
Turbulence forecasting via Neural ODE

Gavin D. Portwood*
 Los Alamos National Laboratory
 Los Alamos, NM 87545
 portwood@lanl.gov

Peetak P. Mitra*†
 Los Alamos National Laboratory
 Los Alamos, NM 87545
 peetak@lanl.gov

Mateus Dias Ribeiro*
 DFKI GmbH
 Kaiserslautern, Germany
 mateus.dias_ribeiro@dfki.de

Tan Minh Nguyen
 NVIDIA Corporation
 Santa Clara, CA 95051
 mn15@rice.edu

Balasubramanya T. Nadiga
 Los Alamos National Laboratory
 Los Alamos, NM 87545
 balu@lanl.gov

Juan A. Saenz
 Los Alamos National Laboratory
 Los Alamos, NM 87545
 juan.saenz@lanl.gov

Michael Chertkov
 Program in Applied Mathematics,
 University of Arizona, AZ 85719
 chertkov@math.arizona.edu

Animesh Garg
 University of Toronto,
 Toronto, ON M5S, Canada
 garg@cs.toronto.edu

Anima Anandkumar
 NVIDIA Corporation
 Santa Clara, CA 95051
 anima@caltech.edu

Andreas Dengel
 DFKI GmbH
 Kaiserslautern, Germany
 andreas.dengel@dfki.de

Richard Baraniuk
 Rice University
 Houston, TX 77005
 richb@rice.edu

David P. Schmidt
 University of Massachusetts
 Amherst, MA 01003
 schmidt@acad.umass.edu

Abstract

Fluid turbulence is characterized by strong coupling across a broad range of scales. Furthermore, besides the usual local cascades, such coupling may extend to interactions that are non-local in scale-space. As such the computational demands associated with explicitly resolving the full set of scales and their interactions, as in the Direct Numerical Simulation (DNS) of the Navier-Stokes equations, in most problems of practical interest are so high that reduced modeling of scales and interactions is required before further progress can be made. While popular reduced models are typically based on phenomenological modeling of relevant turbulent processes, recent advances in machine learning techniques have energized efforts to further improve the accuracy of such reduced models. In contrast to such efforts that seek to improve an existing turbulence model, we propose a machine learning (ML) methodology that captures, de novo, underlying turbulence phenomenology without a pre-specified model form. To illustrate the approach, we consider transient modeling of the dissipation of turbulent kinetic energy—a fundamental turbulent process that is central to a wide range of turbulence models—using a Neural ODE approach. After presenting details of the methodology, we show that this approach out-performs state-of-the-art approaches.

*Authors contributed equally

†Corresponding author: pmitra@umass.edu

1 Introduction

Forecasting the evolution of turbulence is a critical necessity for applications from engineering design to climate modeling. The range of scales involved in a turbulent flow is characterized by the Reynolds number, Re . Realistic turbulent flows often have moderate to large Re and involve a wide range of scales, i.e. large scale separation. Scale separation arises from non-linear interactions between flow components and processes that are often non-local in space and time. The dynamics at different scales are often closely coupled and for a truly predictive numerical framework, need to be well-resolved. Due to the non-linearity of the physics, simulating the entire range of scales is almost always prohibitively expensive in real-world applications, and therein lies the need to formulate reduced order formulations that can accurately predict these multi-scale physics.

Direct numerical simulations (DNS) of the entire range of dynamical scale interactions is possible in a number of idealized canonical flows relevant to applications in weather, and climate modeling. However, in many engineering applications, only the large scale dynamics have practical relevance. A common modeling approach, in such a paradigm, is to then resolve (i.e. directly simulate) dynamics at large scales while modeling the non-linear dynamical interactions between the resolved and the unresolved scale. These models are commonly phenomenological and heuristic in nature, and fitting, or calibration, and validation of these models to DNS data, for such idealized flows, constitutes a major practical hurdle in many fields. This is because it is performed in a cumbersome manner that involves manually iterating the adjustment of parameters until good matches with the DNS data are observed. Given the need to tune these engineering models based on high-fidelity turbulence data, in this manuscript, we develop an automatic methodology for training a family of phenomenology-informed and properly parameterized reduced models.

Recently, Deep Neural Network (DNN)-based approaches related to modeling fluid problems has gained wide attention [1–3]. Prominent among these are approaches based on modeling these dynamical systems as differential equations [4–6]. These methods require replacing the residual networks (ResNet/RNN) with ordinary differential equations (ODE) [4, 7], where Neural ODE (NODE) has emerged as a popular approach. NODE is a supervised machine learning approach that is based on learning the latent space representations of dynamical equations. It is devised in a non-intrusive manner without a pre-specified model form, that makes this framework very appealing for complex, transient, non-linear physics problems as considered here [8–10].

In this work, we look at a simple system of couple ODEs, wherein we examine the ability of the NODE approach to effectively learn the temporal dynamics and interpolate in the parameter space to make predictions for the unseen test case. We are particularly interested in exploring the ability of the continuous-time, generative Latent ODE model [6] within the NODE approach, in which the ODE integration occurs in the latent space.

1.1 Contributions of this work

In this work, we model the temporal evolution of the turbulence prognostics such as kinetic energy, k , and its dissipation rate, ϵ , which has emerged as a common reduced system to dynamically model turbulence [11]. Ground truth data is generated by extracting k and ϵ from a series of DNS datasets, as described in section 2. We then apply the continuous-time NODE framework to model these time series [4]. Compared to a discrete-depth network, such as a recurrent neural network (RNN), a NODE model can learn trajectories which may be sampled at arbitrary frequencies with standard ODE solvers [4–6] and hence is particularly well suited for physics problems.

Results from our experiments indicate that the NODE approach outperforms predictions from existing state-of-the-art models [12] in the settings we consider. We observe NODE models to consistently remain within 1-2% error with respect to the DNS solutions, almost two orders of magnitude less than state-of-the-art models. With these simple but fundamental experiments, we suggest that an automatic approach to training phenomenological reduced models for more complex turbulence phenomena, such as is necessary to integrate into existing climate models, is both practical and effective.

2 Problem Background

A common turbulence model solves for the turbulent kinetic energy k and its dissipation rate ϵ [11], as:

$$\frac{Dk}{Dt} = \nabla \cdot \nu \nabla k + \nabla \cdot \frac{\nu_t}{\sigma_k} \nabla k + \nu_t (\nabla \vec{u} + \nabla \vec{u}^T) : \nabla \vec{u} - \epsilon \quad (1)$$

$$\frac{D\epsilon}{Dt} = \nabla \cdot \nu \nabla \epsilon + \nabla \cdot \frac{\nu_t}{\sigma_\epsilon} \nabla \epsilon + \frac{\epsilon}{k} C_{e1} \nu_t (\nabla \vec{u} + \nabla \vec{u}^T) : \nabla \vec{u} - \frac{\epsilon}{k} C_{e2} \epsilon \quad (2)$$

where ν is the fluid kinematic viscosity, $\sigma_k, \sigma_\epsilon, C_{e1}, C_{e2}$ are model constants, and $\nu_t = C_\mu k^2 / \epsilon$ is the turbulent viscosity where C_μ is another model constant. Here, most constants can be derived from asymptotic analysis in low- or high-Reynolds number limits, where the Reynolds number is a measure of turbulent scale separation, defined,

$$\text{Re} \equiv \frac{k^2}{\nu \epsilon} . \quad (3)$$

At intermediate Reynolds numbers, where turbulence transverse asymptotic limits, there is little theoretical justification for appropriate constants, yet it is well-accepted that C_{e2} has functional dependence on large-scale turbulence properties and Re when $1 < \text{Re} < 100$ [see 12, for a review]. The case of the parameter C_{e2} is particularly interesting, because under assumptions of homogeneity and isotropy, equations (1) and (2) are reduced to two coupled ODEs, wherein C_{e2} is the only model parameter and the system is defined by:

$$\frac{dk}{dt} = -\epsilon \quad \text{and} \quad \frac{d\epsilon}{dt} = -C_{e2} \frac{\epsilon^2}{k} . \quad (4)$$

Despite the simplicity of this reduced model, modeling the parameter C_{e2} as a function of the large-scale fluid state is still an active research area for parameter regimes outside of asymptotic limits. A statistical description of turbulence at large length-scales is the power-law behaviour of kinetic energy \hat{k} at small Fourier wavenumber,

$$\hat{k}(|\xi|) \propto |\xi|^{p_0} \quad \text{for} \quad |\xi| \lesssim \xi_l \quad (5)$$

where \hat{k} is a function of the $L2$ norm of the wavenumber vector ξ , and ξ_l is the integral wavenumber. Indeed, the selection of C_{e2} is known to significantly depend on the parameter p_0 [13, 14], leading to complex and nonlinear functional forms for C_{e2} which adjust the model (4) to arbitrary values of p_0 (such as proposed by [12]). These models are enabled by constants and functional forms derived from data and regression from high-fidelity simulation. Given the dependence of these canonical models on high-fidelity turbulence data, it is reasonable to approach the prognostic modeling problem from a purely data-driven perspective with physics-inspired constraints for efficient and tractable numerical simulation.

3 Models

3.1 Direct Numerical Simulation

In order to evaluate the tractability of modeling turbulent decay by data-driven methods, we develop a series of numerical experiment by simulating the governing equations of fluid motion, the Navier-Stokes equations, directly with robust Fourier spectral method [c.f. 15].

Turbulence is initialized by forcing the low-wavenumber power spectra at various p_0 to an appropriate form via a linear scheme suggested by [16] and for the parameters summarized in Table 1. The parameter space has been designed to investigate Reynolds number regimes where existing models poorly calibrate and a range of p_0 observed in geophysical and engineering flows.

Once the flows are statistically stationary, the forcing is turned off and the turbulence is allowed to decay. The process is repeated for three realization of each parameter combination in order to sample the data variability.

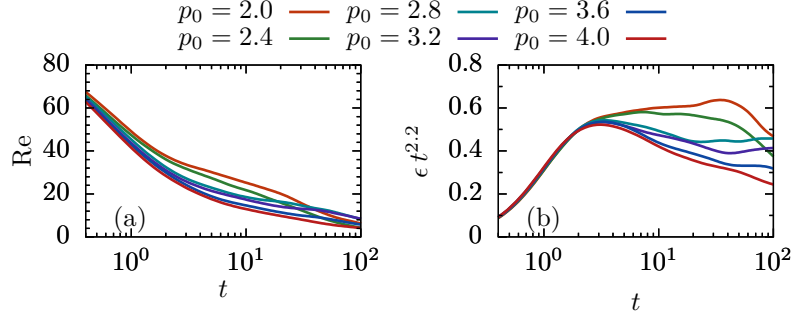


Figure 1: Evolution of (a) the Reynolds number and (b) the kinetic energy dissipation rate, scaled by theoretical decay exponents corresponding to $p_0 = 2$, for single realizations of the cases described in Table 1. Note the strong dependence of the evolution of ϵ on Re and p_0 after $t \approx 2$.

Trajectories of the Reynolds number and ϵ are shown in figure 1, where time has been appropriately non-dimensionalized by a turbulence time-scale. Our objective is to discover a Re -dependent model with p_0 as a parameter which outperforms existing state-of-the-art models.

3.2 Neural ODE for time-series

We use the continuous-time, generative Neural ODE approach called Latent ODE [4, section 5], to model the turbulent kinetic energy dissipation from the DNS data presented in the previous section. In the context of time-series, the model represents each observation by a latent trajectory, z_t , determined by

$$\frac{dz_t}{dt} = f(z_t, \phi) \quad (6)$$

where f specifies the dynamics of the hidden state of the network itself and ϕ is the set of neural network parameters. The model has three different parts to it. First, a recognition recurrent neural network (RNN) reads the observations from the DNS data backwards in time in order to determine an initial latent representation, z_{t_0} for each observation (k and ϵ). Second, this initial latent representation is used as an input to the ODE solver together with a function f , parameterized by a neural network, and is used to obtain latent space observations at all given times:

$$z_{t_1}, z_{t_2}, z_{t_3}, \dots, z_{t_N} = \text{ODESolve}(z_{t_0}, f, \theta_f, t_0, t_1, t_2, \dots, t_N) \quad (7)$$

In the third and final step, the latent space observations are decoded back to data space by another network. Training is performed by optimizing the parameters of the networks at each part (recognition RNN, ODE function, and decoder) in order to minimize the error between input and output or, in the context of a variational autoencoder [17, 18], maximize the evidence lower bound (ELBO).

One of the main advantages of the latent ODE approach in the prediction of time series data lies in the fact that the previously shown function f is time invariant. This means that given a latent state, z_t , one can define a unique latent trajectory. Therefore, a model trained to fit any given set of observations is able to extrapolate the latent trajectory arbitrarily far forwards (forecasting), or backwards in time [6]. Alternatively, one could also train the model on a given set of input parameters and generalize a solution to a new parameter value not seen during training.

In this work, we tested a latent ODE model architecture for deriving a turbulent kinetic energy dissipation model sensible to various p_0 values. The RNN encoder has 25 hidden units and 4 units in the latent space. The function f is parameterized with a fully-connected (FC) layers network with

Case	Re_0	p_0	N
Re100P20	99.85	2.0	5.5×10^6
Re100P24	99.61	2.4	7.1×10^6
Re100P28	100.7	2.8	9.0×10^6
Re100P32	100.2	3.2	1.1×10^7
Re100P36	100.3	3.6	1.4×10^7
Re100P40	100.4	4.0	1.7×10^7

Table 1: Simulation database, each case is simulated with three realizations. The number of datapoints in each simulation is denoted by N , each simulation is run for over 5,000 timesteps.

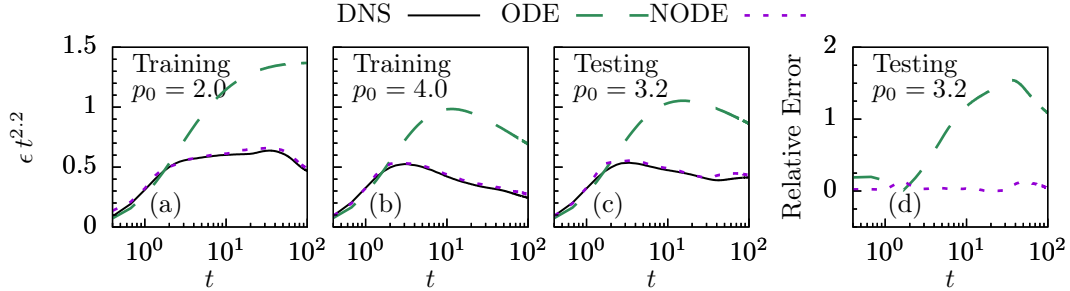


Figure 2: Comparison of models for the evolution of dissipation rate, scaled as in figure 1. Panels (a) and (b) show the purely data-driven model, which accounts for changes in Re and p_0 , to capture the trends of the ground truth much closer than the state-of-the-art analytic model [12]. The same is observed for the testing case in panel (c), whereas panel (d), which quantifies the error with respect to the ground truth data, illustrates that the NODE predictions stay within 1-2% of the DNS solutions.

one hidden layer containing 20 hidden units. The decoder has a similar network, also with 20 hidden units. We train the model with DNS results on a given set of large scale turbulence properties (such as p_0) and attempt to generalize a solution for a new set of unseen test conditions.

4 Experimental Results

Data-driven models are compared to state-of-the-art analytic models [12] in figure 2. Recalling that our objective is to capture transient Re and p_0 dependence in the evolution of turbulence prognostics, we compare against a model which accounts for these two parameters with complex and non-linear sub-models. We have selected an intermediate case Re_{100P32} for testing for the results presented here, but observed qualitatively similar behaviour when other cases have been held-out in cross-validation. The training sets are shown in figures 2a,b, which illustrate the capability of the Neural ODE model applied to time-series data to represent the evolution of the turbulent kinetic energy dissipation rate. The performance of the purely data-driven approach for the testing case is shown in figures 2c,d. We observe that while both models accurately predict the initial period of decay for $t \lesssim 2$, the analytic model over-estimates the dissipation rate for later times. We believe these observations are due to the analytic ODE model poorly capturing low Reynolds number effects in the final stages of decay (see figure 1a). In comparison, the purely data-driven Neural ODE approach accounts for this transient such that it more accurately forecasts the dissipation rate at later times.

5 Conclusions and future applications

We have proposed a well-constrained, fundamental physics problem for the application of data-driven modeling via Neural ODEs. By comparison with state-of-the-art models for the equivalent problem, we have demonstrated that the Neural ODE model more accurately predicts the evolution of the dissipation rate at multiple forcing objectives (as defined by p_0), which is a testament to the generalizability of the framework. This result is encouraging for the development of models for more complex turbulence phenomena, such as model flow configuration relevant to geophysical turbulence as used in climate models, where the evolution of additional prognostics are necessary to model. These prognostics, such as the flux energy between kinetic energy and gravitational potential energy reservoirs, are currently poorly parameterized [19] and outperforming data-driven approaches have an opportunity to immediately benefit such existing models. Another area of future investigation is to explore the interpretability [20–22] of these *black-box* ML models, such they can be better generalized to more complex physics problems.

6 Acknowledgements

The research activities of the authors the from Los Alamos National Laboratory are supported by a LDRD project "Machine Learning for Turbulence (MELT)" #20190059DR.

The data utilized in this study was generated for related research sponsored by the Department of Energy NNSA Advanced Simulation and Computing (ASC) program through the Physics and Engineering Models – Mix & Burn (PEM-M&B) and the Advanced Technology Development and Mitigation – Machine Learning (ATDM-ML) projects.

The authors from the University of Massachusetts would like to acknowledge the ICEnet consortium for the research funding and the GPU resources. Support from the members of the consortium, including NVIDIA Corp., SIEMENS, The MathWorks Inc., Cummins, Convergent Science Inc., & AVL, is appreciated.

This document has been approved by Los Alamos National Laboratory for public release as LA-UR-19-29250

References

- [1] Gavin Portwood, Misha Chertkov, Balasubramanya Nadiga, Juan Saenz, and Daniel Livescu. Physics-informed deep neural networks applied to scalar subgrid flux modeling in a mixed DNS/LES framework. *Bulletin of the American Physical Society*, 2019.
- [2] Gavin Portwood, Juan Saenz, and Daniel Livescu. Autonomous RANS/LES hybrid models with data-driven subclosures. *Bulletin of the American Physical Society*, 2019.
- [3] Peetak Mitra, Mateus Dias Ribeiro, and David Schmidt. A data-driven approach to modeling turbulent flows in an engine environment. *Bulletin of the American Physical Society*, 2019.
- [4] Tian Qi Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. In *Advances in neural information processing systems*, pages 6571–6583, 2018.
- [5] Tan M. Nguyen, Animesh Garg, Richard B. Baraniuk, and Anima Anandkumar. InfoCNF: An efficient conditional continuous normalizing flow with adaptive solvers. *ICML 2019 Workshop on Invertible Neural Networks and Normalizing Flows*, 2019.
- [6] Yulia Rubanova, Ricky TQ Chen, and David Duvenaud. Latent ODEs for irregularly-sampled time series. *arXiv preprint arXiv:1907.03907*, 2019.
- [7] Yifan Sun, Linan Zhang, and Hayden Schaeffer. Neupde: Neural network based ordinary and partial differential equations for modeling time-dependent data. *arXiv preprint arXiv:1908.03190*, 2019.
- [8] Romit Maulik, Arvind Mohan, Bethany Lusch, Sandeep Madireddy, and Prasanna Balaprakash. Time-series learning of latent-space dynamics for reduced-order model closure. *arXiv preprint arXiv:1906.07815*, 2019.
- [9] Omer San, Romit Maulik, and Mansoor Ahmed. An artificial neural network framework for reduced order modeling of transient flows. *Communications in Nonlinear Science and Numerical Simulation*, 77:271–287, 2019.
- [10] Mateus Dias Ribeiro, Gavin D Portwood, Peetak Mitra, Tan Mihn Nyugen, Balasubramanya T Nadiga, Michael Chertkov, Anima Anandkumar, and David P Schmidt. A data-driven approach to modeling turbulent decay at non-asymptotic reynolds numbers. *Bulletin of the American Physical Society*, 2019.
- [11] WP Jones and Brian Edward Launder. The prediction of laminarization with a two-equation model of turbulence. *International journal of heat and mass transfer*, 15(2):301–314, 1972.
- [12] J. B. Perot and S. M. de Bruyn Kops. Modeling turbulent dissipation at low and moderate Reynolds numbers. *J. Turbulence*, 7:1–14, 2006.
- [13] P.G. Saffman. The large-scale structure of homogeneous turbulence. *J. Fluid Mech.*, 27(3):581–593, 1967.
- [14] G. K. Batchelor and A. A. Townsend. Decay of turbulence in the final period. *P. Roy. Soc. Lond. A. Mat.*, 194:527, 1948.

- [15] S. A. Orszag and G. S. Patterson. Numerical simulation of turbulence. In M. Rosenblatt and C. Van Atta, editors, *Statistical Models and Turbulence*, volume 12 of *Lecture Notes in Physics*, pages 127–147. Springer, New York, 1972.
- [16] M. R. Overholt and S. B. Pope. A deterministic forcing scheme for direct numerical simulations of turbulence. *Comput. Fluids*, 27:11–28, 1998.
- [17] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
- [18] Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. *arXiv preprint arXiv:1401.4082*, 2014.
- [19] Gregory N Ivey, Cynthia E Bluteau, and Nicole L Jones. Quantifying diapycnal mixing in an energetic ocean. *J. Geophys. Res.-Oceans*, 123(1):346–357, 2018.
- [20] Xiao Dong and Ling Zhou. Geometrization of deep networks for the interpretability of deep learning systems. *arXiv preprint arXiv:1901.02354*, 2019.
- [21] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. Model-agnostic interpretability of machine learning. *arXiv preprint arXiv:1606.05386*, 2016.
- [22] Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. A survey of methods for explaining black box models. *ACM computing surveys (CSUR)*, 51(5):93, 2019.