Skillful or not the prediction of AI model depends on the enoughness of training data which represents the corresponding physical mechanism

*Mu Mu[1], Guokun Dai[1]

1. Fudan University

The predictive value of AI-based weather and climate models critically depends on their ability to generalize beyond the training dataset. However, how the coverage of physical regimes in training data affects model generalization remains insufficiently understood. In this study, we employ a fully controllable and interpretable framework based on a parameterized Burgers system to systematically investigate the relationship between data characteristics and AI model performance. By varying the nonlinearity and diffusion coefficients, we generate 600 distinct dynamical regimes spanning a wide range of effective Reynolds numbers and flow behaviors. Using this ensemble, we design three targeted experiments in which AI models are trained on subsets of the parameter space and evaluated on unseen regimes characterized by different physical mechanisms. The results show that AI models achieve high skill when interpolating within the training manifold but exhibit pronounced degradation when extrapolating to regimes associated with different nonlinear dynamics or intermittency levels. These findings demonstrate that both the diversity of underlying physical mechanisms and the breadth of parameter-space coverage in training datasets are essential for robust generalization. The results further suggest that AI models trained solely on historical reanalysis data may have limited capability in predicting rare, extreme events that are poorly represented in observations. This study highlights the necessity of supplementing training datasets with simulations from physically consistent numerical models to improve the reliability of AI-based predictions for complex geophysical systems.

Keywords: artificial intelligence, physical mechanism, Reynolds numbers

