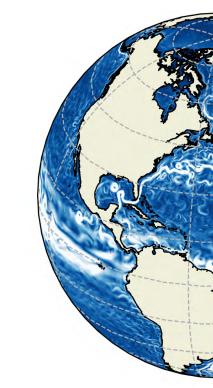
Observations (in situ/satellite)



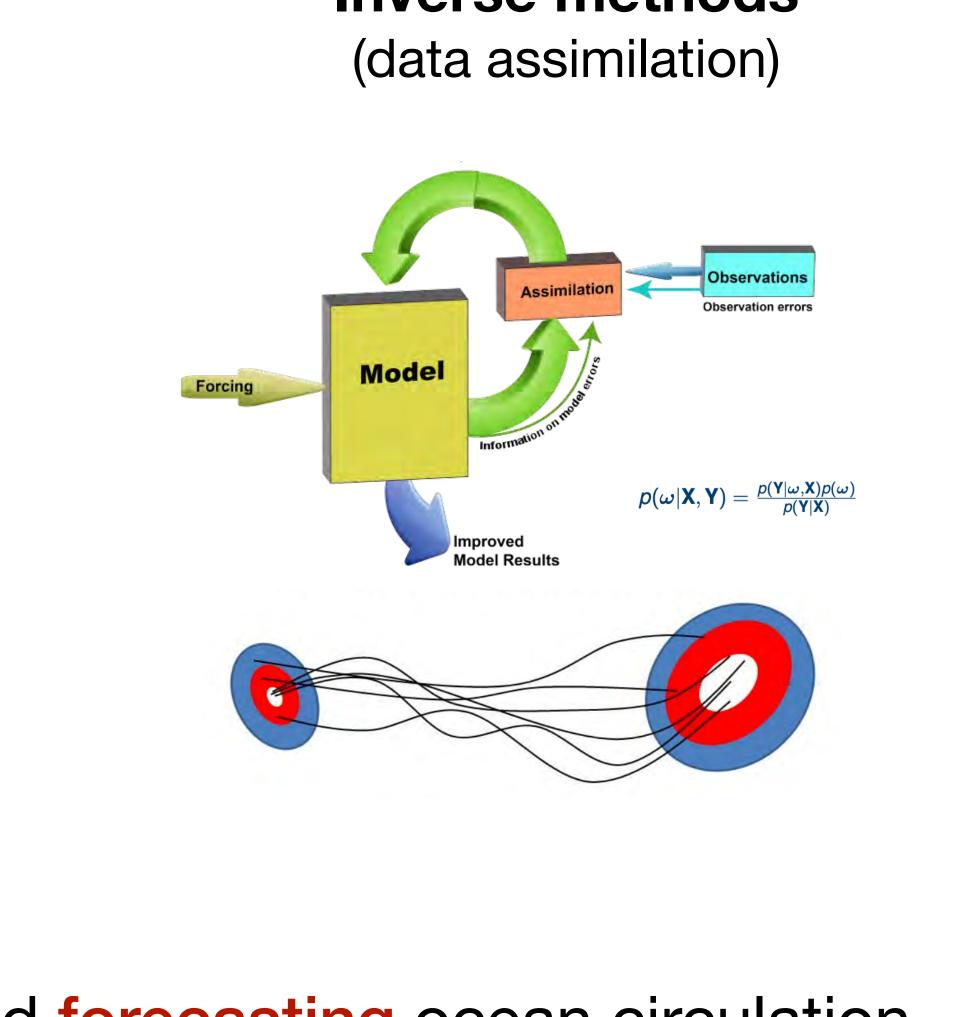


Physical models (ocean circulation models)





Inverse methods (data assimilation)

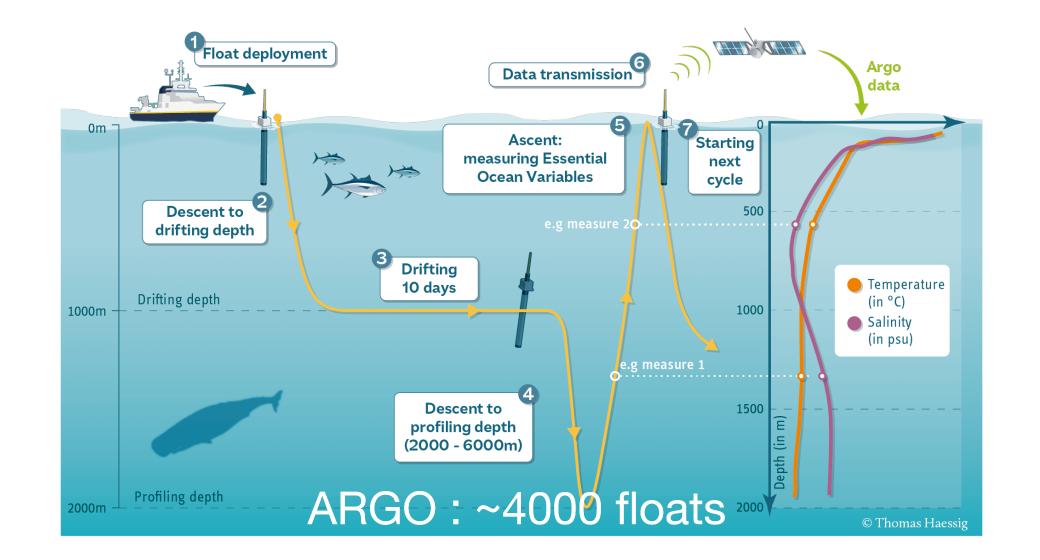


Tools for understanding but also monitoring and forecasting ocean circulation

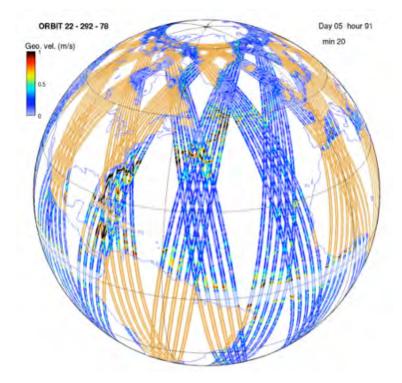
Observations of the ocean

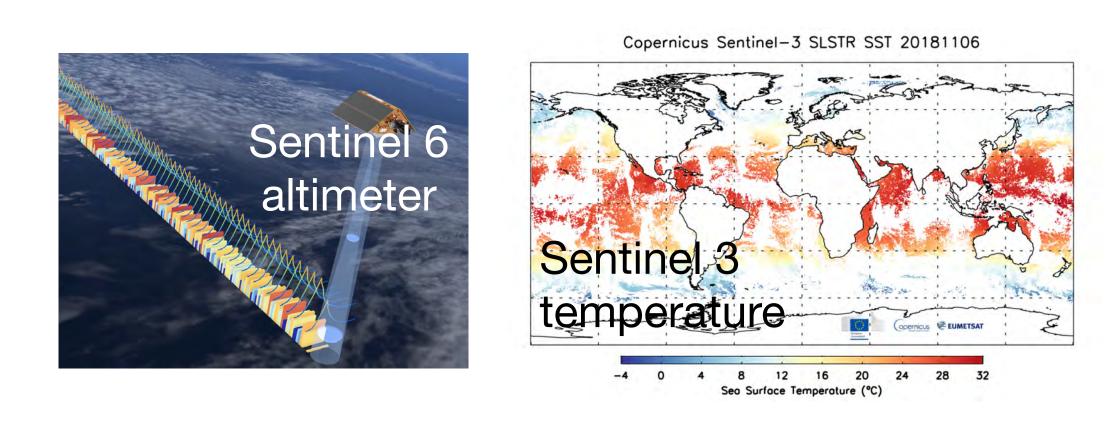
Observations



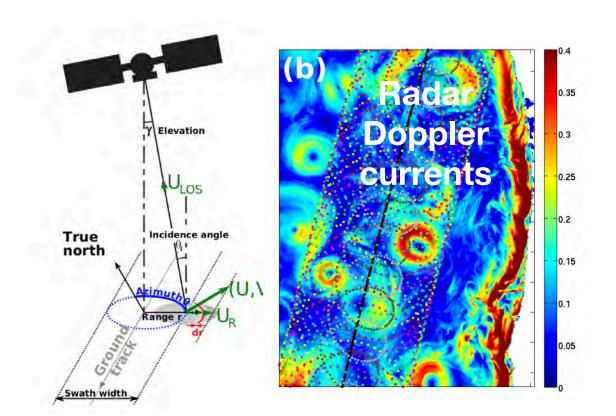


satellite









New platforms

Continuously operated networks

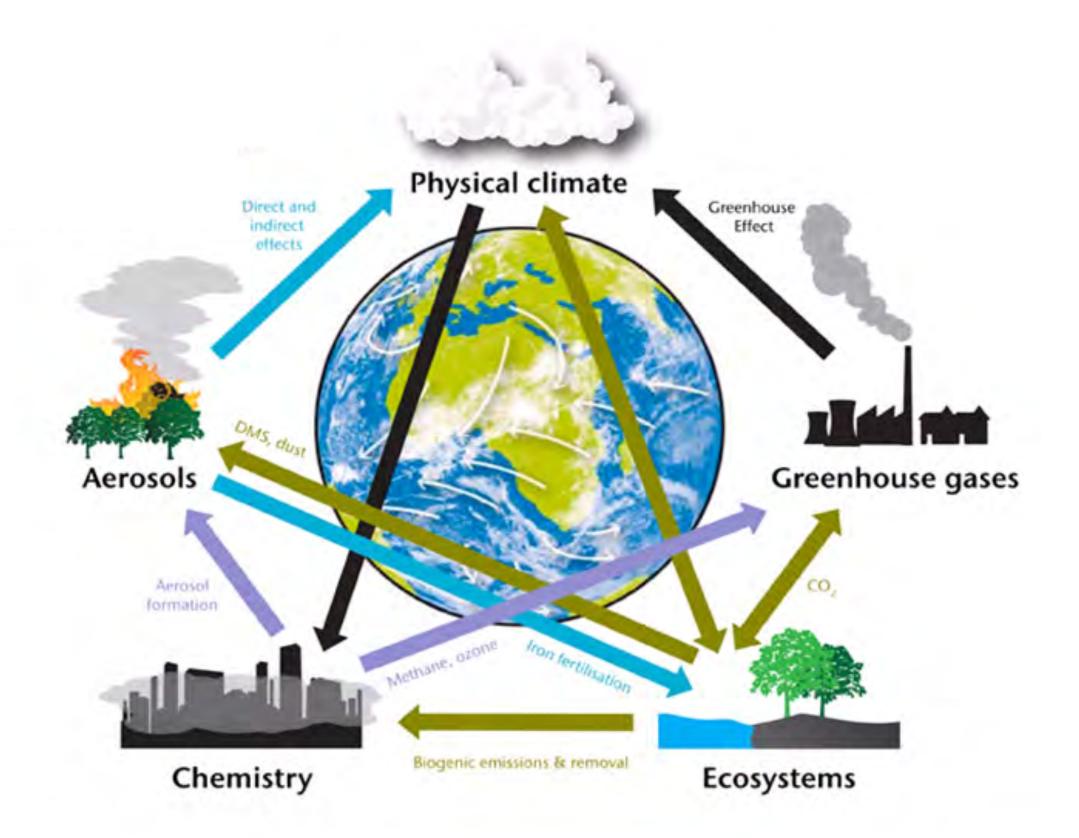




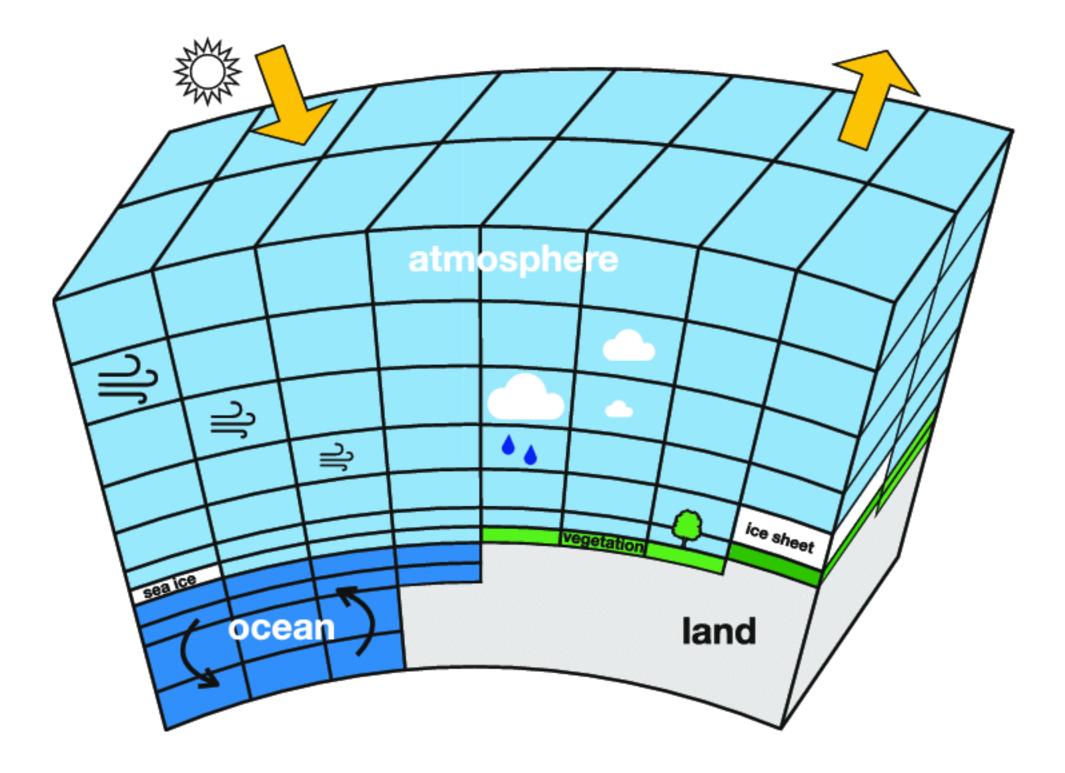
How AI (SciML) is affecting our field ?



Tools are integrated into systems



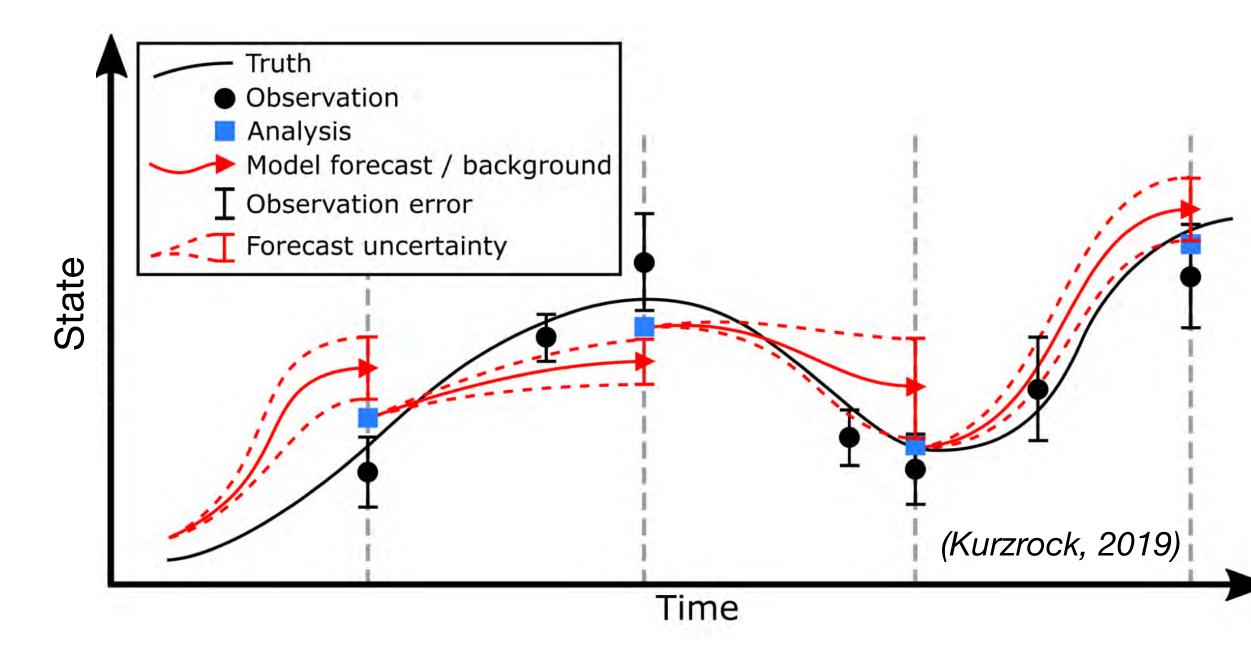
Earth System Models (IPCC)



Combining models of each components of the climate system

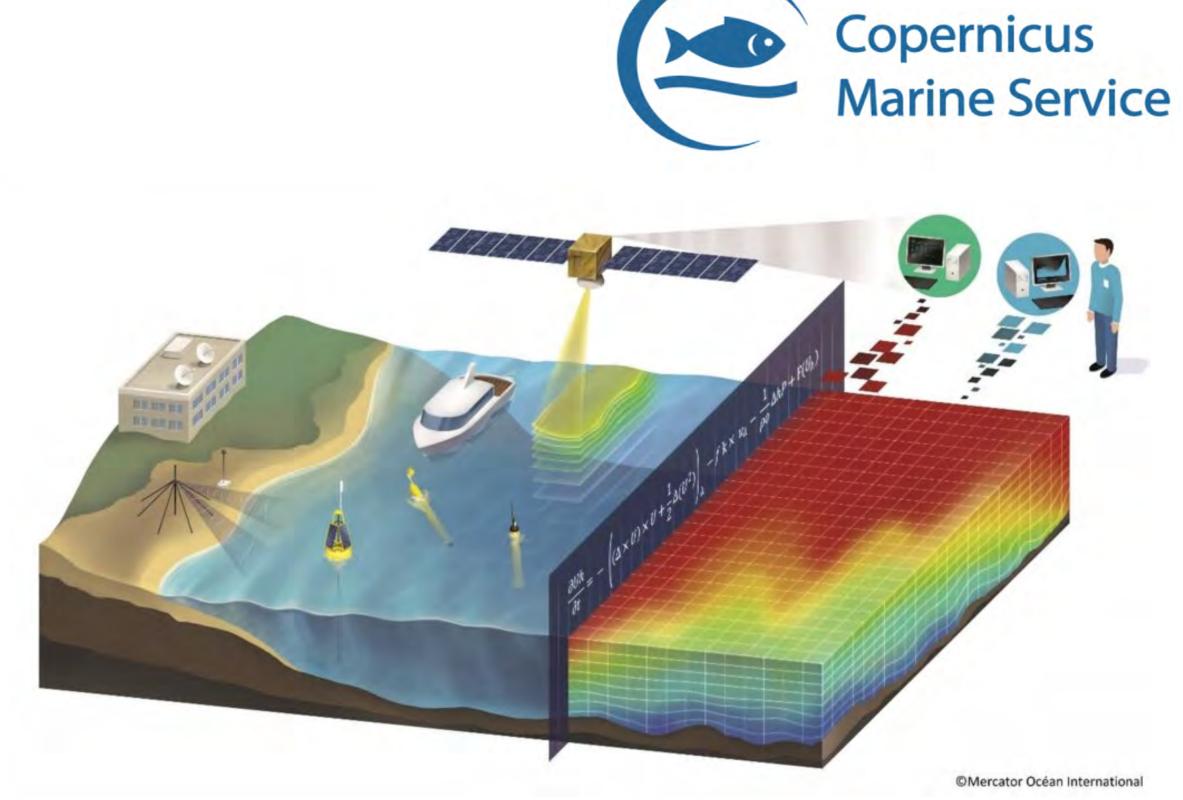


Tools are integrated into systems



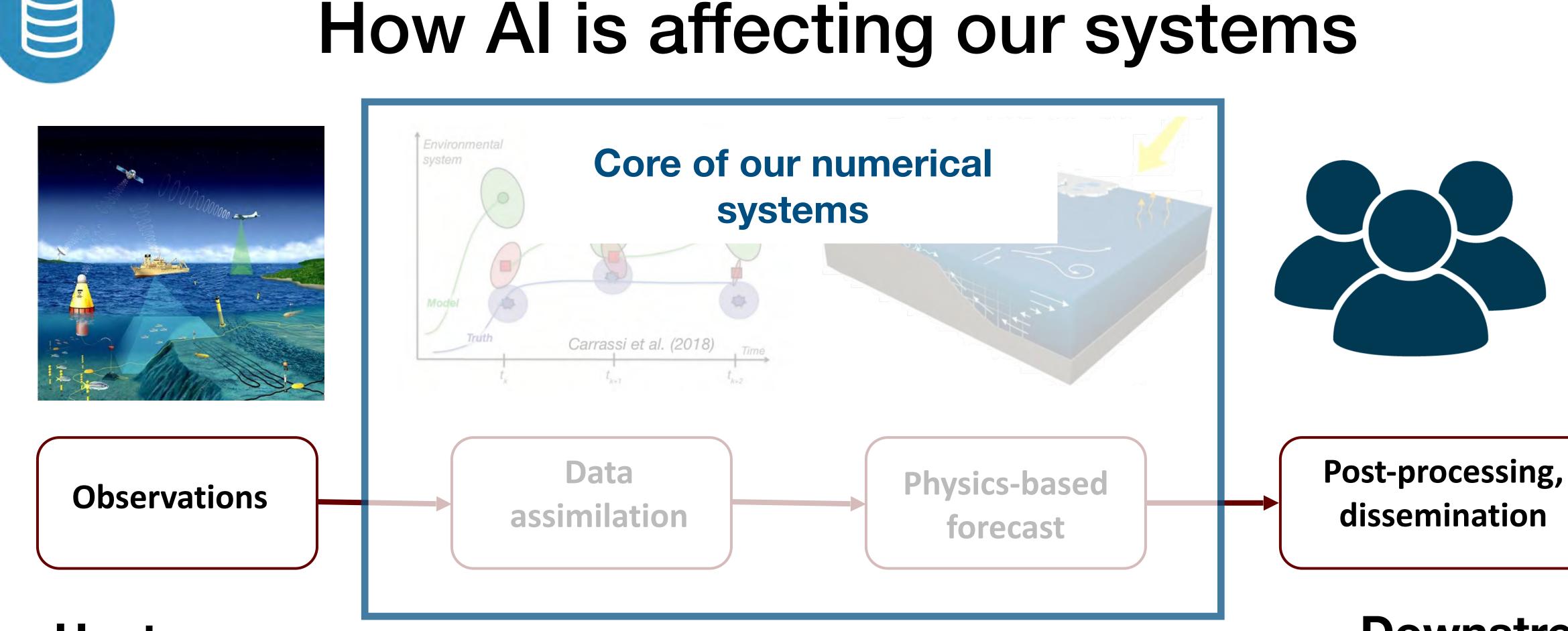
Combining models and observations to produce forecasts





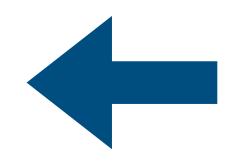
Operational prediction systems (Copernicus)

How Al is affecting our systems

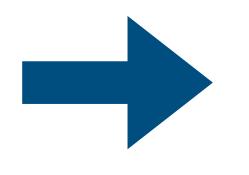


Upstream

denoising, inpainting parameter retrieval quality control



Al, machine learning & data-driven approaches **Downstream**



data fusion, tailored services data mining







Al-based ocean forecasting

MANUSCRIPT 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° 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

[phy INTRODUCTION

2024

Feb

00

[qd-

Q

5

0299

:2402.

Ocean forecasting is critically important for many ma- usually computationally expensive and slow. For example, rine activities. At present, the leading GOFSs (e.g. Mercator a single forecasting simulation in the numerical GOFSs may Ocean Physical System (PSY4) and Real-Time Ocean Fore- take hours on a supercomputer with hundreds of computacast System (RTOFS)) use physics-driven models in fluid tional nodes [2]. Besides, improving the forecasting accuracy mechanics and thermodynamics to predict future ocean of these methods is exceedingly challenging because they motion states and phenomena based on current ocean con- heavily rely on the human cognitive abilities in understandditions [1]. The GOFSs adopt numerical methods that rely on supercomputers to solve the partial differential equa-With the recent advances of Artificial Intelligence (AI) tions of the physical models. Due to their desirable per- techniques, deep learning methods have been widely apformance, they are operationally run in different countries plied in various prediction/forecasting tasks of different worldwide. However, numerical forecasting methods are fields and achieved great success. Particularly, some data-

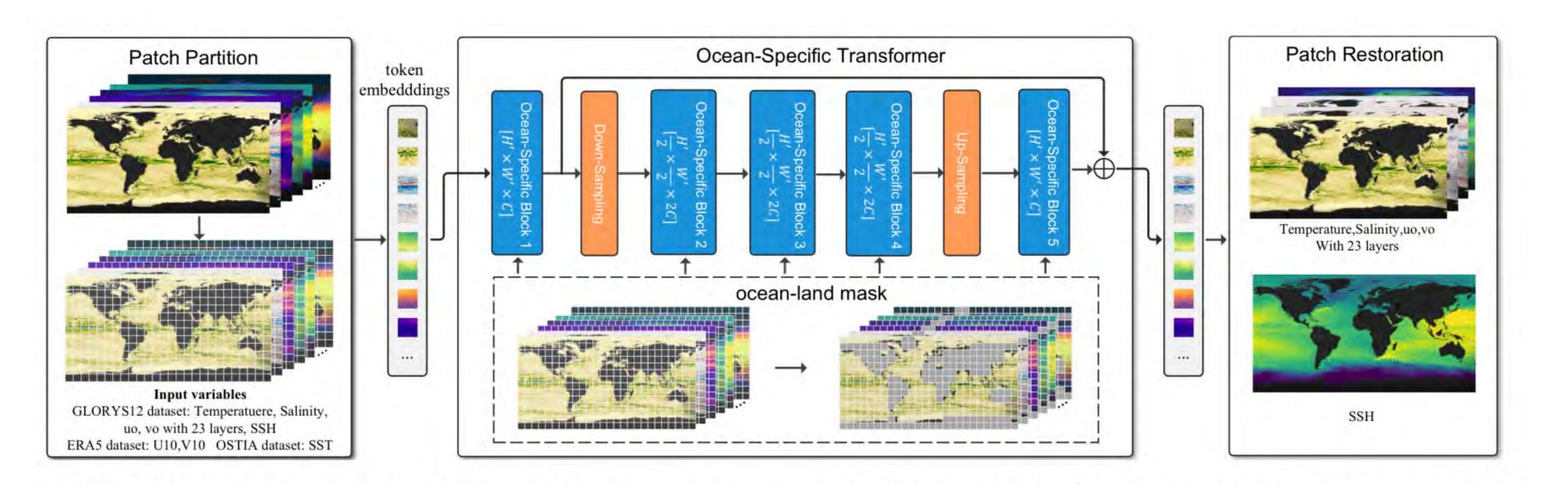
- arX Xiang Wang, Pinqiang Wang, Huizan Wang, Junxing Zhu, Jianbo Xu, Senliang Bao, Ciqiang Luo, Yi Han, Weimin Zhang, Kaijun Ren, Kefeng 410073. Chin.
 - · Renzhi Wang, Senzhang Wang and Jun Yin are with the School of Computer Science and Engineering, Central South University, Changsha
 - Ningzi Hu 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.
 - Guilaua 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,
 - the corresponding authors

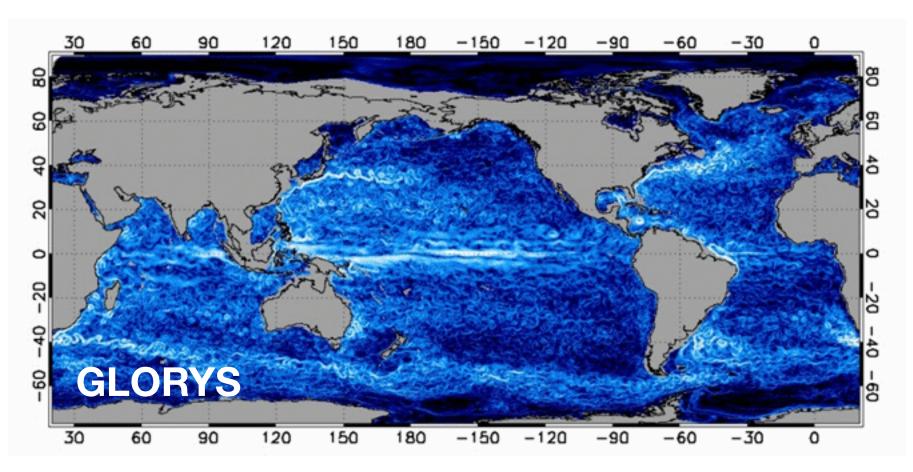
.....

driven AI models have shown the potential in atmosphere weather forecasting like Pangu-Weather [4] and Graph-Cast [5]. They have achieved comparable or even better Deng, and Junqiang Song are with the College of Meteorology and Oceanography, National University of Defense Technology, Changsha prediction results in global medium-range weather fore-(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.

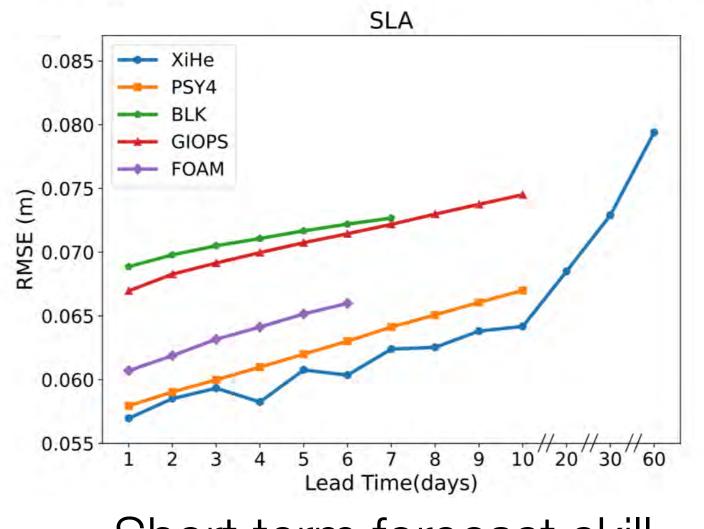
Although data-driven models have achieved promising Beijing 100081, China. Guihua Wang, Huizan Wang, Senzhang Wang, and Weimin Zhang are more accurate and efficient data-driven ocean forecasting results in atmosphere weather forecasting, how to build a model remains an open research issue due to the following

https://arxiv.org/abs/2402.02995 Wang et al. (2024)





Trained from ocean reanalyses



Short term forecast skill



Al-native hybrid climate models

Article

Neural general circulation models for weather and climate

Dmitrii Kochkov^{1,6}, Janni Yuval^{1,6}, Ian Langmore^{1,6}, Peter Norgaard^{1,6}, Jamie Smith^{1,6} https://doi.org/10.1038/s41586-024-07744-y Griffin Mooers¹, Milan Klöwer², James Lottes¹, Stephan Rasp¹, Peter Düben³, Sam Hatfield³, Received: 13 November 2023 Peter Battaglia⁴, Alvaro Sanchez-Gonzalez⁴, Matthew Willson⁴, Michael P. Brenner¹⁵ & Accepted: 15 June 2024 Stephan Hoyer^{1.6} Published online: 22 July 2024 Open access General circulation models (GCMs) are the foundation of weather and climate prediction^{1,2}. GCMs are physics-based simulators that combine a numerical solver Check for updates 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 circula- demonstrating state-of-the-art deterministic forecasts for 1- to 10-day tion models (GCMs) is the basis of weather and climate prediction^{1,2}. weather prediction at a fraction of the computational cost of traditional Over the past 70 years, GCMs have been steadily improved with better models^{3,4}. Machine-learning atmospheric models also require considernumerical methods and more detailed physical models, while exploit- ably less code, for example GraphCast³ has 5,417 lines versus 376,578 ing faster computers to run at higher resolution. Inside GCMs, the lines for the National Oceanic and Atmospheric Administration's FV3 unresolved physical processes such as clouds, radiation and precipi-atmospheric model¹⁵ (see Supplementary Information section A for tation are represented by semi-empirical parameterizations. Tuning details). Nevertheless, machine-learning approaches have noteworthy GCMs to match historical data remains a manual process³, and GCMs retain many persistent errors and biases^{e 8}. The difficulty of reducing limitations compared with GCMs. Existing machine-learning models uncertainty in long-term climate projections⁹ and estimating distribu- have focused on deterministic prediction, and surpass deterministic tions of extreme weather events¹⁰ presents major challenges for climate numerical weather prediction in terms of the aggregate metrics for mitigation and adaptation¹¹. which they are trained^{3,4}. However, they do not produce calibrated Recent advances in machine learning have presented an alter-uncertainty estimates⁴, which is essential for useful weather forecasts¹. native for weather forecasting^{3,4,12,13}. These models rely solely on Deterministic machine-learning models using a mean-squared-error machine-learning techniques, using roughly 40 years of historical loss are rewarded for averaging over uncertainty, producing unrealisdata from the European Center for Medium-Range Weather Forecasts tically blurry predictions when optimized for multi-day forecasts³¹³. (ECMWF) reanalysis v5 (ERA5)¹⁴ for model training and forecast initiali- Unlike physical models, machine-learning models misrepresent derived zation. Machine-learning methods have been remarkably successful, (diagnostic) variables such as geostrophic wind¹⁶. Furthermore Google Research, Mountain View, CA, USA. ³Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. ³European 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. 9 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

1060 | Nature | Vol 632 | 29 August 2024

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

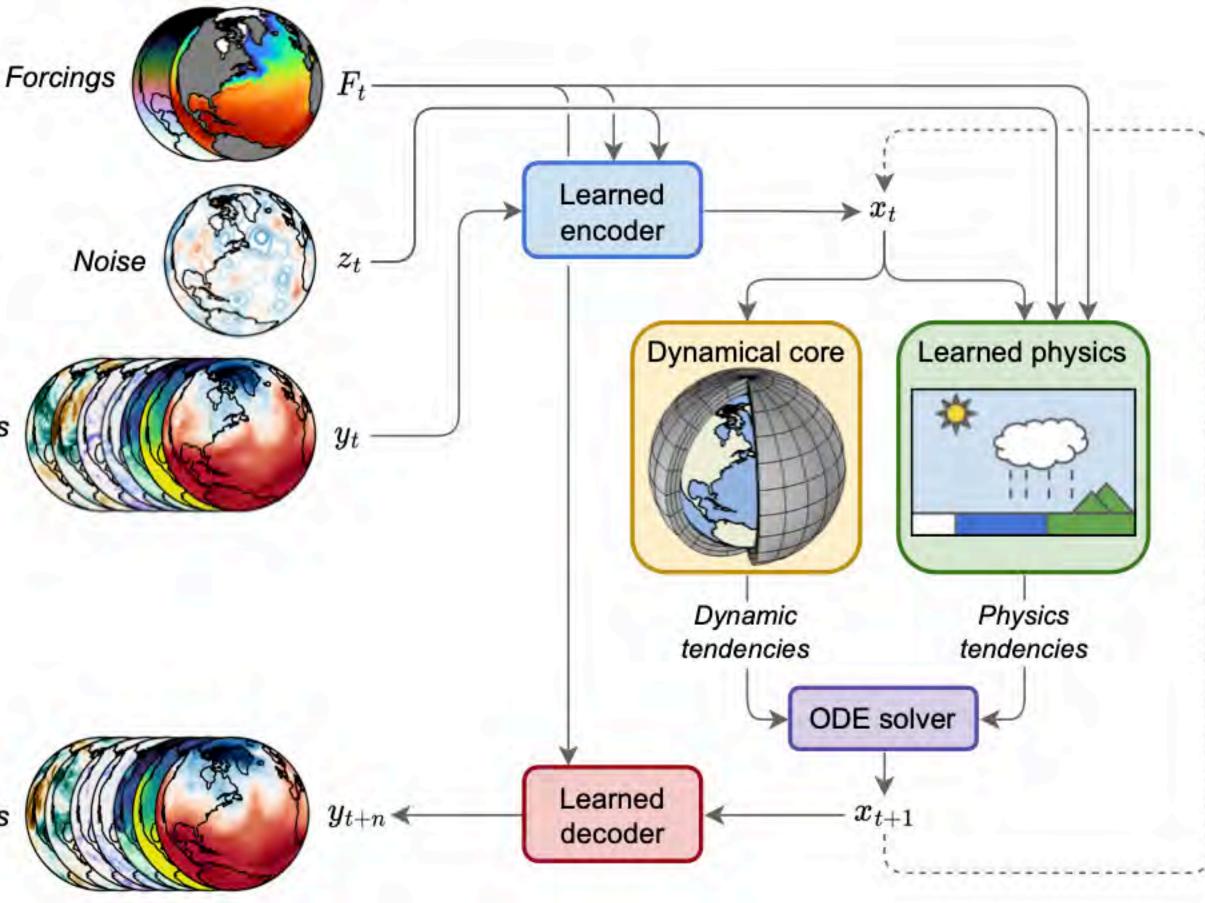
Kochkov et al. (2024)

(a)

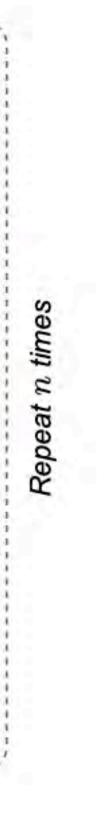
Inputs

Outputs





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

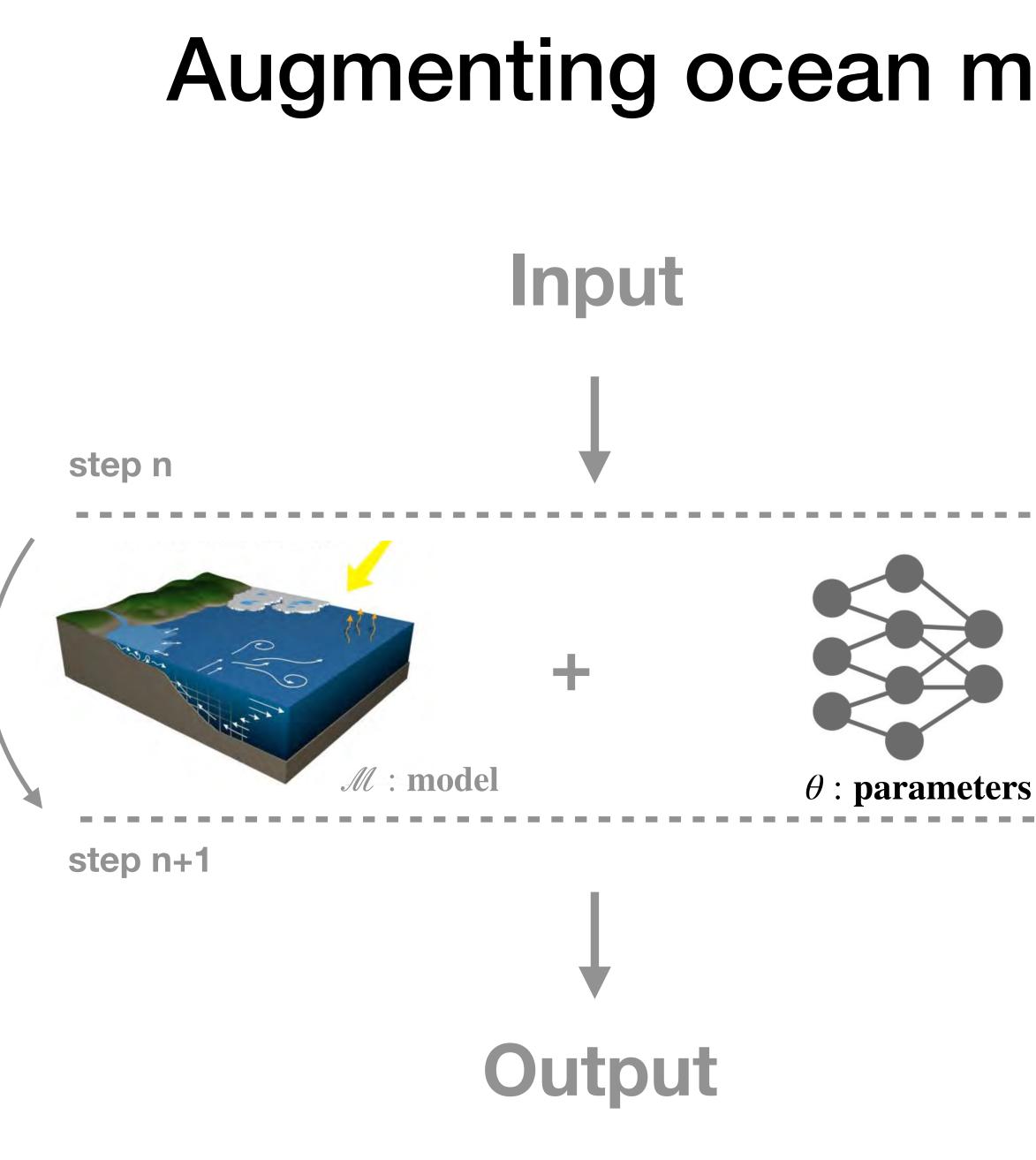






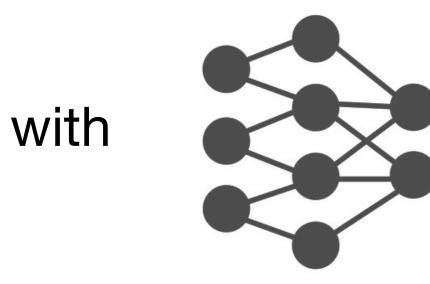
Hybrid models combining physics and ML





The model is augmented with a trainable component

Augmenting ocean models with ML components



 θ : parameters

trained to minimise :

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

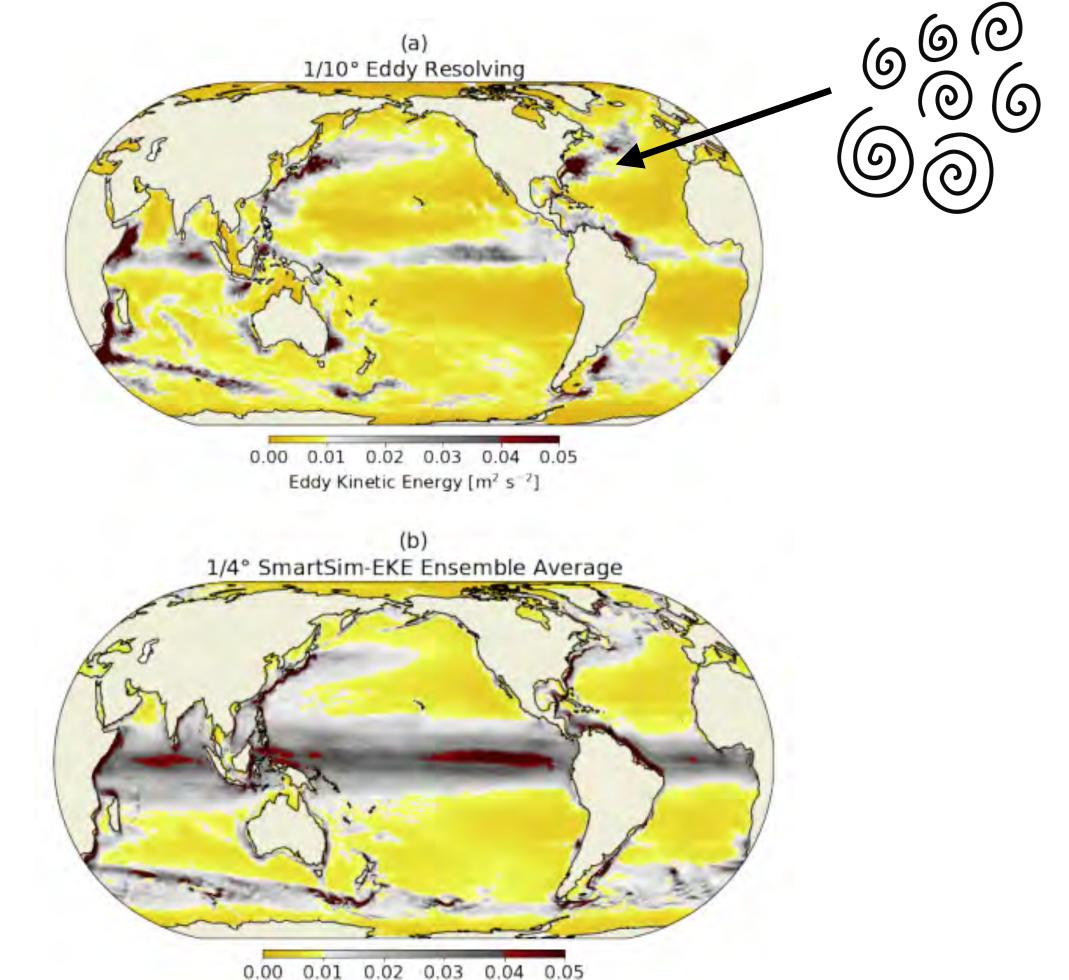
- improving physical consistency
- correcting model errors (vs obs.)
- accelerating execution (x10-100)

NB : does not have to be deterministic





ML for ocean models subgrid physics (1/2)



Eddy Kinetic Energy [m² s⁻²]

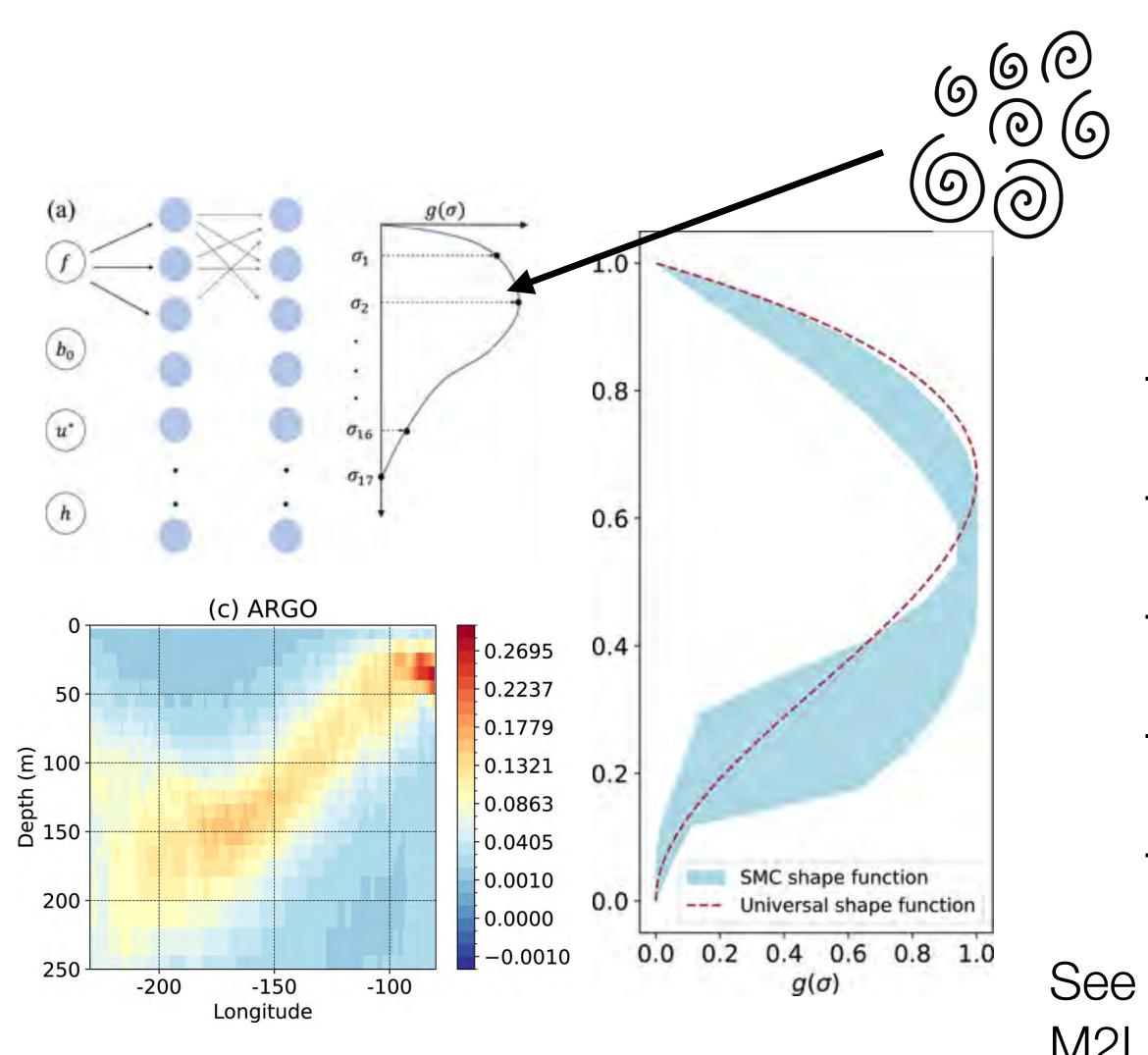
Partee et al. 2022 https://doi.org/10.1016/j.jocs.2022.101707

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)



ML for ocean models subgrid physics (1/2)



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

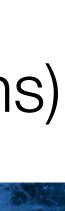
oceanic micro-scale turbulence

- 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)

See for instance : M2LInES consortium

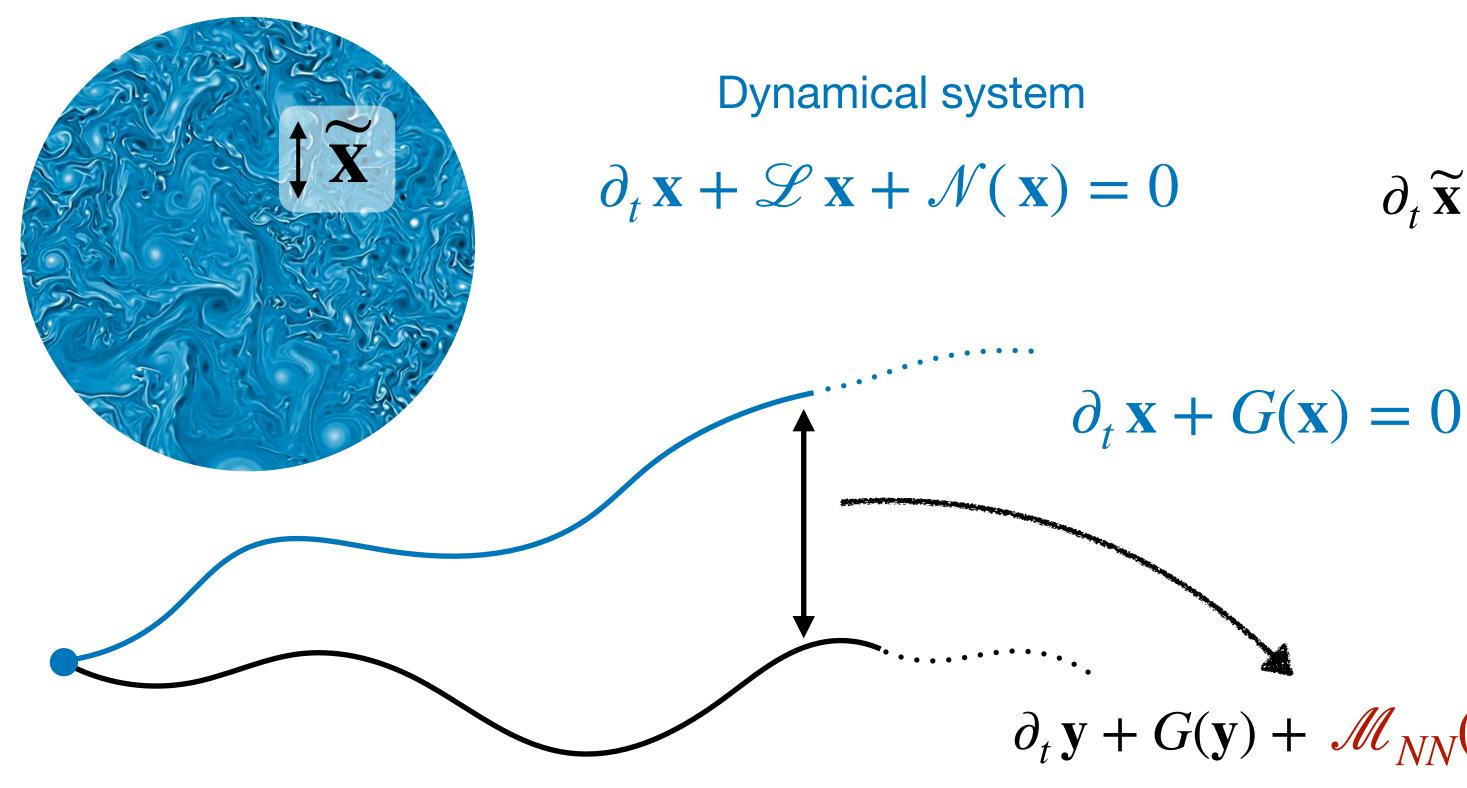
https://m2lines.github.io

M²LINES - Multiscale Machine Learning In Coupled Earth System Modeling





ML for ocean models subgrid physics (2/2)



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

Resolved equations

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

Subgrid closure

 $\mathcal{M}(\widetilde{\mathbf{X}}) \simeq \mathcal{N}(\widetilde{\mathbf{X}}) - \widetilde{\mathcal{N}(\mathbf{X})}$

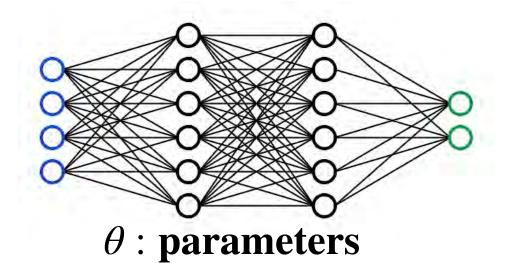
$\partial_t \mathbf{y} + G(\mathbf{y}) + \mathcal{M}_{NN}(\mathbf{y}) = 0$

Frezat et al. (2024) **Gradient-free training**

training model emulator for approx. gradient wrt NN. parameters

Learning the mapping

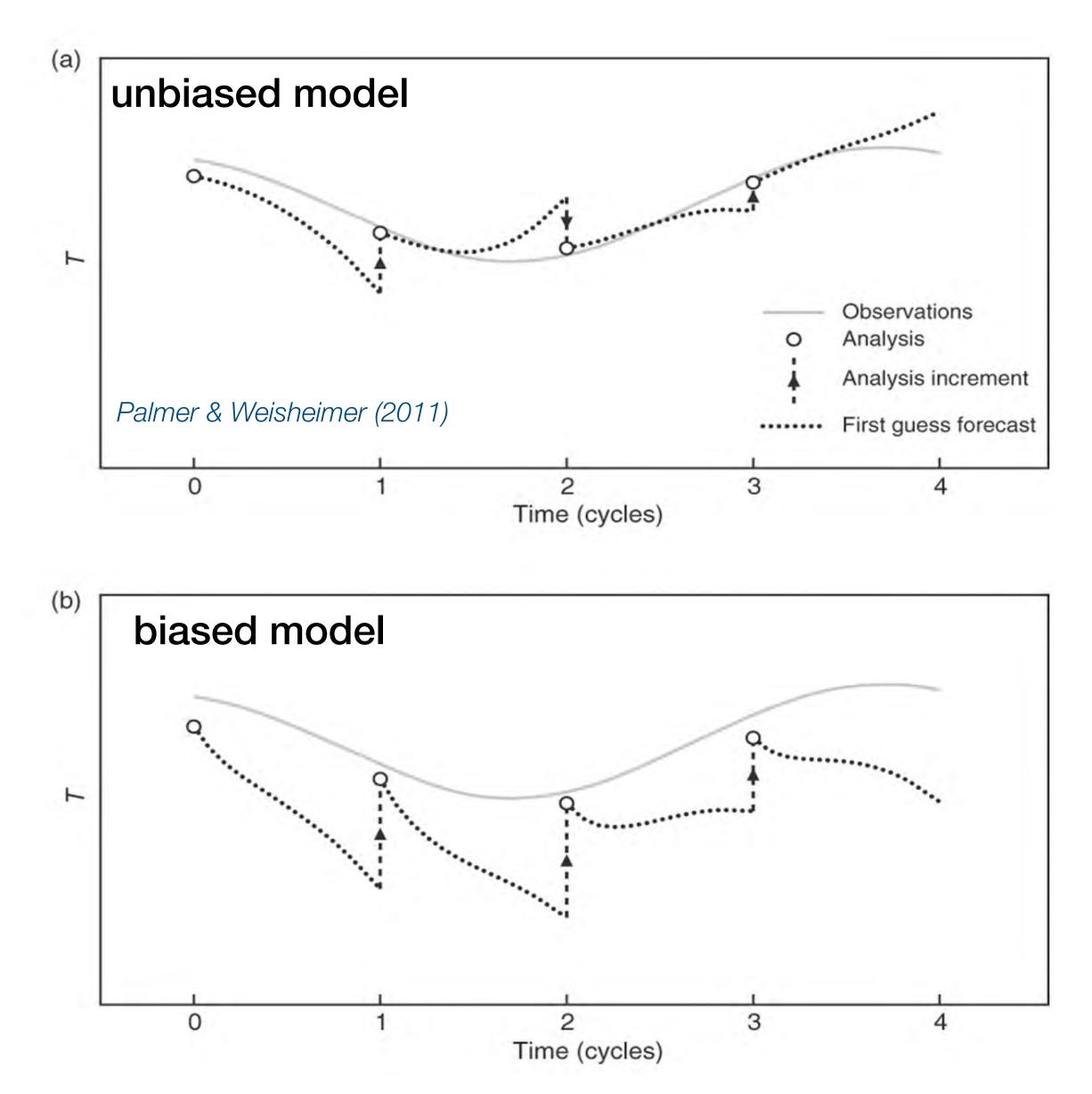
 $\widetilde{\mathbf{x}}(t) \to \mathscr{M}(\widetilde{\mathbf{x}}(t))$

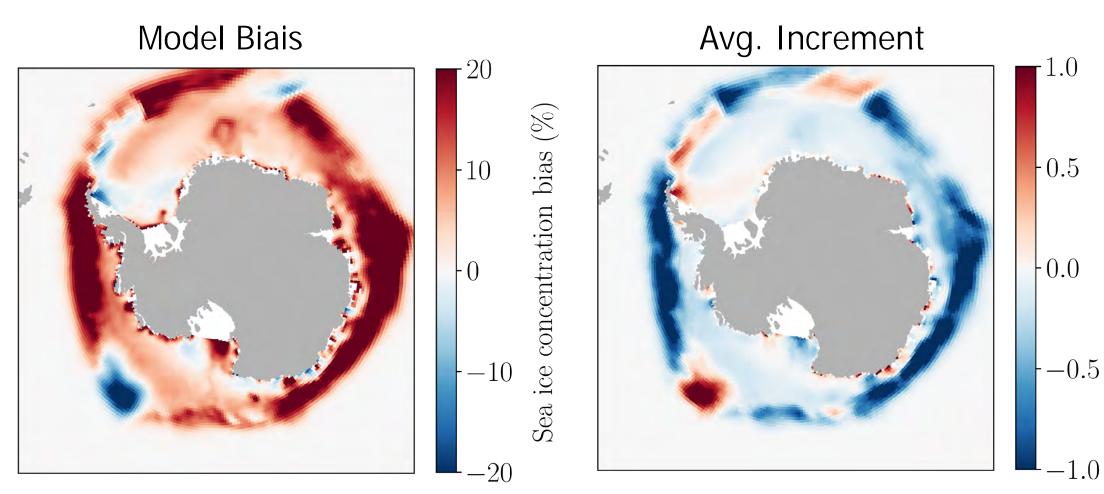


Performance, stability **Generalisation**, interpretability



Learning model error from observations (1/2)





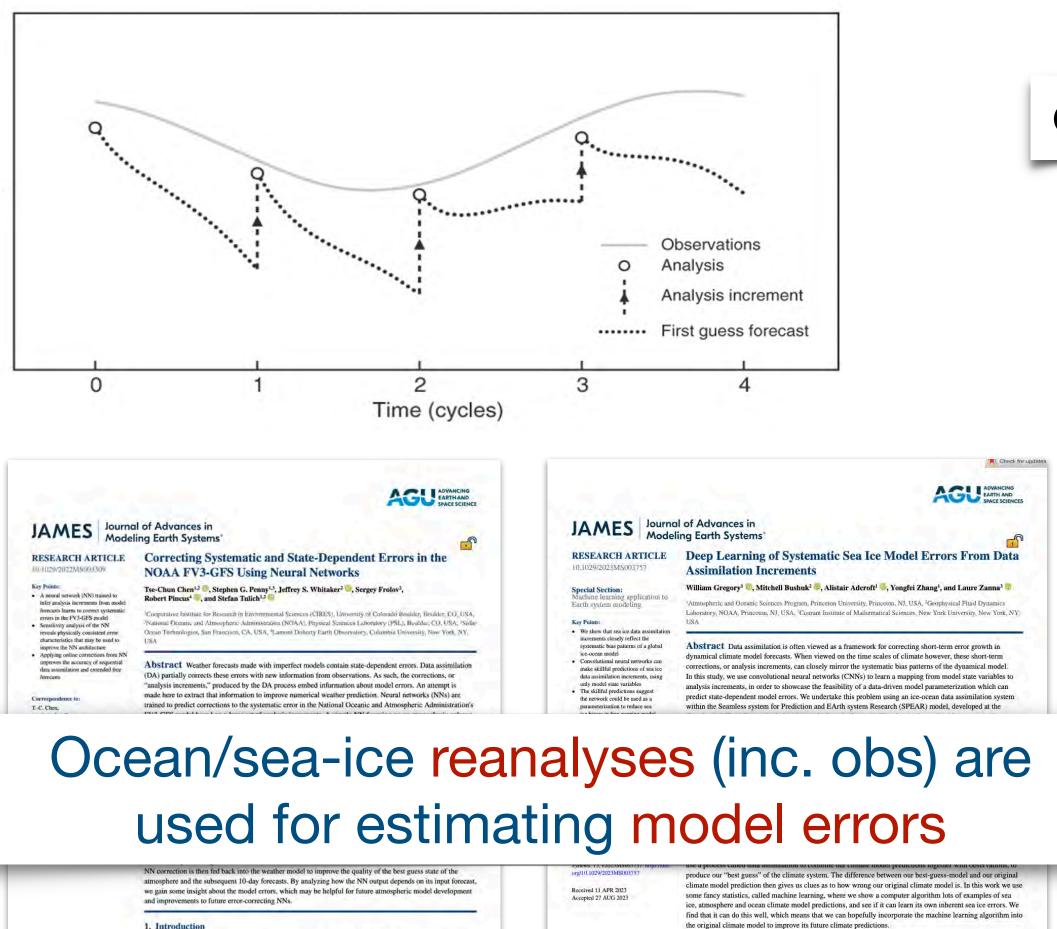
Gregory et al. (2023)

- 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

Analysis increment



Learning model error from observations (2/2)



1. Introduction

© 2023 The Authors, Journal of
Advances in Modeling Earth Systems
published by Wiley Periodicals LLC o
behalf of American Geophysical Union
This is an open access article under
the terms of the Creative Commons
Attribution License, which permits use
distribution and reproduction in any
medium, provided the original work is
properly cited.

subgrid-scale processes, as well as errors in the underlying numerics, leads to systematic biases across the atmosphere, land, sea ice, and ocean. Subsequently, our ability to diagnose and correct these biases ultimately governs

the accuracy of numerical weather and climate predictions on different time scales (Stevens & Bony, 2013). In the context of sea ice for example, much effort has been afforded to the improvement of model physics and subgrid parameterizations through the development of for example, ice thickness distribution (Bitz et al., 2001; Thorndike et al., 1975) and floe-size distribution theory (Horvat & Tziperman, 2015; Rothrock & Thorndike, 1984), surface melt-pond (Flocco et al., 2012), ice drift (Tsamados et al., 2013) and lateral melt parameterizations (M. Smith et al., 2022), as well as sea ice rheology (Dansereau et al., 2016; Hibler, 1979; Ólason et al., 2022). Such studies have shown how the improved representation of sea ice physics produces model simulations which more closely reflect observations in terms of either their mean sea ice volume, drift, or ice thickness distribution. Despite this,

1 of 23

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

ear as forecast progresses, leading to errors that are more difficult to represent.

cantly degrade the forecasting skill.

@ 2022 The Apthony, Jaamal a

properly cited.

CHEN ET AL

oblished by Wiley Periodicals LL

Operational numerical weather prediction (NWP) models are inherently imperfect. Systematic errors result from

approximations in deriving the governing equations, from their numerical implementation, and from conceptus

Even small errors in any component of the NWP model can compound over time to produce errors that signifi-

Systematic errors can be addressed with a wide range of approaches. One approach is to improve the mode

can be improved, say by adopting stochastic parameterizations that account for uncertainty, or by increasing

components-the dynamical core and subgrid scale physics parameterizations. The forecast system as a whole

forecasts can also be further improved by an "offline" post-pr

methods (e.g., Model Output Statistics) or machine learning (ML) methods applied to the model output after the

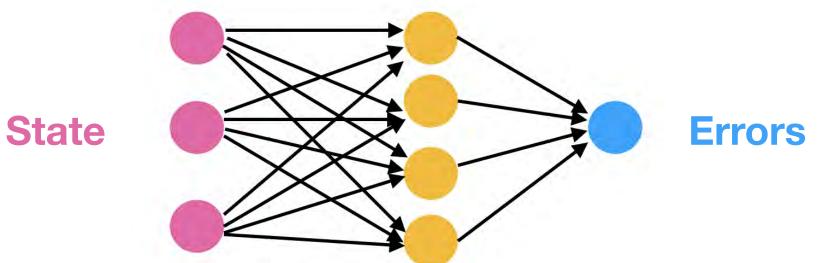
ompletion of model forecast. However, the model errors may be convoluted over time and become more nonlin

aborid scale physical and d

I of 17

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

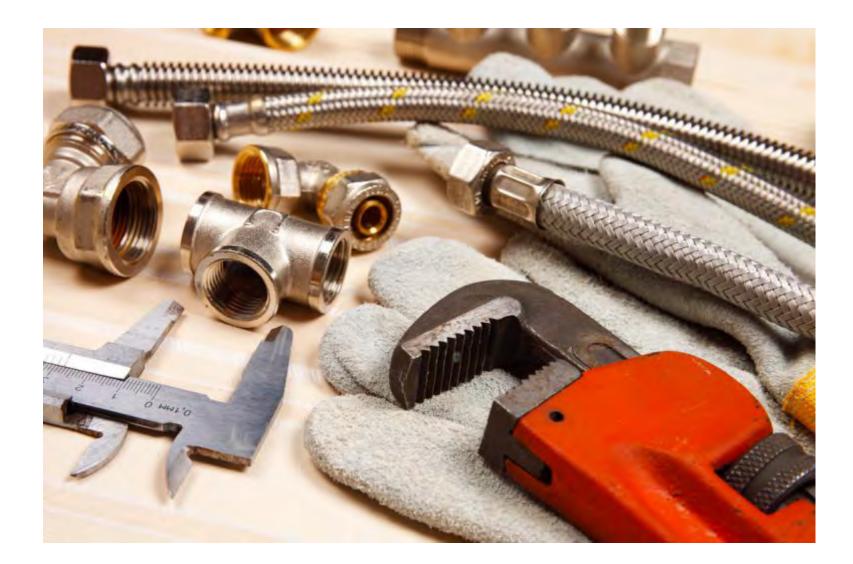
Offline



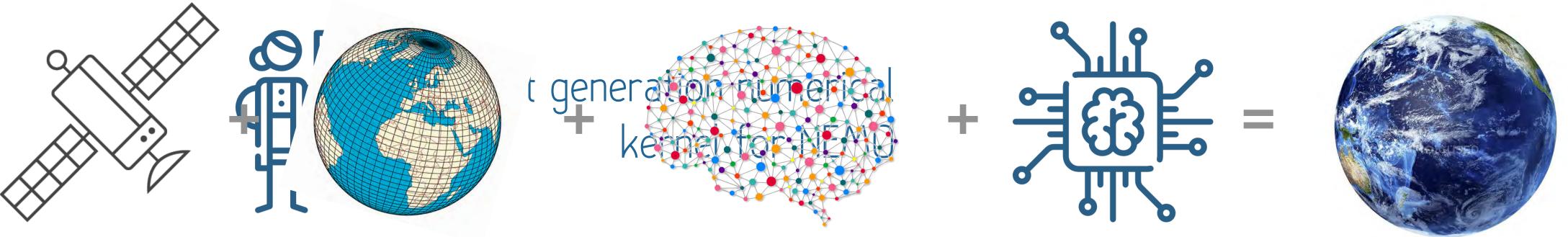
- NN for learning state-dependent biais corrections from analysis increments
- w/ applications in GCMs (atmosphere) and ocean/sea-ice)
- showing success in improving the modeled climate state & forecast skill

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)



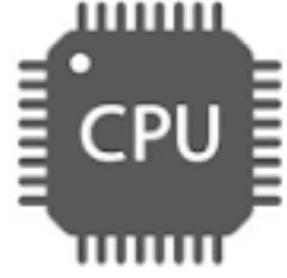
stable, robust, low abstraction languages



high abstraction, fast evolving languages

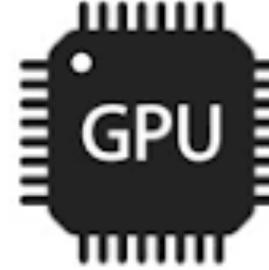


supercomputers



runs only on CPUs



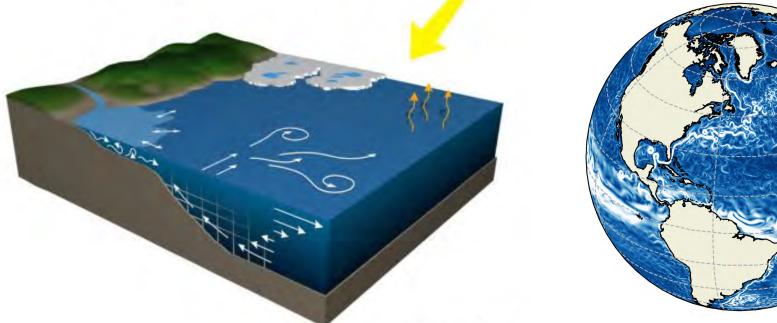


cloud ready

natively runs on GPUs



Interfacing ocean models with DL frameworks (2/3) Input **Ocean circulation models Trainable components** (closures, error corrections) step n PyTorch NEM step n+1





Output

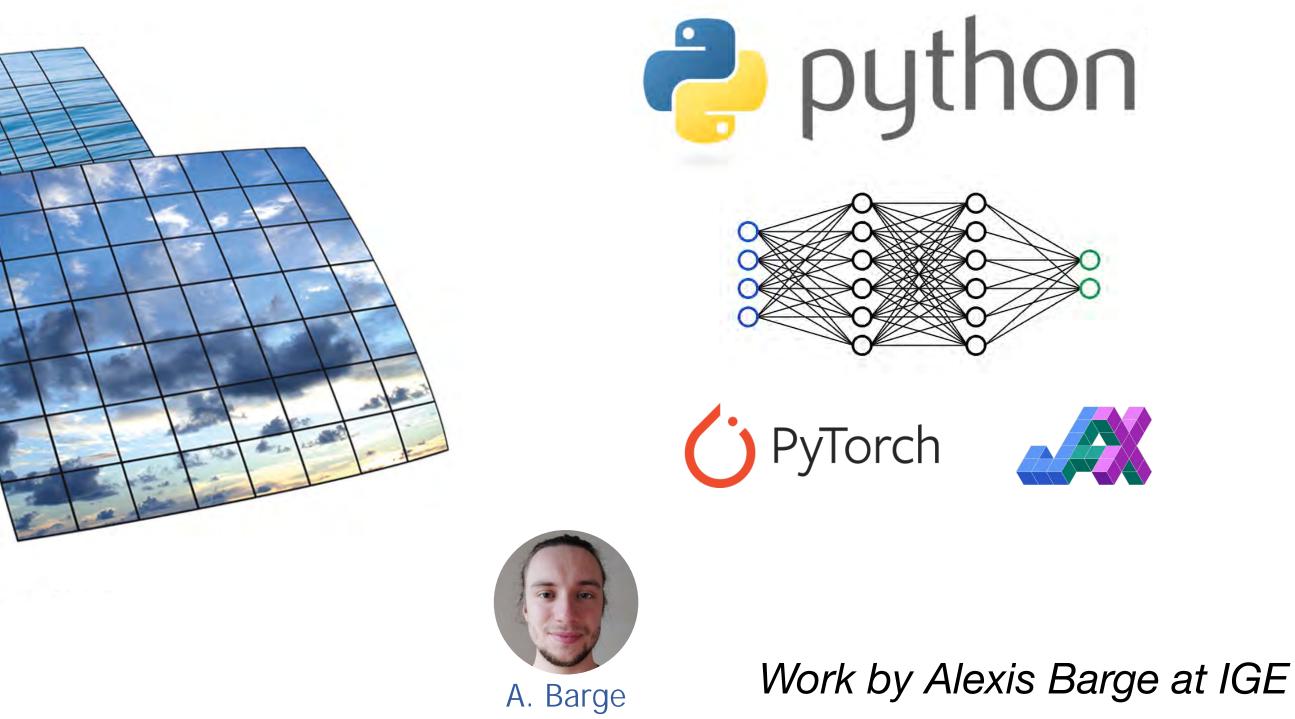




Interfacing ocean models with DL frameworks (3/3)

				The accurate		+ + · · ·	25	0	Da
alexis-barge / eophis	tions 🗄 Projects 🛈 Security	🗠 Insights	Q Type[/] to search	2.	+ - 0	n A	8	C
				⊙ Unwatch 4 •	• • • • • • • • • • • • • • • • • • •	☆ Star 3			
eophis Public	Q. Go to file	1 t	Add file 👻		About	☆ Star 3	*		
cophis Public						Models with Ea			
				<> Code +	About Deploy Inferences	Models with Ea			
Public P main → P 2 Branches ⊙ 0 Tags P alexis-barge fix wrong Logbuffer variab	able name	5b86bl		Code + 3 49 Commits	About Deploy Inferences System simulation ID Readme MIT license	Models with Ea			
 cophis Public P main → P 2 Branches ♥ 0 Tags alexis-barge fix wrong Logbuffer variab docs 	able name add docs	5b86bl		Code + 3 months ago	About Deploy Inferences System simulation	Models with Ea			
 eophis Public P main + P 2 Branches O Tags alexis-barge fix wrong Logbuffer variab docs src tests 	able name add docs fix wrong Logbuffer variable	5b86bl		Code - O 49 Commits 3 months ago last month	About Deploy Inferences System simulation □ Readme Φ MIT license 小 Activity ☆ 3 stars ③ 4 watching	Models with Ea			_
 eophis Public P main P main P 2 Branches O Tags alexis-barge fix wrong Logbuffer variab docs src tests gitignore 	able name add docs fix wrong Logbuffer variable update info for merging	5b86bl		Code + 3 49 Commits 3 months ago last month last month	About Deploy Inferences System simulation □ Readme Φ MIT license 小 Activity ☆ 3 stars ③ 4 watching ♀ 0 forks	Models with Ea			_
 eophis Public P main r Public P main r Public P alexis-barge fix wrong Logbuffer variabe docs src tests jeitignore LICENSE 	able name add docs fix wrong Logbuffer variable update info for merging update gitignore	5b86bl		Code + 3 49 Commits 3 months ago last month last month 3 months ago	About Deploy Inferences System simulation □ Readme Φ MIT license 小 Activity ☆ 3 stars ③ 4 watching	Models with Ea			_
 eophis Public P main P 2 Branches O Tags alexis-barge fix wrong Logbuffer variab docs src 	able name add docs fix wrong Logbuffer variable update info for merging update gitignore add license	5b86bl		Code + 49 Commits 3 months ago last month last month 3 months ago 2 months ago	About Deploy Inferences System simulation 대 Readme 과 MIT license 가 Activity ☆ 3 stars ◇ 4 watching ♀ 0 forks Report repository Releases	Models with Ea			-
 eophis Public P main ~ P 2 Branches O Tags alexis-barge fix wrong Logbuffer variab docs src tests .gitignore LICENSE README.md 	able name add docs fix wrong Logbuffer variable update info for merging update gitignore add license update info for merging	5b86bl		Code + 49 Commits 3 months ago last month last month 3 months ago 2 months ago last month	About Deploy Inferences System simulation D Readme MIT license 사 Activity ☆ 3 stars ③ 4 watching 앟 0 forks Report repository	Models with Ea			_
 eophis Public P main Public P main Public P main Public P alexis-barge fix wrong Logbuffer variable docs src tests .gitignore LICENSE README.md pyproject.toml 	able name add docs fix wrong Logbuffer variable update info for merging update gitignore add license update info for merging	5b86bl		Code • 3 49 Commits 3 months ago last month 3 months ago 2 months ago 2 months ago last month 2 months ago last month 2 months ago	About Deploy Inferences System simulation □ Readme Φ MIT license 小 Activity ☆ 3 stars ③ 4 watching ♀ 0 forks Report repository Releases No releases published	Models with Ea			-
eophis Public P main P 2 Branches P main P 2 Branches alexis-barge fix wrong Logbuffer variab docs docs src tests .gitignore LICENSE README.md pyproject.toml	able name add docs fix wrong Logbuffer variable update info for merging update gitignore add license update info for merging	5b86bl		Code • 3 49 Commits 3 months ago last month 3 months ago 2 months ago 2 months ago last month 2 months ago last month 2 months ago	About Deploy Inferences System simulation Deploy Inferences System simulation Deploy Inferences Market MIT license 사 Activity 다 3 stars Co 4 watching 양 0 forks Report repository Releases No releases published Create a new release	Models with Ea codes			-

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



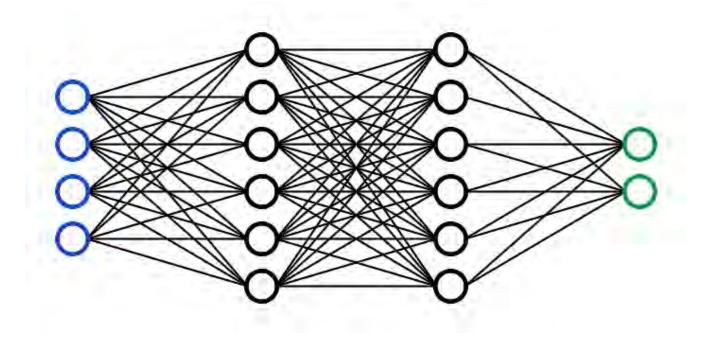
- ASIS : exchange of 3D data between different codes
- phis : simplified deployment of ML models w/ OASIS
- equires some change to the NEMO code
- Key : portability, domain decomposition





The challenge of online training strategies (1/2)online learning

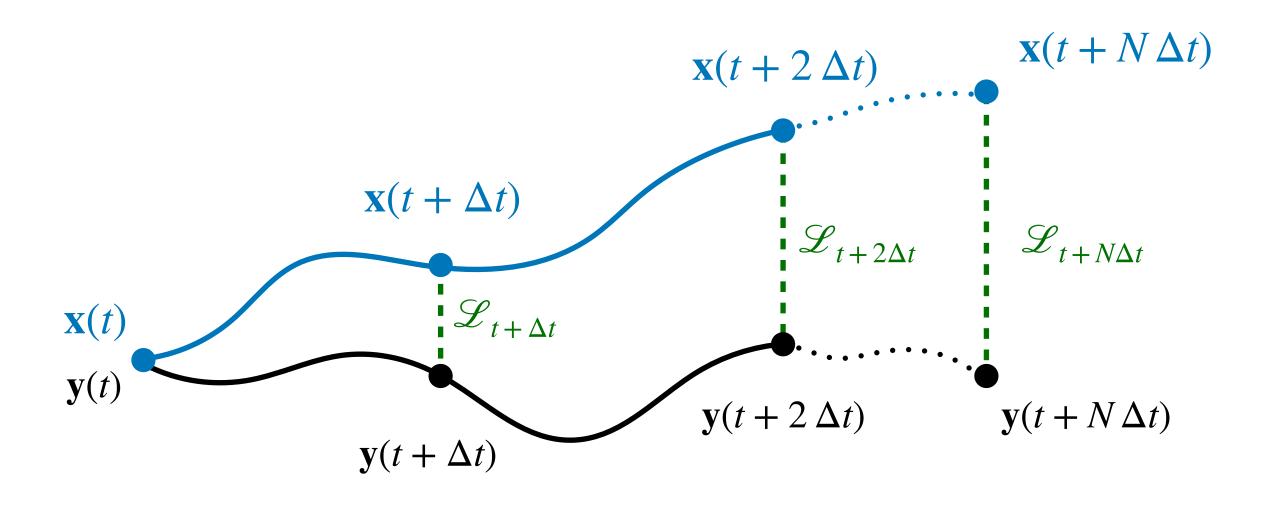
offline learning





from pre-existing data

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



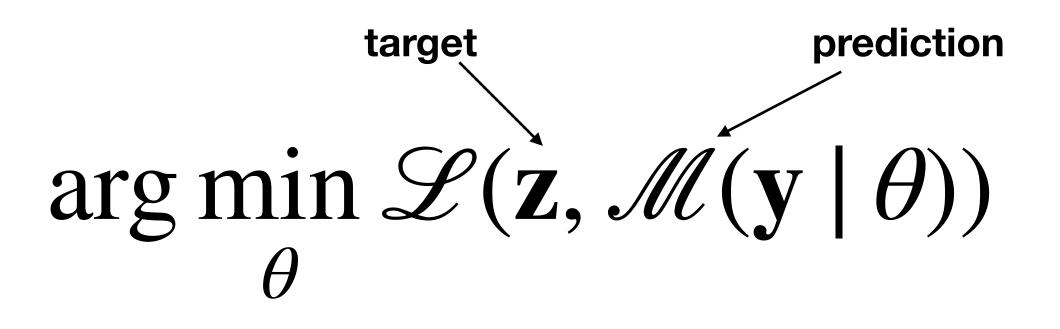
 $\partial_t \mathbf{y} + G(\mathbf{y}) + \mathcal{M}_{NN}(\mathbf{y}) = f$

along a trajectory

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



The challenge of online training strategies (2/2)



For time evolving problems, with

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

The gradient of the loss involves

$$\frac{\partial \mathscr{M}}{\partial \theta} \stackrel{}{=} \frac{\partial E}{\partial \theta} \stackrel{}{=} \frac{\partial (E_m \circ \cdots \circ E_1)}{\partial \theta}$$

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

gradient of the loss

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

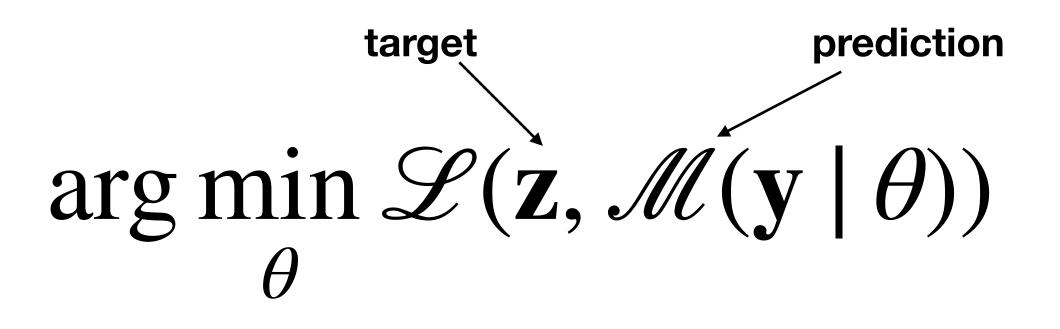
> tricky without Automatic Differenciation (AD) !

 $\partial E_m \quad \partial E_2 \ \partial E_1$ $\frac{\partial E_{m-1}}{\partial E_1} \frac{\partial E_1}{\partial \theta}$





The challenge of online training strategies (2/2)



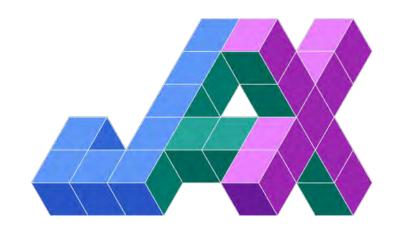
For time evolving problems, with

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

The gradient of the loss involves

$$\frac{\partial \mathscr{M}}{\partial \theta} \stackrel{}{=} \frac{\partial E}{\partial \theta} \stackrel{}{=} \frac{\partial (E_m \circ \cdots \circ E_1)}{\partial \theta}$$

AD is readily available in some language





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

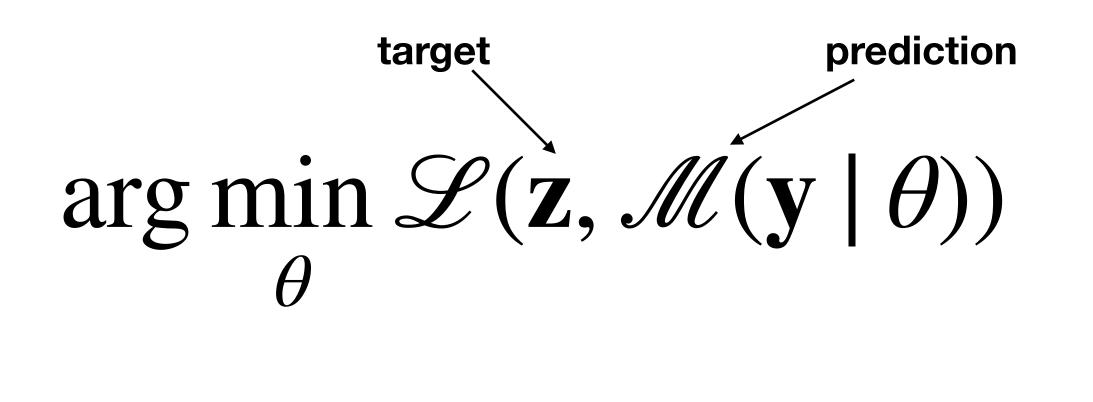
> tricky without Automatic Differenciation (AD) !

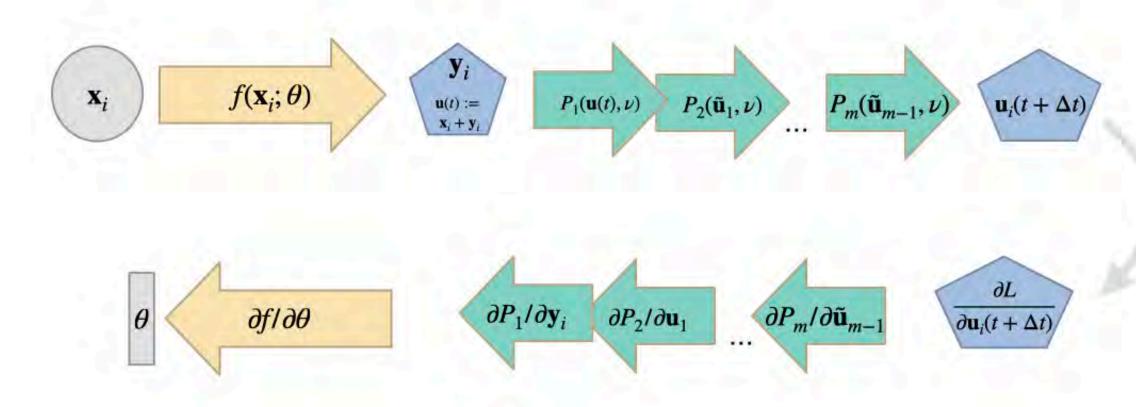
 $\partial E_m \quad \partial E_2 \ \partial E_1$ $\frac{\partial E_{m-1}}{\partial E_{1}} \frac{\partial E_{1}}{\partial \theta}$





The challenge of online training strategies (2/2)

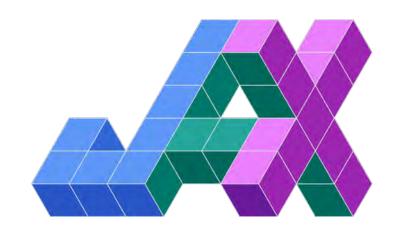




See eg Thuerey et al. 2021

https://arxiv.org/abs/2109.05237

AD is readily available in some language





Supervised or residual loss L



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











Al-native hybrid geoscientific models

Article

Neural general circulation models for weather and climate

Dmitrii Kochkov^{1,6}, Janni Yuval^{1,6}, Ian Langmore^{1,6}, Peter Norgaard^{1,6}, Jamie Smith^{1,6} https://doi.org/10.1038/s41586-024-07744-y Griffin Mooers¹, Milan Klöwer², James Lottes¹, Stephan Rasp¹, Peter Düben³, Sam Hatfield³, Received: 13 November 2023 Peter Battaglia⁴, Alvaro Sanchez-Gonzalez⁴, Matthew Willson⁴, Michael P. Brenner¹⁵ & Stephan Hoyer^{1.6} Accepted: 15 June 2024 Published online: 22 July 2024 Open access General circulation models (GCMs) are the foundation of weather and climate prediction^{1,2}. GCMs are physics-based simulators that combine a numerical solver Check for updates 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 circula- demonstrating state-of-the-art deterministic forecasts for 1- to 10-day weather prediction at a fraction of the computational cost of traditional tion models (GCMs) is the basis of weather and climate prediction^{1,2}. Over the past 70 years, GCMs have been steadily improved with better models^{3,4}. Machine-learning atmospheric models also require considernumerical methods and more detailed physical models, while exploit- ably less code, for example GraphCast³ has 5,417 lines versus 376,578 ing faster computers to run at higher resolution. Inside GCMs, the lines for the National Oceanic and Atmospheric Administration's FV3 unresolved physical processes such as clouds, radiation and precipi-atmospheric model¹⁵ (see Supplementary Information section A for tation are represented by semi-empirical parameterizations. Tuning details). Nevertheless, machine-learning approaches have noteworthy GCMs to match historical data remains a manual process³, and GCMs retain many persistent errors and biases^{e 8}. The difficulty of reducing limitations compared with GCMs. Existing machine-learning models uncertainty in long-term climate projections⁹ and estimating distribu- have focused on deterministic prediction, and surpass deterministic tions of extreme weather events¹⁰ presents major challenges for climate numerical weather prediction in terms of the aggregate metrics for mitigation and adaptation¹¹. which they are trained^{3,4}. However, they do not produce calibrated Recent advances in machine learning have presented an alter-uncertainty estimates⁴, which is essential for useful weather forecasts¹. native for weather forecasting^{3,4,12,13}. These models rely solely on Deterministic machine-learning models using a mean-squared-error machine-learning techniques, using roughly 40 years of historical loss are rewarded for averaging over uncertainty, producing unrealisdata from the European Center for Medium-Range Weather Forecasts tically blurry predictions when optimized for multi-day forecasts³¹³. (ECMWF) reanalysis v5 (ERA5)¹⁴ for model training and forecast initiali- Unlike physical models, machine-learning models misrepresent derived zation. Machine-learning methods have been remarkably successful, (diagnostic) variables such as geostrophic wind¹⁶. Furthermore Google Research, Mountain View, CA, USA. ³Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. ³European 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. 9 These authors contributed equally Dmitrii Kochkov, Janni Yuval. Ian Langmore, Peter Norgaard, Jamie Smith, Stephan Hoyer. 🤤 -mail: dkochkov@google.com; janniyuval@google.com; shoyer@google.com

1060 | Nature | Vol 632 | 29 August 2024

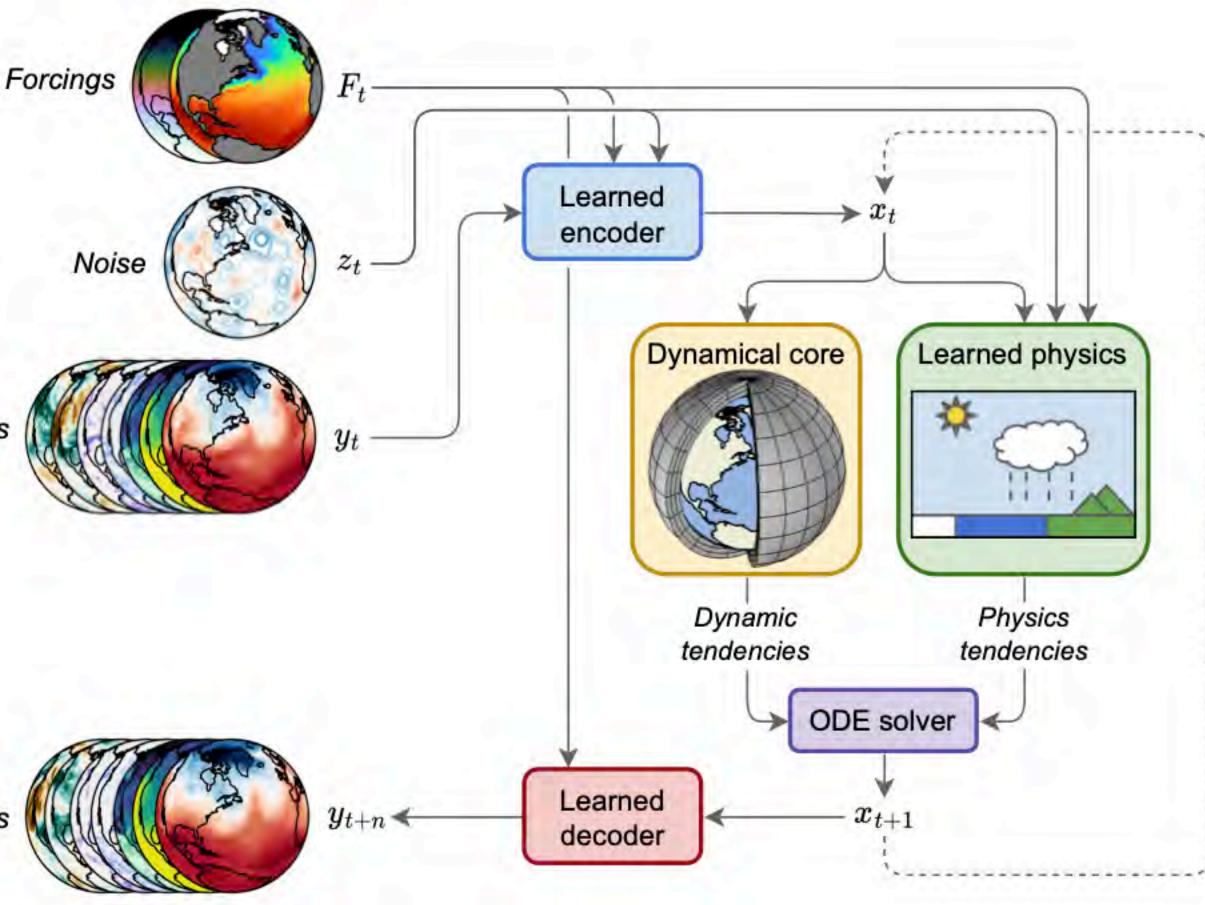
https://doi.org/10.1038/s41586-024-07744-y

Kochkov et al. (2024)

Inputs

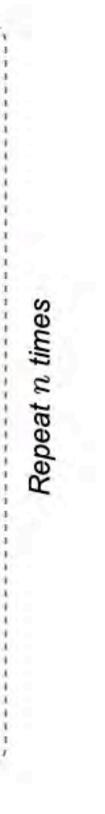
(a)

Outputs

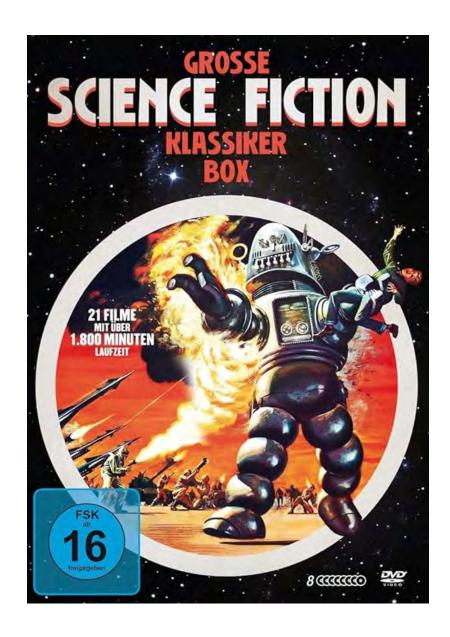




https://github.com/google-research/dinosaur https://github.com/google-research/neuralgcm







Towards Al-native hybrid geoscientific models ?







Al-native hybrid geoscientific models ?

Earth System Observation Data

Ground truth for the validation of process-based models

Physical Equation-driven Earth and Climate Modelling

Main tool for quantifying the Earth's state under ongoing anthropogenic forcing

Contains persistent error sources

Process-based models and neural networks will be coupled as actively learning hybrid models

Irrgang et al. (2021)

Successive research on explainable AI will make hybrid models more physically interpretable Combining the advantages of process-based with machine learning models

Neural Earth System Modelling

Available data pool for neural network training environments

> Earth Data-driven **Machine Learning**

Highly specialized agents that uncover hidden patterns and geophysical quantities

Lack of process knowledge

Hybrid models start to outperform the predictive power of traditional models

Betting harnessing observations & hi-fidelity simulations

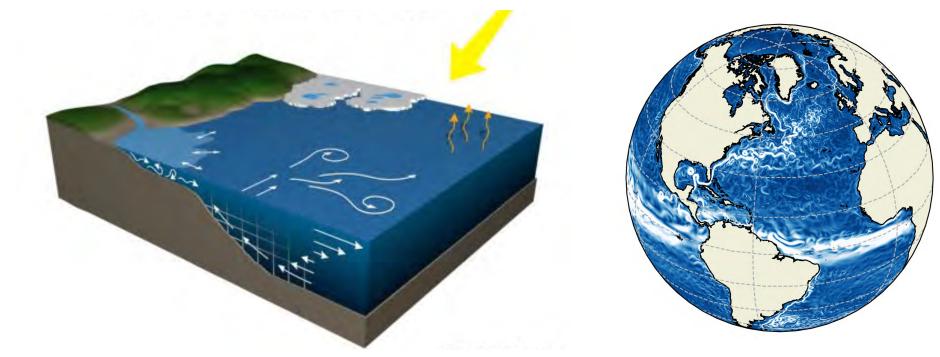
... for optimising

- model parameters
- numerical schemes
- subgrid closures

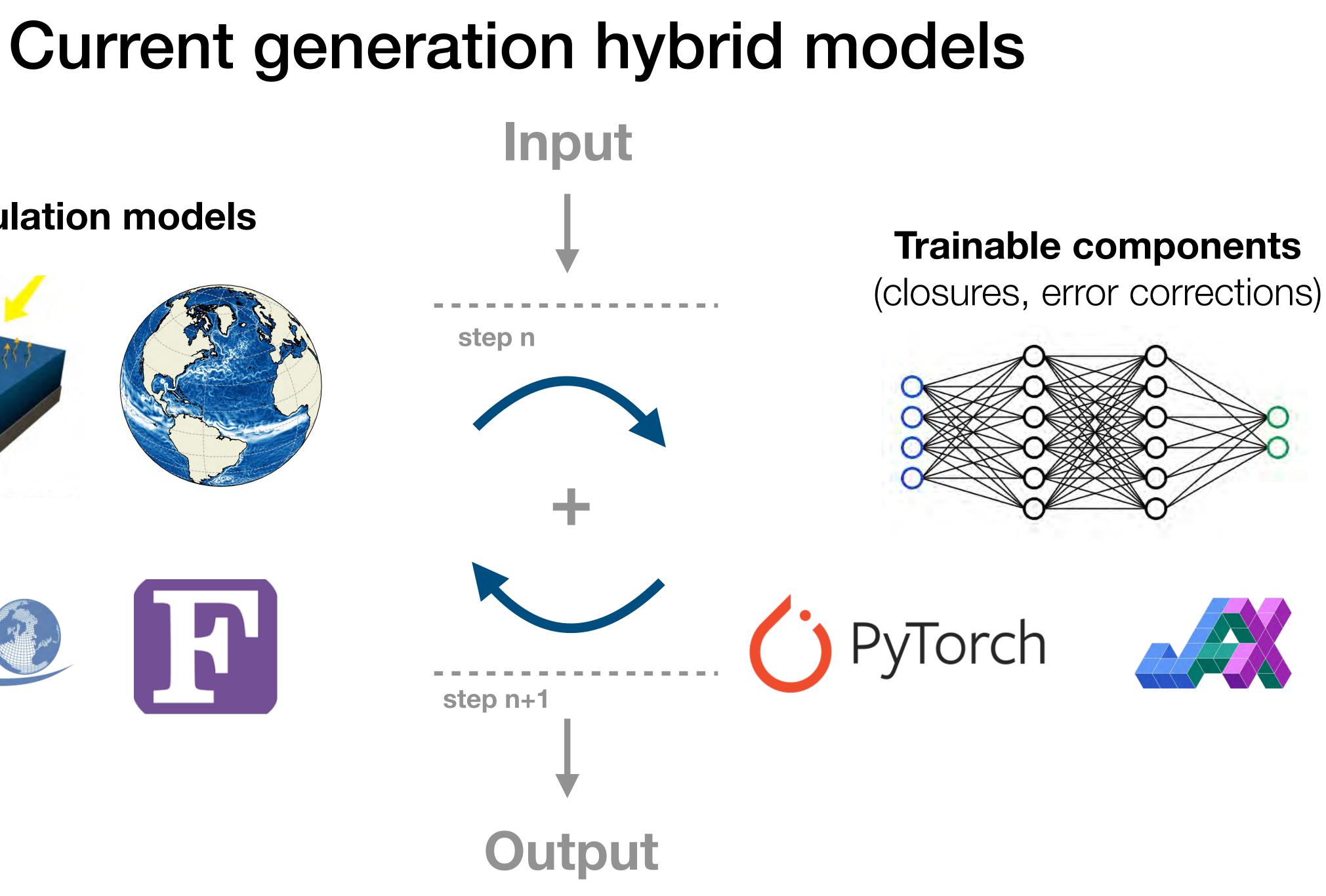
Differentiable programming in earth system models?



Ocean circulation models













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

Veros 1.5.1+51.g4039f76.dirty documentation

veros.readthedocs.io/en/latest/

12

Versatile Ocean Simulation in Pure Python

Veros, the versatile ocean simulator, aims to be the swiss army knife of ocean modeling. It is a fullfledged 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



••• 🖬 - <

=

0

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 mode
- Running Veros
- Enhancing Veros
- Advanced installation
- Using JAX

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

Atmos

Ocean



:0: I

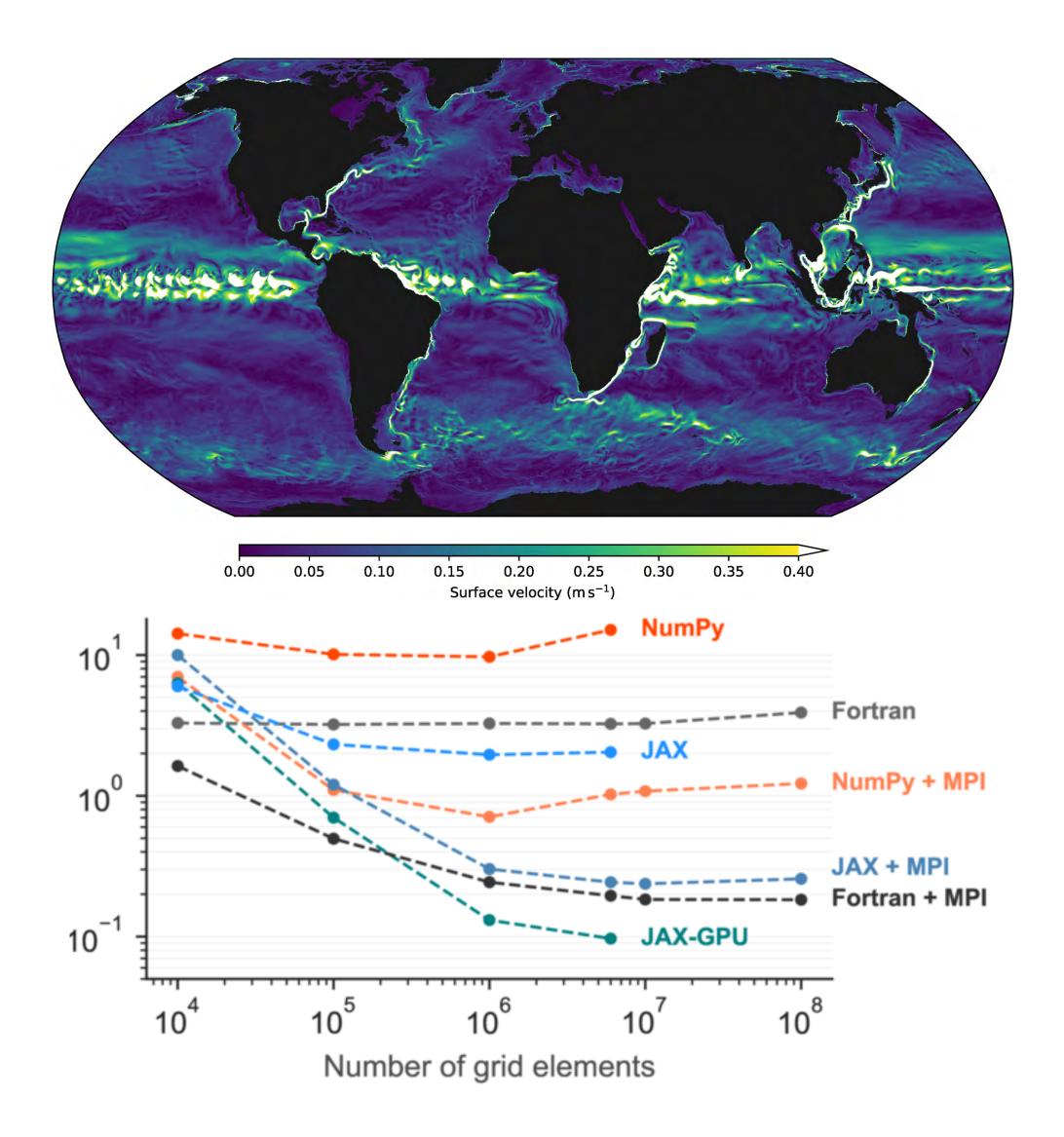
Home							
main · ド9 Branches 0			About No description, website,				
shoyer and Dinosaur aut	hors Conser 🚥 🗸 a7630f3+14 hours a	go 🕲 35 Commits	provided.				
.github/workflows	[dinosaur] add support for Python	3.10 5 months ago	m Readme				
dinosaur	Conservative vertical regridding.	14 hours ago	क Apache-2.0 license				
notebooks	Added Held-Suarez notebook.	5 months ago	- Activity				
gitignore	Initial export of Dinosaur	5 months ago	E Custom properties				
CONTRIBUTING.md	Initial export of Dinosaur	5 months ago	 ☆ 137 stars ⊙ 6 watching 				
LICENSE							
3 README.md	ಳ 8 forks Report repository						
conftest.py	ftest.py Initial export of Dinosaur 5 months ag						
dinosaur-logo.png	saur-logo.png Initial export of Dinosaur 5 months a						
) pyproject.toml	2 months ago	Release dinosaur 1.0. on May 25					
다. README 전 Apache-2.0 licens	se	0 i=	on may 20				
		Packages					
			No packages published				
			Contributors 5				
	F		🌍 🚁 εη 🌏				
			Languages				
			 Jupyter Notebook 89.19 Python 10.9% 				





A new generation of geoscientific models

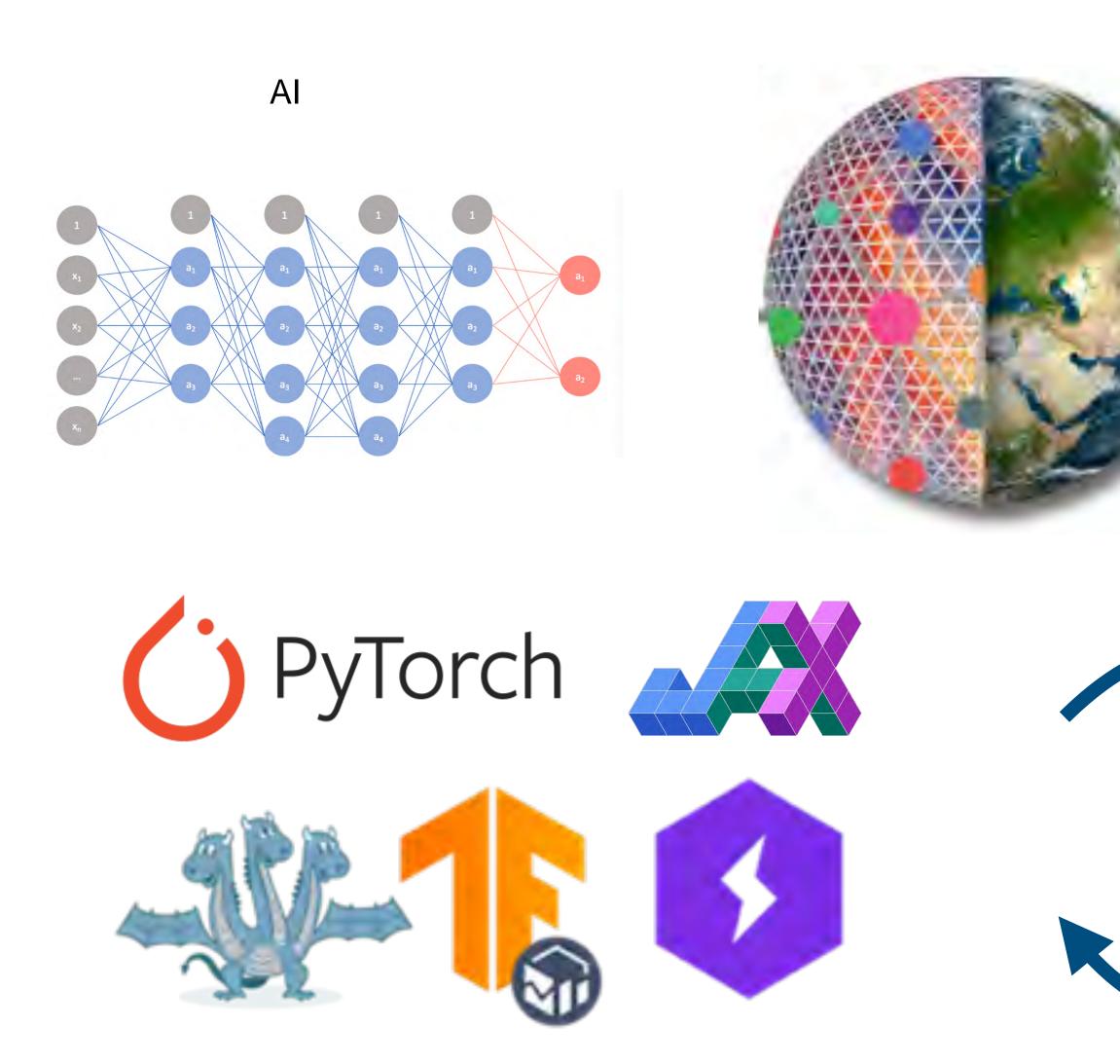
•••• 🖬 - 🤇 ④ 山 + 铂 0 veros.readthedocs.io/en/latest/ r. :0: I = Veros 1.5.1+51.g4039f76.dirty documentation Versatile Ocean Simulation in Pure Python Veros, the versatile ocean simulator, aims to be the swiss army knife of ocean modeling. It is a fullfledged primitive equation ocean model that supports 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 putline 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 mode Running Veros Enhancing Veros Advanced installation Using JAX



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



Allowing a seamless integration with Al



high abstraction, fast evolving languages

Physics $\frac{\mathrm{D}u}{\mathrm{D}t} = \frac{uv\tan\phi}{r} - \frac{uw}{r} + fv - f'w - \frac{c_p\theta}{r\cos\phi}\frac{\partial\Pi}{\partial\lambda} + \mathrm{D}(u),$ $\frac{\mathrm{D}v}{\mathrm{D}t} = -\frac{u^2 \tan \phi}{r} - \frac{vw}{r} - uf - \frac{c_p \theta}{r} \frac{\partial \Pi}{\partial \phi} + \mathrm{D}(v),$ $\delta \frac{\mathrm{D} w}{\mathrm{D} t} = \frac{u^2 + v^2}{r} + u f' - g(r) - c_p \theta \frac{\partial \Pi}{\partial r},$ PyTorch

high abstraction, fast evolving languages



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 Mooers¹, Milan Klöwer², James Lottes¹, Stephan Rasp¹, Peter Düben³, Sam Hatfield³, Peter Battaglia⁴, Alvaro Sanchez-Gonzalez⁴, Matthew Willson⁴, 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 circula- demonstrating state-of-the-art deterministic forecasts for 1- to 10-day tion 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 exploit- ably less code, for example GraphCast³ has 5,417 lines versus 376,578 ing faster computers to run at higher resolution. Inside GCMs, the lines for the National Oceanic and Atmospheric Administration's FV3 unresolved physical processes such as clouds, radiation and precipi-atmospheric model¹⁵ (see Supplementary Information section A for tation are represented by semi-empirical parameterizations. Tuning details). GCMs to match historical data remains a manual process³, and GCMs retain many persistent errors and biases^{e 8}. The difficulty of reducing limitations compared with GCMs. Existing machine-learning models uncertainty in long-term climate projections⁹ and estimating distribu- have focused on deterministic prediction, and surpass deterministic tions of extreme weather events¹⁰ presents major challenges for climate numerical weather prediction in terms of the aggregate metrics for mitigation and adaptation¹¹.

native for weather forecasting^{3,4,12,13}. These models rely solely on Deterministic machine-learning models using a mean-squared-error machine-learning techniques, using roughly 40 years of historical loss are rewarded for averaging over uncertainty, producing unrealisdata from the European Center for Medium-Range Weather Forecasts tically blurry predictions when optimized for multi-day forecasts³¹³. (ECMWF) reanalysis v5 (ERA5)¹⁴ for model training and forecast initiali- Unlike physical models, machine-learning models misrepresent derived zation. Machine-learning methods have been remarkably successful, (diagnostic) variables such as geostrophic wind¹⁶. Furthermore

weather prediction at a fraction of the computational cost of traditional models^{3,4}. Machine-learning atmospheric models also require consider-

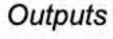
Nevertheless, machine-learning approaches have noteworthy which they are trained^{3,4}. However, they do not produce calibrated Recent advances in machine learning have presented an alter-uncertainty estimates⁴, which is essential for useful weather forecasts¹.

Google Research, Mountain View, CA, USA. ³Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. ³European 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. 9 These authors contributed equally Dmitrii Kochkov, Janni Yuval. Ian Langmore, Peter Norgaard, Jamie Smith, Stephan Hoyer. 🤤 -mail: dkochkov@google.com; janniyuval@google.com; shoyer@google.com

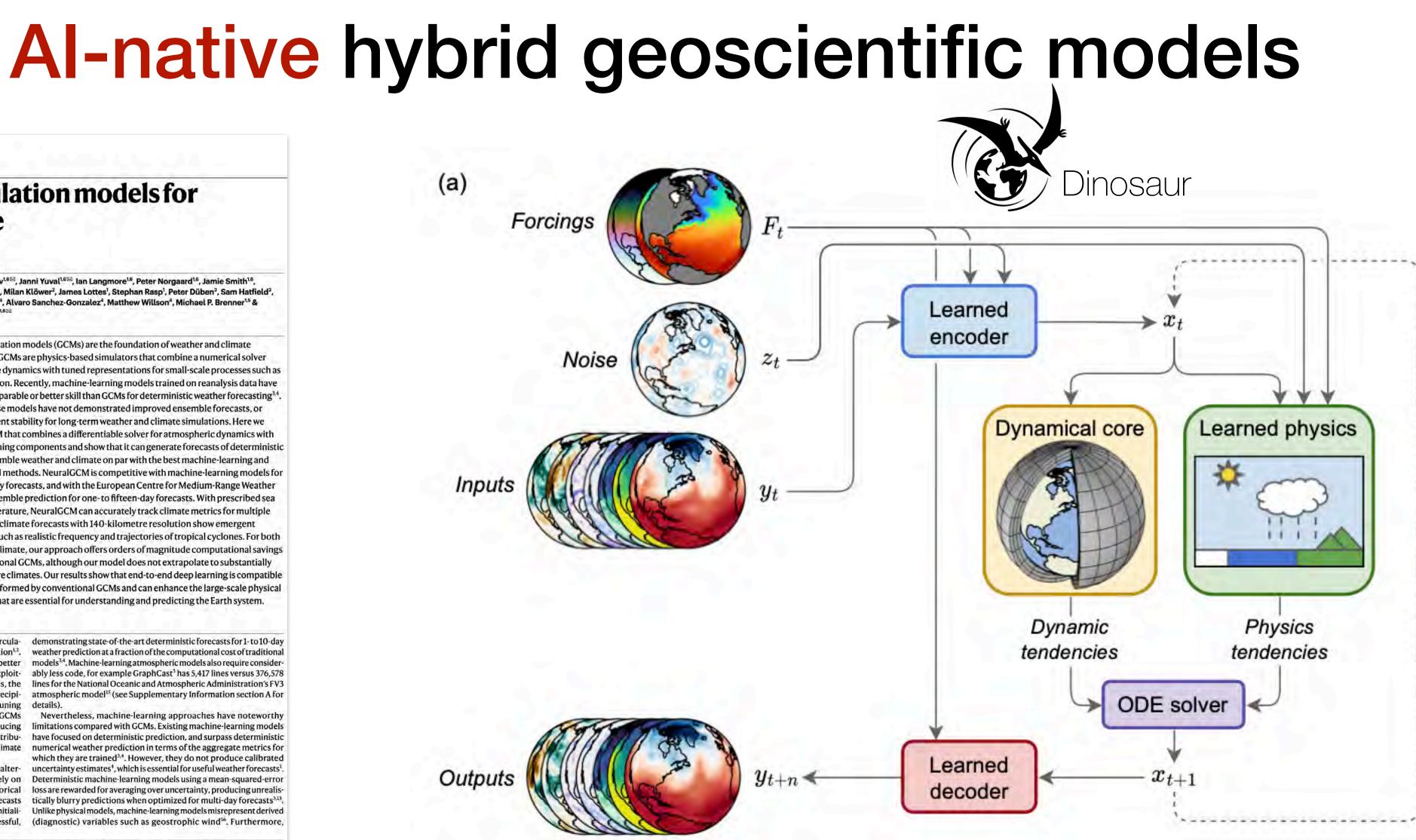
1060 | Nature | Vol 632 | 29 August 2024

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

Kochkov et al. (2024)

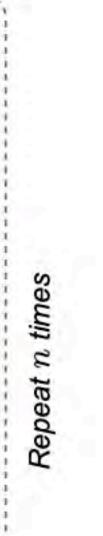


(a)





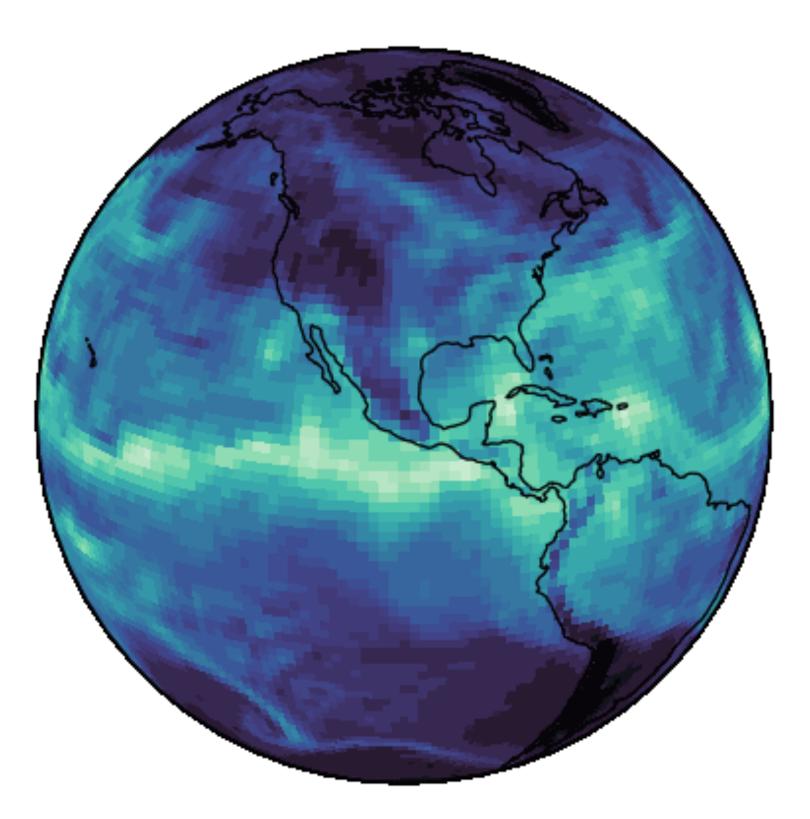
https://github.com/google-research/dinosaur https://github.com/google-research/neuralgcm



Al-native hybrid geoscientific models

Kochkov et al. (2024)

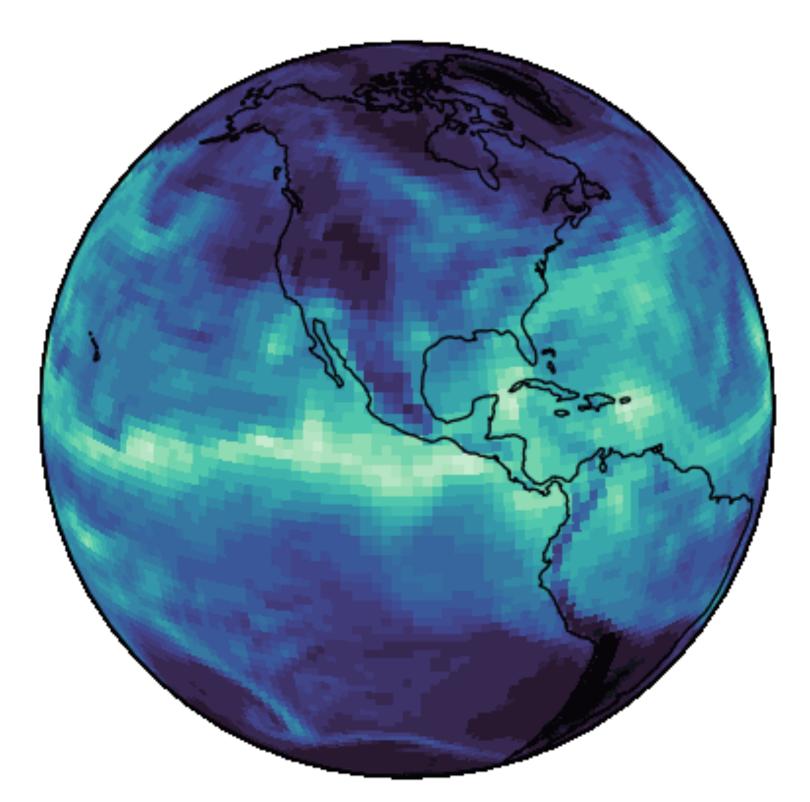
https://arxiv.org/abs/2311.07222





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

Total column water, 0-15 days



NeuralGCM





- Described how this can be done in practice today
- Advocated that a deep recast of our models is needed
- Described upcoming Al-native hybrid models

Integrating model-based products and observations



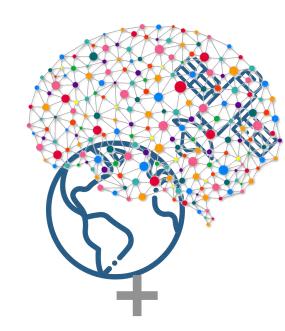


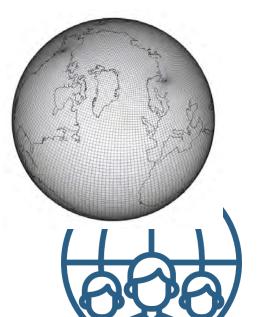
Assessing impact on downstream systems





n





Summary

- Illustrated why we are augmenting models with ML

Exciting time for cross-disciplinary investigations !







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. <u>https://doi.org/10.1029/2022MS003124</u>

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

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







