JℳM RAPIDS

Data-driven nonlinear turbulent flow scaling with Buckingham Pi variables

Kai Fukami[1],†, Susumu Goto[2] and Kunihiko Taira[1]

[1]Department of Mechanical and Aerospace Engineering, University of California, Los Angeles, CA 90095, USA

[2]Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan

Nonlinear machine learning for turbulent flows can exhibit robust performance even outside the range of training data. This is achieved when machine-learning models can accommodate scale-invariant characteristics of turbulent flow structures. This study presents a data-driven approach to reveal scale-invariant vortical structures across Reynolds numbers that provide insights for supporting nonlinear machine-learning-based studies of turbulent flows. To uncover conditions for which nonlinear models are likely to perform well, we use a Buckingham-Pi-based sparse nonlinear scaling to find the influence of the Pi groups on the turbulent flow data. We consider nonlinear scalings of the invariants of the velocity gradient tensor for an example of three-dimensional decaying isotropic turbulence. The present scaling not only enables the identification of vortical structures that are interpolatory and extrapolatory for the given flow field data but also captures non-equilibrium effects of the energy cascade. As a demonstration, the present findings are applied to machine-learning-based super-resolution analysis of three-dimensional isotropic turbulence. We show that machine-learning models reconstruct vortical structures well in the interpolatory space with reduced performance in the extrapolatory space revealed by the nonlinearly scaled invariants. The present approach enables us to depart from labelling turbulent flow data with a single parameter of Reynolds number and comprehensively examine the flow field to support training and testing of nonlinear machine-learning techniques.

Key words: machine learning, isotropic turbulence

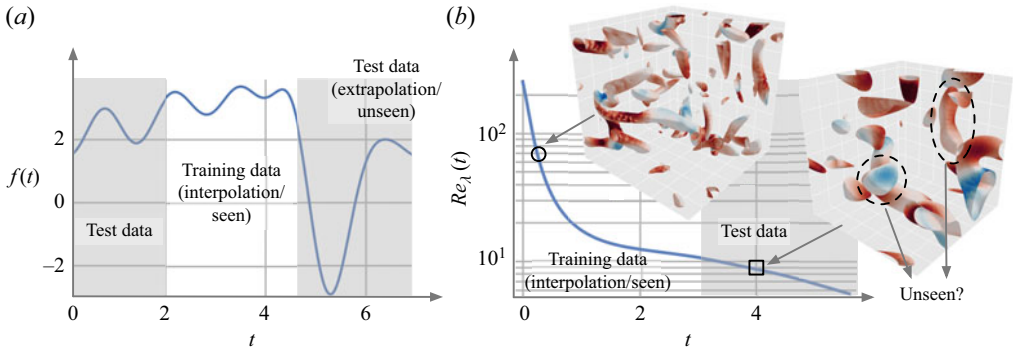† Email address for correspondence: kfukami1@g.ucla.edu

Figure 1. Concept of interpolation and extrapolation. (*a*) A one-dimensional example. (*b*) Evolution of three-dimensional decaying isotropic turbulence over $Re_\lambda(t)$.

1. Introduction

Trained fluid mechanicians can identify similarities in vortical structures for a variety of turbulent flows. Even if there are scale or rotational differences, we can visually extract similar structures due to their geometrical features across spatial and temporal scales. Analogously, recent machine-learning models have capitalized on such structural similarities to achieve reliable performance for the given training data. However, it is also known that machine-learning models often achieve satisfactory performance even in the cases of unseen Reynolds numbers. Examples of models performing successfully beyond the training data include, turbulence modelling (Guan *et al.* 2022), state estimation (Guastoni *et al.* 2021) and super-resolution (Kim *et al.* 2021). This suggests that machine-learning models are able to incorporate scale-invariant properties of the flows while optimizing the output to meet their objectives.

These observations imply that nonlinear machine-learning models are capturing data characteristics in a more holistic manner. Traditionally, interpolation and extrapolation of models have been associated with a particular variable or parameter (Domingos 2012; Marcus 2018), as illustrated in figure 1(*a*). In this example, test data outside of the training data are captured by the temporal variable *t*. In the case of more complex data sets, such as turbulent flows, describing whether a certain type of vortical structure appears in training data or not requires additional considerations, beyond a single parameter such as the Reynolds number.

Given this motivation, we revisit the idea of interpolation and extrapolation in the context of turbulent flow structures. Here, we consider the capturing vortical structures that have certain similarities across training and test data to be 'interpolation' and their structural features to be 'seen'. On the other hand, capturing structures in the test data that do not share similarities with those in the training data are referred to as 'extrapolation' and their structures as 'unseen'. With machine-learning models known to perform well for approximating the nonlinear relationship between the input and output from a large collection of data, the presence of seen (or common) features can provide robust performance under untrained flow situations (Kim *et al.* 2021; Guan *et al.* 2022). The current study presents a data-driven scaling approach to reveal seen/unseen structures of turbulent flow data in machine learning.

To assess rotational and shear similarities between small- and large-length scales in turbulent flows, we examine the flow field data in terms of the invariants of the velocity gradient tensor. In the present analysis, the scaled invariants are found through

sparse nonlinear regression using non-dimensional parameters from the Buckingham Pi theorem (Buckingham 1914). The present approach offers the optimal nonlinear scalings of the invariants, uncovering scale-invariant vortical structures in a turbulent flow field. Analysing the data distribution of the present scaled invariants reveals what types of flow structures are seen and unseen. Furthermore, the findings from this study provide guidance in the choice of machine-learning functions to offer robustness for scale-invariant vortical structures. The present paper is organized as follows. The proposed Buckingham-Pi-based sparse nonlinear scaling of the invariants is introduced in § 2. We demonstrate the current approach for three-dimensional isotropic turbulence in § 3. Concluding remarks are provided in § 4.

2. Methods

Machine-learning models for turbulent flow structures are known to remain accurate beyond the coverage of training data. This is the case especially when the nonlinear machine-learning models have scale and rotational invariances embedded in their formulations. To explain this extended validity of machine-learning models, this study aims to uncover nonlinear scalings that capture the similarities in turbulent vortical structures across a range of Reynolds numbers. The central hypothesis of this study is that the existence of such scalings enables nonlinear machine-learning techniques to perform effectively across different flow fields beyond the range of Reynolds numbers provided in the training data. With these identified scalings, turbulent flow structures that linearly and nonlinearly span across a range of scales may be considered seen beyond the training data due to their structural similarities.

To examine turbulent flows, we consider the invariants of the velocity gradient tensor $A = (\nabla u)^{\mathrm{T}}$, where u represents the velocity, such that the observations are independent of the frame of Chong, Perry & Cantwell (1990). These invariants are $P = \mathrm{trace}(A)$, $Q = [P^2 - \mathrm{trace}(A^2)]/2$ and $R = \det(A)$. For the present study, we consider incompressible turbulent flows, which makes $P = \nabla \cdot u = 0$. The remaining two invariants of Q and R characterize the local rotation and shear, respectively (Ooi *et al.* 1999). According to these invariants, the flow can experience vortex compression ($Q > 0, R < 0$), vortex stretching ($Q > 0, R > 0$), biaxial strain ($Q < 0, R < 0$) and axial strain ($Q < 0, R > 0$) (Davidson 2015). These invariants will be nonlinearly scaled with non-dimensional Pi groups using a data-driven approach.

In this study, we consider three-dimensional incompressible decaying isotropic turbulence. The flow field data are obtained from direct numerical simulation using 64^3 grid points with a Taylor microscale-based Reynolds number of $0.85 \leq Re_\lambda \leq 252$ (Chumakov 2008; Gopalakrishnan Meena & Taira 2021), satisfying $k_{max}\eta \geq 1$, where k_{max} is the maximum resolvable wavenumber of the grid and η is the Kolmogorov length scale, to resolve all important scales of motion. The present simulation is initialized with random velocity fields of Gaussian profiles that satisfy incompressibility and the Kolmogorov spectra for kinetic energy. The flow snapshots are curated after the flow reaches the decaying regime at which we set time t to zero. Note that a large Re_λ at the early stage of decay is due to a small dissipation coefficient, which strongly depends on the flow (Goto & Vassilicos 2009, 2015). We present representative flows with their corresponding invariants on the Q–R plane over time in figure 2. The Reynolds number decreases as vortical structures evolve as their characteristic size increases. Even at a large time when $Re_\lambda = 4.18$, large-scale structures are still observed while the flow is under decay. During this process, invariants Q and R decrease in magnitude while the probability density functions of these invariants remain geometrically similar. This suggests that there
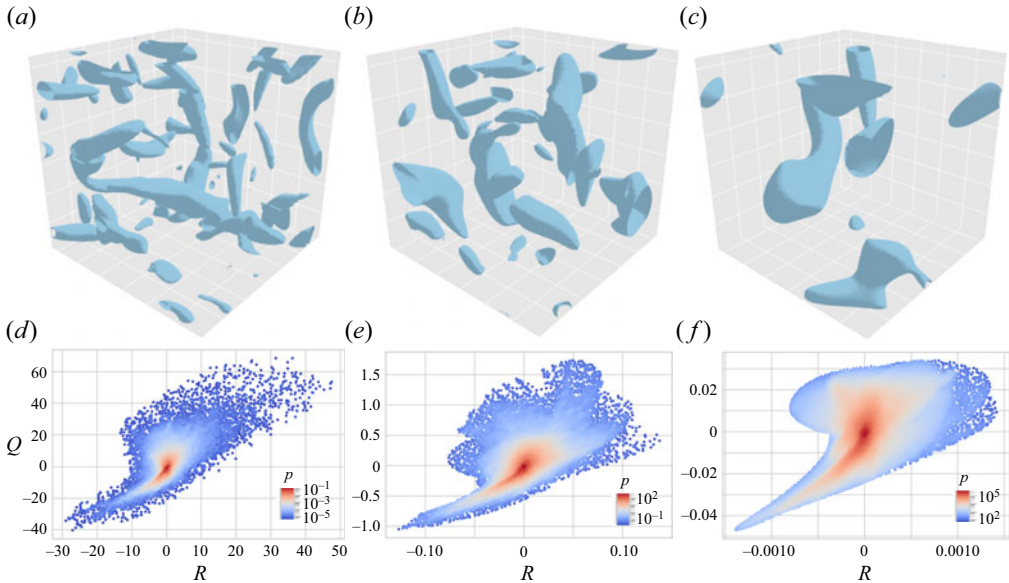
Figure 2. Example flow snapshots with Q–R distributions of three-dimensional decaying turbulence at (a) $Re_\lambda = 214$, (b) 14.6 and (c) 4.18. Each distribution is coloured by density. Turbulent vortices are visualized with (d) $Q = 10$, (e) 0.3 and (f) 0.02.

is some level of scale invariance in the distributions of the turbulent vortical structures over the Q–R plane. That is, the decaying isotropic turbulence holds similar rotational and shear structures whose sizes vary over the Reynolds number.

Linear scaling for the Q–R distributions based on the kinematic viscosity ν and energy dissipation rate ϵ such as $Q/(\epsilon/\nu)$ and $R/(\epsilon/\nu)^{3/2}$ do not completely collapse the distributions, given their long tails especially at high Reynolds number, as presented in figure 3. This seems to be caused by the wide range of Reynolds numbers and strong unsteadiness contained in the present decaying turbulence data, implying that Kolmogorov's similarity hypotheses do not hold in an instantaneous manner. For example, skewness S_R of the linearly scaled distribution for R in figure 2 is 2.74, 0.299 and 0.186 for $Re_\lambda = 214$, 14.6 and 4.18, respectively. These observations suggest that nonlinear scaling needs to be considered to accommodate non-equilibrium effects. Identifying such nonlinear scalings of Q and R distributions with data-driven techniques can also reveal how nonlinear machine learning extracts seen (or common) vortical structures that share similarities with structures outside of the training data sets. Moreover, unseen structures can be captured by uncovering the invariant space that does not overlap for the scaled data.

Let us consider scaling the invariants Q and R in a nonlinear manner using non-dimensional Pi groups from the Buckingham Pi theorem (1914), which distils a number of dimensional parameters into a smaller number of dimensionless groups (Bakarji *et al.* 2022; Xie *et al.* 2022). We assume that the scaling can be obtained through superposing appropriate polynomials of the Pi variables. Denoting these invariants as ϕ (either Q and R), the candidate polynomials are assumed to have the form of

$$\theta_k(\boldsymbol{x}, t) = \Pi_i^m(t)\Pi_j^n(t)\phi(\boldsymbol{x}, t), \quad \text{where } i, j = 1, 2, \ldots, \quad m, n \in \mathbb{Z}, \quad k = 1, \ldots, n_L. \tag{2.1}$$
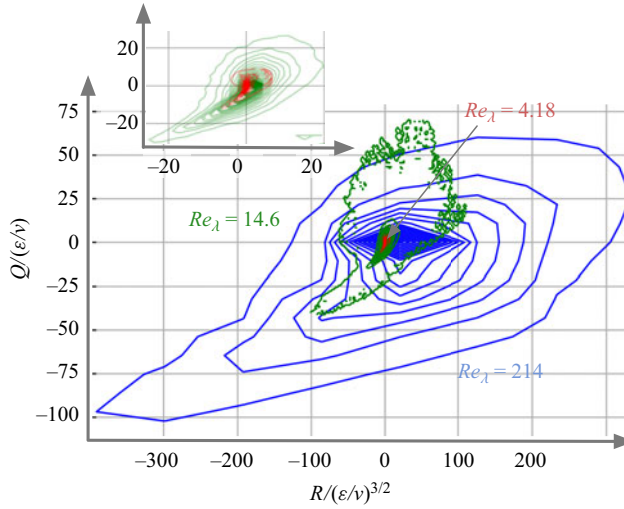
Figure 3. Linearly scaled Q–R distributions based on the kinematic viscosity ν and energy dissipation rate ϵ. A zoom-in view for $Re_\lambda = 14.6$ and 4.18 is also shown.

Here, n_L is the number of library-basis functions. Given these library candidates θ_k, we can express the scaled invariants $\phi^*(\boldsymbol{x}, t)$ as

$$\phi^*(\boldsymbol{x}, t) = \sum_{k=1}^{n_L} a_k \theta_k(\boldsymbol{x}, t), \tag{2.2}$$

with $\boldsymbol{a} = (a_1, a_2, \ldots, a_{n_L})^{\mathrm{T}} \in \mathbb{R}^{n_L}$. Spatiotemporal discretization of this equation yields $\boldsymbol{\Phi}^* = \boldsymbol{\Theta} \boldsymbol{a}$, where

$$\boldsymbol{\Phi}^* = \begin{bmatrix} \boldsymbol{\phi}^*(t_1) \\ \boldsymbol{\phi}^*(t_2) \\ \vdots \\ \boldsymbol{\phi}^*(t_{n_t}) \end{bmatrix} \in \mathbb{R}^{n_x n_t}, \quad \boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{\theta}_1(t_1) & \boldsymbol{\theta}_2(t_1) & \cdots & \boldsymbol{\theta}_{n_L}(t_1) \\ \boldsymbol{\theta}_1(t_2) & \boldsymbol{\theta}_2(t_2) & \cdots & \boldsymbol{\theta}_{n_L}(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\theta}_1(t_{n_t}) & \boldsymbol{\theta}_2(t_{n_t}) & \cdots & \boldsymbol{\theta}_{n_L}(t_{n_t}) \end{bmatrix} \in \mathbb{R}^{n_x n_t \times n_L},$$

$$\tag{2.3a,b}$$

with $\boldsymbol{\phi}(t) = (\phi(x_1, t), \ldots, \phi(x_{n_x}, t))^{\mathrm{T}} \in \mathbb{R}^{n_x}$ and $\boldsymbol{\theta}_k(t) = (\theta_k(x_1, t), \ldots, \theta_k(x_{n_x}, t))^{\mathrm{T}} \in \mathbb{R}^{n_x}$ being the scaled invariant and the library candidate, respectively, discretized in space. Here, the invariants are spatiotemporally stacked into a tall vector. The coefficients $\boldsymbol{a}$ are determined through the sequential threshold least squares method (Brunton, Proctor & Kutz 2016), promoting sparsity of the coefficient matrix in a computationally efficient manner (Fukami *et al.* 2021*b*). Since excessive sparsity promotion leads to a high regression error, the sparsity coefficient in the threshold least square method needs to be carefully tuned (Kaiser, Kutz & Brunton 2018).

For the present decaying turbulence, the Taylor length scale λ can be expressed as the function of the characteristic velocity u (the square root of the spatially averaged kinetic energy), the kinematic viscosity ν, the computational domain size L and viscous dissipation ϵ such that $\lambda = f(u, \nu, L, \epsilon)$. Through the use of the Buckingham Pi theorem, we can find three Pi variables, which namely are $\Pi_1 = u\lambda/\nu$, $\Pi_2 = uL/\nu$ and $\Pi_3 = \epsilon\nu/u^4$. The first Pi variable Π_1 is the Taylor Reynolds number Re_λ. In the present study, we take these Pi variables as a function of time. Given these Pi variables, we can construct

the library candidates for (2.1) as first- and second-order polynomials of the Pi variables such that

$$\{\Pi_1, \Pi_2, \Pi_3, \Pi_1^{-1}, \Pi_2^{-1}, \Pi_3^{-1}, \Pi_1^2, \Pi_2^2, \Pi_3^2, \Pi_1^{-2}, \Pi_2^{-2}, \Pi_3^{-2},$$
$$\Pi_1\Pi_2, \Pi_2\Pi_3, \Pi_3\Pi_1, \Pi_1^{-1}\Pi_2, \Pi_2^{-1}\Pi_3, \Pi_3^{-1}\Pi_1, \ldots, \Pi_i^m\Pi_j^n, \ldots\}, \qquad (2.4)$$

where $|m| + |n| \leq 2$.

With the Buckingham-Pi-based library matrix $\boldsymbol{\Theta}$, we seek coefficients $\boldsymbol{a}$ by maximizing the similarity of data distributions of the invariants over space and time. Here, we utilize the Kullback–Leibler (KL) divergence (Kullback & Leibler 1951) to assess the difference between a probability distribution f_1 and the reference probability distribution f_2. For two probability (data) distributions, the KL divergence is defined as

$$D(f_1\|f_2) \equiv \int_{-\infty}^{\infty} f_1(\boldsymbol{\phi}^*)\log\frac{f_2(\boldsymbol{\phi}^*)}{f_1(\boldsymbol{\phi}^*)}\,\mathrm{d}\boldsymbol{\phi}^*. \qquad (2.5)$$

The minimization of the KL divergence finds the optimal coefficients $\boldsymbol{a}^*$ for maximum similarity of the scaled invariant distributions, which yields an optimization problem of $\boldsymbol{a}^* = \mathrm{argmin}_{\boldsymbol{a}}[D(f_1\|f_2)]$. Below, the data distribution from a snapshot at high Re_λ is used as the reference f_2.

3. Results

We apply the present nonlinear scaling analysis to three-dimensional decaying isotropic turbulence over $0.85 \leq Re_\lambda \leq 252$. The present data comprised 800 snapshots over time and 64^3 grid points in a tri-periodic cubic domain. Here, we aim to (1) uncover the nonlinear influence of characteristic variables on the evolution of the invariants Q and R, (2) identify common (seen) vortical flow features through the scaled invariants and (3) provide guidance on proper training and formulation of machine-learning models.

Based on sparse regression, we find the scaling factors for Q and R to be

$$Q^* = (5.76\Pi_1 + 4.17\Pi_2^{-2} - 3.59\Pi_2\Pi_3)Q, \qquad (3.1)$$

$$R^* = (5.53\Pi_1 - 0.826\Pi_2^{-1} + 4.42\Pi_2^{-2} - 3.44\Pi_2\Pi_3)R. \qquad (3.2)$$

While isotropic turbulence is complex, the nonlinearly scaled invariants turn out to be surprisingly compact in their expressions. For both scaled invariants, we have Π_1 which is the Taylor Reynolds number Re_λ. This reflects the decaying nature, corresponding to the fact that the data spread on the Q–R plane shrinks with decreasing Taylor Reynolds number over time, as shown in figure 4(*a,b*). The present scaling also identifies the influence of $\Pi_2 = uL/\nu$, the box-size-based Reynolds number, revealing that the computational domain size influences the turbulent flow, especially as time advances and vortical structures become comparable in size to the computational domain. This is evident from the inverse and quadratic inverse nature of the scalings. The size of the periodic box affects not only the large-scale vortical motion over the domain but also the energy dissipation of turbulence (Davidson 2015).

Furthermore, the present analysis uncovers the importance of $\Pi_2\Pi_3 = uL/\nu\cdot\epsilon\nu/u^4 = \epsilon L/u^3$, which is the ratio between the instantaneous energy dissipation rate determined by small length scales and the cascading energy flux from the system-size vortices. Thus, $\Pi_2\Pi_3$ plays a similar role to the dissipation coefficient $C_\epsilon \equiv \epsilon l/u^3$, quantifying the degree of non-equilibrium (Vassilicos 2015), i.e. the violation of Kolmogorov's similarity.
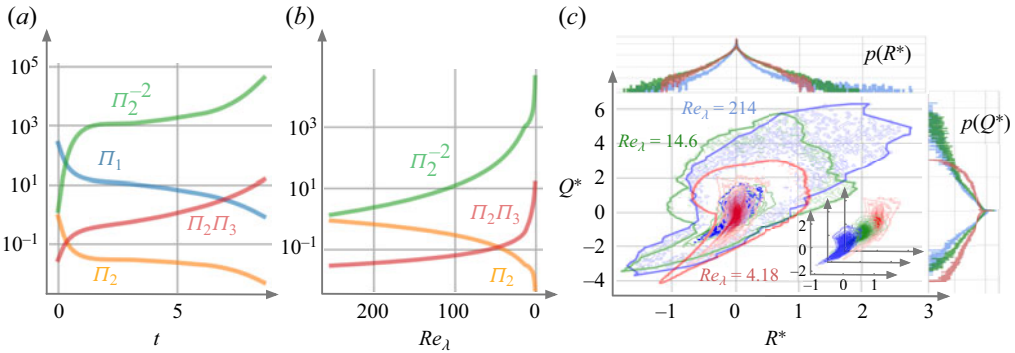
Figure 4. The data-driven Buckingham Pi scaling for three-dimensional decaying turbulence. The time series of the identified non-dimensional variables as a function of (*a*) time and (*b*) the Taylor Reynolds number. (*c*) Scaled Q^* and R^* invariants with their probability density functions. A zoom-in view is also shown as an inset.

Although $\Pi_2\Pi_3$ is constant when the energy flux of large-scale vortices and energy dissipation rate are in balance, that is not the case for the present decaying turbulence (Goto & Vassilicos 2016). The present $\Pi_2\Pi_3$ accounts for the time delay of the energy-cascade process between energy flux and dissipation due to strong nonlinearities in decaying turbulence. This is achieved by correcting non-equilibrium effects in a nonlinear manner. Note that the present formulation provides a similar scaling expression in a range of similar Reynolds numbers even with different initial flow fields since sparse regression captures how the turbulent flow decays over time.

Next, the identified scalings are applied to the original Q and R data distributions, as shown in figure 4(*c*) along with their probability density functions. The identified factors in (3.1) and (3.2) yield the optimal overlap of Q^* and R^* over space and time. The scaled data distribution for the high-Reynolds-number flow spreads over a larger area than that for the low Reynolds number, especially for $Q^* > 0$. This implies high occurrence of vortex stretching and compression at high Reynolds number.

Let us focus on the vortical structures for overlapping and non-overlapping regions of the data over the Q^*–R^* plane, as illustrated in figure 5. For $Q^* > 1$, similar shapes of vortical elements can be observed across a range of length scales. These identified vortical structures are assessed as 'interpolatory' (common) with the present scaling approach. The isosurfaces of $Q^* > 3$ are also visualized in figure 5. These strong rotational elements that barely appear in low-Reynolds-number flows are 'extrapolatory' structures. Moreover, such extrapolatory vortical structures against the low-Re_λ flow field can be seen in the portion of $R^* < -1.2$ and $R^* > 1.2$. The region of $Q < 0$ and $R < 0$ corresponds to biaxial strain, while that of $Q > 0$ and $R > 0$ reflects vortex stretching (Davidson 2015). Hence, the scaled data distribution suggests that these structures caused by strong biaxial strain and vortex stretching correspond to 'extrapolation' for the present turbulence data. While not shown here, the portion of $R^* > -0.5$ and $-3.5 < Q^* < 0.5$ also includes extrapolatory vortical structures. The scaled Q–R data distribution gives insights into turbulent flows, in addition to the identified scaling.

The present Buckingham-Pi-based sparse nonlinear scalings can identify interpolatory and extrapolatory vortical structures in isotropic turbulence. Nonlinear machine-learning models are likely to capture the characteristics and behaviour of what we refer to as the interpolatory structures even in untrained Reynolds number cases. This explains why well-trained machine-learning models may perform well even for test data. However, we
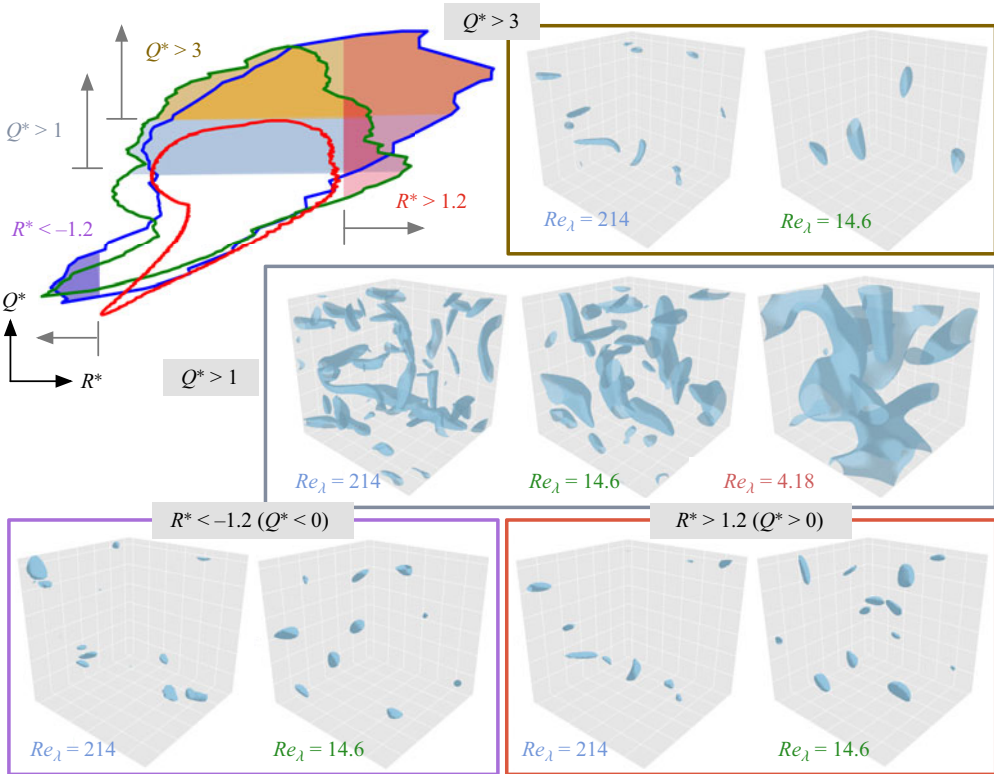
Figure 5. Interpolatory and extrapolatory vortical structures in three-dimensional decaying isotropic turbulence.

caution that when the test data includes extrapolatory structures, machine-learning models are no longer guaranteed to be valid. With regard to these points, classifying interpolatory and extrapolatory structures solely by Re_λ is not encouraged for assessing nonlinear machine-learning models.

Based on the insights from the scaled invariants above, let us consider machine-learning-based super-resolution reconstruction of turbulent flows (Fukami, Fukagata & Taira 2023). Super resolution reconstructs the high-resolution flow field q_{HR} from its low-resolution data q_{LR} with a reconstruction model $\mathcal{F}$ through $q_{HR} = \mathcal{F}(q_{LR})$. Recently developed machine-learning-based super-resolution analysis captures the nonlinear relationship between small- (unresolved) and large-scale (resolved) structures. This study considers the ability of machine-learning-based reconstruction to recognize and reconstruct common turbulent flow structures for a variety of flow field snapshots, even outside of the training data.

For super-resolution reconstruction of turbulent flows, a machine-learning model should be carefully constructed to account for a range of spatial length scales while enforcing rotational and translational invariance of vortical structures. To satisfy these properties, we use the hybrid downsampled skip-connection/multi-scale (DSC/MS) model (Fukami, Fukagata & Taira 2019, 2021a). The DSC/MS model based on convolutional neural networks (CNNs; LeCun et al. 1998) is composed of three main functions: namely, (A) up/downsampling operations, (B) skip connection, and (C) multi-scale filters. The up/downsampling provides robustness against rotation and translation of vortical structures. The skip connection allows communication between the input low-resolution

data and the intermediate output of the DSC model, which is crucial in learning a step-by-step internal process towards the high-resolution output from the low-resolution input while expanding the dimension of the flow field snapshot (He *et al.* 2016). The multi-scale filters apply filtering with a number of different sizes of filter (e.g. three here) in parallel to capture a broad range of scales in turbulent flows. Further details on the present neural network model are in Fukami *et al.* (2019). We will discuss which function inside the present model contributes to gaining robustness for scale invariance later.

Here, we consider two cases of training and testing for super-resolution reconstruction of turbulent flows: (1) a model trained with a low-Re_λ data set and tested with a high-Re_λ data set (low-Re_λ training); (2) a model trained with a high-Re_λ data set and tested with a low-Re_λ data set (high-Re_λ training). We expect that the low-Re_λ training cannot cover the non-overlapping portion of the scaled invariants, which corresponds to extrapolation. The high-Re_λ training, which covers a wide portion of the scaled invariants, may amount to interpolation in terms of the turbulent flow structures. For the present analysis with the two training scenarios, the threshold Re_λ between the low- and high-Re_λ training cases is set to $Re_\lambda = 15$. The velocity vector is used as data attributes q with $\mathcal{F}$ reconstructing the high-resolution flow field on a 64^3 grid from the low-resolution data on a 4^3 grid.

The reconstructed flow fields are shown in figure 6(*a*). The vortical flows are visualized using Q-criteria and coloured by the third invariant R computed from the reconstructed velocities. For the low-Re_λ training case, the model reconstructs the vortical structures for the training Re_λ with an L_2 error of approximately 0.1. In contrast, the visualized Q variable at the high-Re_λ exhibits significant level of error rendering the reconstruction grossly incorrect. Such high errors can be explained using the scaled variables in figure 6(*b*) due to the flow features residing over the non-overlapping region on the present plane. The high error is also observed around the bottom left of figure 6(*b*). This extrapolation region cannot be reconstructed with the low-Re_λ training because these structures are barely seen in the low-Re_λ data sets, as presented in figure 5.

Next, let us consider the high-Re_λ training case. The reconstructed flow from the overlapping case (interpolation) is in agreement with the reference data. In contrast to the extrapolatory low-Re_λ training, reasonable reconstruction can still be achieved at $Re_\lambda = 4.18$. This suggests that the high-Re_λ training data holds insights into a wider range of vortical structures that also appear in the low-Re_λ regime, which is confirmed from the scaled Q^* and R^* data. It is worth pointing out that the scaled invariants at the low Re_λ include some non-overlapping portion (scale-variant structures) in the region of $Q^* < 0$ and $R^* < 0$. This implies that the model trained with only the high-Re_λ regime cannot cover vortical structures of $Q^* < 0$ and $R^* < 0$. For better reconstruction over this regime, we need to include training data that covers $Q^* < 0$ and $R^* < 0$. Such an observation provides guidance on how the vortical flow data should be prepared to enable reliable reconstruction.

We further examine the nonlinearly scaled invariants to assess which inner functions of the machine-learning models contribute to robustness for scale-invariant regression. We here consider six models using the aforementioned functions A, B and C (figure 7*a*):

  (i) the original DSC/MS model (functions A, B and C),
 (ii) the original model without multi-scale filters (functions A and B),
(iii) the original model without the skip connection (functions A and C),
(iv) the original model with up/downsampling only (function A),
 (v) the original model with multi-scale filters only (function C),
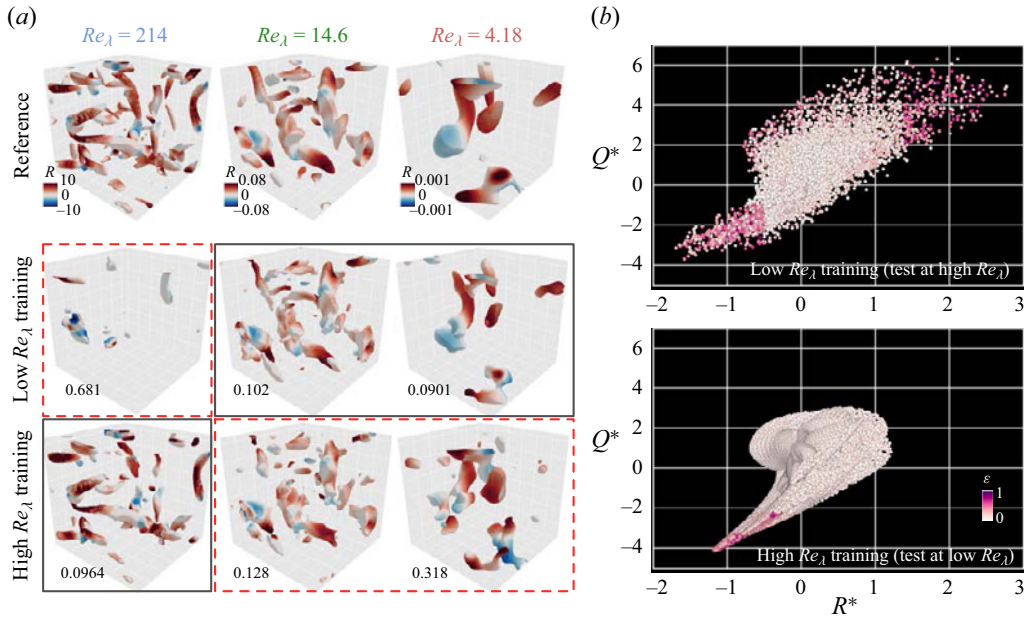(vi) a regular convolutional neural network (without any of the functions above).

Figure 6. (*a*) Super-resolution reconstruction of three-dimensional decaying turbulence. The reconstructed flow fields are visualized with the *Q*-criteria, coloured by *R*. The values underneath each figure represent the L_2 error norm. The grey and red boxes, respectively, highlight snapshots for training and testing the Re_λ regime. (*b*) Scaled Q^* and R^* at test Re_λ's for low- and high-Re_λ training cases, coloured by the spatial L_2 reconstruction error.

These six models consider all possible combinations that can be constructed with the DSC/MS model. We perform the low-Re_λ training for all of these six cases. The corresponding error distributions are presented in figure 7(*b*). For model (ii) which removes the multi-filter function from the baseline model, the error behaviour on the scaled data is similar to that of the original model. A similar observation is seen for models (iii) and (iv) which have no multi-filter functions. Since models (ii–iv) include function A, up/downsampling operations are crucial for robust super-resolution reconstruction of turbulent flows.

This can be further confirmed with model (v) which comprises only the multi-scale filters (function C). The reconstruction error is significantly higher without the up/downsampling operations. Including dimension compression and expansion plays an important role in obtaining robustness for scale-invariant characteristics, which agrees with a number of studies on CNNs for scale invariance characteristics in image science (Jarrett *et al.* 2009; Van Noord & Postma 2017). In contrast, multi-scale filters have secondary importance in reconstructing the flow (model (v)) as evident from its comparison with the results from CNN (model (vi)). While the regular CNN without functions A, B and C returns pixelized flow fields, model (v) achieves qualitative reconstruction. The use of skip connections is also important for successful training to address issues related to the convergence of neural network weights (He *et al.* 2016). These findings suggest that robust turbulent flow reconstruction can be achieved by selecting machine-learning functions based on the implication of scaled invariants.

The current approaches to identify important model functions that accommodate scale invariance could be combined with emerging methods to investigate the role of networks in capturing flow physics, such as the sensitivity analysis (Lee & You 2021; Jagodinski, Zhu
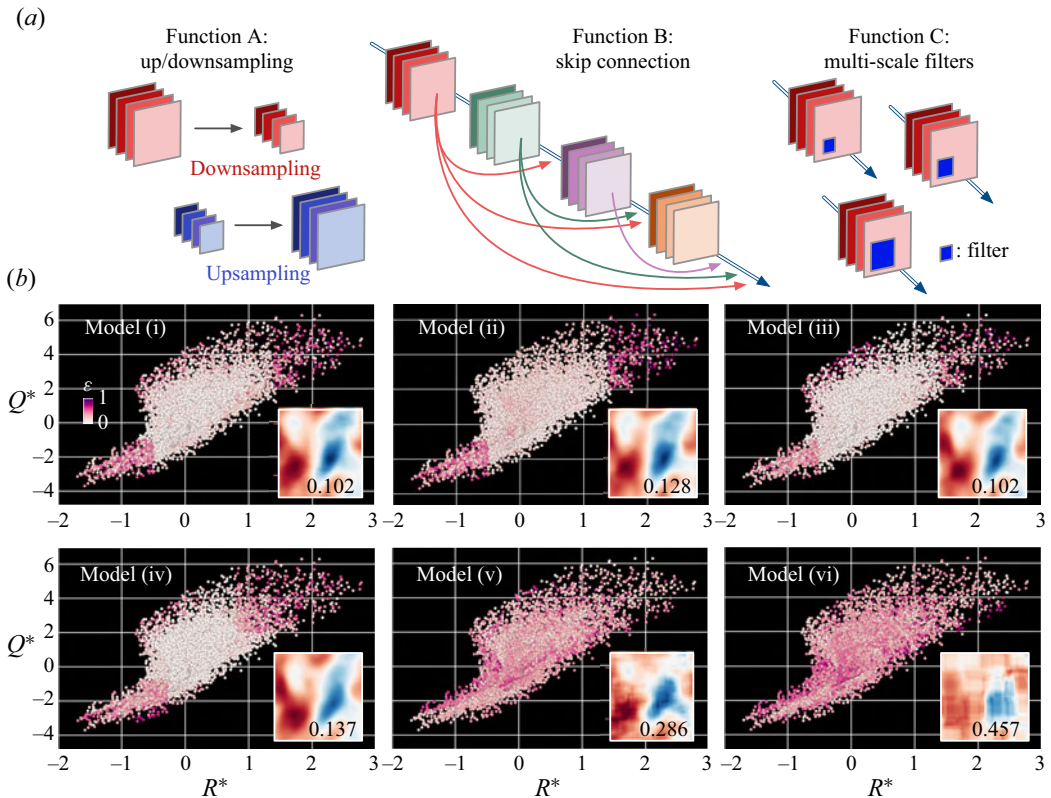
Figure 7. (*a*) The three functions used in the DSC/MS super-resolution model. (*b*) The scaled Q^*–R^* data for six models. The two-dimensional sections of the reconstructed streamwise velocity are also shown with velocity errors.

& Verma 2023) and Fourier analysis of the CNN filters (Subel *et al.* 2023). Along with these techniques, model effectiveness and accuracy should be carefully examined in terms of its construction and the richness of training data for machine-learning applications of turbulent flows.

4. Concluding remarks

For nonlinear machine-learning models for turbulent flows, there is generally not a single parameter that can reveal whether such models are performing an interpolation or extrapolation. This is due to the turbulent flow data containing similarities in flow structures across a range of spatiotemporal scales. These properties contribute to machine-learning models being accurate beyond the coverage of training data in some cases. To shed light on the validity of machine-learning models, we nonlinearly scaled the invariants of the velocity gradient tensor Q and R with non-dimensional parameters using a Buckingham-Pi-theorem-based sparse regression, which maximizes the similarity of invariant data distributions over space and time across Reynolds numbers. As a canonical turbulent flow example, we considered three-dimensional decaying isotropic turbulence for $0.85 \leq Re_\lambda \leq 252$. The present approach found nonlinear scalings that express the influence of the decaying nature of the present flow, domain size of the simulation, and non-equilibrium effects of energy cascade on the invariants Q and R. With the scaled

invariants Q^* and R^*, we were able to determine that the training data lacked flow structures associated with strong biaxial strain and vortex stretching.

We further analysed which types of machine-learning functions contribute to gaining robustness for scale-invariant vortical structures within the context of super-resolution reconstruction. We found that fluid flow reconstruction can be achieved for data of the overlapping portion on the scaled Q^*–R^* plane even under untrained Re_λ with an appropriate construction of machine-learning models. The present findings suggest that transfer learning could be effective for training nonlinear machine-learning models of turbulent flow across Reynolds number provided that extrapolatory structures do not alter the physics significantly (Inubushi & Goto 2020; Guastoni *et al.* 2021). Including fractional exponents for Pi variables would likely enhance the generalizability of the present method for discovering additional nonlinear scaling in turbulent flows. Incorporating nonlinearly scaled invariants into the training process could also support the generalizability of the models. While we can consider the scaled invariants as the input and output of machine-learning models, they could also be incorporated into the loss function of machine-learning optimization. The present procedure to examine nonlinear scalings of turbulent flow structures provides guidance on how to develop robust machine-learning models and compile the necessary training data, enabling us to depart from naïve training and being unaware of the validity of these complex models.

Declaration of interests. The authors report no conflict of interest.

Author ORCIDs.

Ⓑ Kai Fukami https://orcid.org/0000-0002-1381-7322;

Ⓑ Susumu Goto https://orcid.org/0000-0001-7013-7967;

Ⓑ Kunihiko Taira https://orcid.org/0000-0002-3762-8075.

REFERENCES

BAKARJI, J., CALLAHAM, J., BRUNTON, S.L. & KUTZ, J.N. 2022 Dimensionally consistent learning with Buckingham Pi. *Nat. Comput. Sci.* **2**, 834–844.

BRUNTON, S.L., PROCTOR, J.L. & KUTZ, J.N. 2016 Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl Acad. Sci. USA* **113** (15), 3932–3937.

BUCKINGHAM, E. 1914 On physically similar systems; illustrations of the use of dimensional equations. *Phys. Rev.* **4** (4), 345–376.

CHONG, M.S., PERRY, A.E. & CANTWELL, B.J. 1990 A general classification of three-dimensional flow fields. *Phys. Fluids* A **2** (5), 765–777.

CHUMAKOV, S.G. 2008 A priori study of subgrid-scale flux of a passive scalar in isotropic homogeneous turbulence. *Phys. Rev.* E **78** (3), 036313.

DAVIDSON, P.A. 2015 *Turbulence: An Introduction for Scientists And Engineers*. Oxford University Press.

DOMINGOS, P. 2012 A few useful things to know about machine learning. *Commun. ACM* **55** (10), 78–87.

FUKAMI, K., FUKAGATA, K. & TAIRA, K. 2019 Super-resolution reconstruction of turbulent flows with machine learning. *J. Fluid Mech.* **870**, 106–120.

FUKAMI, K., FUKAGATA, K. & TAIRA, K. 2021*a* Machine-learning-based spatio-temporal super resolution reconstruction of turbulent flows. *J. Fluid Mech.* **909**, A9.

FUKAMI, K., FUKAGATA, K. & TAIRA, K. 2023 Super-resolution analysis via machine learning: a survey for fluid flows. *Theor. Comput. Fluid Dyn.* **37**, 421–444.

FUKAMI, K., MURATA, T., ZHANG, K. & FUKAGATA, K. 2021*b* Sparse identification of nonlinear dynamics with low-dimensionalized flow representations. *J. Fluid Mech.* **926**, A10.

GOPALAKRISHNAN MEENA, M. & TAIRA, K. 2021 Identifying vortical network connectors for turbulent flow modification. *J. Fluid Mech.* **915**, A10.

GOTO, S. & VASSILICOS, J.C. 2009 The dissipation rate coefficient of turbulence is not universal and depends on the internal stagnation point structure. *Phys. Fluids* **21** (3), 035104.

GOTO, S. & VASSILICOS, J.C. 2015 Energy dissipation and flux laws for unsteady turbulence. *Phys. Lett.* A **379** (16–17), 1144–1148.

GOTO, S. & VASSILICOS, J.C. 2016 Local equilibrium hypothesis and Taylor's dissipation law. *Fluid Dyn. Res.* **48** (2), 021402.

GUAN, Y., CHATTOPADHYAY, A., SUBEL, A. & HASSANZADEH, P. 2022 Stable a posteriori LES of 2D turbulence using convolutional neural networks: backscattering analysis and generalization to higher *Re* via transfer learning. *J. Comput. Phys.* **458**, 111090.

GUASTONI, L., GÜEMES, A., IANIRO, A., DISCETTI, S., SCHLATTER, P., AZIZPOUR, H. & VINUESA, R. 2021 Convolutional-network models to predict wall-bounded turbulence from wall quantities. *J. Fluid Mech.* **928**, A27.

HE, K., ZHANG, X., REN, S. & SUN, J. 2016 Deep residual learning for image recognition. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778.

INUBUSHI, M. & GOTO, S. 2020 Transfer learning for nonlinear dynamics and its application to fluid turbulence. *Phys. Rev.* E **102** (4), 043301.

JAGODINSKI, E., ZHU, X. & VERMA, S. 2023 Inverse identification of dynamically important regions in turbulent flows using three-dimensional convolutional neural networks. *Phys. Rev. Fluids* **8** (9), 094605.

JARRETT, K., KAVUKCUOGLU, K., RANZATO, M. & LECUN, Y. 2009 What is the best multi-stage architecture for object recognition? In *IEEE 12th ICCV*, pp. 2146–2153.

KAISER, E., KUTZ, J.N. & BRUNTON, S.L. 2018 Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proc. R. Soc.* A **474** (2219), 20180335.

KIM, H., KIM, J., WON, S. & LEE, C. 2021 Unsupervised deep learning for super-resolution reconstruction of turbulence. *J. Fluid Mech.* **910**, A29.

KULLBACK, S. & LEIBLER, R.A. 1951 On information and sufficiency. *Ann. Math. Statist.* **22** (1), 79–86.

LECUN, Y., BOTTOU, L., BENGIO, Y. & HAFFNER, P. 1998 Gradient-based learning applied to document recognition. *Proc. IEEE* **86** (11), 2278–2324.

LEE, S. & YOU, D. 2021 Analysis of a convolutional neural network for predicting unsteady volume wake flow fields. *Phys. Fluids* **33**, 035152.

MARCUS, G. 2018 Deep learning: a critical appraisal. arXiv:1801.00631.

OOI, A., MARTIN, J., SORIA, J. & CHONG, M.S. 1999 A study of the evolution and characteristics of the invariants of the velocity-gradient tensor in isotropic turbulence. *J. Fluid Mech.* **381**, 141–174.

SUBEL, A., GUAN, Y., CHATTOPADHYAY, A. & HASSANZADEH, P. 2023 Explaining the physics of transfer learning in data-driven turbulence modeling. *Proc. Natl Acad. Sci. Nexus* **2** (3), pgad015.

VAN NOORD, N. & POSTMA, E. 2017 Learning scale-variant and scale-invariant features for deep image classification. *Pattern Recognit.* **61**, 583–592.

VASSILICOS, J.C. 2015 Dissipation in turbulent flows. *Annu. Rev. Fluid Mech.* **47**, 95–114.

XIE, X., SAMAEI, A., GUO, J., LIU, W.K. & GAN, Z. 2022 Data-driven discovery of dimensionless numbers and governing laws from scarce measurements. *Nat. Commun.* **13**, 7562.