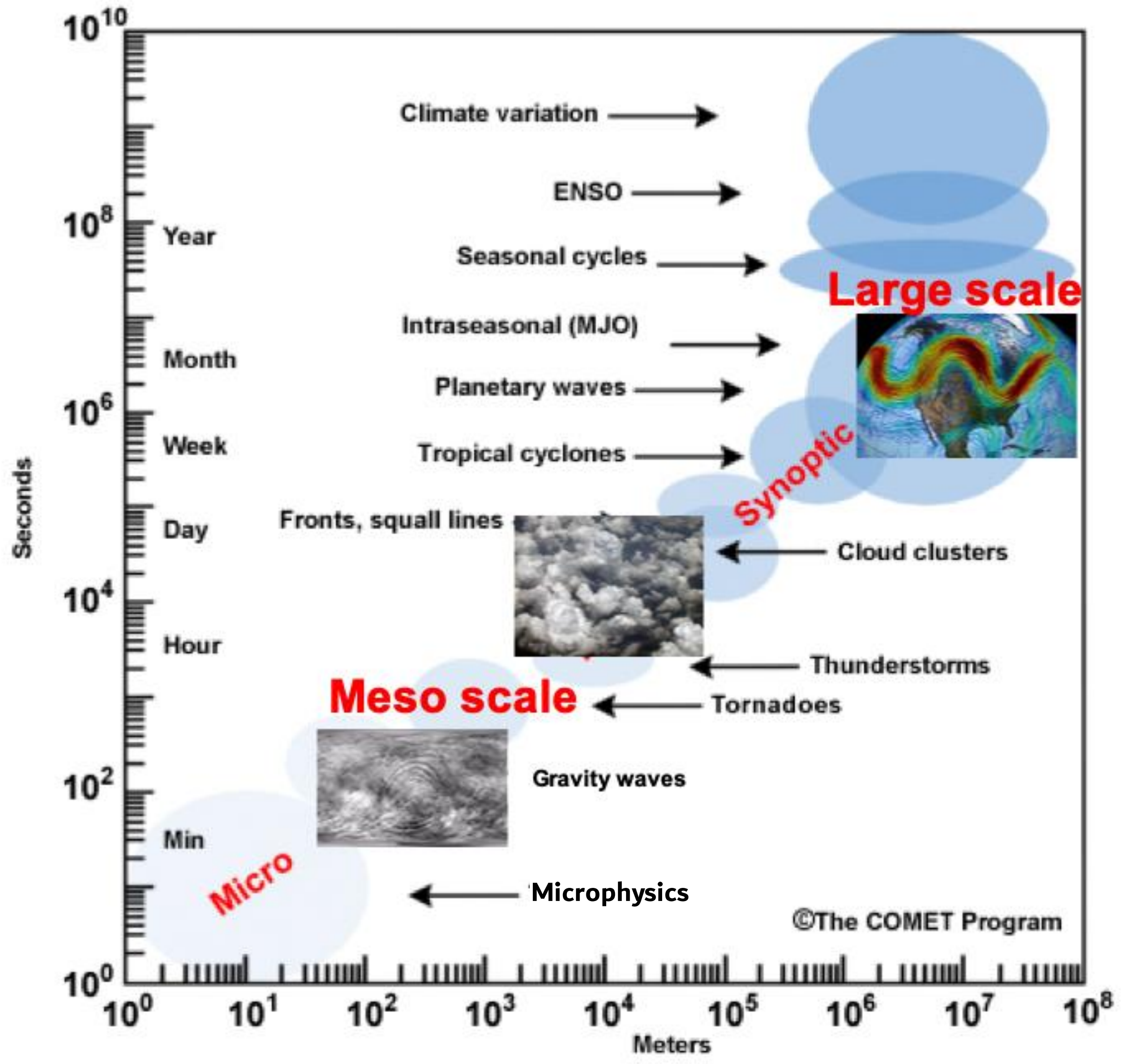
Using transfer learning & backscattering analysis to build ***stable***, ***generalizable***, ***data-driven*** subgrid-scale (SGS) models: A 2D turbulence test case

Pedram Hassanzadeh
Work with: Yifei Guan, Ashesh Chattopadhyay & Adam Subel
Rice University

March 2021

Climate System:

spatio-temporal, *multi-scale*, *multi-physics*, high-dimensional & chaotic …



X: large/slow-scale variables

The main variables of interest

Y: small/fast-scale variables

Influence the spatio-temporal variability of *X*

Traditional approach:

Coarse-resolution numerical solver + physics-based subgrid-scale (SGS) model

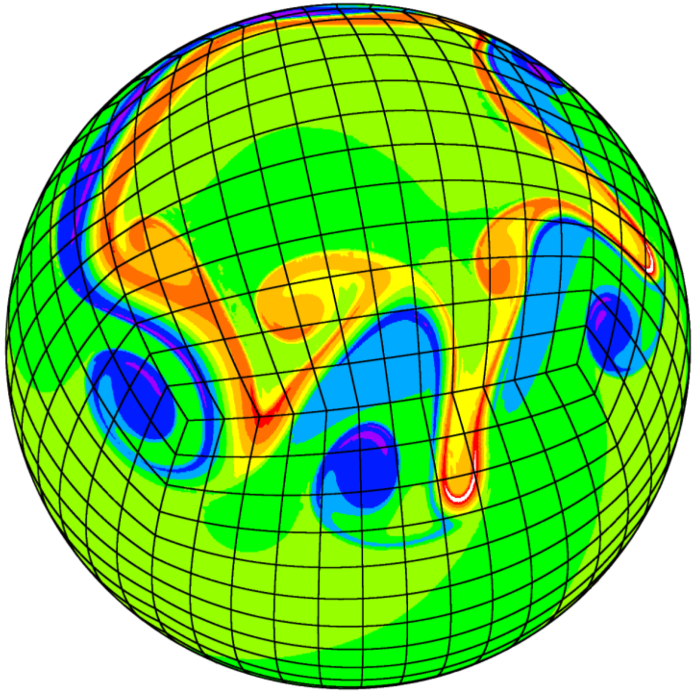
Large-scale processes

$$\dot{\boldsymbol{X}} = \mathbf{F}(\boldsymbol{X})$$

solved numerically at $O(100)$km resolutions

Closure for SGS processes

$$\boldsymbol{Y} = \mathbf{P}(\boldsymbol{X})$$



http://www-personal.umich.edu/~cjablono/

ML-based approach:

Coarse-resolution numerical solver + data-driven subgrid-scale (SGS) model

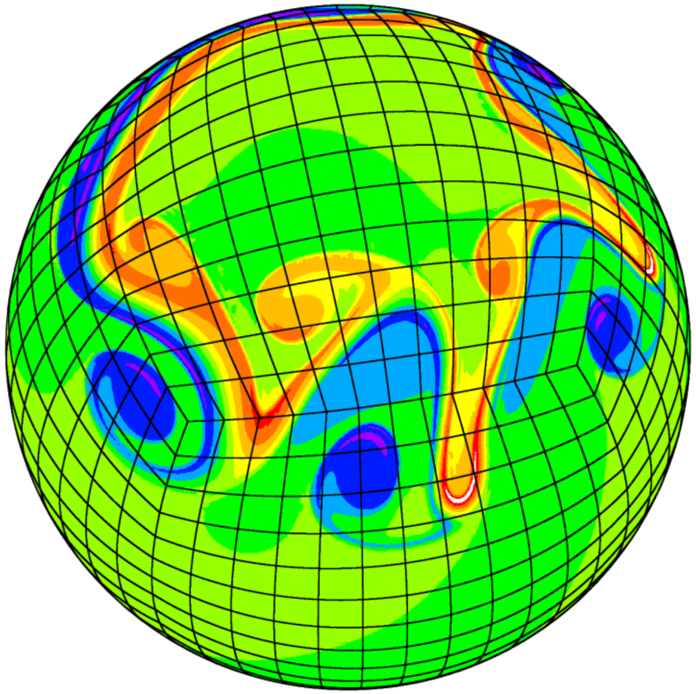
Large-scale processes

$$\dot{\boldsymbol{X}} = \mathbf{F}(\boldsymbol{X})$$

solved numerically at O(100)km resolutions

Data-driven closure for SGS processes

$$\boldsymbol{Y} = \boldsymbol{NN}(\boldsymbol{X})$$



http://www-personal.umich.edu/~cjablono/

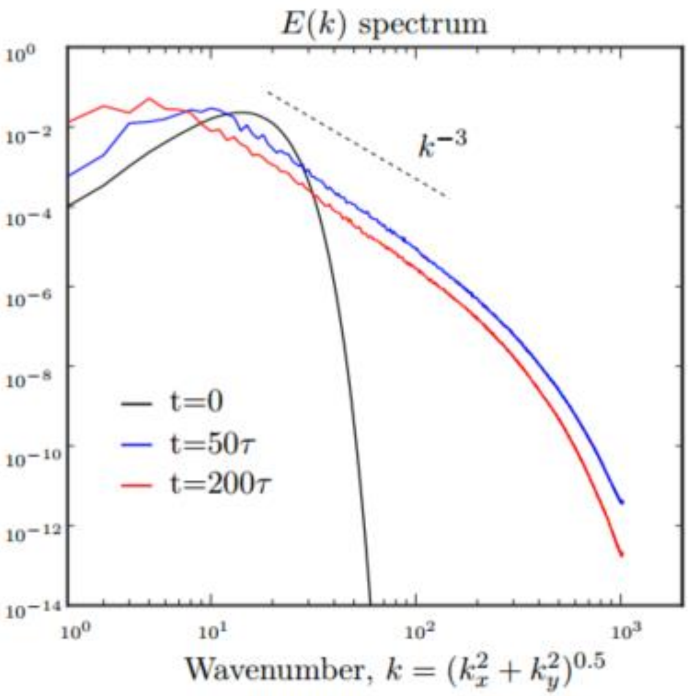
Using ML for weather/climate modeling: Questions, challenges & opportunities

- Best ways to use ML?
- How to choose the ML method?
- Dealing with poor (high-quality) data regimes
- Incorporating physics/PDEs' properties
- Interpretability
- Generalization (i.e., extrapolation)
- Instability: blow-up in coupled (ML+numerical solver) models

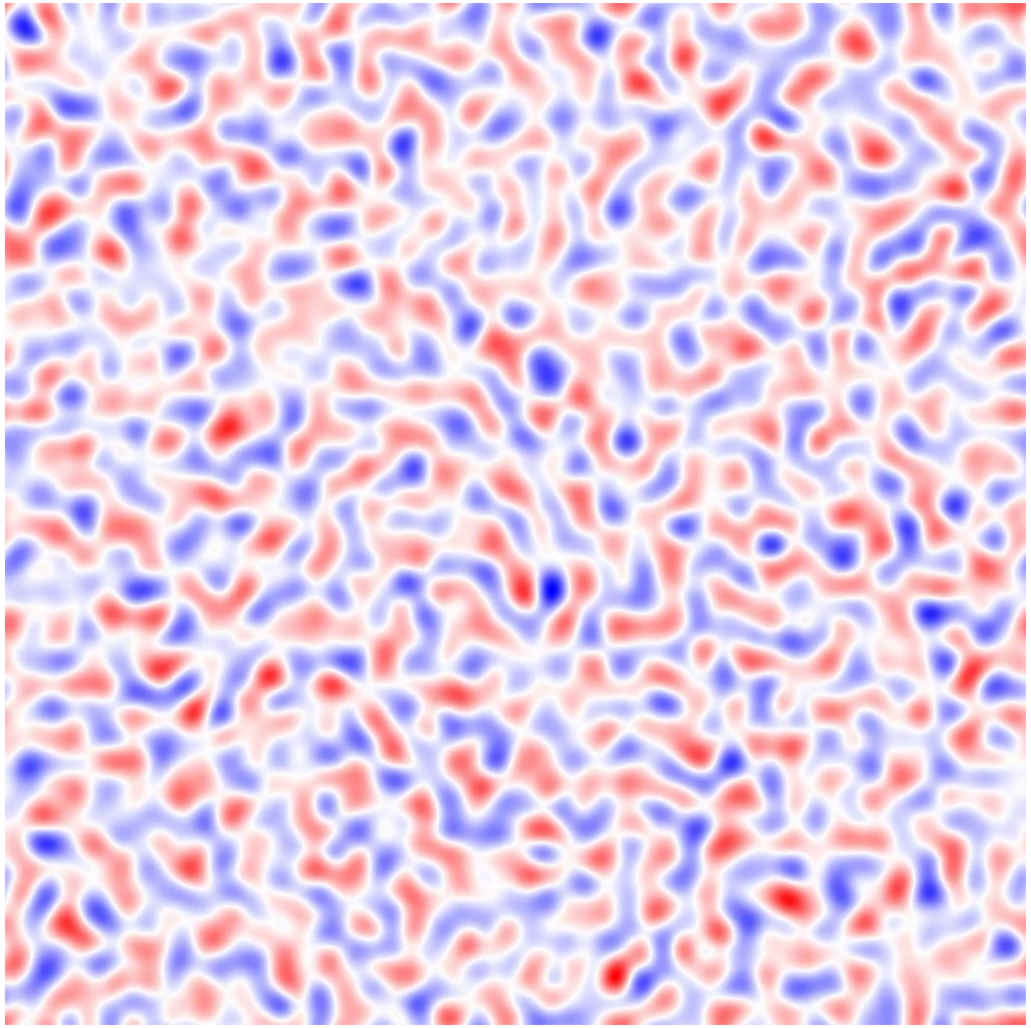
Test case: 2D Turbulence

$$\frac{\partial \omega}{\partial t} + J(\omega, \psi) = \frac{1}{\mathrm{Re}} \nabla^2 \omega$$

$$\nabla^2 \psi = -\omega$$



Direct numerical simulation (DNS)
Re=32000; grid=2048 x 2048



Large-Eddy Simulation (LES)

$$\frac{\partial \omega}{\partial t} + J(\omega, \psi) = \frac{1}{\mathrm{Re}} \nabla^2 \omega$$

$$\frac{\partial \overline{\omega}}{\partial t} + J(\overline{\omega}, \bar{\psi}) = \frac{1}{\mathrm{Re}} \nabla^2 \overline{\omega} + \underbrace{\left[J(\overline{\omega}, \bar{\psi}) - \overline{J(\omega, \psi)}\right]}_{\Pi:\ \text{SGS term}}$$

Guan, Chattopadhyay, Subel & Hassanzadeh, *Stable a posteriori LES of 2D turbulence with convolutional neural networks: backscattering analysis and generalization to higher Re via transfer learning*, under review arXiv: 2102.11400

Re=32000

DNS grid = 2048 x 2048 , time step = Δt

LES grid = 256 x 256, time step = $10\Delta t$

physics-based parameterization: Smagorinsky's model (1963) $\Pi = \nu_e \nabla^2 \overline{\omega}$

data-driven parameterization (DD-P): $\Pi = NN(\overline{\omega}, \bar{\psi})$

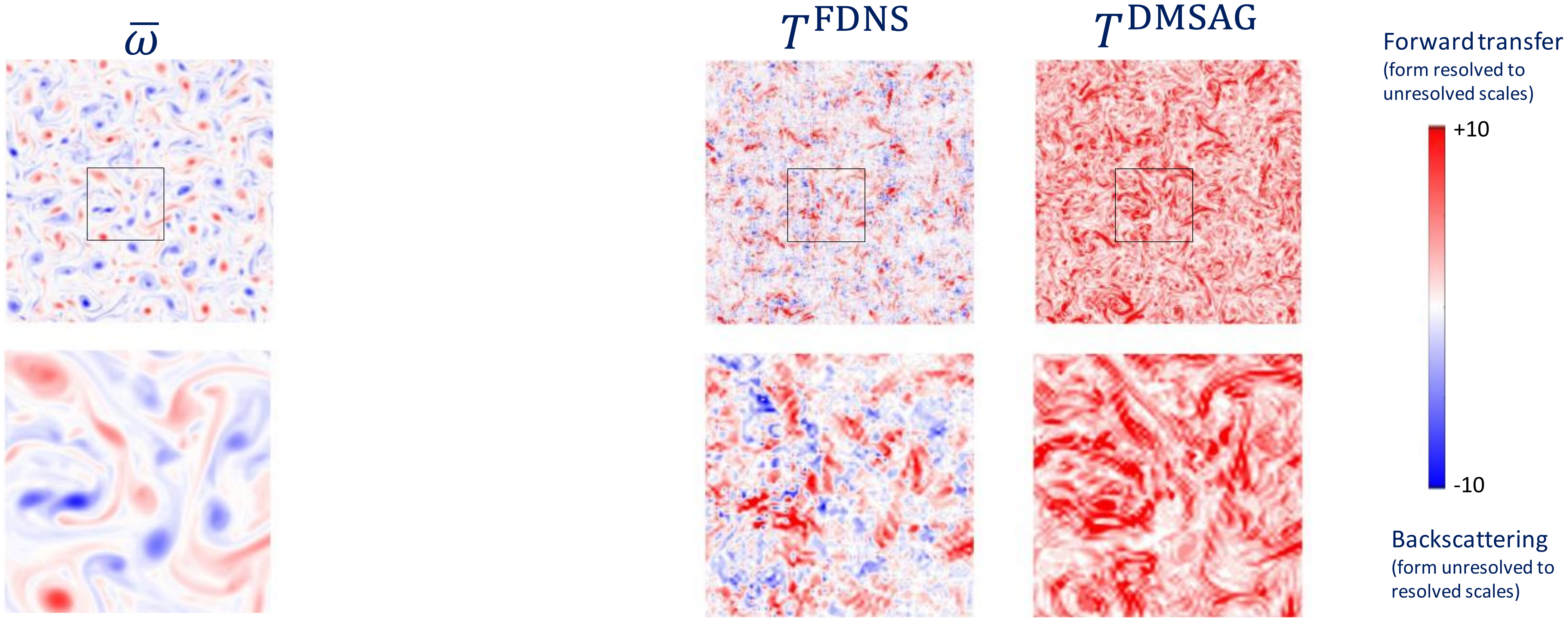
Major shortcoming of many physics-based models: Only diffusive, not accounting for *backscattering*

$T = \Pi\, \nabla^2 \overline{\omega}$

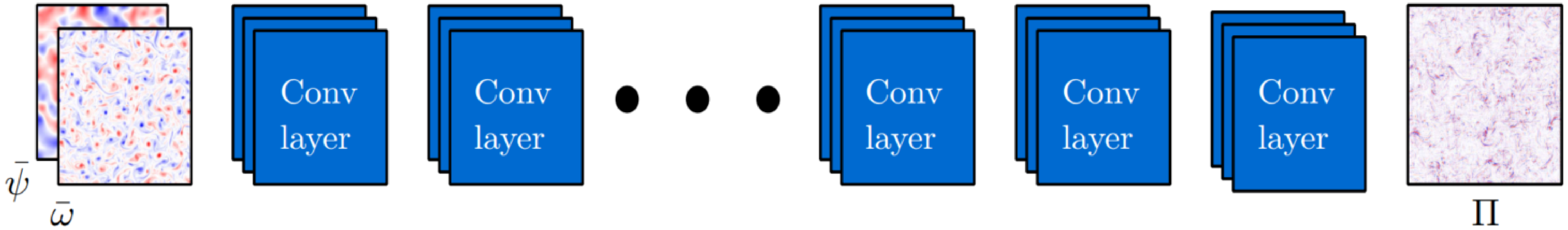
T: subgrid-scale transfer

$T^{\mathrm{DSMAG}} = \upsilon_e\, \nabla^2 \overline{\omega}\, \nabla^2 \overline{\omega} \geq 0$

(Dynamic Smagorinsky, Germano et al. 1991)



Non-local DD-P using CNNs



10 layers (64 filters, 5 x 5) + ReLU + no pooling/upsampling

Training dataset:

From 7 DNS runs started from random initial conditions

Validation dataset:

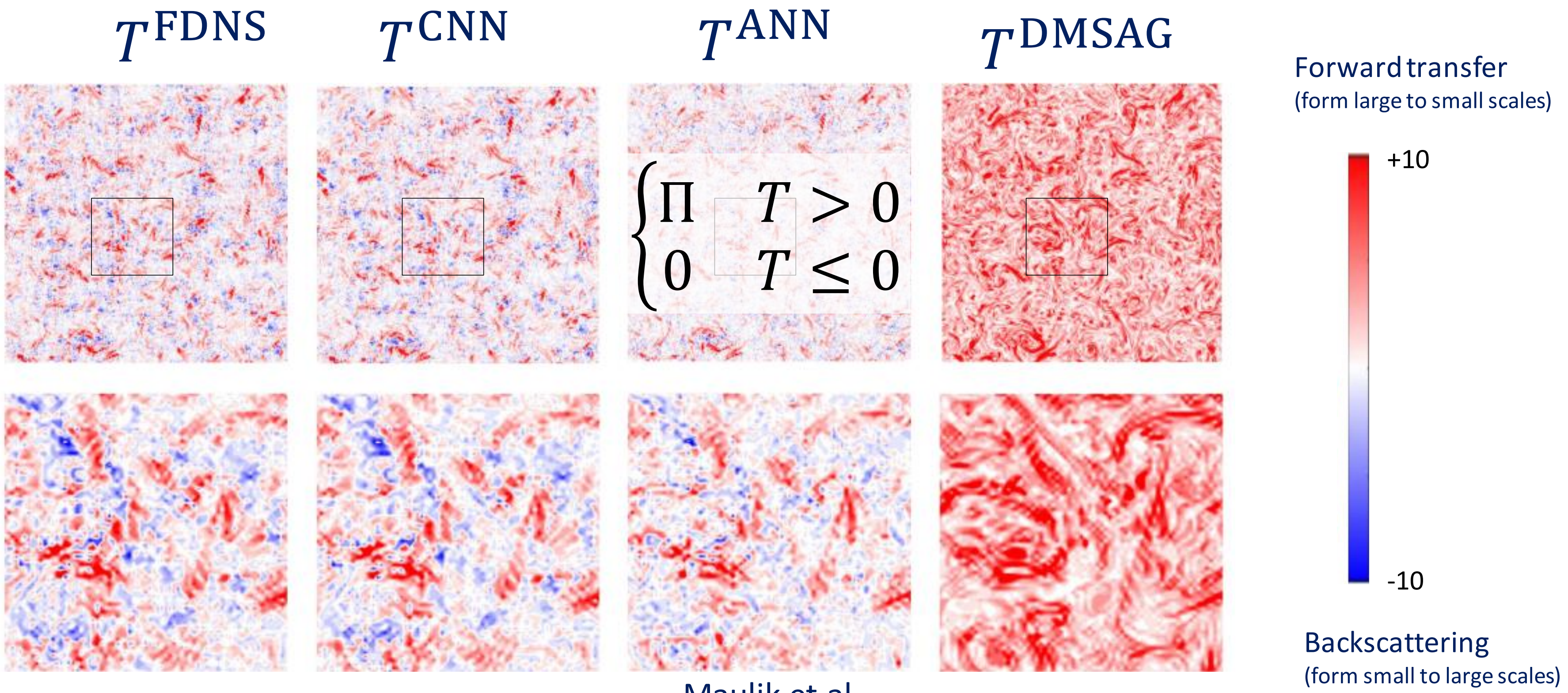
From 3 DNS runs started from random initial conditions

Testing dataset:

From 5 DNS runs started from random initial conditions

A priori (offline) test of DD-P

“online ≠ offline”
Stephan Rasp



Maulik et al.
2019 JFM

	SMAG	DSMAG	ANN	CNN
Correlation coefficient c	0.55	0.55	0.86	0.93

Stability of a posteriori (coupled) LES model?

A priori accuracy of the CNN-based DD-P & the fate of coupled LES run
as a function of the numer of training samples, N

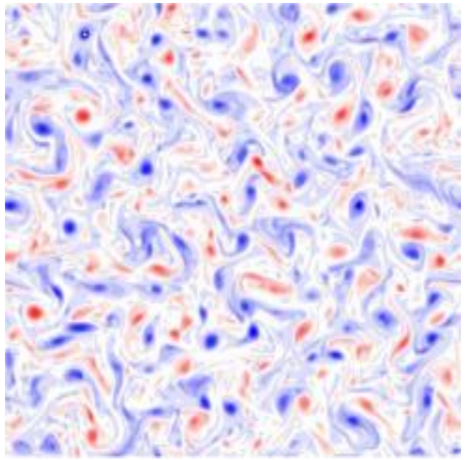
N	500	1000	10000	30000	50000
c	0.78	0.83	0.90	0.92	0.93
$c_{T>0}$: diffusion	0.84	0.89	0.93	0.93	0.96
$c_{T<0}$: backscatter	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Fate	Unstable	Unstable	Unstable	Stable	Stable

- Backscattering is harder to learn data drivenly when the training set is small

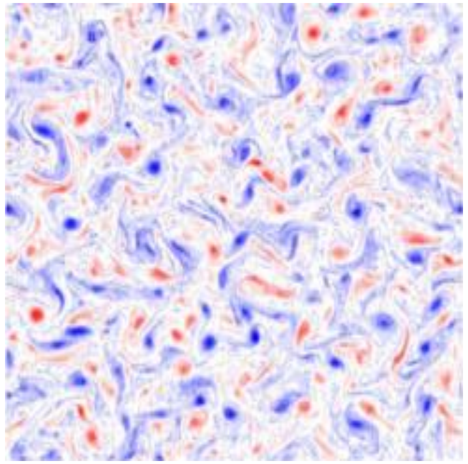
- Speculation: Disproportionally low accuracy for backscattering is the reason for instabilities

Accuracy of *a posteriori* (online) LES with DD-P

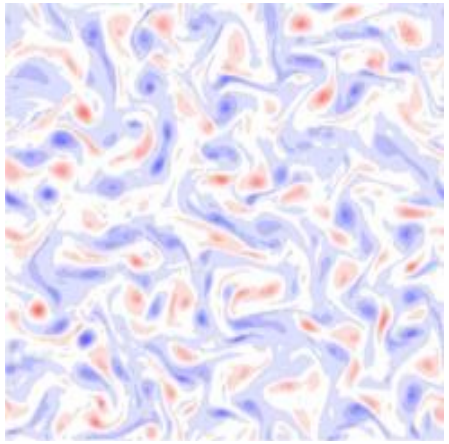
FDNS



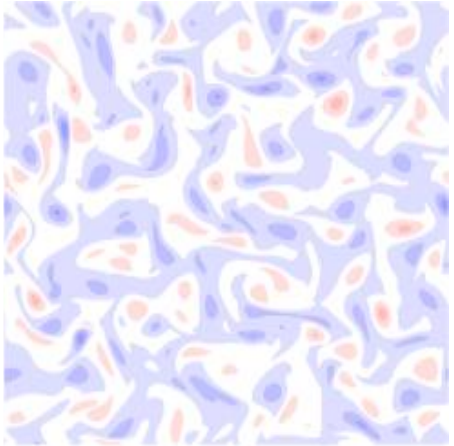
LES-CNN

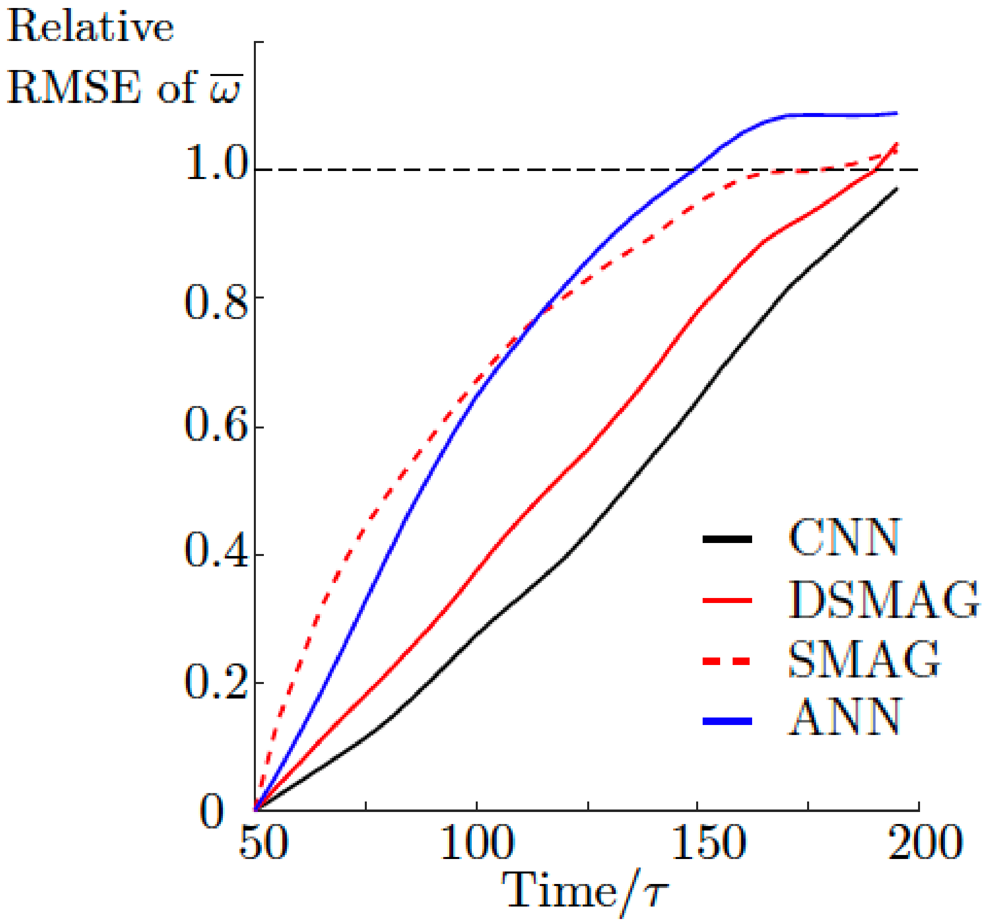


LES-ANN

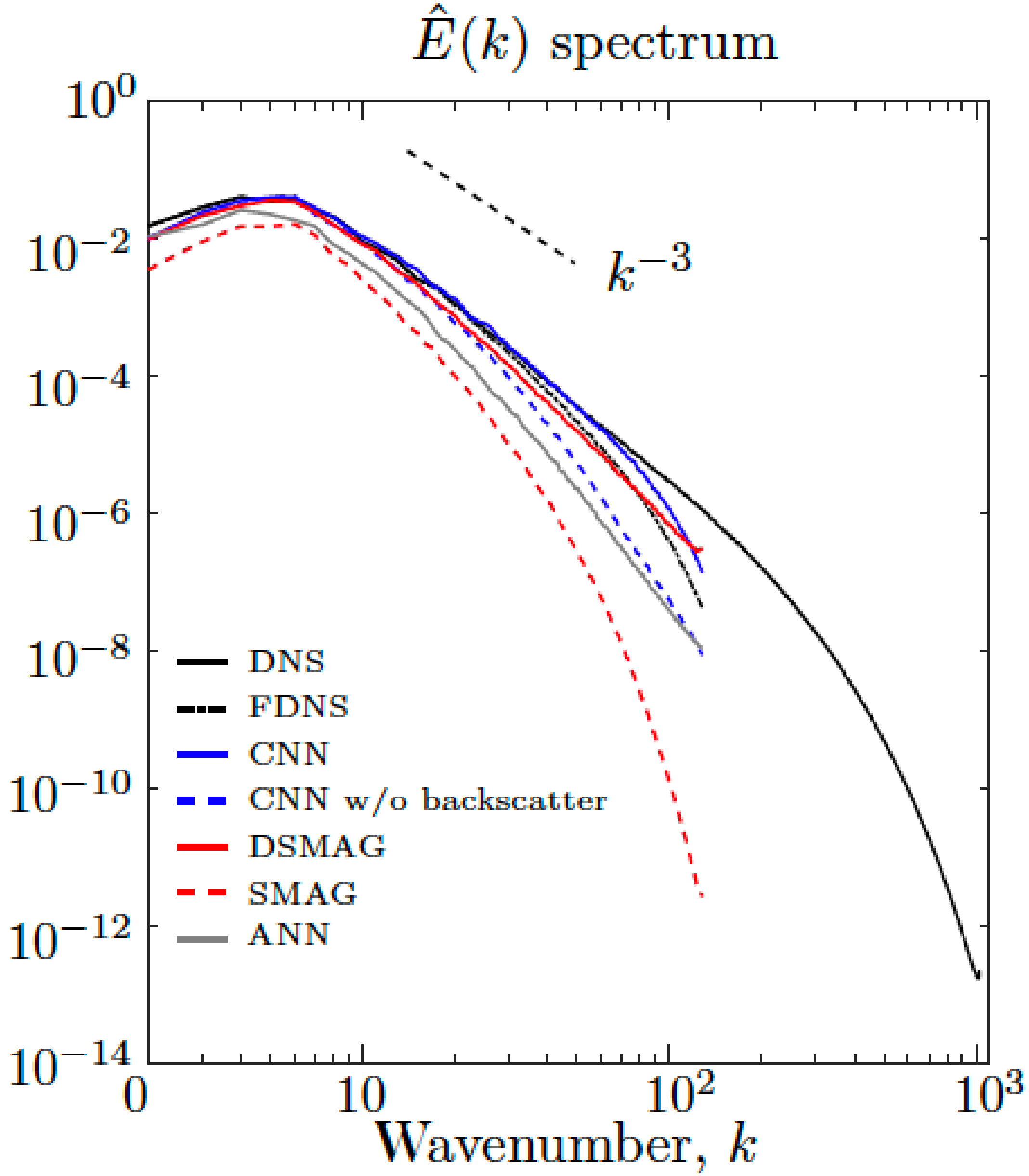


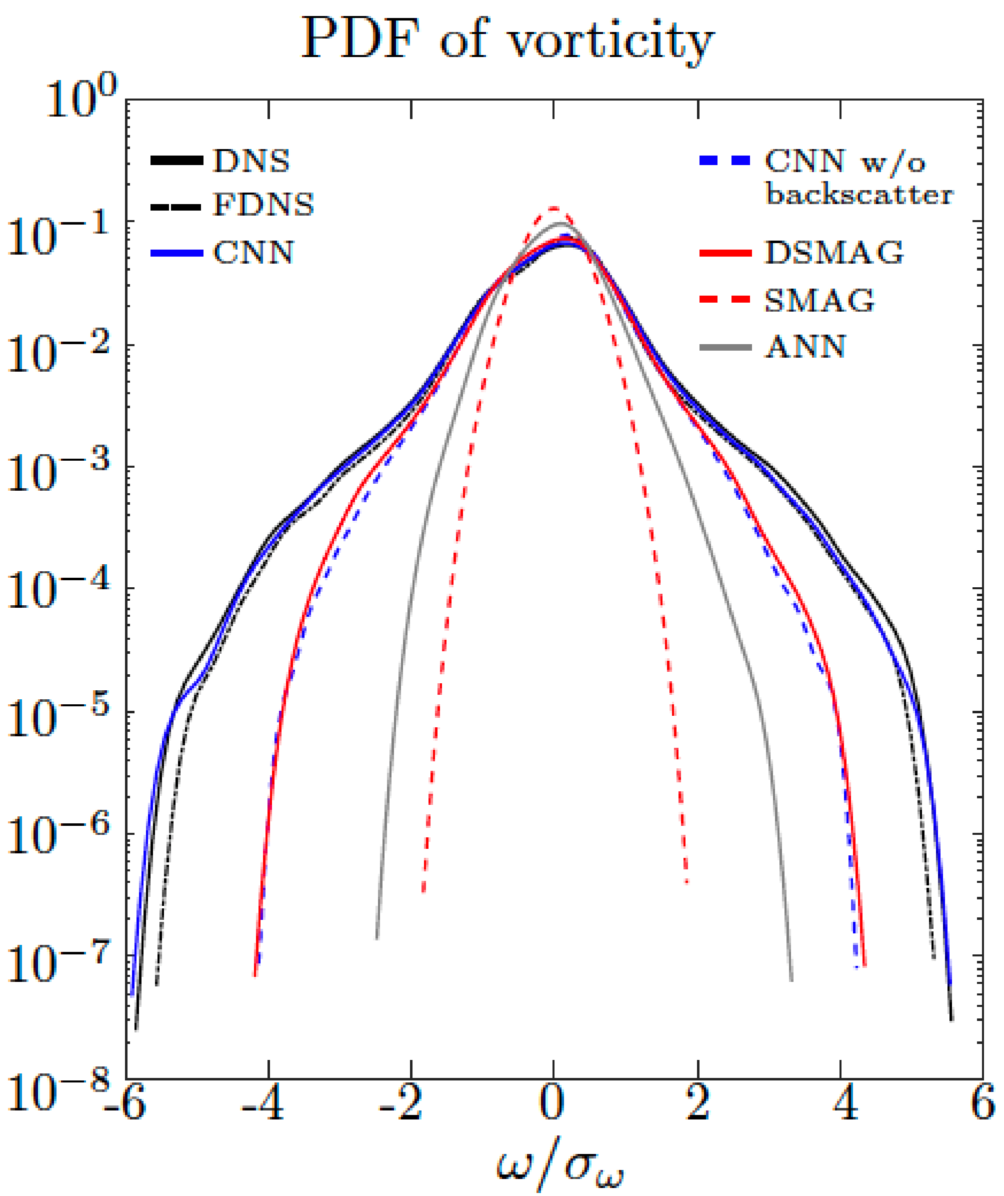
LES-SMAG



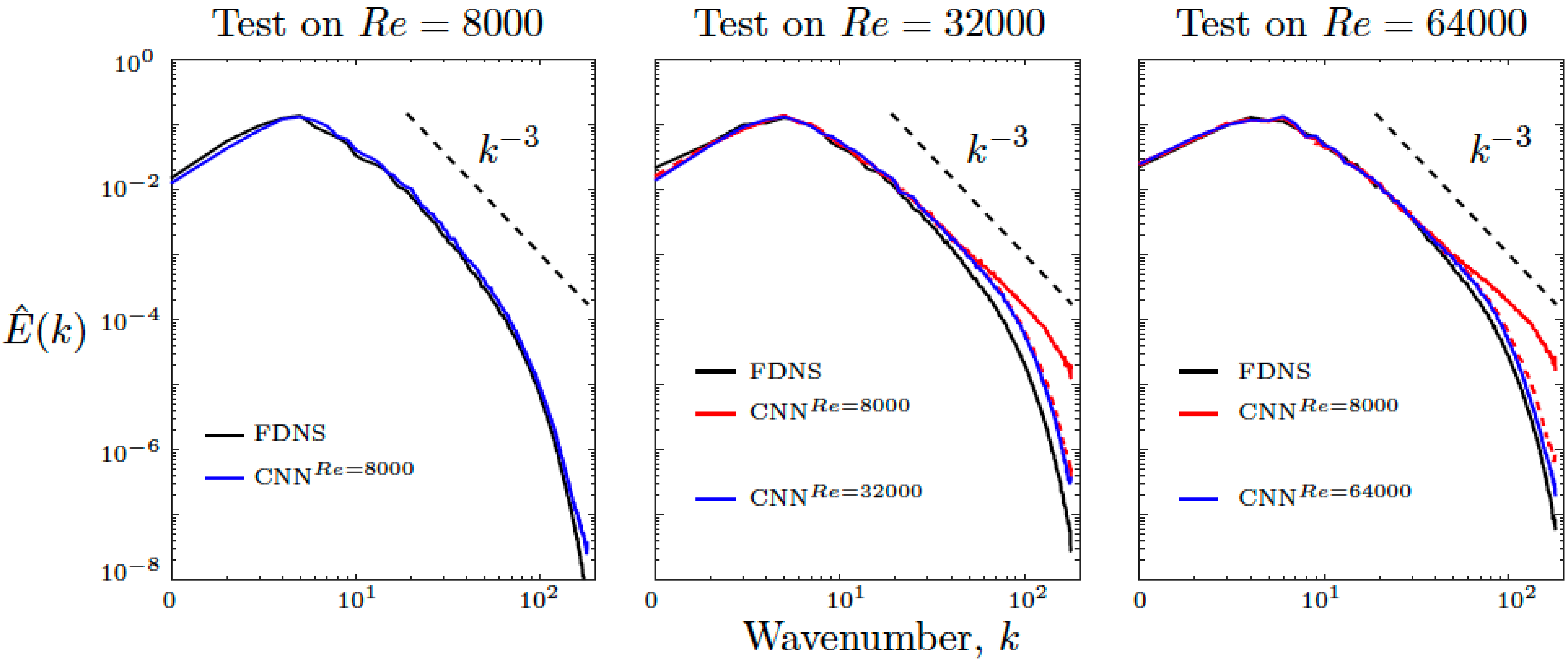


Accuracy of *a posteriori* (online) LES with DD-P



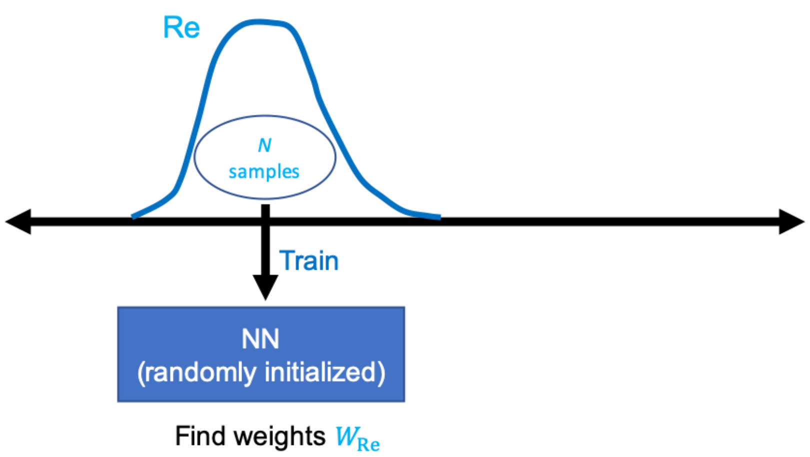


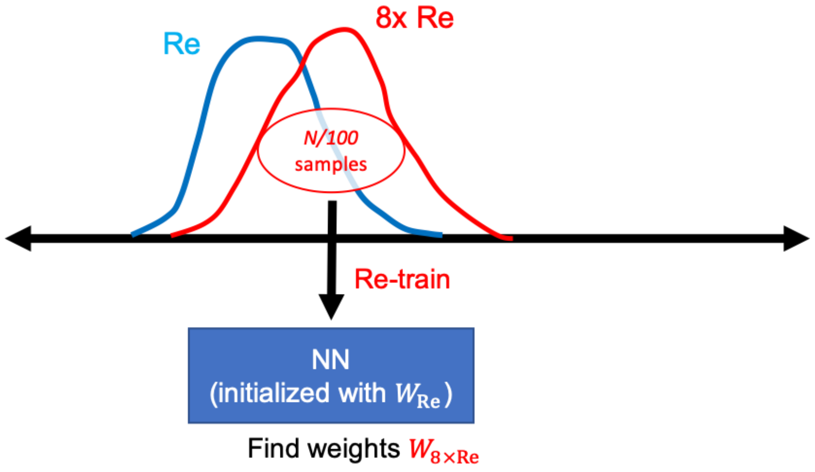
DD-P does not generalize to higher *Re*



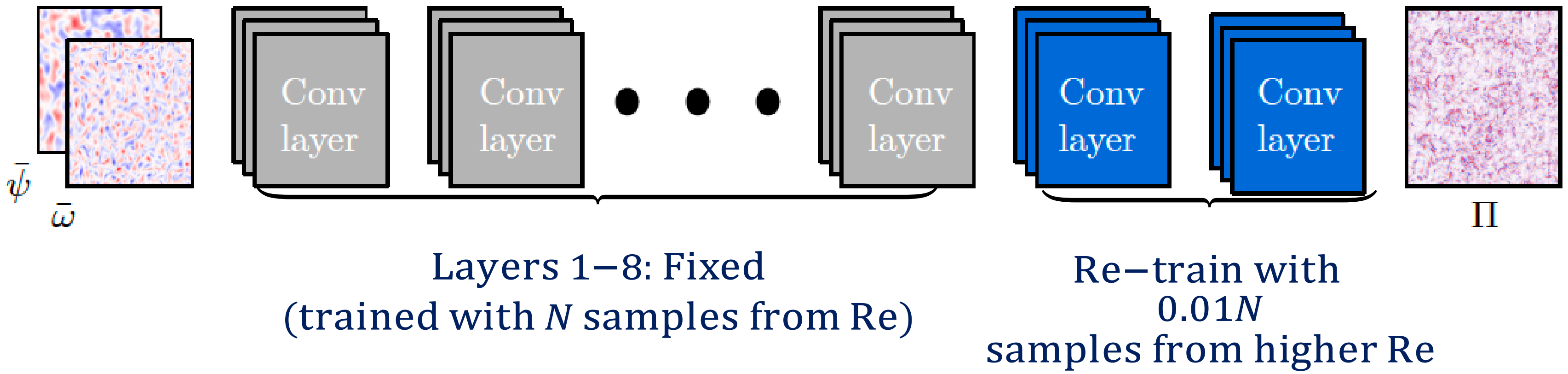
LES resolution: 256 x 256

Generalization to higher *Re* via transfer learning

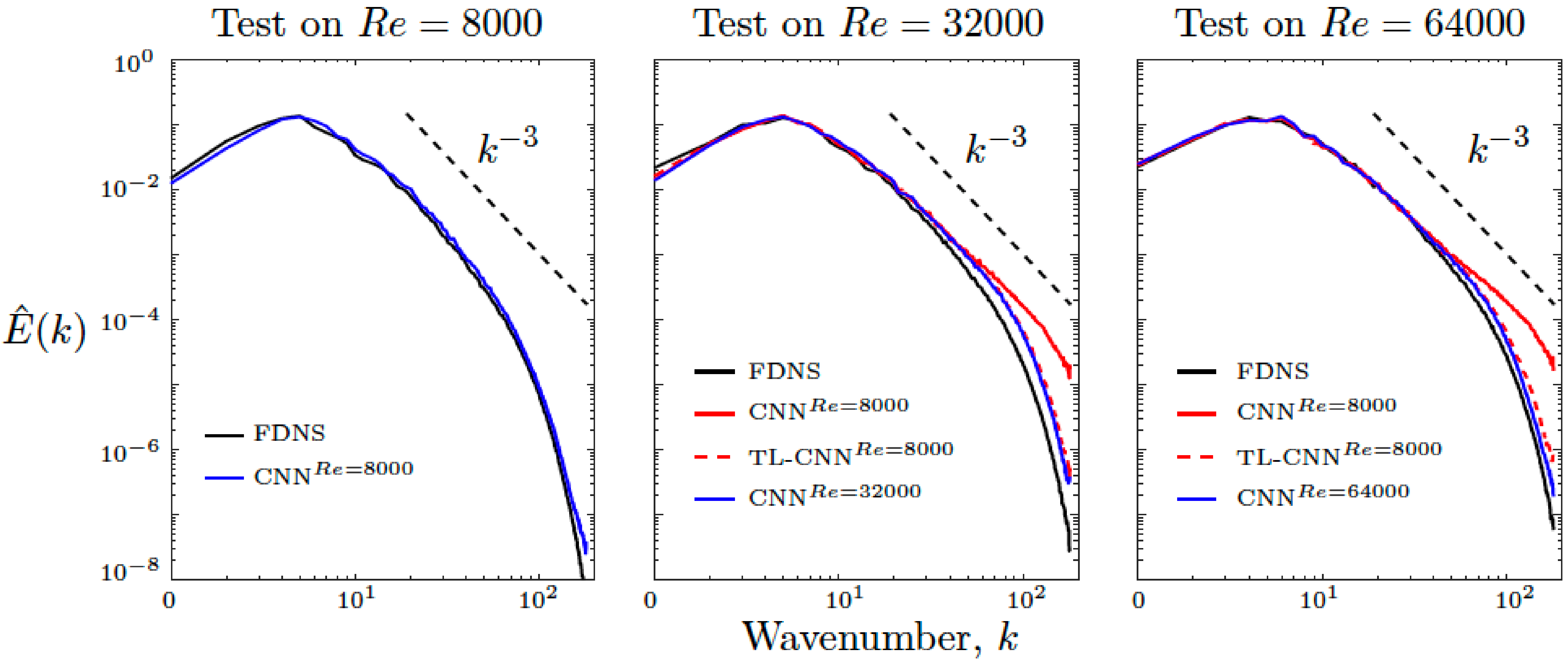




Chattopadhyay, Subel & Hassanzadeh, *Data-driven super-parameterization using deep learning: Experimentation with multi-scale Lorenz 96 systems and transfer learning*. J. Advances in Modeling Earth Systems (2020)

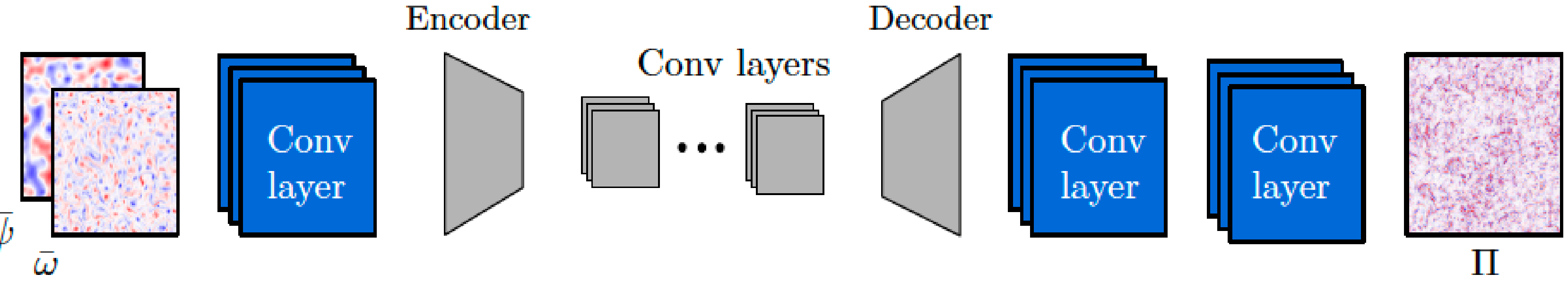


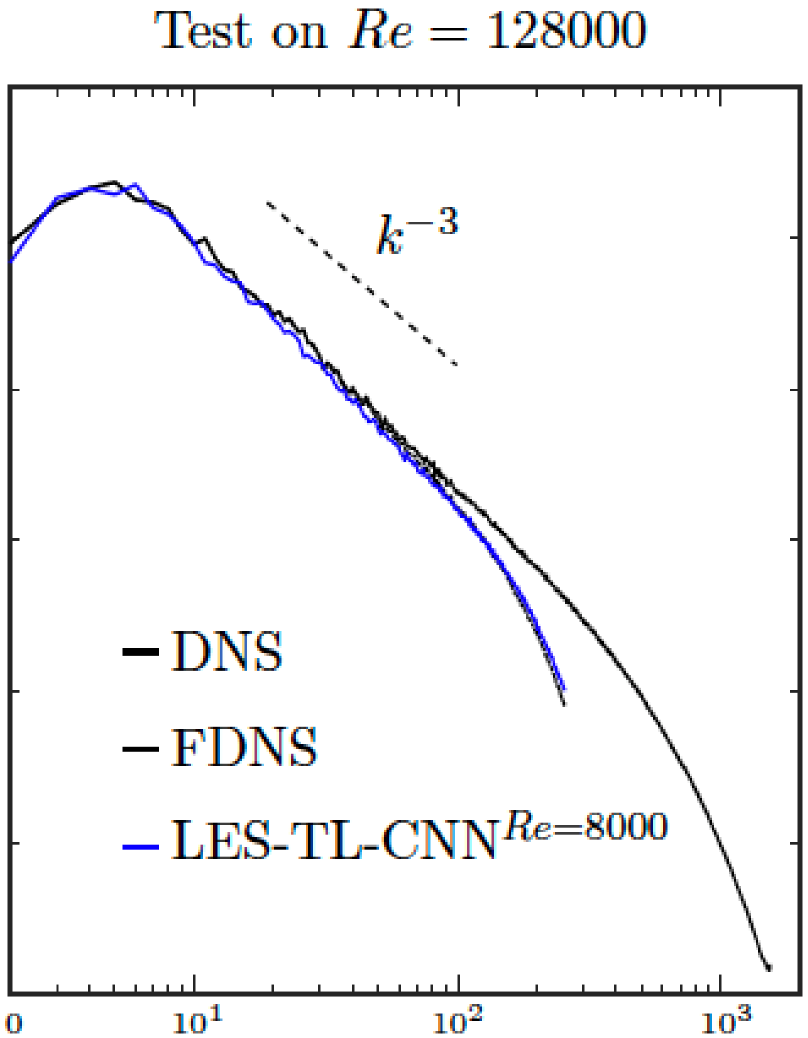
Generalization to higher *Re* via transfer learning



LES resolution: 256 x 256

Generalization to higher *Re* & different grid resolution via transfer learning + auto-encoder





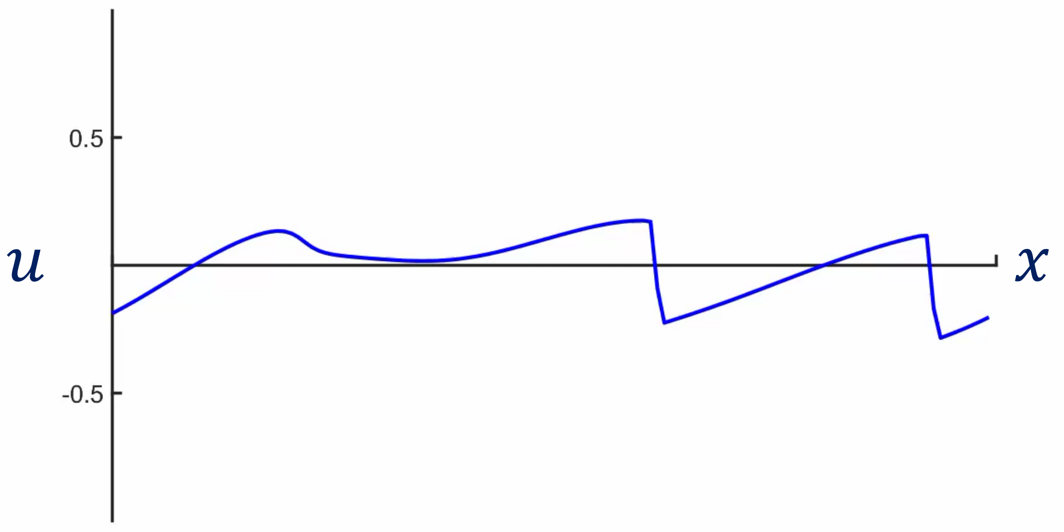
	Re	*N*	grid resolution
Base CNN	8000	50,000	256 x 256
Transfer learned CNN	128000	500	512 x 512

1D Stochastically forced Burgers turbulence

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial uu}{\partial x} = \frac{1}{\mathrm{Re}}\frac{\partial^2 u}{\partial x^2} + f(x,t)$$

$$\frac{\partial \bar{u}}{\partial t} + \frac{1}{2}\frac{\partial \bar{u}\bar{u}}{\partial y} = \frac{1}{\mathrm{Re}}\frac{\partial^2 \bar{u}}{\partial y^2} + \bar{f} + \Pi(y)$$

$$\Pi(y) = \mathrm{ANN}(\bar{u})$$



Subel, Chattopadhyay, Guan & Hassanzadeh, *Data-driven subgrid-scale modeling of forced Burgers turbulence using deep learning with generalization to higher Reynolds numbers via transfer learning*, Physics of Fluids (2021)

- **Stable** *a posteriori* LES after data augmentation
- **Generalization** to 10x Re using transfer learning

Stable, accurate & generalizable SGS modeling for LES

Takeaway:

- **Stability: *might* require large training sets**
- **Transfer learning: large training sets required only from a base system**

- Reduce the required size of the training set
  - Data augmentation
  - Include physics constraints
- Better understanding of the relationship between a priori accuracy & a posteriori stability
- Online training?
- Add memory to the SGS model
- Further explore the power of transfer learning (e.g., between setups)
- More complex test cases

PHILOSOPHICAL TRANSACTIONS A

royalsocietypublishing.org/journal/rsta

Review

Check for updates

Cite this article: Kashinath K *et al.* 2021 Physics-informed machine learning: case studies for weather and climate modelling. *Phil. Trans. R. Soc. A* **379**: 20200093.
https://doi.org/10.1098/rsta.2020.0093

Accepted: 24 November 2020

Physics-informed machine learning: case studies for weather and climate modelling

K. Kashinath[1], M. Mustafa[1], A. Albert[1,2], J-L. Wu[1,3], C. Jiang[1,4], S. Esmaeilzadeh[5], K. Azizzadenesheli[6], R. Wang[1,7], A. Chattopadhyay[1,8], A. Singh[1,2], A. Manepalli[1,2], D. Chirila[9], R. Yu[7], R. Walters[10], B. White[2], H. Xiao[11], H. A. Tchelepi[5], P. Marcus[4], A. Anandkumar[3,12], P. Hassanzadeh[8] and Prabhat[1]

Papers on data-driven forecasting

http://pedram.rice.edu/publications/

Chattopadhyay A., Nabizadeh E. & Hassanzadeh P., Analog forecasting of extreme-causing weather patterns using deep learning, *Journal of Advances in Modeling Earth Systems*, 2020

Chattopadhyay A., Mustafa M., Hassanzadeh P., Bach E. & Kashinath K., Towards physically consistent data-driven weather forecasting: Integrating data assimilation with equivariance-preserving deep spatial transformers, *under review at Geoscientific Model Development*

Chattopadhyay A., Hassanzadeh P. & Subramanian D., Data-driven prediction of a multi-scale Lorenz 96 chaotic system using machine learning methods: Reservoir computing, artificial neural network, and long short-term memory network, *Nonlinear Processes in Geophysics*, 2020