Learning non-linear spatio-temporal dynamics with convolutional Neural ODEs

Carnegie Mellon University

Varun Shankar[1], Gavin D. Portwood[2], Arvind T. Mohan[2], Peetak P. Mitra[3], Christopher Rackauckas[4], Lucas A. Wilson[5], David Schmidt[3], Venkat Viswanathan[1]

[1]Carnegie Mellon University, [2]Los Alamos National Laboratory, [3]University of Massachusetts, [4]Massachusetts Institute of Technology, [5]Dell Technologies

NEURAL INFORMATION PROCESSING SYSTEMS

Motivation

- Turbulent flow modeling is still challenging
  - Multi-scale, non-linear, non-local
- Turbulence models to approximate the small scales
  - Still expensive for many practical applications
- Can we use physics-informed data-driven approach to learn flow dynamics
  - Goal: model that is much cheaper than conventional techniques
  - Retain important features of interest

Background

Governing Equations

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = \frac{1}{Re}\nabla^2 \mathbf{u} - \nabla p + \mathbf{f}; \;\; \nabla \cdot \mathbf{u} = 0$$

- Approximate dynamics of discretized velocity field with neural network

$$\frac{\partial \mathbf{u}}{\partial t} = f(\nabla, \mathbf{u}, \ldots)$$

$$\frac{d\boldsymbol{u}}{dt} = g_\theta(\nabla, \boldsymbol{u}, \ldots)$$

Machine Learning Concepts

- Neural ODE
  - Derivative approx. w/ NN
  - Backwards ODE to calculate parameter gradients

$$\frac{d\boldsymbol{u}}{dt} = g_\theta(\boldsymbol{u})$$

- Convolution gradients
  - Dynamics are a function of spatial grads
  - We already have a tool for calculating this

$$\frac{d^2\phi}{dx^2} \approx \frac{\phi_{i+1} - 2\phi_i + \phi_{i-1}}{\Delta x^2}$$

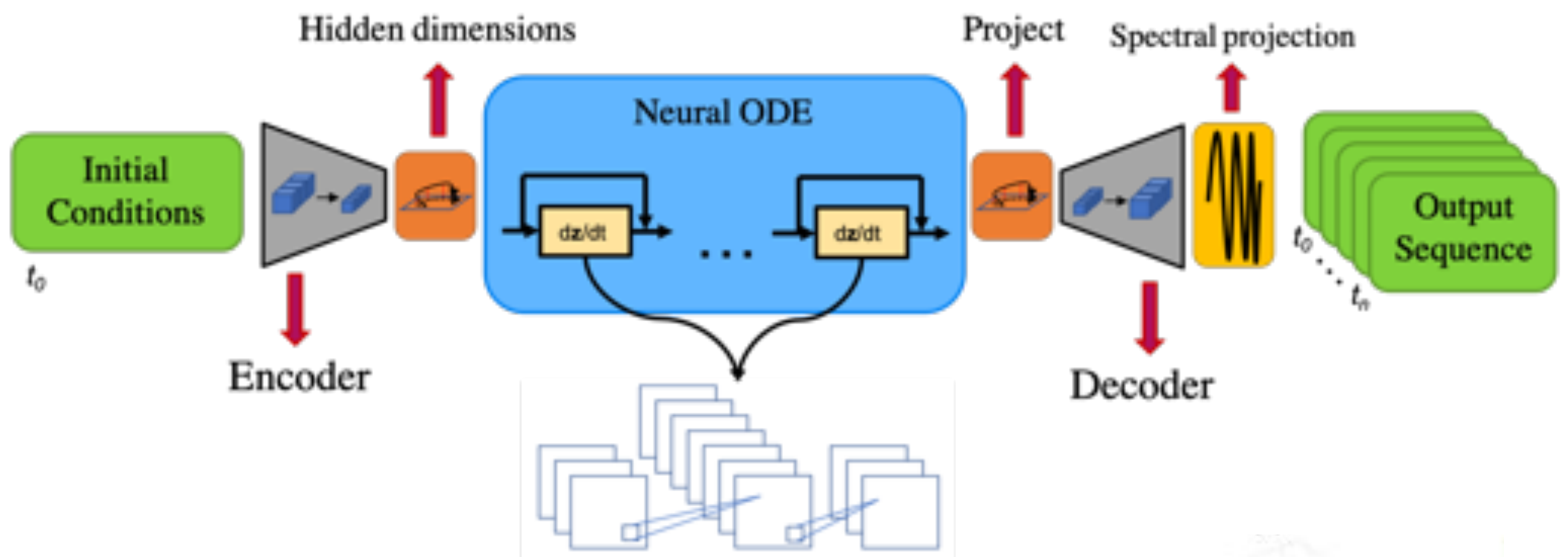
[1 -2 1]

1D kernel

- g_θ is a CNN

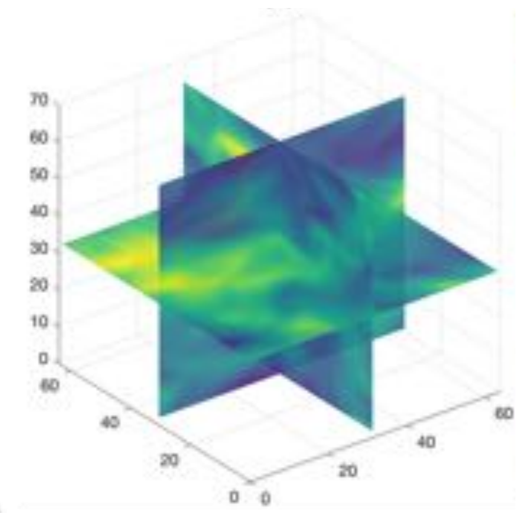
Methods

Overall Architecture

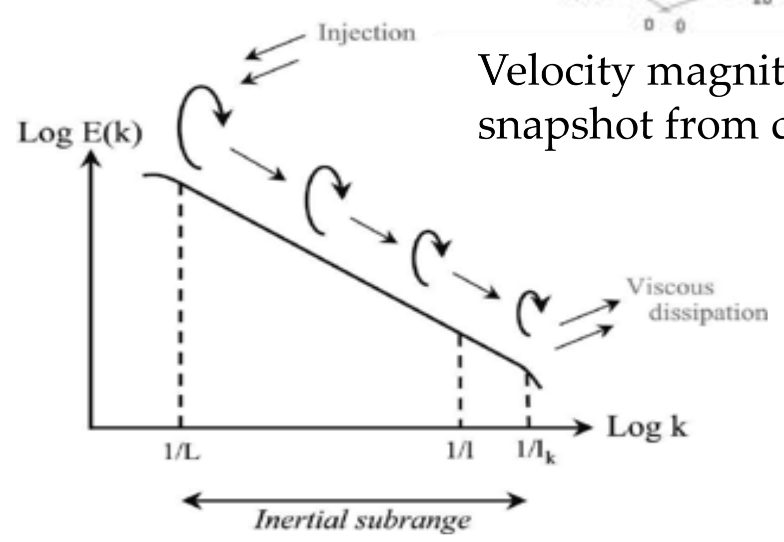


Dataset and Metrics

- Homogeneous Isotropic Turbulence
  - Close coupling of scales
  - Incompressible
  - Forced
  - Stationary
- Energy spectra
  - Function of wavenumber
- Velocity PDFs
  - $\partial u_i / \partial x_i$
- Kinetic energy
  - ½ RMS of velocity fluctuations

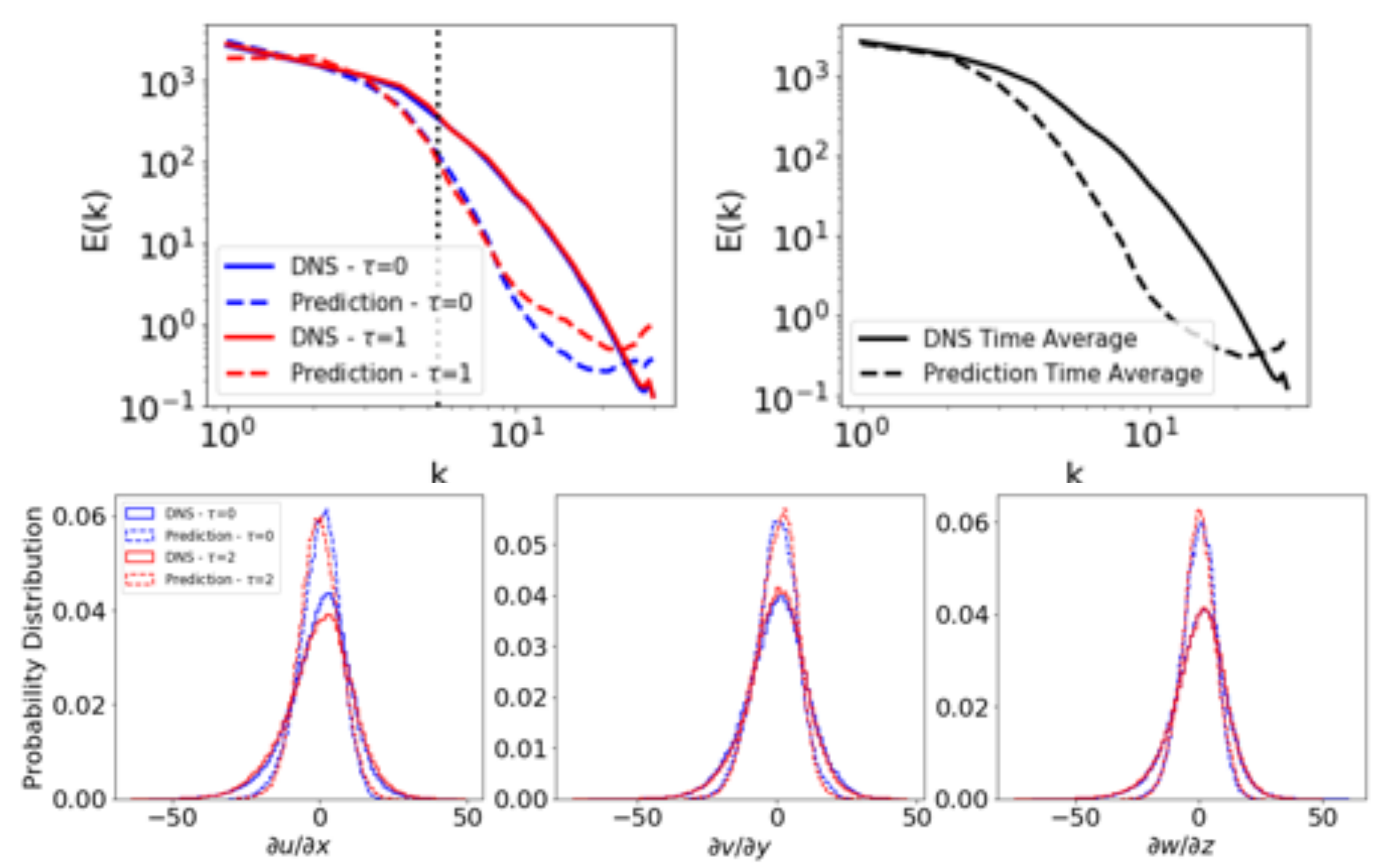


Velocity magnitude snapshot from dataset

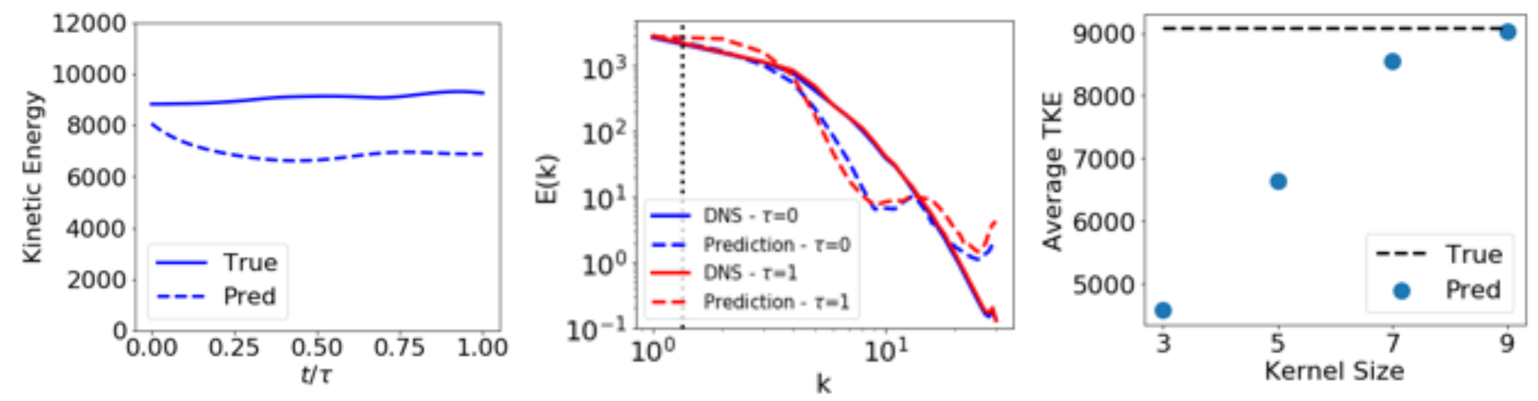


Sketch of typical turbulent energy spectrum[1]

Results



- Energy spectra: good agreement at low wavenumbers → large eddies
- Stationarity observed across snapshots (spectra, PDFs)



- *Left*: kinetic energy over 1 integral time scale
  - Underpredicts true value
  - After initial loss, TKE stabilizes to constant
- *Middle*: spectra snapshots for model with greater compression ratio (smaller latent space)
  - Poorer agreement, artifacts at the highest wavenumbers
- *Right*: average energy for various ConvNODE kernel sizes
  - Kernel size related to "differencing order"
  - Computational costs and tradeoffs

Conclusions

- ConvNODE architecture paired with latent space encoder/decoder can express large-scale turbulent dynamics
- Stationarity and stability observed in predictions
- Compression ratio and neural ODE kernel size identified as parameters that impact energy spectra and avg. energy

Acknowledgements

Mobility21 A USDOT NATIONAL UNIVERSITY TRANSPORTATION CENTER · DELL · NSF · XSEDE · NVIDIA

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE 1745016 awarded to VS.

[1]Reprinted from Wikimedia Commons, 2017.