Machine Learning Methods for Atmosphere, Ocean, and Climate Science

Lecture 4: Climate Research Applications of ML

Mathematical modeling of Climate, Ocean, and Atmosphere processes
International Centre for Theoretical Sciences, TIFR, Bengaluru, India

Aman Gupta

Lecture 1

- Parametric estimation
- Introduction to deep neural networks
- The training algorithm

Lecture 2

- The PyTorch library
- Implementing Artificial Neural Nets in PyTorch

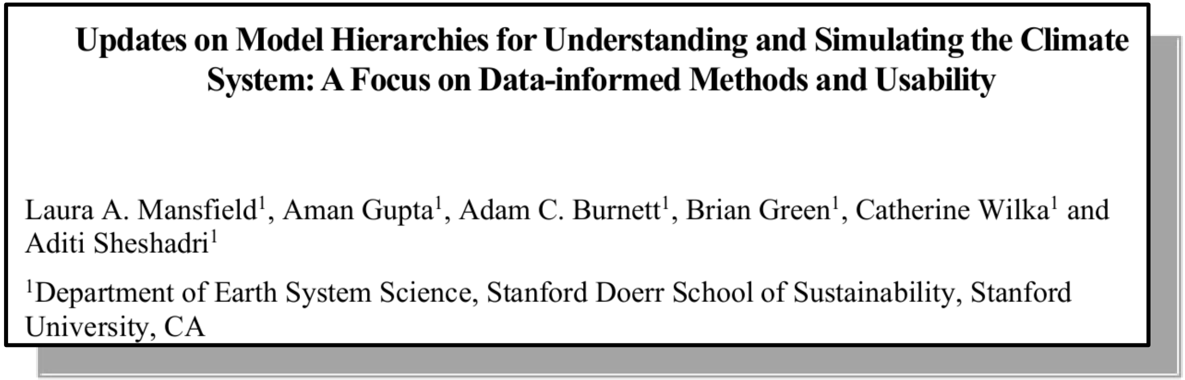
PyTorch

Lecture 3

- Implementing Convolutional Neural Networks in PyTorch

Lecture 4

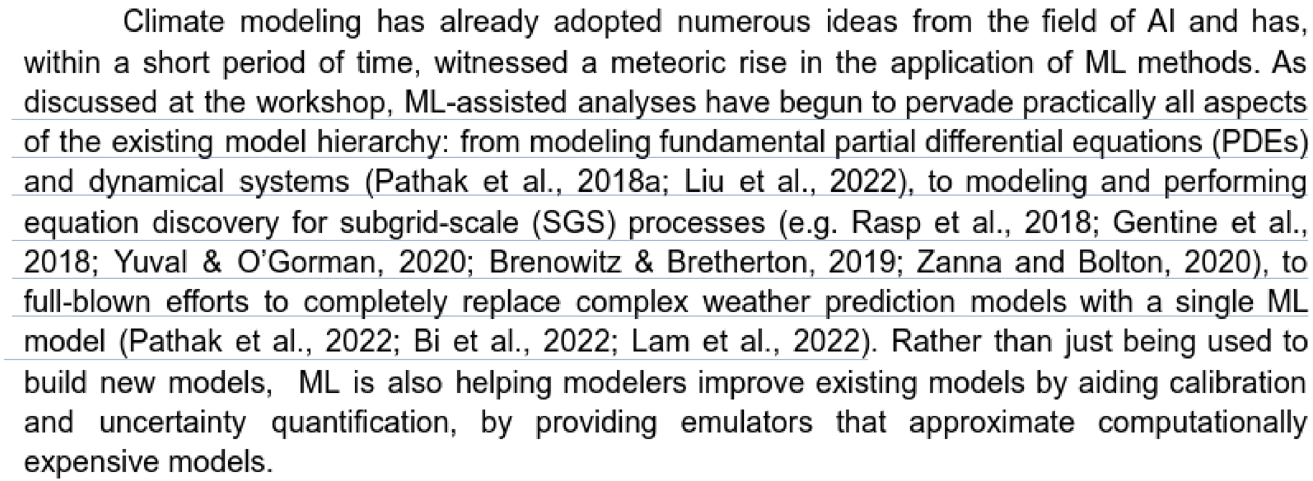
- Applications of ML in climate science



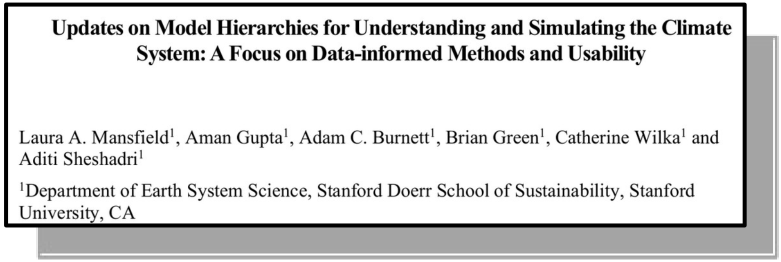
Updates on Model Hierarchies for Understanding and Simulating the Climate System: A Focus on Data-informed Methods and Usability

Laura A. Mansfield[1], Aman Gupta[1], Adam C. Burnett[1], Brian Green[1], Catherine Wilka[1] and Aditi Sheshadri[1]

[1]Department of Earth System Science, Stanford Doerr School of Sustainability, Stanford University, CA



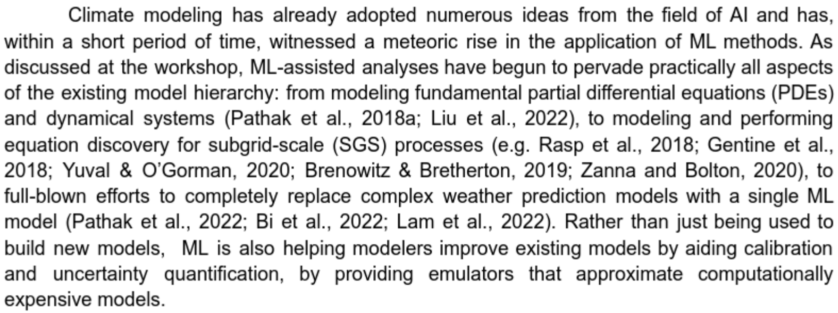
Climate modeling has already adopted numerous ideas from the field of AI and has, within a short period of time, witnessed a meteoric rise in the application of ML methods. As discussed at the workshop, ML-assisted analyses have begun to pervade practically all aspects of the existing model hierarchy: from modeling fundamental partial differential equations (PDEs) and dynamical systems (Pathak et al., 2018a; Liu et al., 2022), to modeling and performing equation discovery for subgrid-scale (SGS) processes (e.g. Rasp et al., 2018; Gentine et al., 2018; Yuval & O'Gorman, 2020; Brenowitz & Bretherton, 2019; Zanna and Bolton, 2020), to full-blown efforts to completely replace complex weather prediction models with a single ML model (Pathak et al., 2022; Bi et al., 2022; Lam et al., 2022). Rather than just being used to build new models, ML is also helping modelers improve existing models by aiding calibration and uncertainty quantification, by providing emulators that approximate computationally expensive models.



Updates on Model Hierarchies for Understanding and Simulating the Climate System: A Focus on Data-informed Methods and Usability

Laura A. Mansfield[1], Aman Gupta[1], Adam C. Burnett[1], Brian Green[1], Catherine Wilka[1] and Aditi Sheshadri[1]

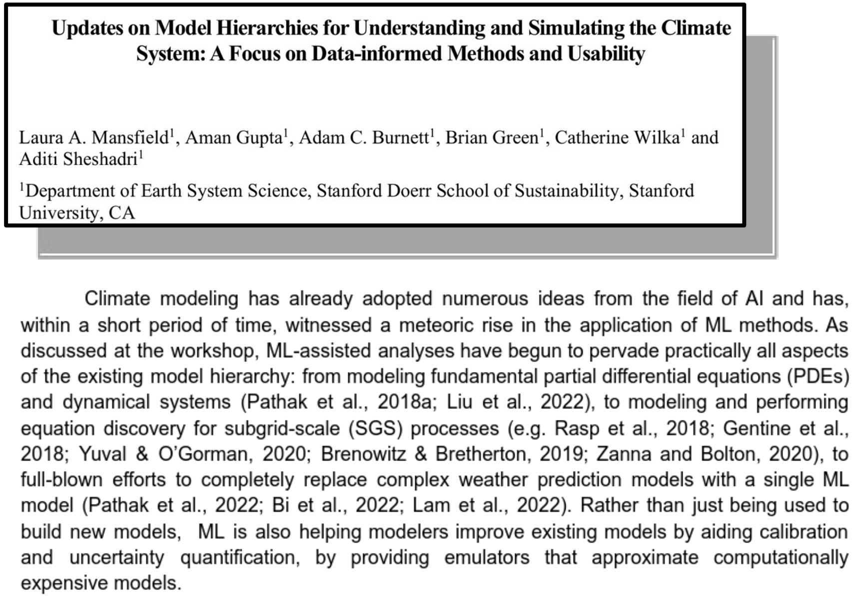
[1]Department of Earth System Science, Stanford Doerr School of Sustainability, Stanford University, CA



Climate modeling has already adopted numerous ideas from the field of AI and has, within a short period of time, witnessed a meteoric rise in the application of ML methods. As discussed at the workshop, ML-assisted analyses have begun to pervade practically all aspects of the existing model hierarchy: from modeling fundamental partial differential equations (PDEs) and dynamical systems (Pathak et al., 2018a; Liu et al., 2022), to modeling and performing equation discovery for subgrid-scale (SGS) processes (e.g. Rasp et al., 2018; Gentine et al., 2018; Yuval & O'Gorman, 2020; Brenowitz & Bretherton, 2019; Zanna and Bolton, 2020), to full-blown efforts to completely replace complex weather prediction models with a single ML model (Pathak et al., 2022; Bi et al., 2022; Lam et al., 2022). Rather than just being used to build new models, ML is also helping modelers improve existing models by aiding calibration and uncertainty quantification, by providing emulators that approximate computationally expensive models.

Interpretability

Generalizability



Updates on Model Hierarchies for Understanding and Simulating the Climate System: A Focus on Data-informed Methods and Usability

Laura A. Mansfield[1], Aman Gupta[1], Adam C. Burnett[1], Brian Green[1], Catherine Wilka[1] and Aditi Sheshadri[1]

[1]Department of Earth System Science, Stanford Doerr School of Sustainability, Stanford University, CA

Climate modeling has already adopted numerous ideas from the field of AI and has, within a short period of time, witnessed a meteoric rise in the application of ML methods. As discussed at the workshop, ML-assisted analyses have begun to pervade practically all aspects of the existing model hierarchy: from modeling fundamental partial differential equations (PDEs) and dynamical systems (Pathak et al., 2018a; Liu et al., 2022), to modeling and performing equation discovery for subgrid-scale (SGS) processes (e.g. Rasp et al., 2018; Gentine et al., 2018; Yuval & O'Gorman, 2020; Brenowitz & Bretherton, 2019; Zanna and Bolton, 2020), to full-blown efforts to completely replace complex weather prediction models with a single ML model (Pathak et al., 2022; Bi et al., 2022; Lam et al., 2022). Rather than just being used to build new models, ML is also helping modelers improve existing models by aiding calibration and uncertainty quantification, by providing emulators that approximate computationally expensive models.

Interpretability

+

Generalizability

=

Trustworthiness!

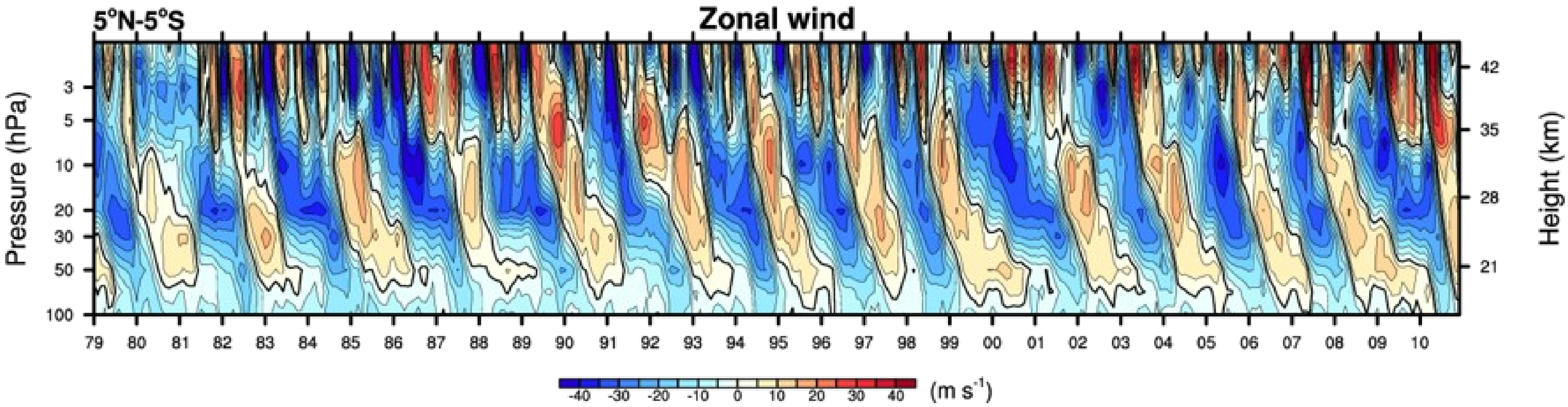
Data-driven Model Parameterizations

Equation Discovery using ML

ML in Weather Forecasting

Data-driven Model Parameterizations

Gravity Waves Significantly Influence Stratospheric Variability



- Tropical stratosphere dominated by oscillating wind patterns with a period of ~28 months, called the **Quasi-Biennial Oscillation (QBO)**. Can have influence over tropical convective systems like ENSO, Madden-Julian Oscillation.

- **Driven mostly by convectively generated gravity waves**, which are not completely resolved in most climate models

Gravity Waves Significantly Influence Stratospheric Variability

Geophysical Research Letters®

RESEARCH LETTER

10.1029/2022GL098174

Key Points:

- Neural networks trained on one annual cycle accurately emulate a physics-based gravity wave parameterization (GWP) when coupled to a climate model
- Although trained on only one phase of the Quasi-Biennial Oscillation, the

Machine Learning Gravity Wave Parameterization Generalizes to Capture the QBO and Response to Increased CO_2

Zachary I. Espinosa[1], **Aditi Sheshadri**[1], **Gerald R. Cain**[2], **Edwin P. Gerber**[3], **and Kevin J. DallaSanta**[4,5]

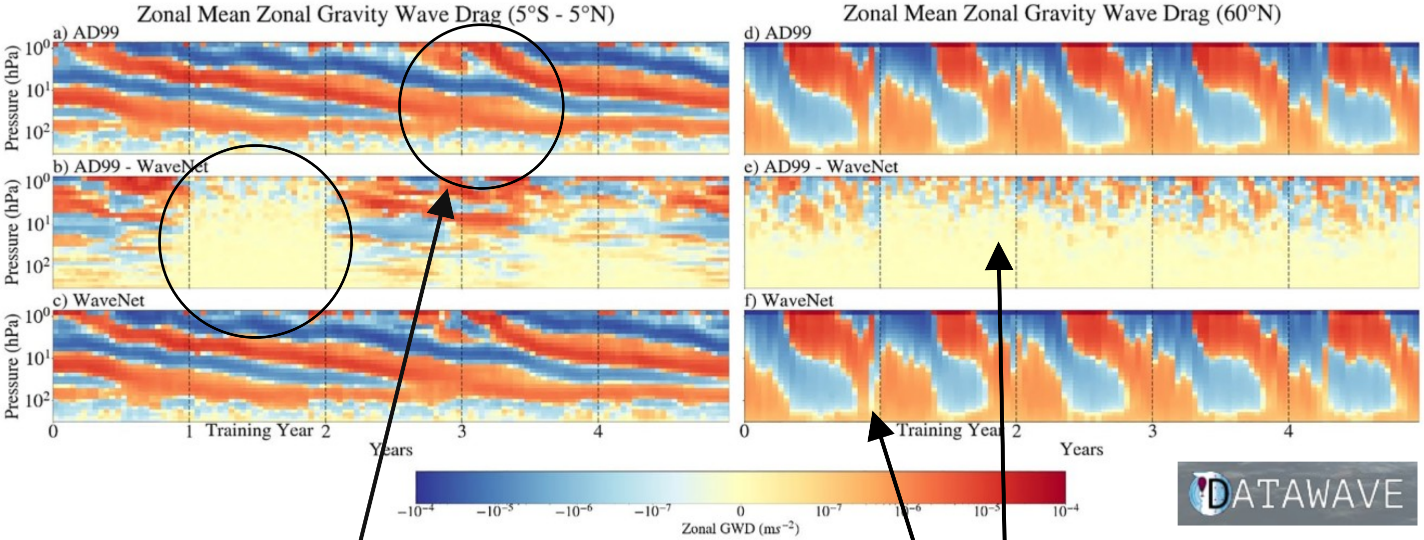
[1]Department of Earth System Science, Stanford University, Stanford, CA, USA, [2]Department of Computer Science, Stanford University, Stanford, CA, USA, [3]Courant Institute of Mathematical Sciences, New York University, New York, NY, USA, [4]NASA Goddard Institute for Space Studies, New York, NY, USA, [5]Universities Space Research Association, Columbia, MD, USA

- Tropical stratosphere dominated by oscillating wind patterns with a period of ~28 months, called the Quasi-Biennial Oscillations (QBO). Can have influence over tropical convective systems like ENSO, Madden-Julian Oscillation.
- Driven mostly by convectively generated gravity waves, which are not completely resolved in most climate models

WaveNet: an ML Emulator for Atmospheric Gravity Waves

Equatorial gravity wave drag

Northern polar region drag



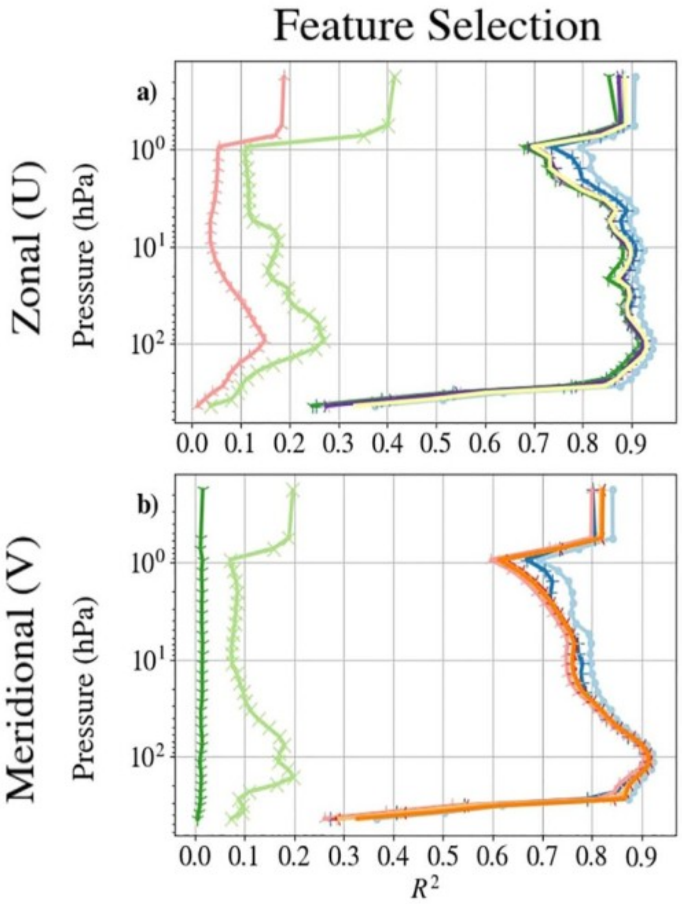
DATAWAVE

Simulates the westerly phase of the QBO well despite being trained upon the easterly QBO winds!

Vanilla ANN, 9 layers, trained for 200 epochs, variable learning rate, logcosh error, Adam optimizer, no regularization, minibatch size = 1024

- Trained on just one year of atmospheric data **(features)** and GW parameterization data **(labels)** from an intermediate complexity climate model.
- Generalizes well to identify longer period signals and out-of-sample data points for four test years

Offline tests: comparable performance using less features and data

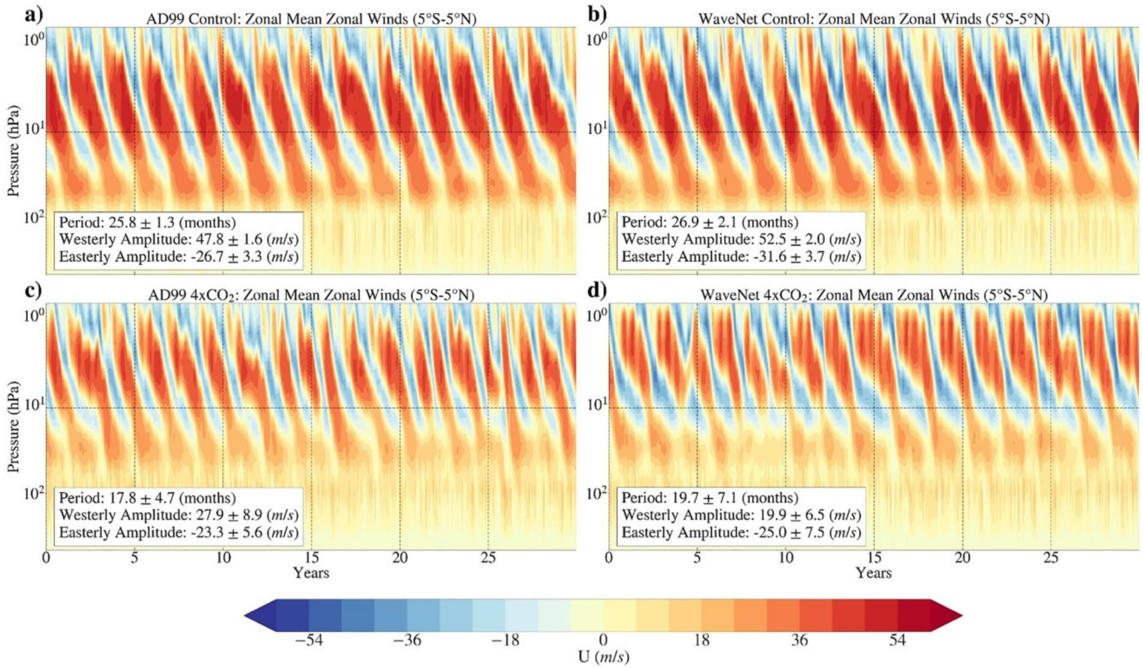


*Higher R^2 means higher prediction score

Robustness to feature omission: only training on the winds sufficient to retain 96% of the prediction skill of the full model

Robustness to training data: training on 1/4th of the years retains 98% of the prediction skill of the full NN

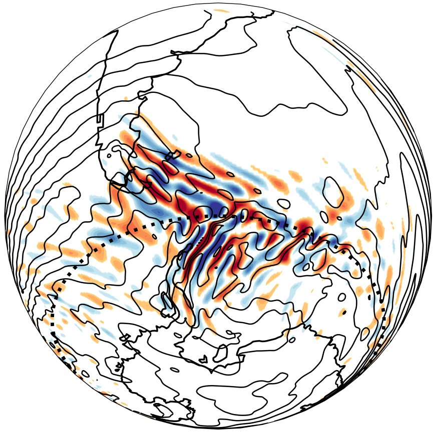
Online tests: consistent results for QBO evolution under climate change

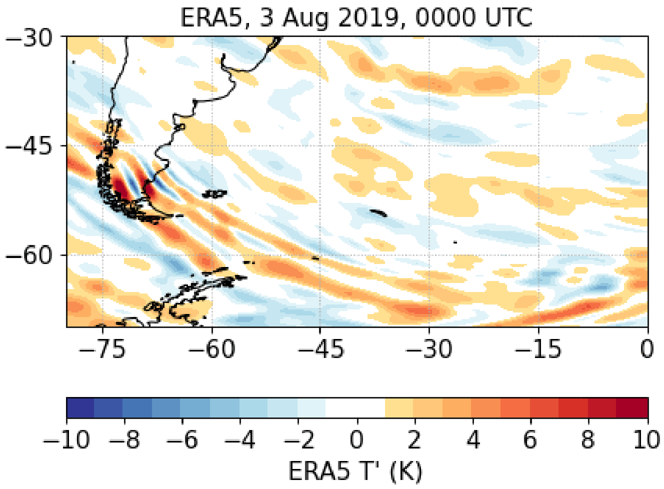


- WaveNet plugged into the Fortran climate model code. Model integrated for 50 years with original gravity wave parameterization and with WaveNet

- Both parameterizations and WaveNet predict a reduction in the QBO period and weakening of the maximum windspeeds

Towards Nonlocal GW Parameterizations





Properties desired:

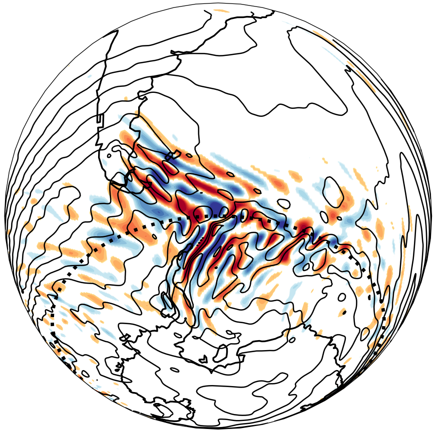
1) Lateral propagation

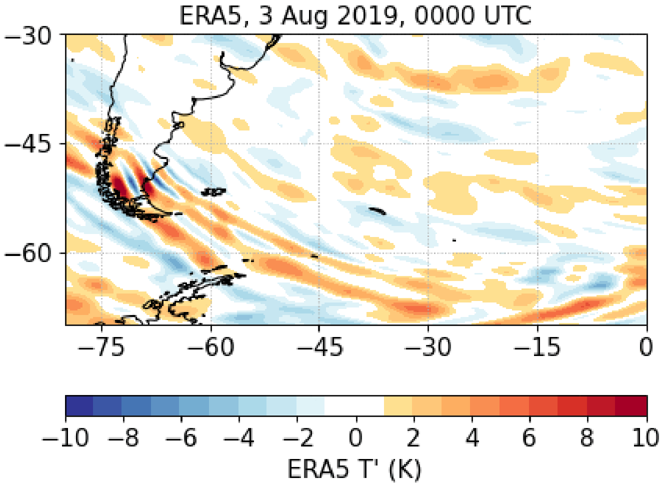
2) Refraction

3) Transience

(Gupta et al. (2021); Gupta et al. *in prep*)

Towards **Nonlocal GW Parameterizations**





Properties desired:

1) Lateral propagation

2) Refraction

3) Transience

WaveNet → **Nonlocal emulators**

WaveNet

Trained on parameterizations
Limited training data
Incomplete wave

Nonlocal emulators

Trained on HighRes data
Extend training periods
Complete physics

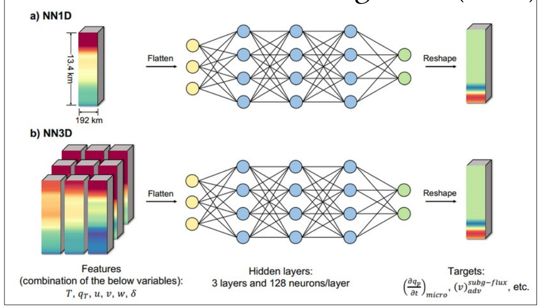
(Gupta et al. (2021); Gupta et al. *in prep*)

Towards Nonlocal GW Parameterizations

Nonlocal emulators based on Recurrent Neural Networks

Trained on global 1km simulations from ECMWF
Extended training periods
Complete physics

Wang et al. (2022)



Training using non-local columns

- ✓ Lateral propagation
- ✓ Refraction

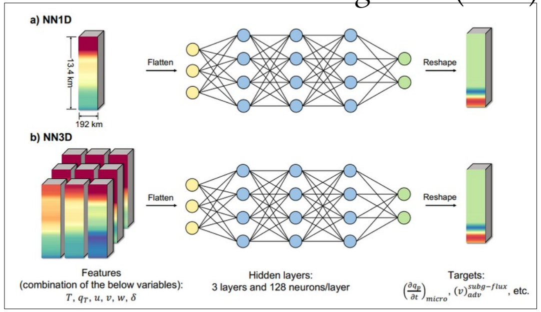
(Gupta et al. (2021); Gupta et al. *in prep*)

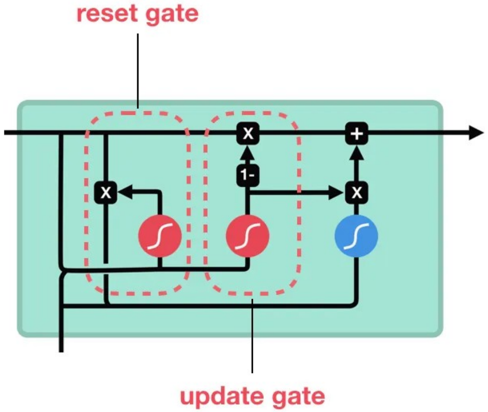
Towards **Nonlocal GW Parameterizations**

Nonlocal emulators based on Recurrent Neural Networks

Trained on global 1km simulations from ECMWF
Extended training periods
Complete physics

Wang et al. (2022)





Training using non-local columns

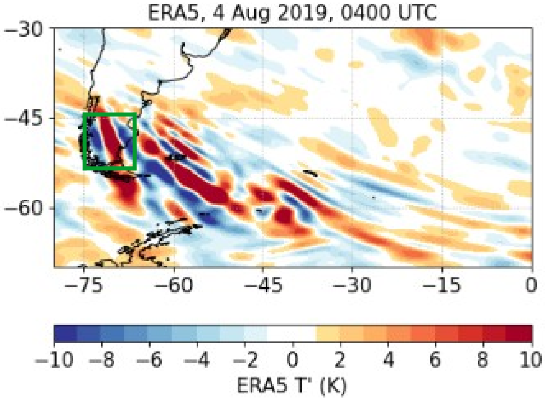
- ✓ Lateral propagation
- ✓ Refraction

Implement transience using Recurrent Networks

- ✓ Transience

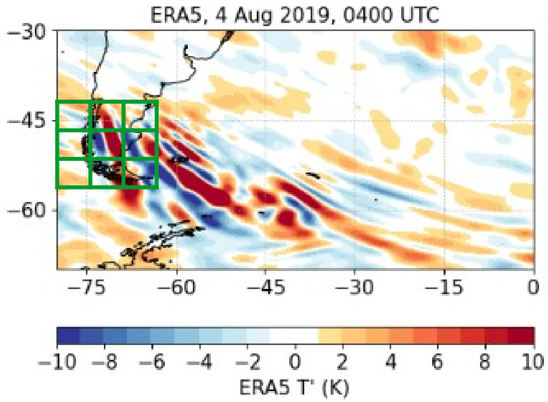
(Gupta et al. (2021); Gupta et al. *in prep*)

Towards **Nonlocal GW Parameterizations**



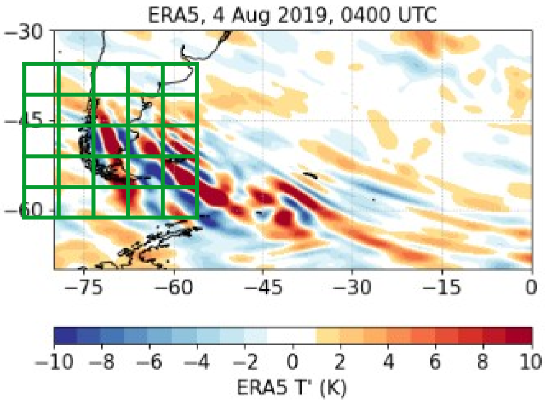
(Gupta et al. (2021); Gupta et al. *in prep*)

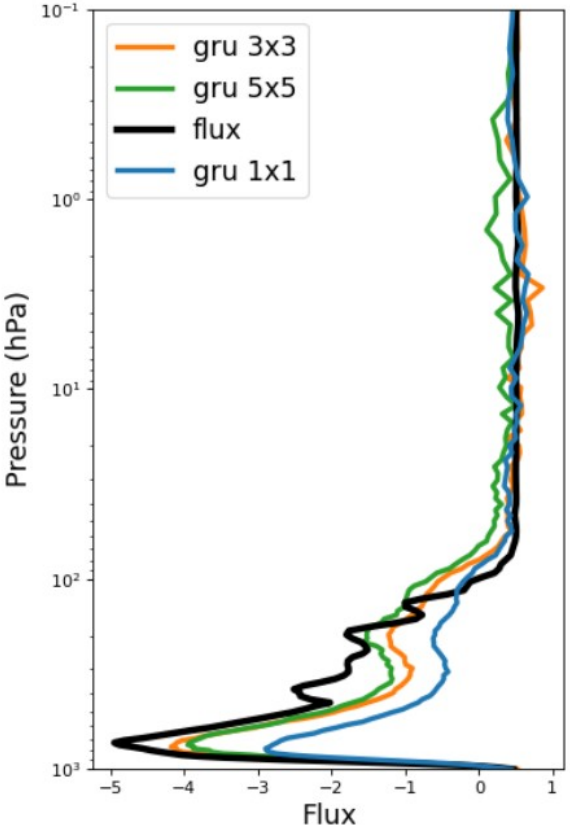
Towards **Nonlocal GW Parameterizations**



(Gupta et al. (2021); Gupta et al. *in prep*)

Towards Nonlocal GW Parameterizations

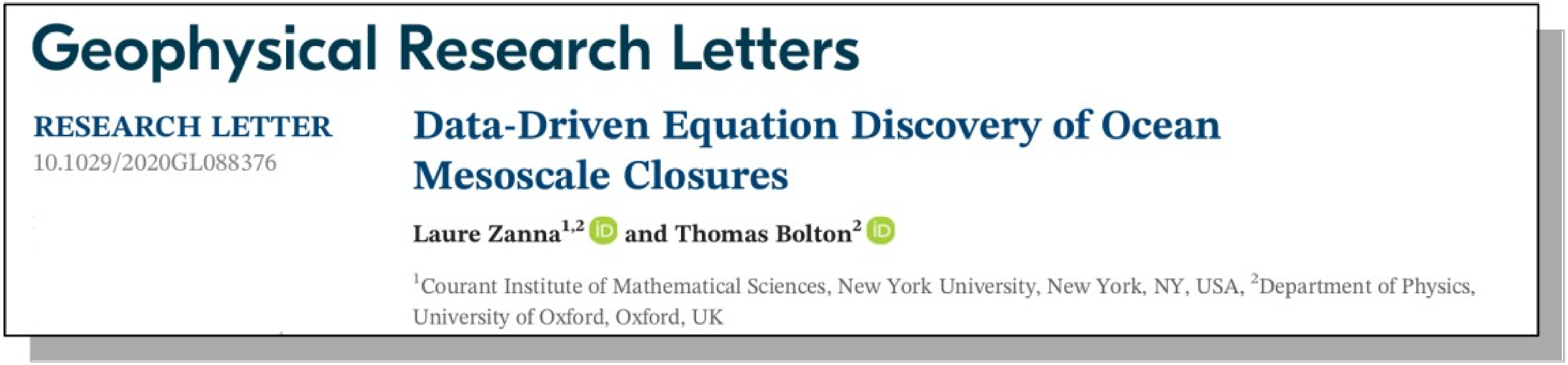




(Gupta et al. (2021); Gupta et al. *in prep*)

Equation Discovery

Mesoscale Dynamics are Important, but are not Resolved in Ocean Models



Geophysical Research Letters

RESEARCH LETTER
10.1029/2020GL088376

Data-Driven Equation Discovery of Ocean Mesoscale Closures

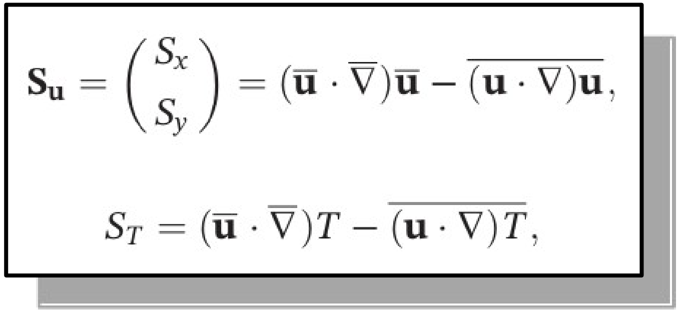
Laure Zanna[1,2] and Thomas Bolton[2]

[1]Courant Institute of Mathematical Sciences, New York University, New York, NY, USA, [2]Department of Physics, University of Oxford, Oxford, UK

- Subgrid-scale turbulent fluxes at scales 10-100 km important for global heat, oxygen, and tracer transport
- Interaction on small-scales with large-scales not well understood
- These scale not typically resolved in ocean models

- Parameterizations spuriously dissipate kinetic energy, affecting large-scale currents
- Use simple quadratic closures and hyperdiffusion to parameterize subgrid scales

Using ML to find Equations that Govern Mesoscale Ocean Dynamics



$$\mathbf{S_u} = \begin{pmatrix} S_x \\ S_y \end{pmatrix} = (\overline{\mathbf{u}} \cdot \overline{\nabla})\overline{\mathbf{u}} - \overline{(\mathbf{u} \cdot \nabla)\mathbf{u}},$$

$$S_T = (\overline{\mathbf{u}} \cdot \overline{\nabla})T - \overline{(\mathbf{u} \cdot \nabla)T},$$

u = (u,v): horizontal winds, T: temperature
overbar: 30 km Gaussian filtering

> “Given some spatiotemporal data set of the subgrid eddy forcing, we uncover an equation that could have produced that dataset.”

(Zanna and Bolton (2020), GRL)

Using ML to find Equations that Govern Mesoscale Ocean Dynamics

$$\mathbf{S_u} = \begin{pmatrix} S_x \\ S_y \end{pmatrix} = (\overline{\mathbf{u}} \cdot \overline{\nabla})\overline{\mathbf{u}} - \overline{(\mathbf{u} \cdot \nabla)\mathbf{u}},$$

$$S_T = (\overline{\mathbf{u}} \cdot \overline{\nabla})\overline{T} - \overline{(\mathbf{u} \cdot \nabla)T},$$

u = (u,v): horizontal winds, T: temperature
overbar: 30 km Gaussian filtering

> "Given some spatiotemporal data set of the subgrid eddy forcing, we uncover an equation that could have produced that dataset."

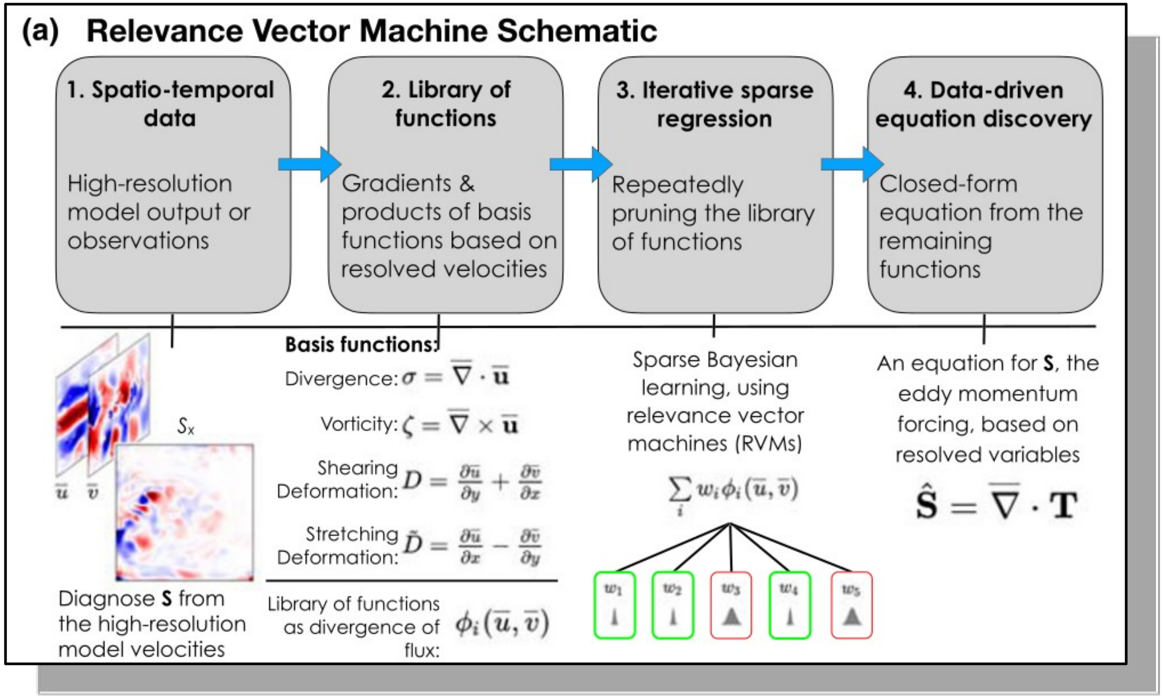
(a) Relevance Vector Machines (RVMs)
iterative regression
possible to physically interpret the results

(b) Convolutional Neural Networks (CNNs)

(Zanna and Bolton (2020), GRL)

Using ML to find Equations that Govern Mesoscale Ocean Dynamics

Barotropic Model: highly idealized, 3.75 km, 10 years of data



(a) Relevance Vector Machine Schematic

1. Spatio-temporal data

High-resolution model output or observations

2. Library of functions

Gradients & products of basis functions based on resolved velocities

3. Iterative sparse regression

Repeatedly pruning the library of functions

4. Data-driven equation discovery

Closed-form equation from the remaining functions

Diagnose **S** from the high-resolution model velocities

Basis functions:

Divergence: $\sigma = \overline{\nabla} \cdot \overline{\mathbf{u}}$

Vorticity: $\zeta = \overline{\nabla} \times \overline{\mathbf{u}}$

Shearing Deformation: $D = \frac{\partial \overline{u}}{\partial y} + \frac{\partial \overline{v}}{\partial x}$

Stretching Deformation: $\tilde{D} = \frac{\partial \overline{u}}{\partial x} - \frac{\partial \overline{v}}{\partial y}$

Library of functions as divergence of flux: $\phi_i(\overline{u}, \overline{v})$

Sparse Bayesian learning, using relevance vector machines (RVMs)

$\sum_i w_i \phi_i(\overline{u}, \overline{v})$

An equation for **S**, the eddy momentum forcing, based on resolved variables

$\hat{\mathbf{S}} = \overline{\nabla} \cdot \mathbf{T}$

Assume a library of basis functions:
- Start with u and v and their gradients

- Obtain vorticity and divergence as a more reasonable basis

- Use vorticity, divergence, shear terms as the improved basis, ϕ_is to create a library of funcitons (products and derivatives). Compute these quantities using the high resolution data

- Compute the weights w_is for functions, and prune to get the basis that best captures the subgrid scale fluxes

- Obtain subgrid scale closures:

$$\Sigma_i \; w_i \phi_i(\bar{u}, \bar{v})$$

(Zanna and Bolton (2020), GRL)

Using ML to find Equations that Govern Mesoscale Ocean Dynamics

RVM reveals the expression:

$$\boxed{\hat{\mathbf{S}}_{\mathbf{u}}^{BT} = \begin{pmatrix} w_0(\zeta^2)_x - w_1(\zeta D)_x + w_2(\zeta \tilde{D})_y \\ w_3(\zeta^2)_y + w_4(\zeta D)_y + w_5(\zeta \tilde{D})_x \end{pmatrix},}$$

$$\begin{aligned} w_0 &= -4.096 \times 10^8 \ \mathrm{m}^2 \\ w_1 &= -5.483 \times 10^8 \ \mathrm{m}^2 \\ w_2 &= -4.384 \times 10^8 \ \mathrm{m}^2 \\ w_3 &= -4.100 \times 10^8 \ \mathrm{m}^2 \\ w_4 &= -6.332 \times 10^8 \ \mathrm{m}^2 \\ w_5 &= -4.815 \times 10^8 \ \mathrm{m}^2 \end{aligned}$$

Pruning the basis –
Just one basis function captures ~20% of the variance
Three functions capture ~50% of the variance

(Zanna and Bolton (2020), GRL)

Using ML to find Equations that Govern Mesoscale Ocean Dynamics

RVM reveals the expression:

$$\boxed{\hat{\mathbf{S}}_{\mathbf{u}}^{BT} = \begin{pmatrix} w_0(\zeta^2)_x - w_1(\zeta D)_x + w_2(\zeta\tilde{D})_y \\ w_3(\zeta^2)_y + w_4(\zeta D)_y + w_5(\zeta\tilde{D})_x \end{pmatrix},}$$

$$\begin{aligned} w_0 &= -4.096 \times 10^8 \ \mathrm{m}^2 \\ w_1 &= -5.483 \times 10^8 \ \mathrm{m}^2 \\ w_2 &= -4.384 \times 10^8 \ \mathrm{m}^2 \\ w_3 &= -4.100 \times 10^8 \ \mathrm{m}^2 \\ w_4 &= -6.332 \times 10^8 \ \mathrm{m}^2 \\ w_5 &= -4.815 \times 10^8 \ \mathrm{m}^2 \end{aligned}$$

$$\xrightarrow{\text{approximate form}}$$

Approximate form for subgrid scale momentum transport

$$\boxed{\hat{\mathbf{S}}_{\mathbf{u}}^{BT} \approx \kappa_{BT}\overline{\nabla} \cdot \begin{pmatrix} \zeta^2 - \zeta D & \zeta\tilde{D} \\ \zeta\tilde{D} & \zeta^2 + \zeta D \end{pmatrix}.}$$

$$\kappa_{BT} \approx \Sigma_i w_i = -4.87 \times 10^8 \ \mathrm{m}^2$$

Pruning the basis –
Just one basis function captures ~20% of the variance
Three functions capture ~50% of the variance

(Zanna and Bolton (2020), GRL)

Using ML to find Equations that Govern Mesoscale Ocean Dynamics

RVM reveals the expression:

$$\boxed{\hat{\mathbf{S}}_{\mathbf{u}}^{BT} = \begin{pmatrix} w_0(\zeta^2)_x - w_1(\zeta D)_x + w_2(\zeta\tilde{D})_y \\ w_3(\zeta^2)_y + w_4(\zeta D)_y + w_5(\zeta\tilde{D})_x \end{pmatrix},}$$

$$\begin{aligned} w_0 &= -4.096 \times 10^8 \text{ m}^2 \\ w_1 &= -5.483 \times 10^8 \text{ m}^2 \\ w_2 &= -4.384 \times 10^8 \text{ m}^2 \\ w_3 &= -4.100 \times 10^8 \text{ m}^2 \\ w_4 &= -6.332 \times 10^8 \text{ m}^2 \\ w_5 &= -4.815 \times 10^8 \text{ m}^2 \end{aligned}$$

Pruning the basis –
Just one basis function captures ~20% of the variance
Three functions capture ~50% of the variance

approximate form →

Approximate form for subgrid scale momentum transport

$$\boxed{\hat{\mathbf{S}}_{\mathbf{u}}^{BT} \approx \kappa_{BT}\overline{\nabla} \cdot \begin{pmatrix} \zeta^2 - \zeta D & \zeta\tilde{D} \\ \zeta\tilde{D} & \zeta^2 + \zeta D \end{pmatrix}.}$$

$$\kappa_{BT} \approx \Sigma_i w_i = -4.87 \times 10^8 \text{ m}^2$$

Connect to the baroclinic form ↓

$$\boxed{\hat{\mathbf{S}}_{\mathbf{u}}^{BC} \approx \kappa_{BC}\overline{\nabla} \cdot \begin{pmatrix} -\zeta D & \zeta\tilde{D} \\ \zeta\tilde{D} & \zeta D \end{pmatrix} + \mathbf{I}\,\frac{1}{2}\kappa_{BC}\overline{\nabla}(\zeta^2 + D^2 + \tilde{D}^2),}$$

Similarly for temperature,

$$\boxed{\hat{S}_T = \kappa_T\overline{\nabla} \cdot \begin{pmatrix} -\overline{u}_y\overline{u}_z - \overline{v}_y\overline{v}_z \\ \overline{u}_x\overline{u}_z + \overline{v}_x\overline{v}_z \end{pmatrix}.}$$

(Zanna and Bolton (2020), GRL)

RVMs match the true variability pretty well in offline tests*

$$\mathbf{S_u} = \begin{pmatrix} S_x \\ S_y \end{pmatrix} = (\overline{\mathbf{u}} \cdot \overline{\nabla})\overline{\mathbf{u}} - \overline{(\mathbf{u} \cdot \nabla)\mathbf{u}},$$

$$S_T = (\overline{\mathbf{u}} \cdot \overline{\nabla})T - \overline{(\mathbf{u} \cdot \nabla)T},$$

Idealized simulation

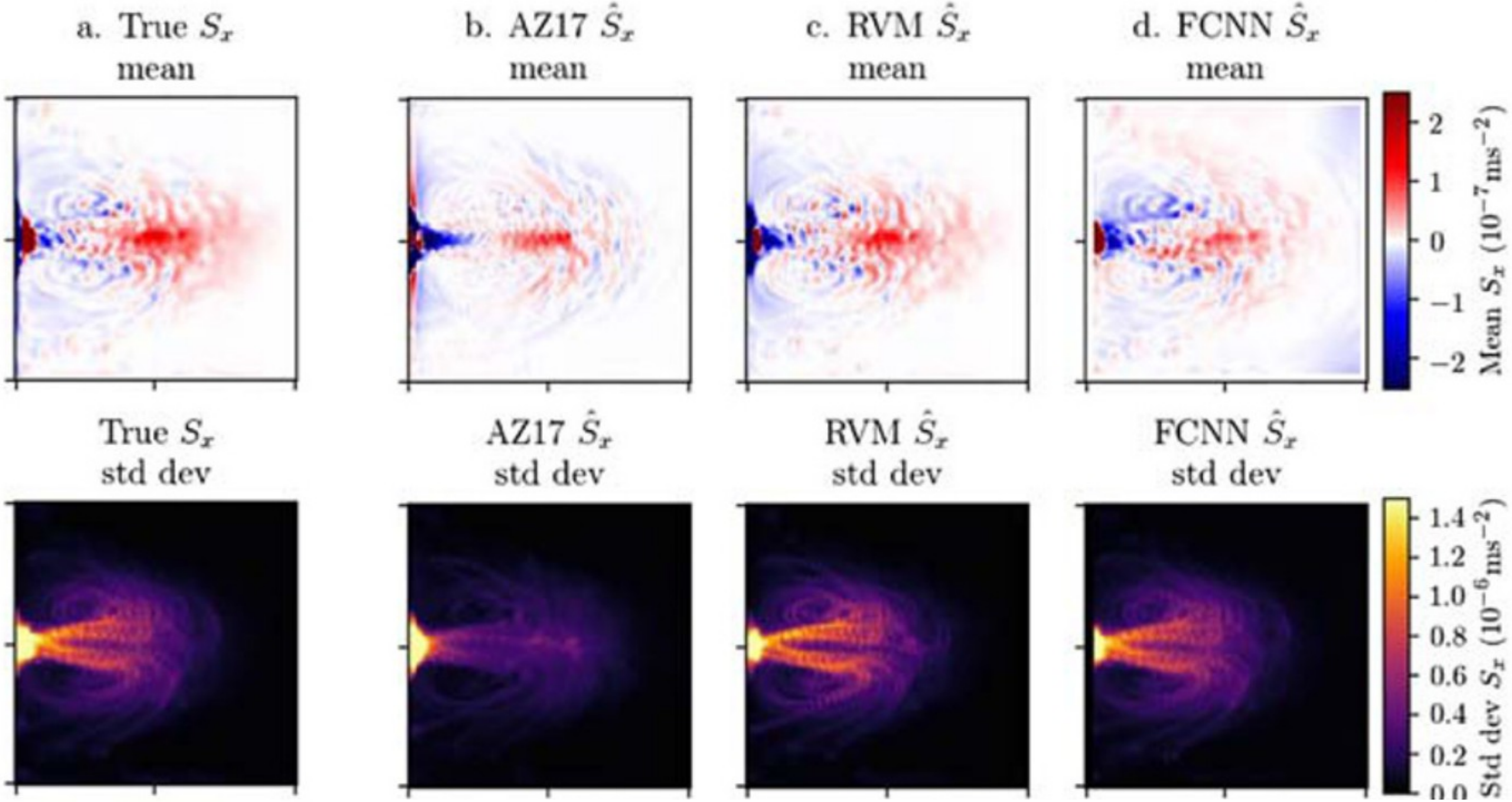
Established parameterization

RVMs

CNNs

Mean zonal momentum fluxes

Standard deviation of momentum fluxes



(Zanna and Bolton (2020), GRL)

ML in Weather Forecasting

FourCastNet: Data-driven Weather Forecasting

FOURCASTNET: A GLOBAL DATA-DRIVEN HIGH-RESOLUTION WEATHER MODEL USING ADAPTIVE FOURIER NEURAL OPERATORS

A PREPRINT

NVIDIA

Jaideep Pathak
NVIDIA Corporation
Santa Clara, CA 95051

Shashank Subramanian
Lawrence Berkeley
National Laboratory
Berkeley, CA 94720

Peter Harrington
Lawrence Berkeley
National Laboratory
Berkeley, CA 94720

Sanjeev Raja
University of Michigan
Ann Arbor, MI 48109

Ashesh Chattopadhyay
Rice University
Houston, TX 77005

Morteza Mardani
NVIDIA Corporation
Santa Clara, CA 95051

Thorsten Kurth
NVIDIA Corporation
Santa Clara, CA 95051

David Hall
NVIDIA Corporation
Santa Clara, CA 95051

Zongyi Li
California Institute of Technology
Pasadena, CA 91125
NVIDIA Corporation
Santa Clara, CA 95051

Kamyar Azizzadenesheli
Purdue University
West Lafayette, IN 47907

Pedram Hassanzadeh
Rice University
Houston, TX 77005

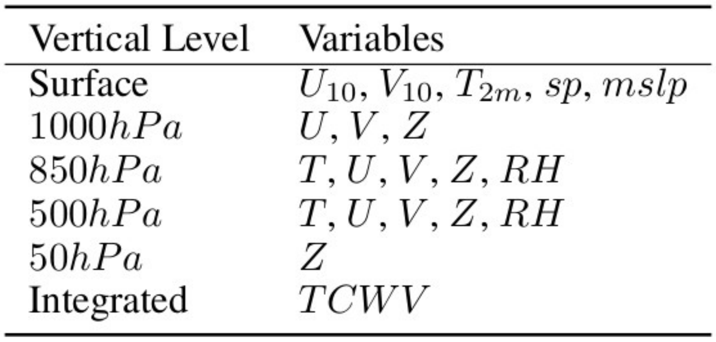
Karthik Kashinath
NVIDIA Corporation
Santa Clara, CA 95051

Animashree Anandkumar
California Institute of Technology
Pasadena, CA 91125
NVIDIA Corporation
Santa Clara, CA 95051

A complex semi-supervised ML model built using Fourier Neural Operators and Vision Transformer

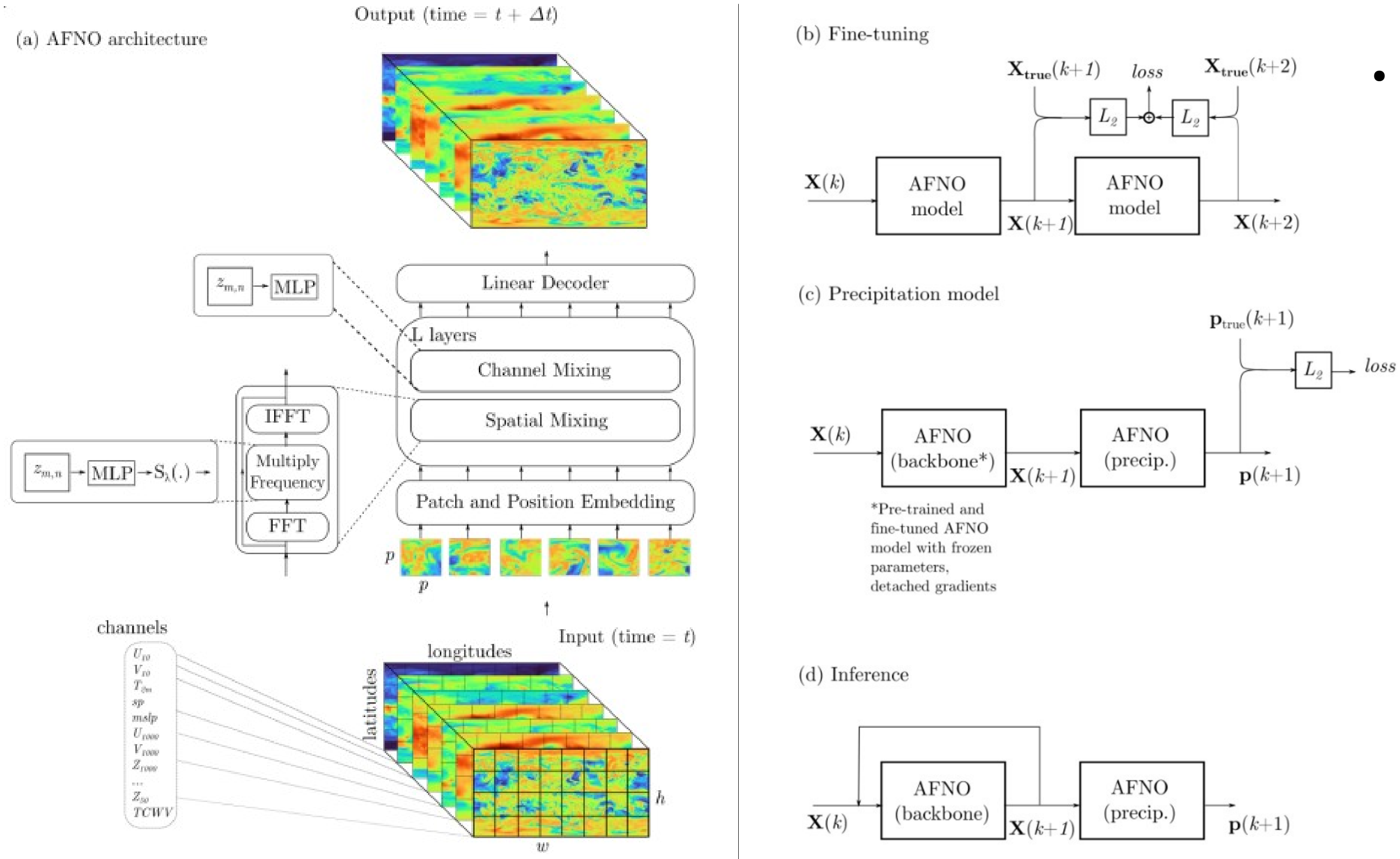
FourCastNet: Data-driven Weather Forecasting

- Pre-training stage | Fine-tuning stage | Inference
- Trained on 6-hourly ERA5 reanalysis on pressure levels from 1979-2015. 0.25° resolution. 2016-17 validation data. 2018-2020 testing data.
- A separate AFNO model to forecast total precipitation: an additional convolutional layer and ReLU() activation to enforce non-negative outputs
- Model trained with cosine learning rates for a total of 130 epochs



Vertical Level	Variables
Surface	$U_{10}, V_{10}, T_{2m}, sp, mslp$
$1000hPa$	U, V, Z
$850hPa$	T, U, V, Z, RH
$500hPa$	T, U, V, Z, RH
$50hPa$	Z
Integrated	$TCWV$

FourCastNet’s Model Architecture

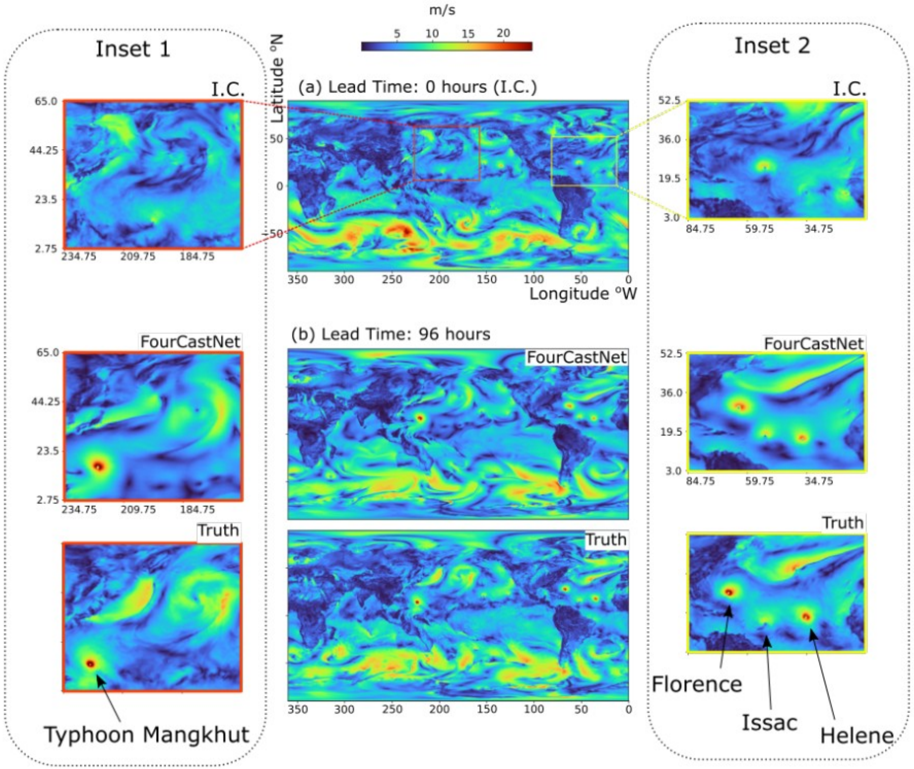


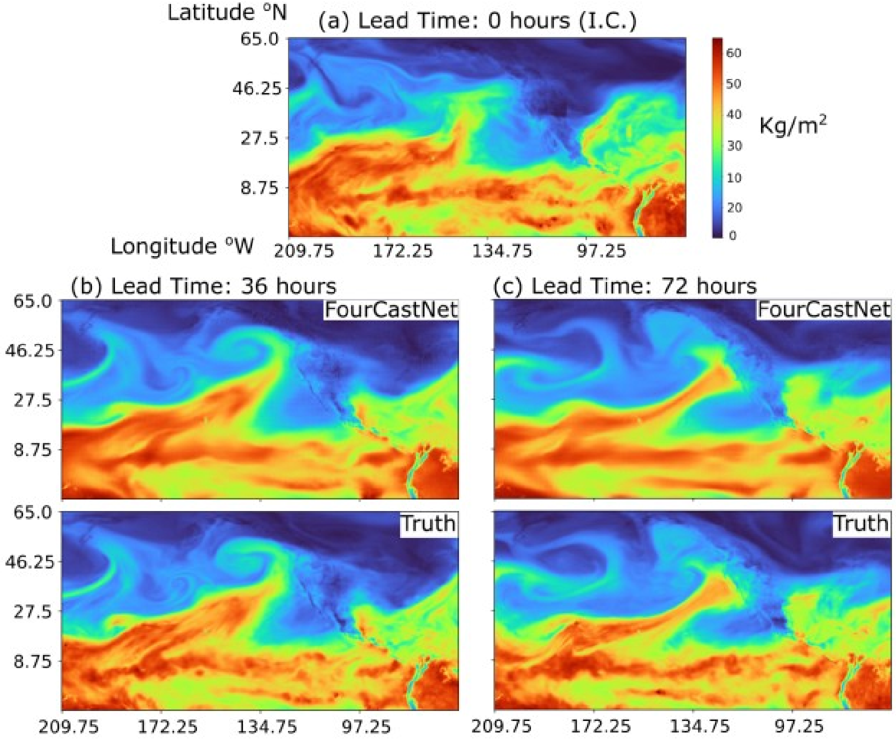
- Pretty complex model!

  - create patches
  - form patch embedding
  - input to the attention layer
  - optimize attention
  - decode the image to get forecast
  - forecast precipitation as a diagnostic

$\mathrm{C\ x\ H\ x\ W} \longrightarrow (\mathrm{N\ x\ P^2\ x\ C}) \longrightarrow \mathrm{N\ x\ D}$ (using a linear projection)

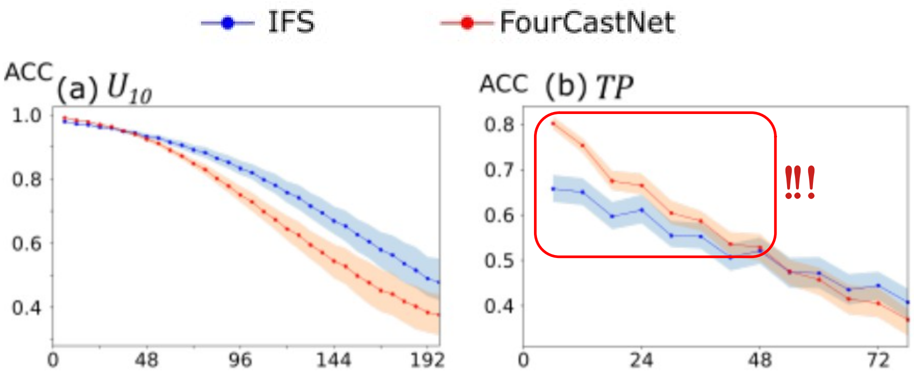
FourCastNet provides impressive Hurricane and AR forecasts over short lead times





Small-scales captured better than by CNNs

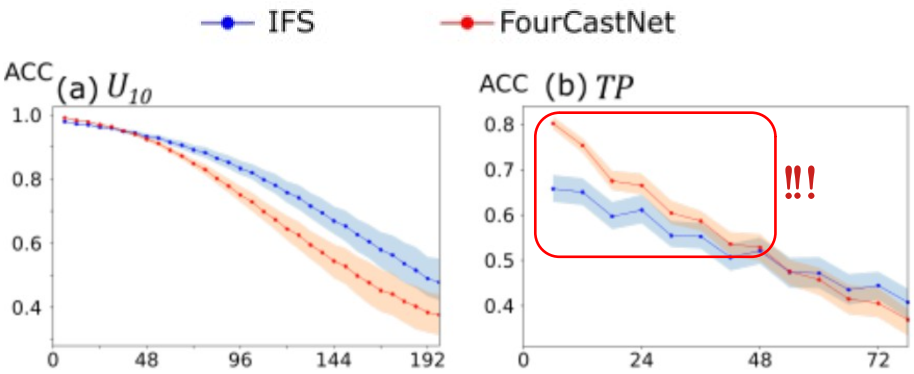
FourCastNet outperforms Weyn et al.'s DLWP model

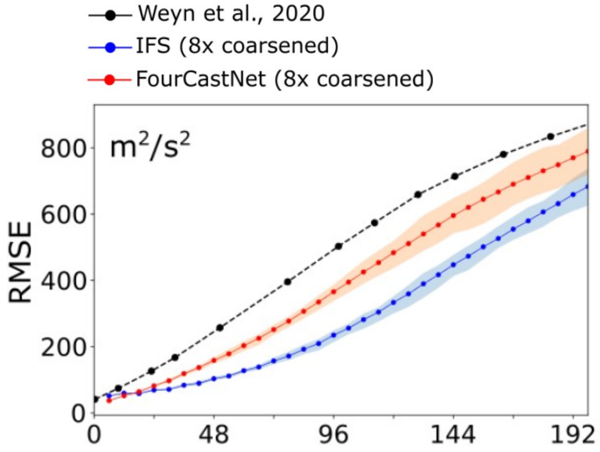


ACC: Anomaly correlation coefficient. Higher ACC means higher skill

On 0-48 lead times, **FourCastNet provides better total precipitation forecasts than IFS**

FourCastNet outperforms Weyn et al.'s DLWP model

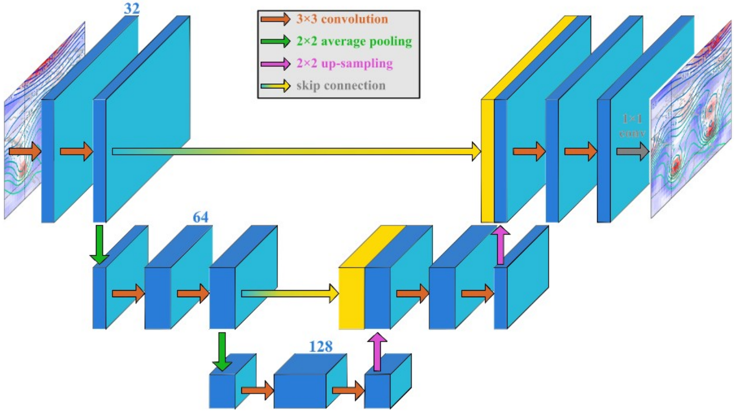


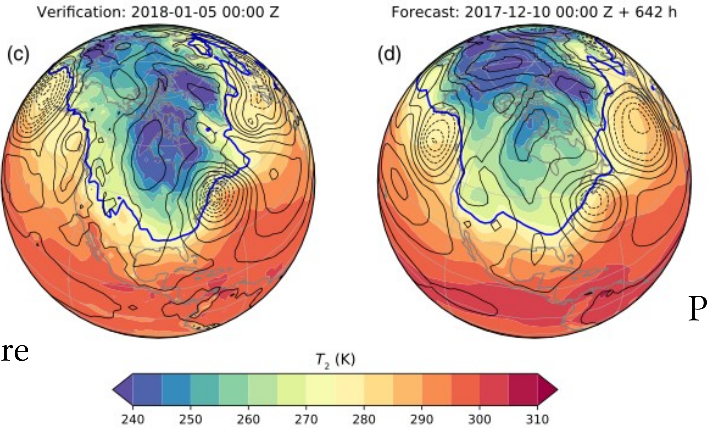


ACC: Anomaly correlation coefficient. Higher ACC means higher skill

On 0-48 lead times, **FourCastNet provides better total precipitation forecasts than IFS**

Brief outline of Weyn et al.'s DLWP



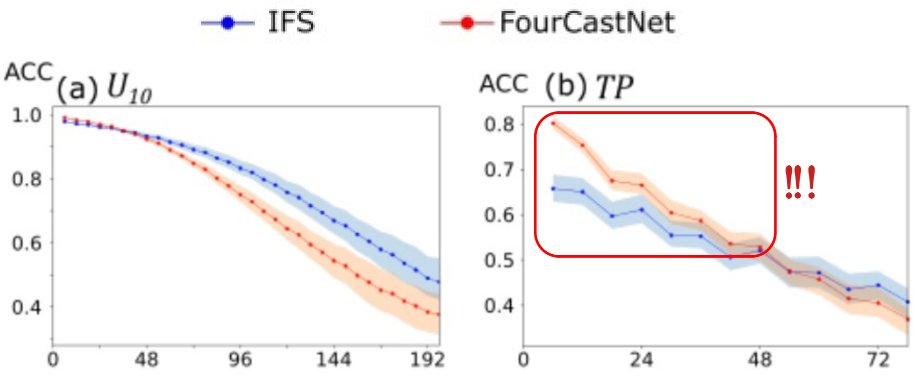


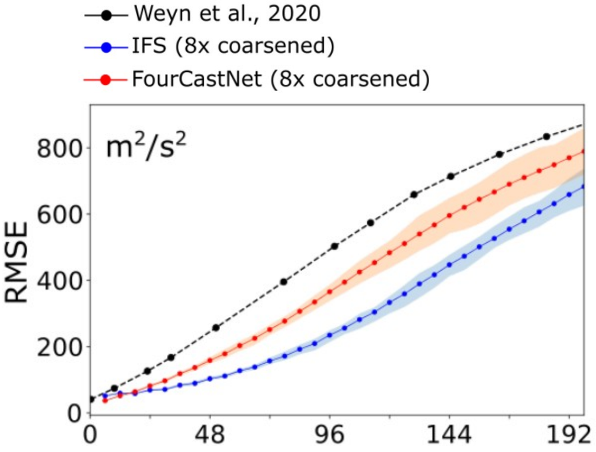
Observed
surface
temperature

Prediction

- Weyn et al (2020)'s Deep Learning Weather Prediction model uses convolutions (FCNN) on a cubed-sphere grid to predict weather on short and even sub-seasonal timescales. It is based on the **U-Net architecture**
- Trained on 2° resolution ERA5 data and four prognostic variables: 500 hPa geopotential height, 1000 hPa geopotential height, 300-700 hPa geopotential thickness, and 2m surface temperature.
- Improved to have 6 predicted variables and produce a 320-ensemble S2S forecast

FourCastNet outperforms Weyn et al.'s DLWP model





ACC: Anomaly correlation coefficient. Higher ACC means higher skill

On 0-48 lead times, **FourCastNet provides better total precipitation forecasts than IFS**

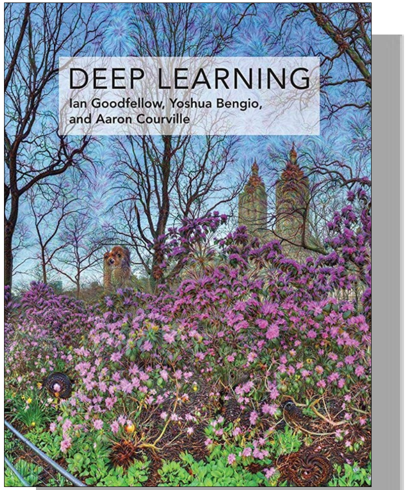
A factor 45,000 speedup and 12,000x lower energy footprint in generating forecasts using a larger ensemble size. 100 ensemble size vs traditional 50 used by ECMWF's IFS.

Scalable: model memory requirements ~10 GB for a batch size=1

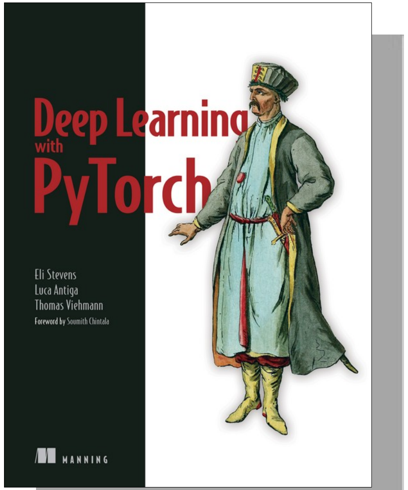
Forecasts using the pre-trained model can be generated using a laptop

Additional Resources for Learning ML

BOOKS



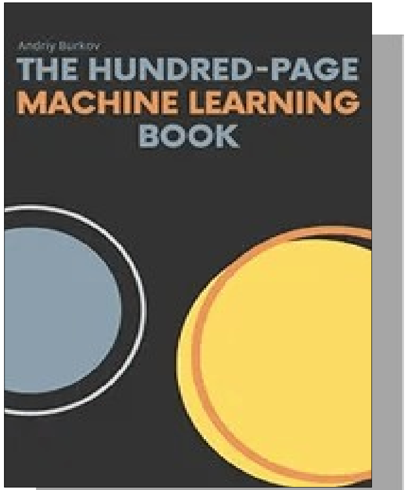
DEEP LEARNING
Ian Goodfellow, Yoshua Bengio,
and Aaron Courville



Deep Learning
with
PyTorch

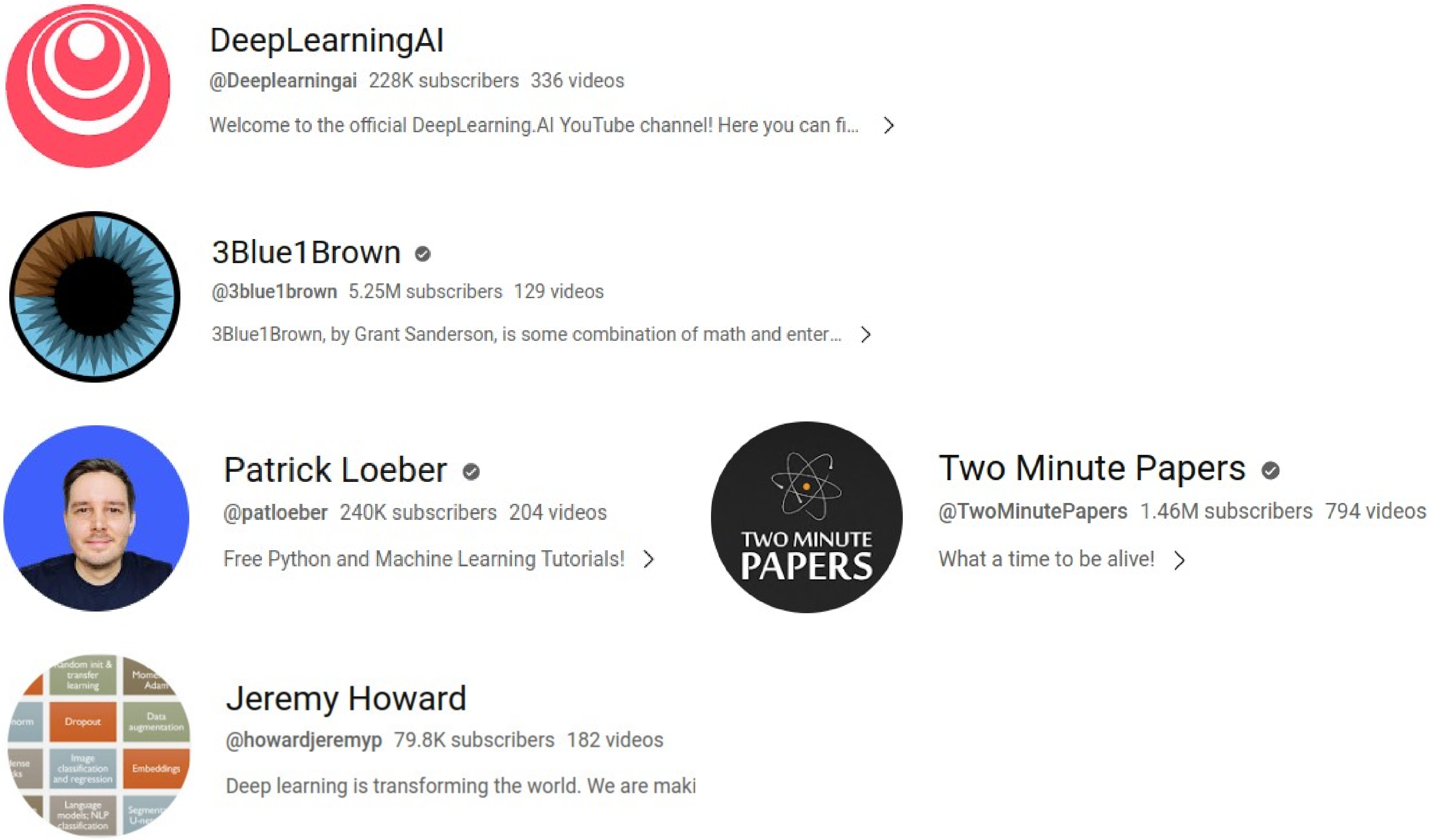
Eli Stevens
Luca Antiga
Thomas Viehmann
Foreword by Soumith Chintala

MANNING



Andriy Burkov
THE HUNDRED-PAGE
MACHINE LEARNING
BOOK

YouTube



DeepLearningAI

@Deeplearningai 228K subscribers 336 videos

Welcome to the official DeepLearning.AI YouTube channel! Here you can fi...

3Blue1Brown

@3blue1brown 5.25M subscribers 129 videos

3Blue1Brown, by Grant Sanderson, is some combination of math and enter...

Patrick Loeber

@patloeber 240K subscribers 204 videos

Free Python and Machine Learning Tutorials!

Two Minute Papers

@TwoMinutePapers 1.46M subscribers 794 videos

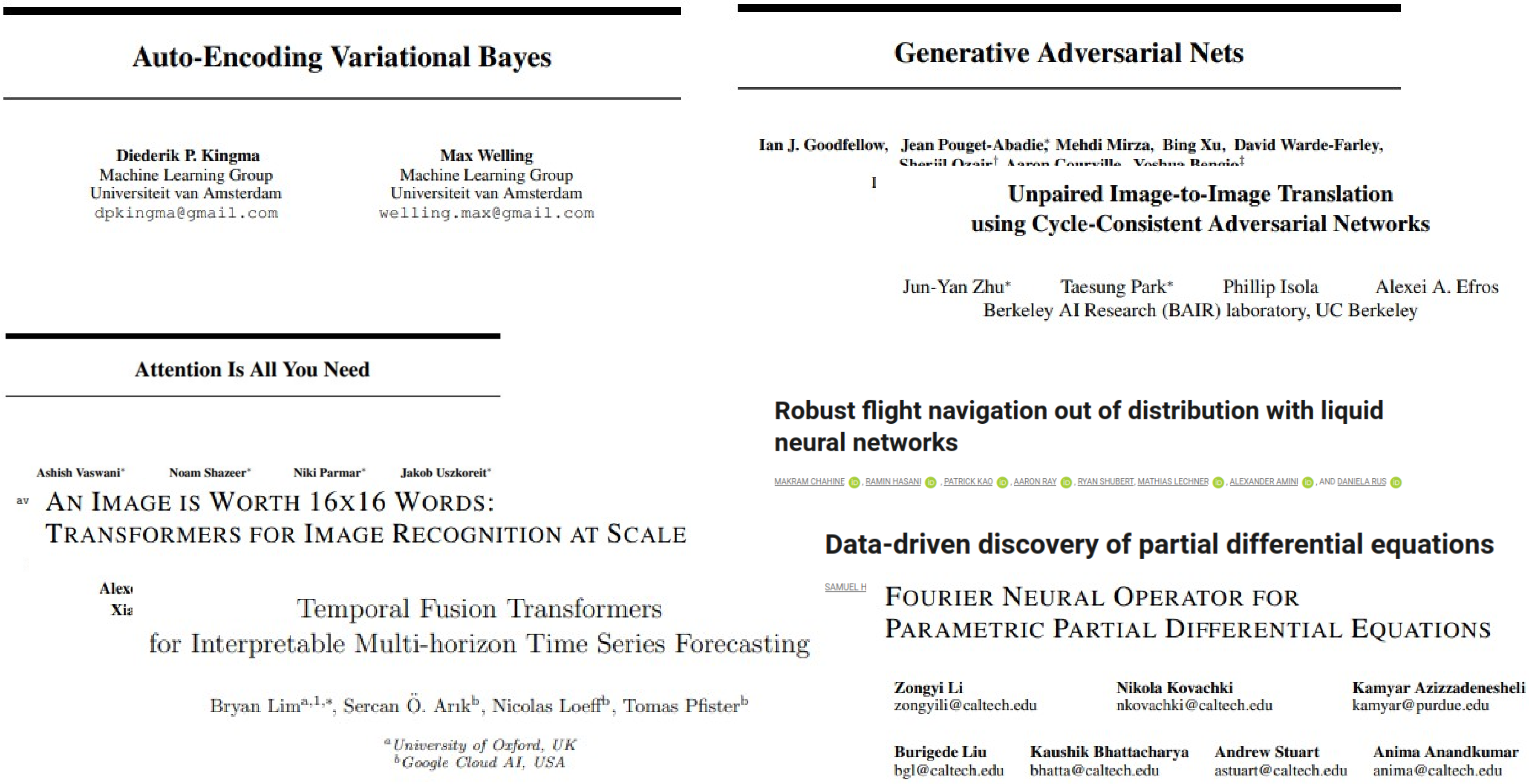
What a time to be alive!

Jeremy Howard

@howardjeremyp 79.8K subscribers 182 videos

Deep learning is transforming the world. We are maki

PAPERS



Auto-Encoding Variational Bayes

Diederik P. Kingma
Machine Learning Group
Universiteit van Amsterdam
dpkingma@gmail.com

Max Welling
Machine Learning Group
Universiteit van Amsterdam
welling.max@gmail.com

Attention Is All You Need

Ashish Vaswani* **Noam Shazeer*** **Niki Parmar*** **Jakob Uszkoreit***

av

AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE

Alex
Xia

Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting

Bryan Lim[a,1,*], Sercan Ö. Arık[b], Nicolas Loeff[b], Tomas Pfister[b]

[a] *University of Oxford, UK*
[b] *Google Cloud AI, USA*

Generative Adversarial Nets

Ian J. Goodfellow, Jean Pouget-Abadie*, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair[†], Aaron Courville, Yoshua Bengio[‡]

I

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

Jun-Yan Zhu* Taesung Park* Phillip Isola Alexei A. Efros
Berkeley AI Research (BAIR) laboratory, UC Berkeley

Robust flight navigation out of distribution with liquid neural networks

MAKRAM CHAHINE, RAMIN HASANI, PATRICK KAO, AARON RAY, RYAN SHUBERT, MATHIAS LECHNER, ALEXANDER AMINI, AND DANIELA RUS

Data-driven discovery of partial differential equations

SAMUEL H

FOURIER NEURAL OPERATOR FOR PARAMETRIC PARTIAL DIFFERENTIAL EQUATIONS

Zongyi Li
zongyili@caltech.edu

Nikola Kovachki
nkovachki@caltech.edu

Kamyar Azizzadenesheli
kamyar@purdue.edu

Burigede Liu
bgl@caltech.edu

Kaushik Bhattacharya
bhatta@caltech.edu

Andrew Stuart
astuart@caltech.edu

Anima Anandkumar
anima@caltech.edu

Data-driven Model Parameterizations

Using ANNs to create global emulators for gravity waves using parameterizationd or nonlocal high resolution climate data, and testing on new climate scenarios

Equation Discovery using ML

Relevance Vector Machines to obtain closed-form solutions for subgrid scale momentum and temperature fluxes for barotropic and baroclinic models that capture a bulk of the subgrid scale variance

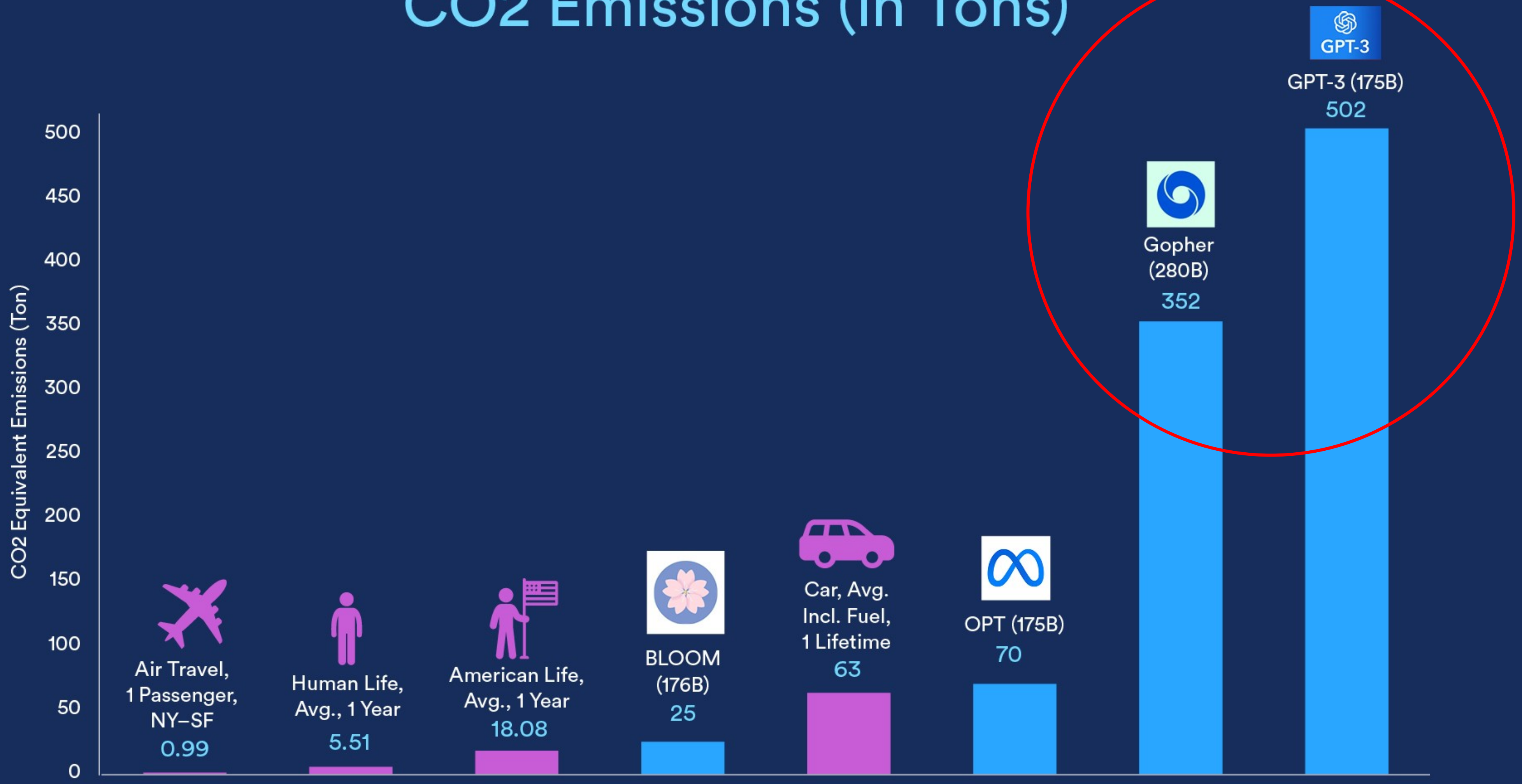
ML in Weather Forecasting

Novel architectures using vision transformer + adaptive Fourier neural operators, and UNet CNN architecture to forecast surface weather with commendable skill

Thank You! I hope this helped. :)

Feel free to reach out: ag4680@stanford.edu

CO2 Emissions (in Tons)



Source: Luccioni et al., 2022; Strubell et al., 2019 | Chart: 2023 AI Index Report