

Prediction of equivalent sand-grain size and identification of drag-relevant scales of roughness – a data-driven approach

Jiasheng Yang¹, Alexander Stroh¹, Sangseung Lee^{2,3}, Shervin Bagheri⁴, Bettina Frohnäpfel¹ and Pourya Forooghi^{5,†}

¹Institute of Fluid Mechanics, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany

²Department of Mechanical Engineering, Inha University, Incheon 22212, Republic of Korea

³Applied AI Center for Thermal and Fluid Research, Inha University, Incheon 22212, Republic of Korea

⁴Flow Centre, Department of Engineering Mechanics, Royal Institute of Technology, 100 44 Stockholm, Sweden

⁵Department of Mechanical & Production Engineering, Aarhus University, 8200 Aarhus C, Denmark

(Received 28 April 2023; revised 27 August 2023; accepted 11 October 2023)

Despite decades of research, a universal method for prediction of roughness-induced skin friction in a turbulent flow over an arbitrary rough surface is still elusive. The purpose of the present work is to examine two possibilities; first, predicting equivalent sand-grain roughness size k_s based on the roughness height probability density function and power spectrum (PS) leveraging machine learning as a regression tool; and second, extracting information about relevance of different roughness scales to skin-friction drag by interpreting the output of the trained data-driven model. The model is an ensemble neural network (ENN) consisting of 50 deep neural networks. The data for the training of the model are obtained from direct numerical simulations (DNS) of turbulent flow in plane channels over 85 irregular multi-scale roughness samples at friction Reynolds number $Re_\tau = 800$. The 85 roughness samples are selected from a repository of 4200 samples, covering a wide parameter space, through an active learning (AL) framework. The selection is made in several iterations, based on the informativeness of samples in the repository, quantified by the variance of ENN predictions. This AL framework aims to maximize the generalizability of the predictions with a certain amount of data. This is examined using three different testing data sets with different types of roughness, including 21 surfaces from the literature. The model yields overall mean error 5%–10% on different testing data sets. Subsequently, a data interpretation technique, known as layer-wise relevance propagation, is applied to measure the contributions of different roughness

† Email address for correspondence: forooghi@mpe.au.dk

wavelengths to the predicted k_s . High-pass filtering is then applied to the roughness PS to exclude the wavenumbers identified as drag-irrelevant. The filtered rough surfaces are investigated using DNS, and it is demonstrated that despite significant impact of filtering on the roughness topographical appearance and statistics, the skin-friction coefficient of the original roughness is preserved successfully.

Key words: turbulent flows

1. Introduction

Surface degradation in flow-related engineering applications can take various forms, such as wearing or fouling, resulting in roughness on the solid surfaces. The most significant effect of surface roughness, in a practical sense, is an increase in the skin-friction drag under turbulent flow conditions. As an example, the uncertainties in the prediction of roughness-induced skin friction on ship hulls subjected to bio-fouling can cause multiple billion dollars of energy waste every year (Schultz *et al.* 2011; Chung *et al.* 2021). Understandably, study of turbulent flow over rough surfaces has been an active area of research for nearly a century (Nikuradse 1933; Schlichting 1936; Perry, Schofield & Joubert 1969; Krogstad, Antonia & Browne 1992; Raupach 1992; Bhaganagar, Kim & Coleman 2004; Busse, Thakkar & Sandham 2017; Jouybari *et al.* 2022).

The seminal work by Nikuradse (1933) has provided the following researchers with a common ‘currency’ to measure roughness-induced drag; that is, equivalent sand-grain roughness size, k_s , defined as the sand-grain size in Nikuradse’s experiments producing the same skin-friction coefficient as a rough surface of interest in the fully rough regime. Equivalent sand-grain size is related to the downward shift in the logarithmic region of the inner-scaled mean velocity profile observed widely on rough walls, which is referred to as roughness function ΔU^+ (Hama 1954). In the fully rough regime, where the skin-friction coefficient is independent of Reynolds number,

$$\Delta U^+ = \frac{1}{\kappa} \ln(k_s^+) + B - 8.5, \quad (1.1)$$

where κ is the von Kármán constant, and B is the smooth-wall log-law intercept (Jiménez 2004).

One must note that k_s is, by definition, a flow variable and not a geometric one. As a result, for any ‘new’ rough surface, it needs to be determined through a (physical or high-fidelity numerical) experiment in which the skin-friction drag is measured. Obviously, such an exercise is not practical in many applications; therefore, a great amount of effort in the past few decades has been devoted to determining k_s of an arbitrary roughness *a priori*, i.e. based merely on its geometry (see e.g. van Rij, Belnap & Ligrani 2002; Flack & Schultz 2010; Chan *et al.* 2015; Forooghi *et al.* 2017; Thakkar, Busse & Sandham 2017; Flack, Schultz & Barros 2020). A comprehensive description of these efforts can be found in the reviews by Chung *et al.* (2021) and Flack & Chung (2022). Essentially, they can be summarized as attempts to regress correlations between k_s (or ΔU^+) and a few statistical parameters of roughness geometry based on available data. Some widely used parameters in this context are skewness of roughness height probability density function (p.d.f.) (Flack & Schultz 2010), effective (or mean absolute) slope (Napoli, Armenio & DeMarchis 2008), and correlation length of rough surface geometry (Thakkar *et al.* 2017).

As a result of increased computational capacities in recent years, direct numerical simulations (DNS) have become a source of data for development of accurate roughness correlations as pointed out by Flack (2018). In this regard, the idea of DNS in minimal channels, proposed by Chung *et al.* (2015), has enabled characterizing larger numbers of roughness samples with a certain computational resource. Availability of more data, on the one hand, has opened the door to utilization of machine learning (ML) based regression tools, and on the other hand, enables inclusion of more roughness information (beyond only a few parameters) as the input to such tools. The latter point is particularly important since there is increasing evidence that both statistical and spectral information on roughness geometry are required for prediction of flow response to a multi-scale roughness (Alves Portela, Busse & Sandham 2021). In this regard, it has been shown that k_s for multi-scale random roughness can be determined nearly uniquely with a combined knowledge of roughness height p.d.f. and its power spectrum (PS) (Yang *et al.* 2022, 2023).

The first ML-based ‘data-driven’ tool for prediction of k_s has been reported recently by Jouybari *et al.* (2021). These authors used deep neural network and Gaussian process regression to train models with 17 inputs, including widely used roughness parameters and their products. The training data for their model are obtained from DNS of flow over certain types of artificially generated roughness, which were also used to evaluate the model. Lee *et al.* (2022) used a neural network similar to that of Jouybari *et al.* (2021) and showed that improvements in predictive performance can be achieved if the network is ‘pre-trained’ on existing empirical correlations. While these pioneering works deliver promising results, the data-driven approach arguably has the potential to realize truly universal models, which can generalize beyond a certain class of roughness. The present work is an attempt to explore this potential. To this end, a model is trained on a wide variety of multi-scale irregular roughness samples, selected based on an adaptive approach (explained shortly), which is aimed at enhancing the universality of the predictions. This is evaluated using ‘unseen’ roughness from different testing data sets with different natures. Moreover, unlike the previous efforts, the present model incorporates the complete p.d.f. and PS of roughness as inputs rather than a finite set of predetermined parameters.

Considerable attention has been paid in recent literature to the multi-scale nature of realistic roughness and the significance of its ‘spectral content’. It has been suggested that beyond a certain threshold, large roughness wavelengths may impact the roughness-induced drag less significantly (Barros, Schultz & Flack 2018; Yang *et al.* 2022). While parametric studies of roughness PS (Anderson & Meneveau 2011; Barros *et al.* 2018) or Fourier filtering (Busse, Lützner & Sandham 2015; Alves Portela *et al.* 2021) can shed light on this matter, in the present work we explore the possibility of evaluating directly contributions of different roughness scales utilizing the information embedded in the data-driven model developed in this study. This is motivated by the fact that the (discretized) PS is a direct input to the model, which hints at the potential to extract information about the role of different wavelengths through interpretation of the model.

In order to train the data-driven roughness model, k_s for several roughness samples should be determined. This is referred to as ‘labelling’ those samples, borrowing the term from the ML terminology. Moreover, each roughness sample along with its k_s value is called a training ‘data point’. One should note that labelling is a computationally expensive process due to the need to perform DNS. In dealing with such scenarios, ML methods classified under active learning (AL) – also known as query-based learning (Abe & Mamitsuka 1998) or optimal experimental design (Fedorov 1972) in different contexts – have been proven particularly advantageous (Zhu *et al.* 2005; Settles & Craven 2008; Bangert *et al.* 2021). In AL, selection of the training data is navigated in a way such that

the information gain from a certain amount of available data is maximized (Settles 2009). The ‘informativeness’ of a potential data point is commonly measured by the uncertainty in its prediction, which needs to be determined without labelling, e.g. through the standard deviation of the predictive distribution of a Bayesian model (Gal & Ghahramani 2016) or the variation of the predictions among a number of individual models (Raychaudhuri & Hamey 1995).

Two major AL categories can be identified in the literature (Lang & Baum 1992; Lewis & Gale 1994; Angluin 2004). The methods based on membership query synthesis expand an existing data set by creating and labelling new samples that the model is most curious about. In contrast, the methods based on pool-based sampling utilize a ‘bounded’ unlabelled data set (also called a repository) \mathcal{U} , select and label the most informative samples from \mathcal{U} , and include them in the labelled training data set \mathcal{L} . In the present work, pool-based sampling is deemed more suitable as it can prevent creating unrealistic samples (Lang & Baum 1992). Moreover, identification of the most informative samples follows a query-by-committee (QBC) strategy (Seung, Opper & Sompolinsky 1992), in which variance in the outputs of an ensemble of individual models (the committee) is the basis for the next query. A detailed description of the implemented QBC is provided in § 2.

In summary, the present work aims to answer two questions; first, whether ‘universal’ data-driven predictions of k_s can be approached using a complete statistical-spectral representation of roughness (i.e. with p.d.f. and PS as inputs). We leverage AL to facilitate achieving this goal. The second question is whether and how the information embedded in a data-driven model can provide insight on the contributions of different roughness scales to the added drag. Following this introduction, the roughness generation approach, DNS and the ML methodology are described in § 2. In § 3, first the results and performance of the model are discussed, then the analysis of drag-relevant scales is presented. Section 4 summarizes the main conclusions.

2. Methodology

2.1. Roughness repository

The (unlabelled) roughness ‘repository’ \mathcal{U} is constructed by a collection of 4200 artificial irregular rough surfaces. These surfaces are generated through a mathematical roughness generation method where the PS and p.d.f. of each roughness can be prescribed (Pérez-Ràfols & Almqvist 2019). For creation of the present repository, p.d.f. and PS are parametrized, as described shortly, and their parameters are varied randomly within a realistic range to generate a variety of roughness samples while imitating the random nature of roughness formation in practical applications.

In total, three types of p.d.f. – namely, Gaussian, Weibull and bimodal – are used, and for each new roughness added to the repository, one type is randomly selected. The Weibull distribution of random variable k – here the roughness height – follows

$$f_W(k) = K\beta^K k^{(K-1)} \exp(-(\beta k)^K), \quad (2.1)$$

where the shape parameter $0.7 < K < 1.7$ is selected randomly with $\beta = 1.0$. In the present notation, k denotes the local roughness height as a function of wall-parallel coordinates (x, z) . The bimodal distribution is obtained by combining two Gaussian distributions through (Peng & Bhushan 2000)

$$f_B(k) = f_G(k|0, 1) + f_G(k|\mu, \sigma) - f_G(k|0, 1)f_G(k|\mu, \sigma), \quad (2.2)$$

where $f_G(x|\mu, \sigma)$ is the p.d.f. of the Gaussian distribution with randomized mean $0 < \mu < 0.5$ and randomized standard deviation $0 < \sigma < 0.5$. The p.d.f. variable k is then scaled from 0 to the roughness peak-to-trough height $k_t = \max(k) - \min(k)$, whose value is determined randomly in the range $0.06 < k_t/H < 0.18$, where H is the channel half-height.

The PS of the roughness samples in the repository is controlled by two randomized parameters, namely the roll-off length L_r (Jacobs, Junge & Pastewka 2017) and the power-law decline rate θ_{PS} (Lyashenko, Pastewka & Persson 2013), whose values are selected in the ranges $0.1 < L_r/(\log(\lambda_0/\lambda_1)) < 0.6$ and $-3 < \theta_{PS} < -0.1$. Here, λ_0 and λ_1 represent the upper and lower bounds of the PS, or the largest and smallest wavelengths forming the roughness topography. Random perturbations are added to the PS to achieve higher randomness in PS. The lower bound of the roughness wavelength is set to $\lambda_1 = 0.04H$ to ensure that the finest structures can be discretized by an adequate number of grid points. The upper bound of the roughness wavelength λ_0 is selected randomly in the range $0.5H < \lambda_0 < 2H$. As will be discussed later, the roughness sample size as well as the simulation domain size should both be adjusted to accommodate this wavelength.

Eventually, 4200 separate pairs of p.d.f. and PS are generated using the described random process, each leading to one rough surface added to the repository \mathcal{U} . A representation of the parameter space covered by these samples is illustrated in § 3.1. Moreover, examples of the generated samples can be seen in Appendix A.

2.2. Direct numerical simulations

Direct numerical simulations are employed to solve the turbulent flow over selected rough surfaces from the repository in a plane channel driven by a constant pressure gradient. Each simulation leads to determination of the k_s value for the respective roughness sample – a practice referred to as ‘labelling’ in this paper. The DNS are performed with a pseudo-spectral Navier–Stokes solver SIMSON (Chevalier *et al.* 2007). Fourier and Chebyshev series are employed for the discretization in wall-parallel and wall-normal directions, respectively. Time integration is carried out using a third-order Runge–Kutta method for the advective and forcing terms, and a second-order Crank–Nicolson method for the viscous terms. The roughness representation in the fluid domain is based on the immersed boundary method (IBM) of Goldstein, Handler & Sirovich (1993). The code and the IBM have been validated previously and used in several publications in the past (Forooghi *et al.* 2018a; Vanderwel *et al.* 2019; Yang *et al.* 2022). The solved Navier–Stokes equation gives

$$\nabla \cdot \mathbf{u} = 0, \tag{2.3}$$

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot (\mathbf{u}\mathbf{u}) = -\frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{u} - \frac{1}{\rho} P_x \mathbf{e}_x + \mathbf{f}_{IBM}, \tag{2.4}$$

where $\mathbf{u} = (u, v, w)^T$ is the velocity vector, and P_x is the mean pressure gradient in the flow direction added as a constant and uniform source term to the momentum equation to drive the flow. Moreover, p , \mathbf{e}_x , ρ , ν and \mathbf{f}_{IBM} denote pressure fluctuation, streamwise unit vector, density, kinematic viscosity and external body force term due to the IBM, respectively. Periodic boundary conditions are applied in the streamwise and spanwise directions. The friction Reynolds number is defined as $Re_\tau = u_\tau(H - k_{md})/\nu$, where $u_\tau = \sqrt{\tau_w/\rho}$ and $\tau_w = -P_x(H - k_{md})$ are the friction velocity and the wall shear stress, respectively. Here, H and $H - k_{md}$ are channel half-height without and with roughness,

λ_0/H	L_x/H	L_z/H	L_y/H	N_x	N_z	N_y	Δ_x^+	Δ_z^+	$\Delta_{y,min}^+$	$\Delta_{y,max}^+$
$0.5 \geq \lambda_0/H \geq 0.6$	1.8	0.6	2	400	160	451	3.6	3.0	0.02	5.7
$0.6 > \lambda_0/H \geq 0.7$	2.1	0.7	2	576	144	451	2.9	3.9	0.02	5.7
$0.7 > \lambda_0/H \geq 0.8$	2.4	0.8	2	480	192	451	4.0	3.3	0.02	5.7
$0.8 > \lambda_0/H \geq 0.9$	2.7	0.9	2	640	256	451	3.4	2.8	0.02	5.7
$0.9 > \lambda_0/H \geq 1.0$	3.0	1.0	2	640	256	451	3.8	3.1	0.02	5.7
$1.0 > \lambda_0/H \geq 1.1$	3.3	1.1	2	720	288	451	3.7	3.1	0.02	5.7
$1.1 > \lambda_0/H \geq 1.2$	3.6	1.2	2	720	288	451	4.0	3.3	0.02	5.7
$1.2 > \lambda_0/H \geq 1.3$	3.9	1.3	2	800	320	451	3.9	3.3	0.02	5.7
$1.3 > \lambda_0/H \geq 1.4$	4.2	1.4	2	960	384	451	3.5	2.9	0.02	5.7
$1.4 > \lambda_0/H \geq 1.5$	4.5	1.5	2	960	384	451	3.8	3.1	0.02	5.7
$1.5 > \lambda_0/H \geq 1.6$	4.8	1.6	2	960	384	451	4.0	3.3	0.02	5.7
$1.6 > \lambda_0/H \geq 1.7$	5.1	1.7	2	1080	432	451	3.8	3.1	0.02	5.7
$1.7 > \lambda_0/H \geq 1.8$	5.4	1.8	2	1200	480	451	3.6	3.0	0.02	5.7
$1.8 > \lambda_0/H \geq 1.9$	5.7	1.9	2	1200	480	451	3.8	3.2	0.02	5.7
$1.9 > \lambda_0/H \geq 2.0$	6.0	2.0	2	1200	480	451	4.0	3.3	0.02	5.7

Table 1. Simulation set-ups.

k_{md} being the mean (meltdown) roughness height. In the present work, all simulations are performed at $Re_\tau = 800$.

Due to the high computational demand of many DNS, the concept of DNS in minimal channels (Chung *et al.* 2015; MacDonald *et al.* 2016) is adopted for the considered simulations. Recently, Yang *et al.* (2022) showed the applicability of this concept for flow over irregular roughness subject to certain criteria. Accordingly, a roughness function over a rough surface can be predicted accurately by a comparison of mean velocity profiles in smooth and rough minimal channels if the size of the channels satisfies the following conditions:

$$L_z^+ \geq \max \left(100, \frac{\tilde{k}^+}{0.4}, \lambda_0^+ \right), \quad L_x^+ \geq \max (1000, 3L_z^+, \lambda_0^+). \quad (2.5a,b)$$

Here, L_z and L_x are the spanwise and streamwise extents of the minimal channel, respectively, λ_0 is the largest wavelength in the roughness spectrum, and \tilde{k} is the characteristic physical roughness height. The plus superscript indicates viscous scaling hereafter. The above condition suggests that the minimal channel size of each roughness should be determined based on λ_0 (which in practice defines the most strict constraint). As described before, λ_0 is known for each generated roughness sample. Table 1 summarizes the simulation set up for all DNS based on the respective λ_0 value. Due to the different sizes of the simulation domains, the chosen numbers of grid points differ according to the mesh size, but in all cases, $\Delta_{x,z}^+ \leq 4$. In wall-normal directions, cosine stretching mesh is adopted for the Chebyshev discretization. The mesh independence is confirmed in a set of additional tests.

For each investigated roughness, ΔU^+ is determined from the offset in the logarithmic velocity profile comparing corresponding rough and smooth DNS. Notably, when plotting mean velocity profiles, zero-plane displacement y_0 is applied in order to achieve parallel velocity profiles in the logarithmic layer, where y_0 is determined as the moment centroid of the drag profile on the rough surface following Jackson’s method (Jackson 1981). It is worth noting that in the extensive literature on rough wall-bounded turbulent flows,

various definitions of y_0 have been proposed, and furthermore, the choice of virtual wall position can affect the predicted rough-wall shear stress τ_w and thus the resulting k_s value (Chan-Braun, García-Villalba & Uhlmann 2011). Therefore, it is important to recognize this as a possible source of uncertainty, and take into account the definitions of τ_w and y_0 when comparing data from different sources.

It is also important to determine if the flow has reached the fully rough regime in each simulation. To this end, ΔU^+ is combined with (1.1) to yield a testing value of k_s^+ . Then, following the threshold adopted by Jouybari *et al.* (2021), a roughness with $k_s^+ \geq 50$ is deemed to be in the fully rough regime, and all samples not matching this criterion are excluded from the training or testing process. The selected threshold $k_s^+ \geq 50$ is somewhat lower than the common threshold of $k_s^+ \geq 70$ (Flack & Schultz 2010) and thus may introduce into the database some data points with limited transitionally rough behaviour. This threshold is, however, chosen deliberately as a trade-off to maximize the number of training data given the limited computational resources. One should note that an increase in the threshold value of k_s^+ while maintaining the same parameter space would be possible by increasing Re_τ . This would, however, lead to an obvious compromise in the final performance of the model by reducing the number of training data points at a given computational cost.

Overall, 85 roughness samples are DNS-labelled and eventually included in the labelled data set \mathcal{L} to train the final AL-based model. The procedure for selection of these training samples is explained in detail in the following. Eight out of the 85 labelled samples are located in the range of $50 \leq k_s^+ \leq 70$. We observe that incorporating these samples into the training process improves model performance. This improvement in the model performance can be attributed both to the incorporation of more informative samples according to AL as well as to the regularization effect of data diversity introduced by including transitionally rough training samples, which makes the model more robust and mitigates over-fitting (Bishop 1995; Reed & Marks 1999).

2.3. Machine learning

The ML model in the present work is constructed in a QBC fashion by building an ensemble neural network (ENN) model consisting of 50 independent neural networks (NNs) with identical architecture as the ‘committee’ members. Similar to the methods proposed by Raychaudhuri & Hamey (1995) and Burbidge, Jem & King (2007), the prediction uncertainty of the ENN model is defined as the variance of the predictions among the members, σ_{k_r} .

The workflow of the AL framework is sketched in figure 1. Two collections of roughness samples are included in the framework. These are the (unlabelled) repository \mathcal{U} and the (labelled) training data set \mathcal{L} . As a starting point in the AL framework, 30 samples are selected randomly from the repository, labelled (i.e. their k_s is calculated) through DNS, and used to train a first ENN model, which is referred to as the ‘base model’. This preliminary base model is subsequently improved throughout multiple AL iterations. In each AL iteration, approximately 20 new roughness samples from \mathcal{U} are DNS-labelled and added to \mathcal{L} for training of the ENN. These are the samples in \mathcal{U} with the highest prediction variances according to the most recent ENN. This QBC strategy leads to an effective exploration of the repository and adding the new data at the most uncertain regions of the parameter space.

The function of the ENN model is to regress the (dimensionless) equivalent sand-grain roughness $k_r = k_s/k_{99}$, and to calculate the variance of the predictions σ_{k_r} , as a basis for

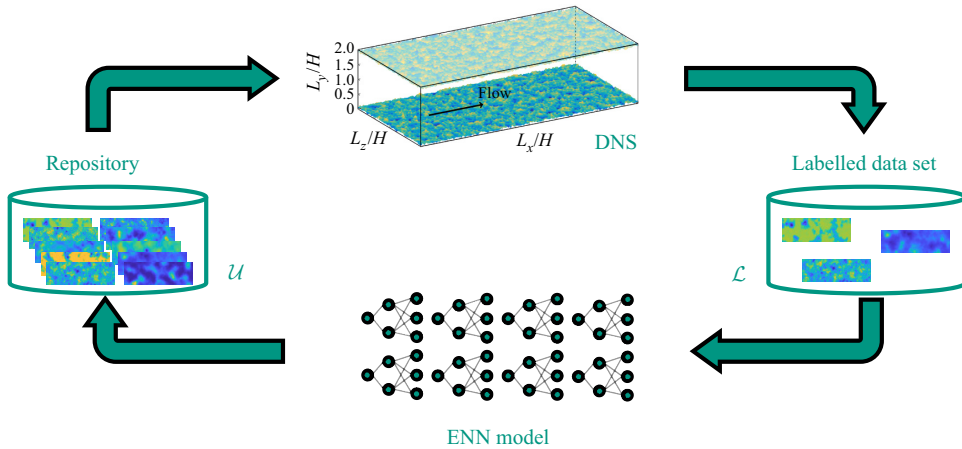


Figure 1. Schematic of the AL framework.

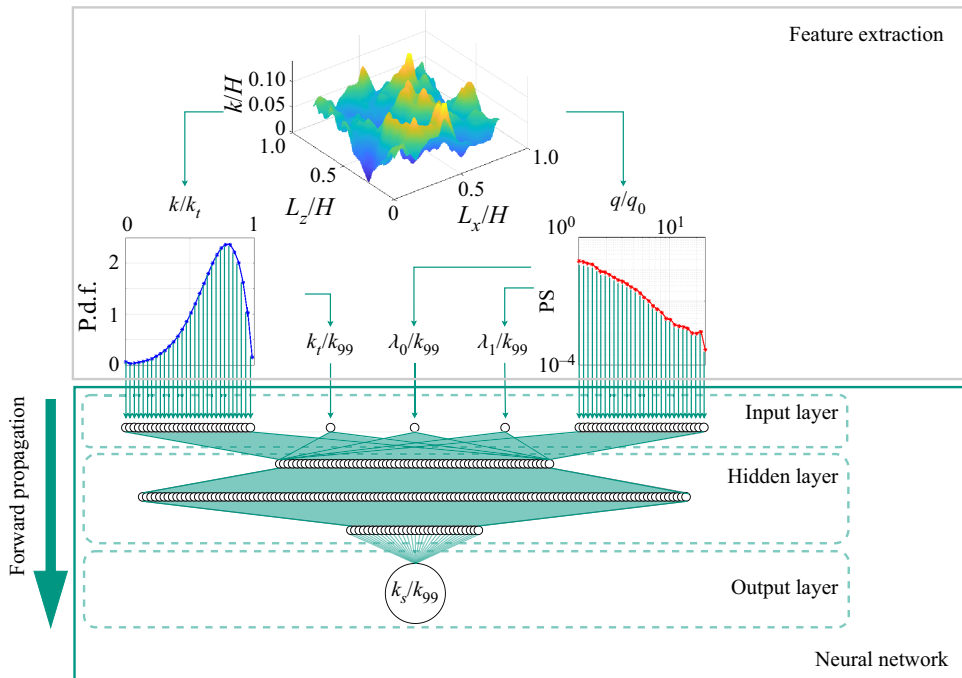


Figure 2. Schematic of a single NN in an ENN.

QBC (k_{99} is the 99 % confidence interval of the roughness p.d.f., which is used as the representative physical scale of roughness height in this paper). The ENN is composed of multiple NNs with similar structures that is shown in figure 2. The input vector I of the NN contains the discretized roughness p.d.f. and PS along with three additional characteristic features of the rough surface, i.e. k_t/k_{99} and the normalized largest and smallest roughness wavelength $\lambda_0^* = \lambda_0/k_{99}$ and $\lambda_1^* = \lambda_1/k_{99}$, respectively. The input elements in I that represent the roughness p.d.f. and PS are obtained by discretizing equidistantly the roughness p.d.f. and PS each into 30 values within the height range $0 < k < k_t$ and the wavenumber range $2\pi/\lambda_1 > 2\pi/\lambda > 2\pi/\lambda_0$. Each NN in the ensemble is constructed

with one input layer with 63 (3 + 30 + 30) input elements, three hidden layers with 64, 128 and 32 nonlinear neurons with rectified linear units (ReLU) activation ($\max\{0, x\}$), and one linear neuron in the output layer. The optimal number of neurons at each layer is determined through a grid search of a range of numbers that achieves the lowest model prediction error on \mathcal{T}_{inter} . The L2-regularization is applied to the loss function. Adaptive momentum estimation (Adam) is employed to train the model. The final prediction of the ENN is defined as the mean prediction over the 50 NNs, namely $\mu_{k_r} = \sum_{i=1}^{50} \hat{k}_{r,i}/50$, where \hat{k}_r represents the prediction of a single NN, and the index i indicates the index of the NN. The prediction variance is calculated as $\sigma_{k_r} = \sqrt{\sum_{i=1}^{50} (\hat{k}_{r,i} - \mu_{k_r})^2/50}$. It is worth noting that each NN in the ENN model is trained individually based on 90% of the randomly selected samples in the labelled data set \mathcal{L} , while the rest of the samples are used for validation. The initial weights of the neurons in each NN are assigned randomly at the beginning of the training process. In such a way, the diversity among the QBC members is ensured, which is an important factor in determining the generalization of the ENN model (Melville & Mooney 2003). It is important to note that the current ensemble members used in the model are deterministic NNs, and the uncertainty of the training data from DNS is assumed to be minimal. However, when considering experimental training data, where (aleatoric) uncertainties arise from possible measurement errors, the performance of the current ENN approach may be compromised due to its limited capability in handling such uncertainties. In these scenarios, the utilization of probabilistic models – such as Bayesian NNs – may be more suitable as they allow for the explicit incorporation of measurement uncertainties.

2.4. Testing data sets

In the present work, three distinct testing data sets are introduced to evaluate the model performance and its universality. The difference among the data sets lies in the nature and origin of the samples that they contain. The first data set, \mathcal{T}_{inter} , is composed of 20 samples chosen randomly from \mathcal{U} that have never been seen by the model during the training process.

Despite the fact that the employed roughness generation method can generate irregular, multi-scale surfaces resembling realistic roughness, we test the model separately for additional rough surfaces extracted from scanning of naturally occurring roughness, which form the second testing data set, $\mathcal{T}_{ext,1}$. There are five samples in this ‘external’ data set. These include roughness generated by ice accretion (Velandia & Bansmer 2019), deposit in internal combustion engine (Forooghi *et al.* 2018b), and a grit-blasted surface (Thakkar *et al.* 2017). In addition to that, we test the model against a second external data set, $\mathcal{T}_{ext,2}$, which contains irregular roughness samples from the database provided by Jouybari *et al.* (2021). In this data set, many roughness samples are generated by placing ellipsoidal elements of different sizes and orientations on a smooth wall, making them rather distinct from the type of roughness used to train the model. We separate this testing data set from the other two as it contains a specific type of artificial roughness.

3. Results

3.1. Assessment of the AL framework

In this subsection, we explore if the AL framework enhances the training behaviour of the model. To do so, we compare a model trained with AL-selected data points to one trained

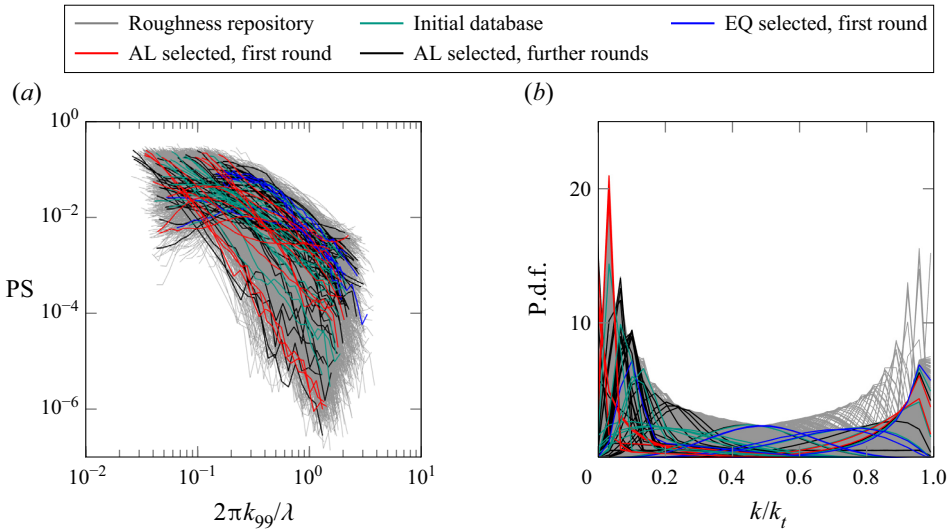


Figure 3. Plots of (a) PS and (b) p.d.f. of 4200 roughness samples in the roughness repository (grey). The samples selected for training are distinguished with different colours. While the AL model tends to explore the PS and p.d.f. domain, the EQ model contains samples that are placed closely to the known initial database.

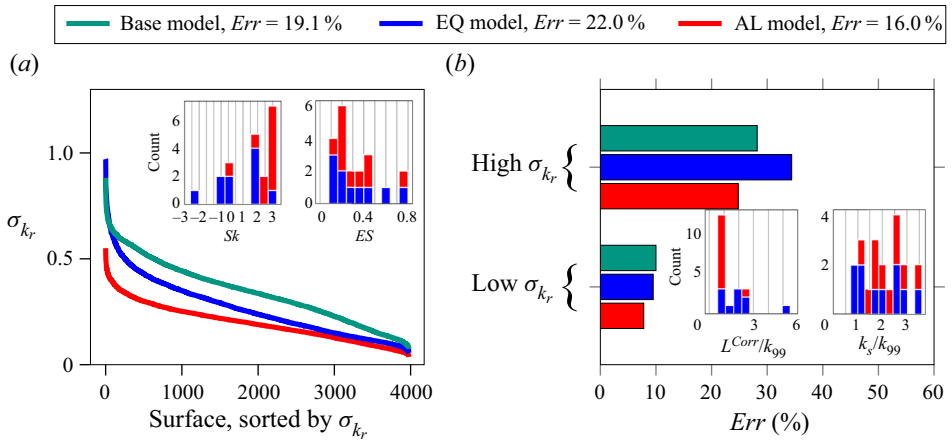


Figure 4. (a) Prediction variance σ_{k_r} obtained by three different models for all the samples in repository \mathcal{U} . (b) The average error obtained by the three models for 10 high-variance samples and 10 low-variance samples in \mathcal{T}_{inter} (sorted based on the variance of the base model). The total averaged errors are displayed in the legend. Insets show the distribution of the statistical parameters as well as the corresponding k_r of the new samples with AL and EQ sampling strategies with identical colour code.

with an arbitrary selection of data points. To avoid the computational cost of running many eventually unused DNS, the comparison is made for only one AL iteration. Figure 3 shows all p.d.f. and PS pairs contained in the repository \mathcal{U} (grey) and those randomly selected for the initial base model (green), as well as those selected for further training (other colours). The wide range of available roughness can be understood from the area covered by grey curves. As explained before, once the base model is trained using the initial randomly selected data set, it is used to determine which samples from the repository

\mathcal{U} should be selected for the next round of training. In figure 4(a), the green line shows the prediction variance σ_{k_r} of all roughness samples in the repository based on the base model. Here, the abscissa is the sample number sorted from high to low σ_{k_r} values. According to the AL framework, the samples selected for the next round are the ones with the largest σ_{k_r} . These are shown in red in figure 3. For comparison, a second sampling strategy (denoted as EQ) is employed in which the same number of samples as in AL are selected, but they are distributed equidistantly along the abscissa of figure 4(a). These samples are shown in blue in figure 3. It is observed clearly in figure 3 that the AL model explores surfaces that are least similar to those in the initial data set (green) and tend to cover the entire repository, with a higher weight given to the marginal cases. Furthermore, the parameter distribution as well as the corresponding k_r values of the selected roughness by means of AL and EQ is compared in the insets of figure 4. It can be seen that both the AL and EQ models generally prioritize selecting samples within the waviness regime, i.e. effective slope $ES < 0.35$ (Napoli *et al.* 2008). This preference may arise from the fact that the resulting drag in the waviness regime ($ES < 0.35$) is sensitive to changes in ES (Schultz & Flack 2009). Conversely, beyond this regime ($ES > 0.35$), the resulting ΔU^+ saturates in relation to increasing ES , making these samples less interesting for both labelling strategies. On the other hand, the AL model particularly tends to sample the roughness with positive skewness and low correlation length. This can similarly be a result of the roughness effect being highly sensitive to the variations in roughness statistics within these ranges of parameters, which is in line with previous findings (Schultz & Flack 2009; Busse & Jelly 2023).

Subsequently, two separate models are trained based on the AL and EQ strategies. These models are applied separately to determine the variance of prediction for roughness in the repository, and the results are depicted in figure 4(a) using red and blue lines. It is evident from the results that both the AL and EQ models generally reduce the prediction variance. However, a more substantial decline in the values of σ_{k_r} is achieved by the AL model. This is the expected behaviour as AL is designed to reduce the prediction uncertainty by targeting regions of the parameter space where the uncertainty is the largest. Interestingly, some increase in σ_{k_r} of the EQ model can be observed for a number of samples with very high σ_{k_r} , which can be a sign that the performance of the EQ model in the ‘difficult’ tasks deteriorates as it is not trained well for those tasks due to ineffective selection of its training data. Moreover, the prediction errors (calculated based on correct k_s values of testing data set \mathcal{T}_{inter} obtained by DNS) are illustrated in figure 4(b). The averaged prediction errors, Err , achieved by the base model, the AL model and the EQ model for the entire \mathcal{T}_{inter} are 19.1 %, 16.0 % and 22.0 %, respectively. While the AL model yields a meaningful reduction in Err , the overall performance of the EQ model deteriorates, possibly due to the over-fitting, which in our case refers to the condition where the model is trained to fit a limited number of relatively similar data points so precisely that its ability to extrapolate on dissimilar testing data is degraded (Hastie, Tibshirani & Friedman 2009). To better analyse this observation, the testing data set \mathcal{T}_{inter} is split evenly into two subsets according to their σ_{k_r} , namely the high- and low-variance subsets. The Err values for both high- and low-variance subsets are illustrated in the figure. It is clear that while the EQ strategy improves the model performance for the already low-variance test data, its error increases for high-variance test data, which can be taken as an indication of over-fitting as described above. The AL sampling strategy, in contrast, seems to protect the model from over-fitting – especially in the circumstance of a small training data set – hence the error is reduced for both high- and low-certainty test data as a result of effective selection of training data.

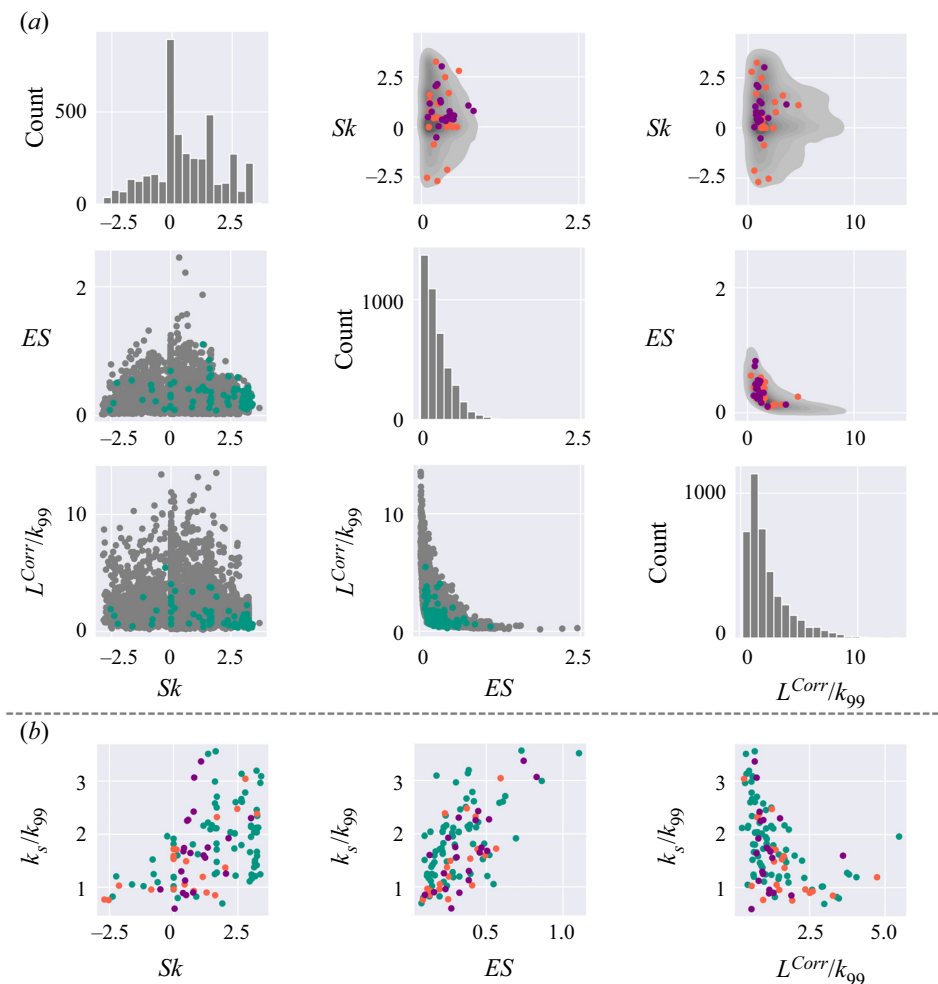


Figure 5. (a) Pair plots of roughness statistics. Lower left: the distributions of the samples in \mathcal{U} (grey) and \mathcal{L} (green). Diagonal: histograms of single roughness statistics in \mathcal{U} . Upper right: joint probability distributions of statistics overlaid by test data in \mathcal{T}_{inter} (orange) and $\mathcal{T}_{ext,1\&2}$ (purple). (b) Values of $k_r = k_s/k_{99}$ obtained from DNS (ground truth) as a function of the selected statistics. Colour code is the same as in (a).

3.2. Performance of the final model

Having demonstrated the advantage of AL over random sampling, three additional AL iterations are carried out. The distributions of the PS and p.d.f. of the selected roughness from the second to the fourth AL iterations are displayed in figure 3 with black lines. A number of roughness maps from each AL round are also displayed in Appendix A.

The total number of data points for training of the model after four iterations adds up to 85; these are the data that form \mathcal{L} . The scatter plots of some widely investigated roughness parameters in \mathcal{L} as well as in the unlabelled repository \mathcal{U} are displayed in the lower left part of figure 5(a). In the figure, $k(x, z)$ is the elevation map of the roughness, $Sk = 1/(Sk_{rms}^3) \int_S (k - k_{md})^3 dS$ represents the skewness, where S is the wall-projected surface area, and $k_{md} = (1/S) \int_S k dS$ is the meltdown height of the roughness. The effective slope is defined as $ES = (1/S) \int_S |\partial k/\partial x| dS$. Here, L^{Corr} is the correlation length representing the horizontal separation at which the roughness height autocorrelation function drops

Predicting k_s and identifying drag-relevant roughness scales

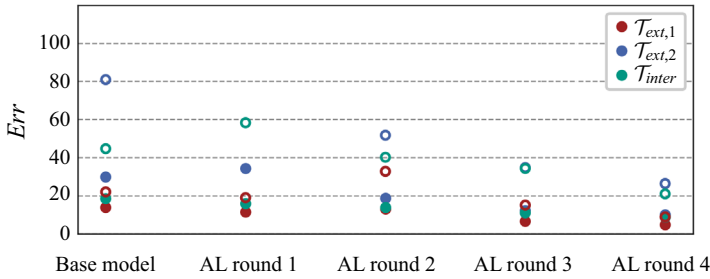


Figure 6. The arithmetically averaged Err (%) as well as maximum Err of the model after different training rounds on each of the testing data sets \mathcal{T}_{inter} , $\mathcal{T}_{ext,1}$ and $\mathcal{T}_{ext,2}$. The mean Err is represented with a closed circle, while the maximum Err is displayed with an open circle of corresponding colour. The maximum Err for $\mathcal{T}_{ext,2}$ at AL round 1 is out of the plot range.

under 0.2. An inverse correlation can be observed between L^{Corr} and ES , which is expected as roughness with larger dominant wavelength tends to have lower mean slope. The distribution of other statistics in \mathcal{U} appears to be reasonably random.

For the sake of comparison, additionally the test data are represented in the upper right part of figure 5(a), with orange (for \mathcal{T}_{inter}) and purple (for $\mathcal{T}_{ext,1\&2}$) symbols. It is worth noting that only the roughness samples that locate in the fully rough regime at the currently investigated Re_τ are included in \mathcal{L} and shown in the figure. Figure 5(b) shows the values of k_r (from DNS) against the three roughness statistics for all labelled data in the training and testing data sets. As can be observed clearly in the figure, while equivalent sand-grain roughness shows some general correlation with each of these statistics (increasing with Sk and ES , decreasing with L^{Corr}), the collapse of data is far from perfect. Clearly, no roughness statistics can capture entirely the effect of an irregular multi-scale roughness topography on drag, which is essentially a motivation behind seeking an NN-based model to find the functional relation between k_s and a higher-order representation of roughness (here p.d.f. and PS).

Eventually, the final model is trained on the entire labelled data set \mathcal{L} . The mean and maximum error values achieved by this model on all three testing data sets, as well as those errors after each training round, are displayed separately in figure 6. The figure shows a generally decreasing trend in both mean and maximum error as the model is trained progressively for more AL rounds, despite some exceptions to the general trend in the first two rounds when the number of data points is low. It is notable that the AL model is particularly successful in bringing down the maximum error, and hence can be considered reliable over a wide range of scenarios.

One should mention that the model performs consistently well for three different testing data sets with different natures. While the data set \mathcal{T}_{inter} covers an extensive parameter space – hence containing more extreme cases – it is generated employing the same method as the training data. Therefore, to avoid a biased evaluation of the model, two ‘external’ testing data sets from literature are also included. The data set $\mathcal{T}_{ext,2}$ is believed to be particularly challenging for the model, since it is formed by roughness generated artificially using discrete elements (Jouybari *et al.* 2021), which is fundamentally different from the target roughness of this study. Nevertheless, the final model yields very similar errors for all data sets; what can be taken as an indication of its generalizability. The averaged errors of the final model within the data sets \mathcal{T}_{inter} , $\mathcal{T}_{ext,1}$, and $\mathcal{T}_{ext,2}$ are approximately 9.3%, 5.2% and 10.2%, respectively.

It is crucial to acknowledge that the present model is developed under the assumption of statistical surface homogeneity. However, when reaching beyond this assumption, the presence of surface heterogeneity introduces additional complexity to the problem that cannot be represented adequately by the current training samples. As a consequence, the effect of heterogeneous roughness structures (Hinze 1967; Stroh *et al.* 2020) cannot be accounted for adequately by the current model.

3.3. Data-driven exploration of drag-relevant roughness scales

The fact that naturally occurring roughness usually has a multi-scale nature with continuous spectrum is well established (Sayles & Thomas 1978). How spectral content of roughness affects skin-friction drag, and whether a certain range of length scales dominates it, are, however, questions receiving attention more recently (Anderson & Meneveau 2011; Mejia-Alvarez & Christensen 2010; Barros *et al.* 2018; Medjnoun *et al.* 2021). In this sense, Busse *et al.* (2015) applied low-pass Fourier filtering to a realistic roughness and observed no significant effect on skin-friction drag when the filtered wavelengths were lower than a certain threshold. On the other hand, Barros *et al.* (2018) used high-pass filtering and suggested that very large length scales may not contribute significantly to drag. Alves Portela *et al.* (2021) examined three filtered surfaces, each maintaining one-third of the original spectral content associated with large, intermediate or small scales. In all cases, the filtered scales were shown to include ‘drag-relevant’ information. While both lower and higher limits of drag-relevant scales (if they exist) can be a matter of discussion, the present study focuses mainly on the latter. Possibly related to that question, Schultz & Flack (2009) documented the equivalent sand-grain size of pyramid-like roughness with wavelengths higher (hence lower effective slopes) than a certain value not to scale in the same way as those with smaller wavelengths. These authors coined the term ‘wavy’ for the high-wavelength roughness behaviour. Later, Yuan & Piomelli (2014a) revealed that the wavy regime may emerge at a different threshold (in terms of effective slope) in a multi-scale roughness compared to the single-scale pyramid-like roughness. Recently, Yang *et al.* (2022) showed that the spectral coherence of roughness topography and time-averaged drag force on a rough wall drops at large streamwise wavelengths, which, in line with the finding of Barros *et al.* (2018), suggests decreasing drag relevance of large scales.

In the present work, we are particularly interested to explore the possibility of extracting the drag-relevant scales from the knowledge embedded in the data-driven model. In doing so, we employ the layer-wise relevance propagation (LRP) technique (Bach *et al.* 2015), which has proven successful previously in other contexts as a way to interpret decisions of NN models (Samek *et al.* 2017; Arras *et al.* 2017). LRP is an instance-based technique, which can be used to quantify the contribution of each input feature (here points in discretized p.d.f. and PS) to the output of the model (here $k_r = k_s/k_{99}$) for a single test case (here a roughness sample). According to this technique, the contribution score (or relevance) of neuron j at each layer of the deep NN can be expressed as

$$R_j = \sum_l \left(\frac{a_j w_{jl}}{\sum_j a_j w_{jl}} \right) R_l, \quad (3.1)$$

where R_l is the contribution score of neuron l in the subsequent layer. In (3.1), w and a are the weight and activation of the neuron that are obtained when the model is used to predict one instance (here the k_r for the roughness sample of interest). Note that in our NN,

Predicting k_s and identifying drag-relevant roughness scales

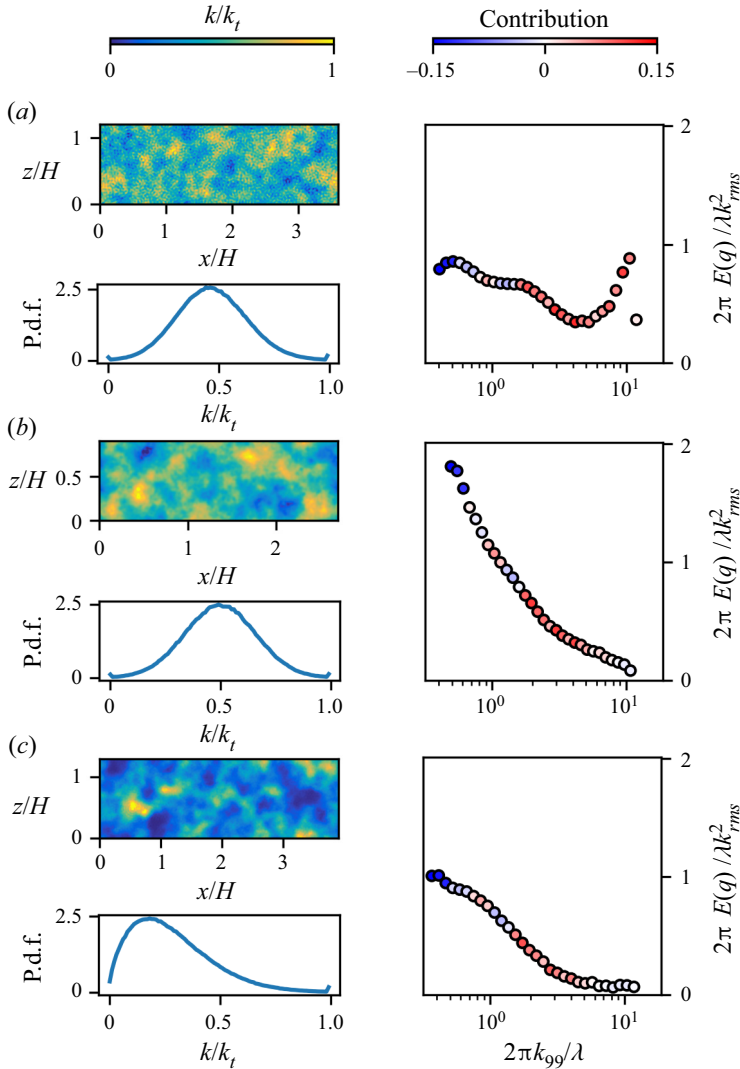


Figure 7. Height maps, p.d.f.s and discretized colour-coded pre-multiplied roughness height PS of three exemplary samples (a) A, (b) B, and (c) C. The spectra are coloured by the LRP contribution scores.

the last layer corresponds to the predicted output, and the first layer to the input roughness information. For better interpretability, we assign the value 1 to the contribution score (or relevance) of the output neuron. As a result, the sum of contribution scores of all inputs must be 1. Note that the contribution scores shown in this section are averaged over the 50 NN members.

In order to extract drag-relevant scales, we consider the following idea. A wavelength that does not affect k_s (which is a measure of added drag) still contributes to an increasing variance of the roughness height, and hence k_{99} . Therefore, the related output of the NN, which is the ratio k_s/k_{99} , is decreased. An input that decreases the output shows a negative LRP contribution score. With that in mind, figure 7 shows three exemplary roughness samples (named A, B and C) and their discretized PS. Each discrete wavenumber in a PS is an input to the model, thus has a contribution score, which is indicated using the specified

	k_{md}/H	k_{99}/H	Sk	ES	L^{Corr}/H
Roughness A original	0.043	0.076	0.149	0.520	0.149
Roughness A filtered	0.052	0.065	0.115	0.519	0.092
Roughness B original	0.041	0.069	0	0.146	0.198
Roughness B filtered	0.024	0.041	0.170	0.128	0.102
Roughness C original	0.026	0.078	0.885	0.129	0.234
Roughness C filtered	0.034	0.056	0.307	0.122	0.126

Table 2. Statistical properties of selected surfaces A, B and C.

colour code. The spectra are shown in pre-multiplied form, and the p.d.f. of each roughness is also displayed. Samples with both Gaussian and non-Gaussian p.d.f.s are included. It is observed in [figure 7](#) that the small wavenumbers (i.e. large wavelengths) generally have more negative contribution scores, which is in accordance to the suggestion of Barros *et al.* (2018). Indeed, the most negative contributions belong consistently to the largest wavelengths for all samples. On the other hand, smaller wavelengths generally show larger contribution scores, but the trend is not monotonic. This might indicate that drag-relevant scales reside within a certain range of the spectral content.

In order to examine whether or not negative LRP contribution score indeed indicates drag irrelevance, we apply high-pass filtering to the samples in [figure 7](#), and examine the resulting roughness using DNS under the same conditions as for the original roughness. The position of the filter is chosen to be the largest wavelength with non-positive contribution score (a three-point moving average is applied to smooth the LRP scores beforehand). [Figure 8](#) shows the height map of original versus filtered samples, the spectra with filter positions, and the inner-scaled mean velocity profiles before and after filtering for samples A, B and C. Some statistical properties of all original and filtered samples are also displayed in [table 2](#). It is clear from [figure 8](#) that the velocity profiles of original and filtered samples collapse very well in the logarithmic region and beyond, which obviously leads to similar values of the roughness function and the drag coefficient. This observation lends support to the hypothesis that the large roughness scales beyond a threshold do not have a meaningful contribution to the added drag, and that LRP analysis can be a data-driven route to identifying those scales *a priori*. One obvious application of this finding can be in selection of sampling size for the investigations of roughness effect. In practice, it is not always possible to obtain roughness samples that are large enough to encompass the full spectrum of scales. However, once the range of drag-relevant scales is covered completely by a roughness sample, a miscalculation due to a limited sample size can be avoided.

Interestingly, in all samples shown in [figure 8](#), the filtered scales have a significant contribution to the roughness height variance based on the pre-multiplied roughness spectra. This is also reflected in the significant decrease in roughness height k_{99} and L^{Corr} in [table 2](#), as anticipated. Additionally, based on the three observed cases, the reductions in the ES values are found proportional to the filtered fraction of the PS. Other statistical parameters also undergo changes due to filtering, while obviously none of these changes is relevant in determining the drag. It is worth noting that in addition to the roughness height k_{99} and ES , other drag-determining quantities, such as Sk , undergo a general reduction for roughness C. According to some existing empirical correlations (e.g. the correlations

Predicting k_s and identifying drag-relevant roughness scales

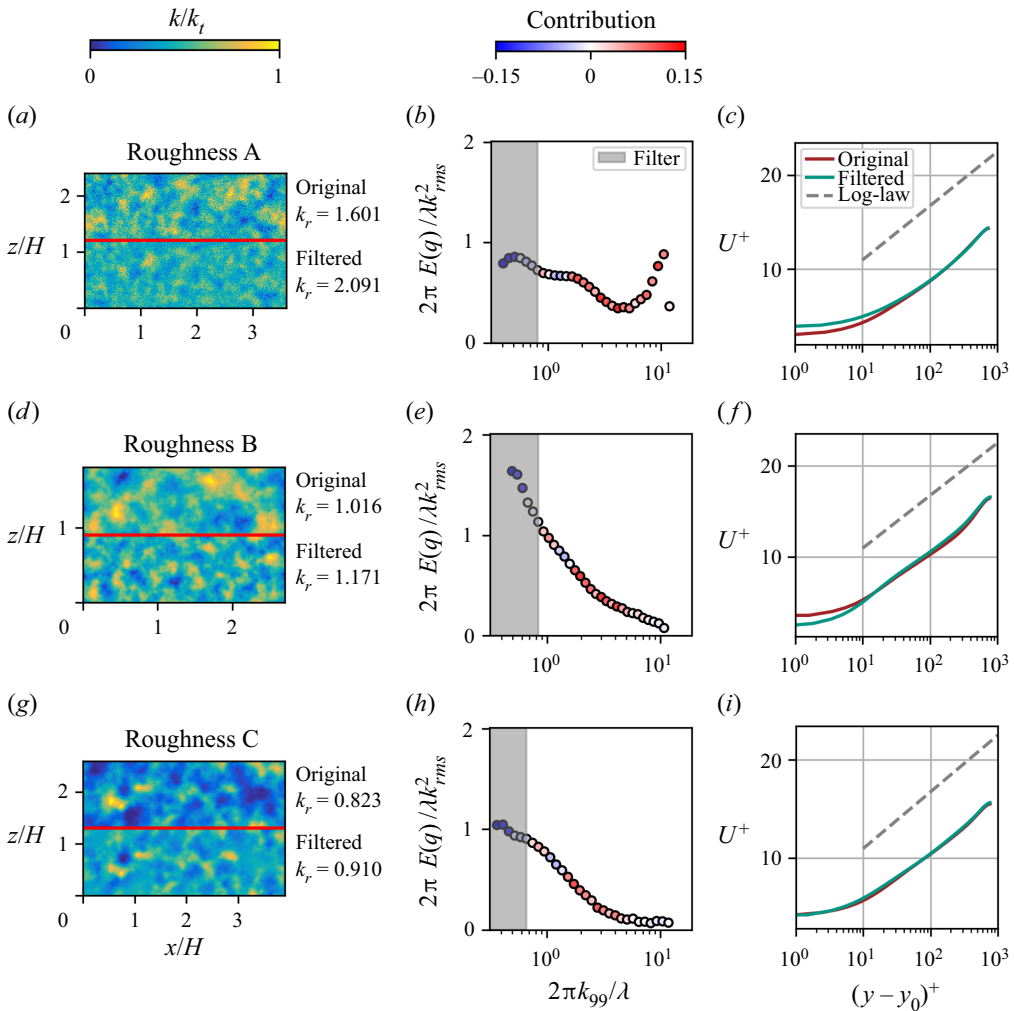


Figure 8. (a,d,g) The original and high-pass filtered roughness, (b,e,h) the pre-multiplied roughness height PS with the filtered scales indicated by grey shading, and (c,f,i) the inner-scaled mean velocity profiles out of DNS on the original and filtered roughness. Note that the DNS are carried out in minimal channels.

proposed by Chan *et al.* 2015; Forooghi *et al.* 2017; Flack *et al.* 2020), the simultaneous reduction in these statistics should lead to a lower k_s . This is, however, not the case in reality based on the DNS results, which can be reminiscent of the suggestion by Barros *et al.* (2018) that a high-pass filtering is necessary if predictive correlations are to capture the correct trend between k_s and the roughness statistics. This also provides an indication for the hypothesis that while statistical parameters can correlate the equivalent sand-grain size of irregular roughness to some degree, only a combined statistical–spectral approach can fully capture the physics of roughness-induced drag.

Furthermore, it is observed in figure 8 that the mean velocity profiles of original and filtered can exhibit some deviation very close to the wall. These deviations can be attributed to the altered volume occupied by roughness close to the wall, as reflected by their k_{md} and Sk values. However, these do not seem to have a significant influence beyond the region occupied by roughness.

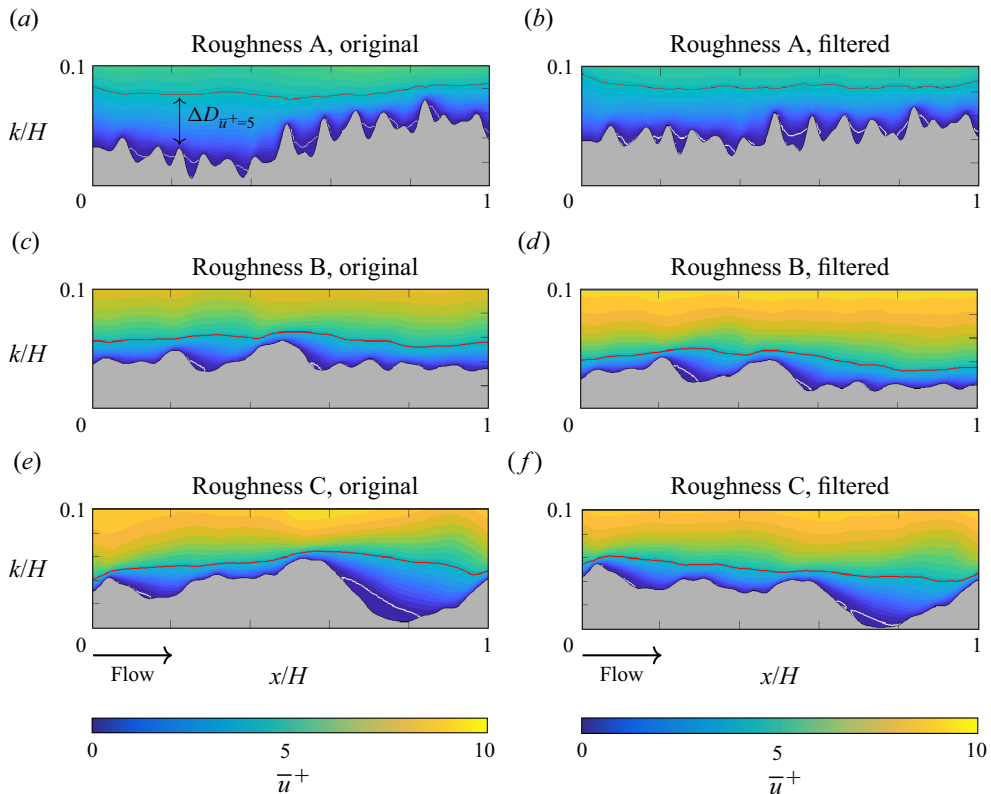


Figure 9. Time-averaged streamwise velocity distribution \bar{u}^+ in selected z -normal planes for the original and filtered cases A–C. The overlaid white contour lines mark the regions of reversed flow ($\bar{u} < 0$). The blanketing layer (iso-contours of $\bar{u}^+ = 5$) is displayed with red contour lines. The grey colour represents the rough structures. The calculation of blanketing layer depth $\Delta D_{\bar{u}^+=5}$ is illustrated schematically in (a). (a) roughness A, original, (b) roughness A, filtered, (c) roughness B, original, (d) roughness B, filtered, (e) roughness C, original and (f) roughness C, filtered.

To shed further light on why the filtered large scales do not contribute to added drag, exemplary x – y planes of the time-averaged streamwise velocity field are examined in figure 9. The overlaid white contour lines are iso-contours of streamwise mean velocity $\bar{u} = 0$, which mark the regions of reversed flow. As expected, roughness A exhibits relatively frequent flow recirculation due to larger local surface slope. In contrast, the occurrences of flow separation over roughnesses B and C seem to be less frequent, which could be linked to the waviness characteristics (Schultz & Flack 2009) and less dominant form drag as a result of the low surface slope. When comparing the flow fields over filtered and original roughness, it is evident that the locations of flow recirculation are the same, and filtering has a minimal impact on the extent of reversed flow regions. Moreover, in figure 9, red contours are used to show the blanketing layer, which, following Busse *et al.* (2017), is defined as the flow region confined by iso-surfaces of $\bar{u}^+ = 5$. On a smooth wall, the blanketing layer would be identical to the viscous sub-layer, while on a rough wall, it can be an indication of how the near-wall flow adapts to the roughness topography. Similarly to the observations in Busse *et al.* (2017), one can observe in figure 9 that the blanketing layers in the cases shown do not follow the small roughness scales and steep roughness patterns. This behaviour can be recognized better if the ‘depth’ of the blanketing layer, i.e. $\Delta D_{\bar{u}^+=5}(x, z) = y_{\bar{u}^+=5}(x, z) - k(x, z)$, is considered.

Predicting k_s and identifying drag-relevant roughness scales

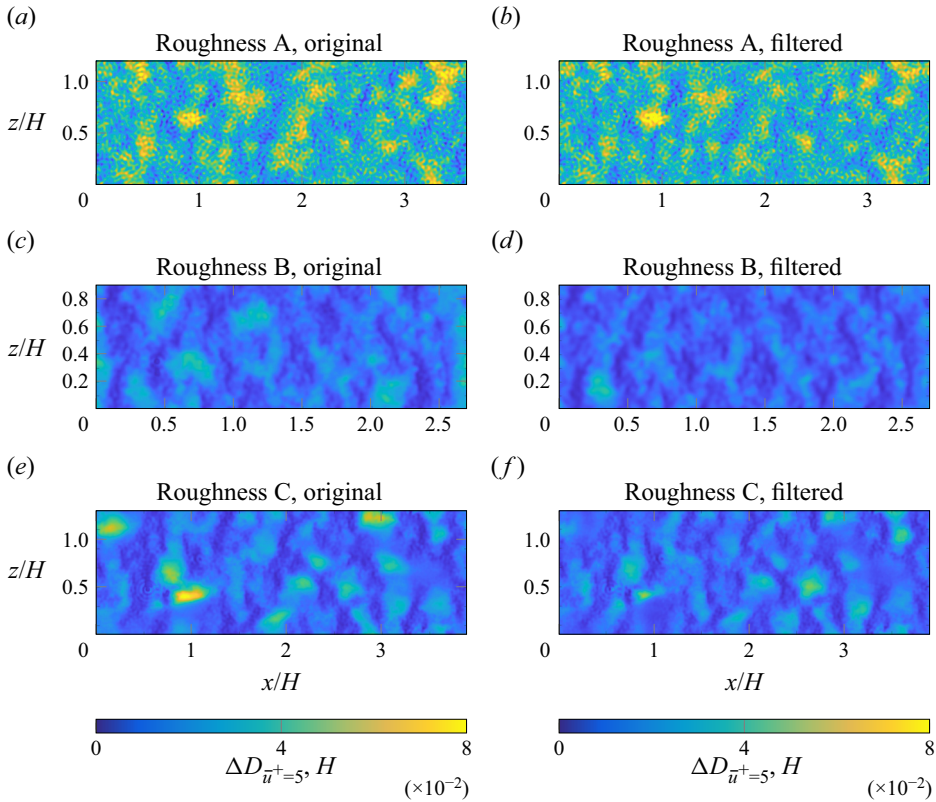


Figure 10. Blanketing layer depth $\Delta D_{\bar{u}^+ = 5}(x, z) = y_{\bar{u}^+ = 5}(x, z) - k(x, z)$ measured from the rough surface for the original and filtered cases A–C.

The maps of $\Delta D_{\bar{u}^+ = 5}(x, z)$ are shown in figure 10, where a visual inspection reveals relative insensitivity of the blanketing layers to the smaller scales of roughness topography (which appear when the roughness height is subtracted from the $\bar{u}^+ = 5$ iso-contour height). Interestingly, in the same figure, a fair level of similarity is observed between the $\Delta D_{\bar{u}^+ = 5}(x, z)$ maps of the corresponding original and filtered cases. This can be a hint that the blanketing layer has adapted to the filtered scales. This idea is examined in Appendix B through a spectral analysis of $\Delta D_{\bar{u}^+ = 5}(x, z)$ for cases A–C. Based on this analysis, one might be able to hypothesize that the drag-irrelevant roughness scales are indeed those to which the blanketing layer can adapt. As a final remark, a relation between the drag and blanketing layer depth is physically plausible as a change in this depth is generally accompanied by modifications in local flow phenomena (flow separation, strong changes in local velocity gradient on the wall, etc.) that can be linked to added drag.

3.4. Turbulent statistics over original and filtered roughness

In the previous subsection, we used an LRP analysis of the trained model to identify which roughness scales contribute to the added skin-friction drag. While ΔU^+ is arguably the most important flow statistic in the practical sense, due to its relation to drag, roughness also affects higher-order flow statistics, particularly in the so-called ‘roughness sub-layer’ (Chung *et al.* 2021). In the present study, we focus specifically on comparing the turbulent

and dispersive stresses over pairs of unfiltered and filtered samples A, B and C from § 3.3 as the main means of momentum transfer away from the wall.

The velocity fluctuations in rough channels can be decomposed into turbulent and time-averaged spatial fluctuations following the triple decomposition of the velocity field proposed by Raupach (1992):

$$u_i(x, y, z, t) = \langle \bar{u}_i \rangle(y) + \tilde{u}_i(x, y, z) + u'_i(x, y, z, t). \quad (3.2)$$

Here, $\langle \bar{u}_i \rangle(y)$ is the time-averaged (overbar) and x - z -plane-averaged (angle brackets) velocity, $\tilde{u}_i(x, y, z) = \bar{u}_i(x, y, z) - \langle \bar{u}_i \rangle(y)$ is the spatial variation of the time-averaged velocity, and $u'_i(x, y, z, t)$ is the space- and time-dependent turbulent fluctuation. Extrinsic plane-averaging is utilized in the present calculation of statistics, i.e. the solid regions are included in the averaging procedure with zero velocity (similar to e.g. Yuan & Piomelli 2014b; Stroh *et al.* 2020). Based on the above decomposition, local turbulent stresses $\overline{u'_i u'_j}(x, y, z)$ can be interpreted as measures of momentum transfer due to turbulent fluctuations. Analogous to the local turbulent stresses, one can define the dispersive stresses $\langle \tilde{u}_i \tilde{u}_j \rangle(y)$ as the momentum transfer due to roughness-induced spatial fluctuations. Furthermore, double-averaged turbulent stresses are calculated through spatial averaging of the local turbulent stresses, i.e. $\langle \overline{u'_i u'_j} \rangle(y)$.

The comparison of turbulent stresses for the three considered rough surfaces A, B and C in filtered and unfiltered states are shown in the near-wall region $(y - y_0)^+ < 200$ in figure 11 (red colour). Only a minor difference between original and filtered roughness can be observed for the turbulent Reynolds stresses. For the $\langle \overline{u' u'} \rangle^+$ component, the peak values are comparable, although for samples B and C, filtering increases the peak value slightly. Roughness has been shown previously to damp inner-scaled streamwise turbulent stress that can be related to the suppression of elongated near-wall turbulent structures (Yuan & Piomelli 2014b; Forooghi *et al.* 2018a). This effect seems not to be affected significantly by elimination of drag-irrelevant large roughness scales. Moreover, the agreement can be observed for the other two normal turbulent stresses as well as the shear stress ($\langle \overline{w' w'} \rangle^+$ is not shown for the sake of brevity). Excellent agreement of wall-normal turbulent stresses is reminiscent of the suggestion by Orlandi & Leonardi (2008) that the roughness function is related to this component of turbulent stress. Furthermore, the collapse of shear stress profiles is an indication of similarity in the vertical mean momentum transport due to turbulence. The agreement of these components thus contributes to the concordance of the mean velocity profiles in the log-layer.

For the dispersive stresses, it is apparent that the only component affected by filtering of roughness is the streamwise normal component $\langle \tilde{u} \tilde{u} \rangle^+$, for which the peak values are reduced by filtering. It is worth mentioning that same trend (reduction of the $\langle \tilde{u} \tilde{u} \rangle^+$ peak values, and agreement of other dispersive stresses) can be observed if an intrinsic averaging approach is used (not shown for brevity). Despite the possible shift of the zero-plane y_0 after filtering, the peak of $\langle \tilde{u} \tilde{u} \rangle^+$ is observed consistently at the vicinity of the respective zero-planes, i.e. at $(y - y_0) \approx 0$. The discernible reduction in this peak value suggests a less pronounced inhomogeneity of the mean streamwise velocity when larger wavelengths are filtered. Arguably, the large-scale undulations present in the original roughness lead to large-scale variations in mean velocity, resulting in greater flow inhomogeneity, as also pointed out by Yuan & Jouybari (2018).

Despite the fact that the values of dispersive shear stress are small in all cases, a comparison among the three cases shown can provide certain insight into the roughness-flow interactions. As depicted in figures 11(c,f,i), roughness A exhibits a

Predicting k_s and identifying drag-relevant roughness scales

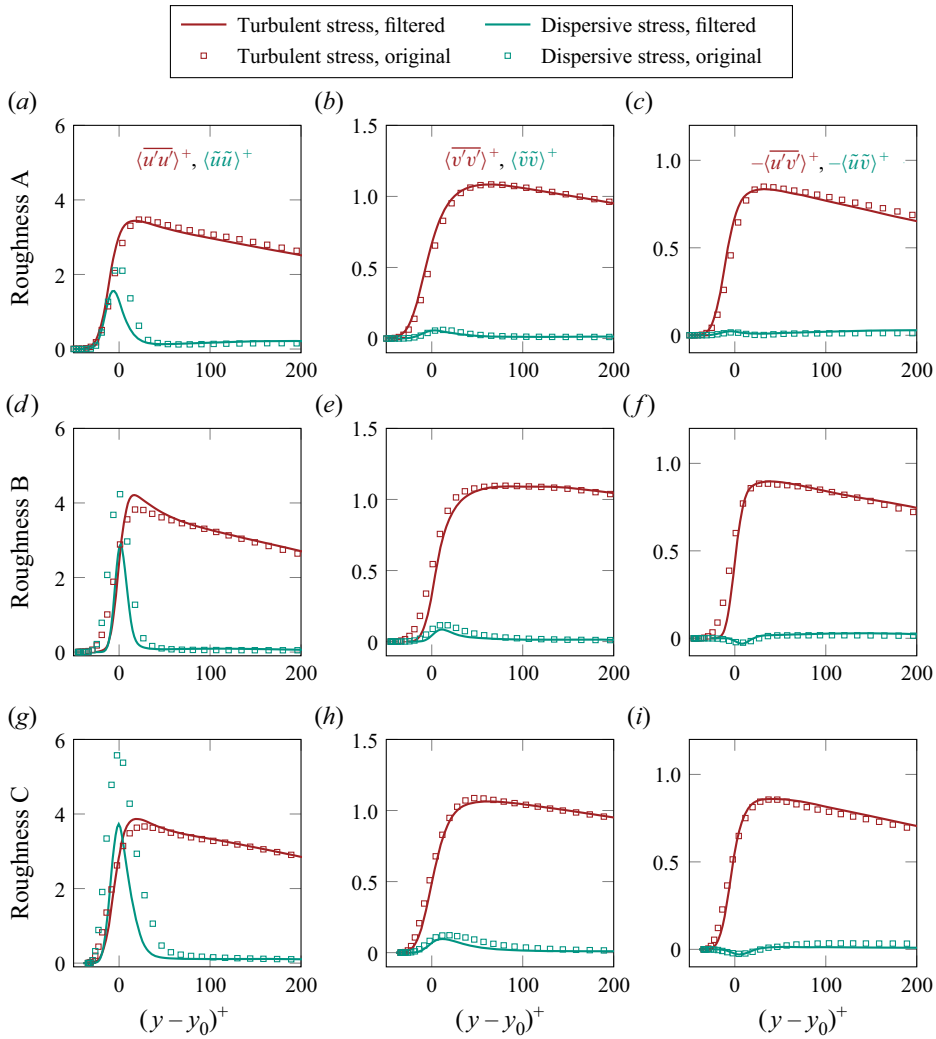


Figure 11. Double-averaged turbulent and dispersive stresses for roughnesses A, B and C.

positive $-\langle \tilde{u}\tilde{v} \rangle^+$ peak, whereas roughnesses B and C display negative peaks. Such a negative sign can be attributed to the ‘waviness effect’ since a wavy structure (one with relatively low slope) causes an acceleration of the mean flow on the windward side, and a deceleration on the leeward side (Alves Portela *et al.* 2021). Positive $-\langle \tilde{u}\tilde{v} \rangle^+$, on the other hand, can be linked to recirculation behind steep roughness elements (Yuan & Jouybari 2018). This is in line with the fact that roughness A has a much larger ES compared to the other two. The collapse of dispersive shear stress profiles in figure 11 shows that none of these behaviours is affected by the applied filtering.

The results shown so far indicate that the streamwise dispersive stress is the only second-order one-point velocity statistic affected by filtering of drag-irrelevant scales. This, however, does not modify the shear stress profile as discussed above. To elaborate this finding further, joint p.d.f.s of local dispersive motions in the wall-parallel plane $y = y_0$ are calculated for all three samples, along with their filtered counterparts, and shown

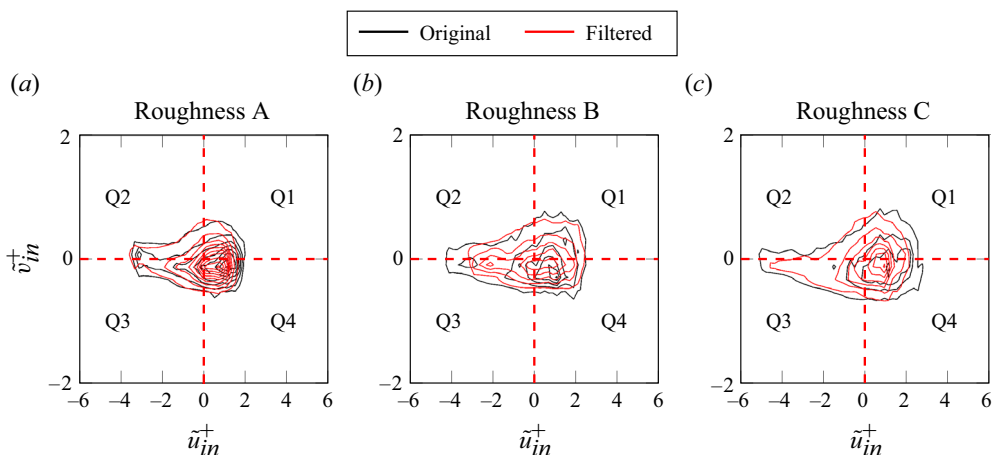


Figure 12. Joint p.d.f.s of \tilde{u}_{in}^+ and \tilde{v}_{in}^+ at plane $y = y_0$, values in roughness excluded. Contour lines range from 0.05 to 1.55, with step 0.1. Subscript *in* indicates being a result of intrinsic averaging.

in figure 12. Here, intrinsic averaging is used, meaning that the areas inside roughness are excluded for calculation of the dispersive velocities shown in the joint p.d.f. The subscript *in* denotes intrinsic averaging. Following the idea of quadrant analysis (Wallace, Eckelmann & Brodkey 1972), the $\tilde{u}_{in}^+ - \tilde{v}_{in}^+$ plane is divided into four quadrants, Q1–Q4, based on the signs of \tilde{u}_{in}^+ and \tilde{v}_{in}^+ . While the joint p.d.f.s look relatively similar before and after high-pass filtering, it is observed that filtering results in contours shrinking along the \tilde{u}_{in}^+ -axis. This is in line with the reduction of peak values of streamwise dispersive components discussed before. Notably, the joint p.d.f. retains its near-symmetry with respect to the \tilde{v}_{in}^+ -axis, which means that reduction of extreme \tilde{u}_{in}^+ fluctuations shows no preference in the direction of momentum transfer. This results in the similar shape of the contours apart from horizontal stretching. An obvious outcome is that the shear stress profiles are unaffected by modifications in \tilde{u}_{in}^+ .

4. Conclusions

In this study, we present a new approach for predicting the normalized equivalent sand-grain height $k_r = k_s/k_{99}$ of homogeneous irregular roughness based on roughness p.d.f. and PS utilizing a machine learning ENN model. The model is developed within the AL framework to reduce effectively the required amount of training data. This framework searches for roughness samples with the highest prediction variances σ_{k_r} in an unlabelled repository \mathcal{U} of 4200 samples. Eventually, a labelled data set \mathcal{L} comprising 85 AL-selected samples is constricted and utilized to derive the ENN model. The significant improvement of the learning efficiency of the model through AL is demonstrated by comparing it with a non-AL approach. Furthermore, it is observed that the employment of AL serves to mitigate effectively the deleterious effects of over-fitting, as evidenced by the observed general drop in prediction error. The mean prediction errors of the final AL-ENN model for an internal testing data set \mathcal{T}_{inter} , as well as two external data sets containing both realistic and artificially generated roughness, $\mathcal{T}_{ext,1}$ and $\mathcal{T}_{ext,2}$, are 9.3 %, 5.2 % and 10.2 %, respectively. The consistently good predictions for testing data with different natures can be taken as a sign that a universal model is approached.

Moreover, novel physical insights on the interactions between roughness and turbulent flow are sought by exploring the information embedded in the data-driven model. To this end, the LRP technique is employed to evaluate the contributions of different wavenumbers in the discretized roughness PS towards the predicted value k_r . The PS content identified with a positive contribution according to the LRP is interpreted as ‘drag-relevant’. Subsequently, high-pass filtering is used to exclude the drag-irrelevant scales, and based on the DNS results for exemplary cases, it is observed that despite the considerable variations in the roughness appearance and statistics, the mean velocity profiles of these high-pass filtered samples collapse well into the original samples in the logarithmic layer, thus having the same k_s values. The LRP-identified drag-irrelevant structures are studied further through an analysis of the behaviour of the blanketing layers over filtered and original roughness. Similarity is observed when maps of blanketing layer ‘depth’ $\Delta D_{\bar{u}^+ = 5}$ of filtered and original roughness are compared. This can indicate that the blanketing layer can adapt to the drag-irrelevant scales. Furthermore, turbulent and dispersive stresses over original and filtered roughness are compared; it is shown that the turbulent stresses are not affected meaningfully by removal of the drag-irrelevant structures. Agreement is observed for both turbulent and dispersive shear stresses, which indicates identical momentum transport patterns in the wall-normal direction over original and filtered roughness. The sole effect of filtering observed on one-point second-order velocity statistics is the reduced streamwise dispersive stress, which can be an indication of less inhomogeneity of mean flow over the filtered roughness. Finally, the joint p.d.f.s of streamwise and wall-normal dispersive velocity components are compared for original and filtered roughness, and it is observed that the probability contours are generally similar, with the streamwise component having a smaller extent in the filtered case. No strong change of preference towards a certain quadrant results from filtering.

In summary, according to the present results, it can be stated that use of roughness height p.d.f. and PS as the model inputs, combined with an AL framework for exploring the vast parameter space, has the potential for developing universal roughness predictive models. Additionally, the present work shows a clear potential to extract physical information from the data-driven models through interpretation techniques, here the LRP.

The LRP-based analysis presented in this work is obviously a first step towards utilizing data-driven models beyond merely predictive tools in the context of rough-wall turbulence. Future investigations can explore other avenues to extract knowledge on the roughness–turbulence interactions from such models. The present LRP-based analysis can also be investigated further towards more rigorous criteria for identifying the drag-irrelevant structures. Furthermore, the LRP-based filtering can be a basis for developing more accurate empirical correlations incorporating solely the drag-relevant structures. Finally, one should note that the present work merely focuses on roughness topographies of isotropic and homogeneous nature. Extension towards more general anisotropic and/or heterogeneous roughness is another obvious direction for future research.

Funding. J.Y. gratefully acknowledges partial financial support from the Friedrich und Elisabeth Boysen-Foundation (BOY-151). P.F. gratefully acknowledges financial support from the Aarhus University Research Foundation (starting grant AUFF-F-2020-7-9). S.L. and S.B. sincerely appreciate financial support from the Swedish Energy Agency under grant no. 51554-1. This work was performed on the supercomputer HoreKa and the storage facility LSDF funded by the Ministry of Science, Research and the Arts Baden-Württemberg and by the Federal Ministry of Education and Research.

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

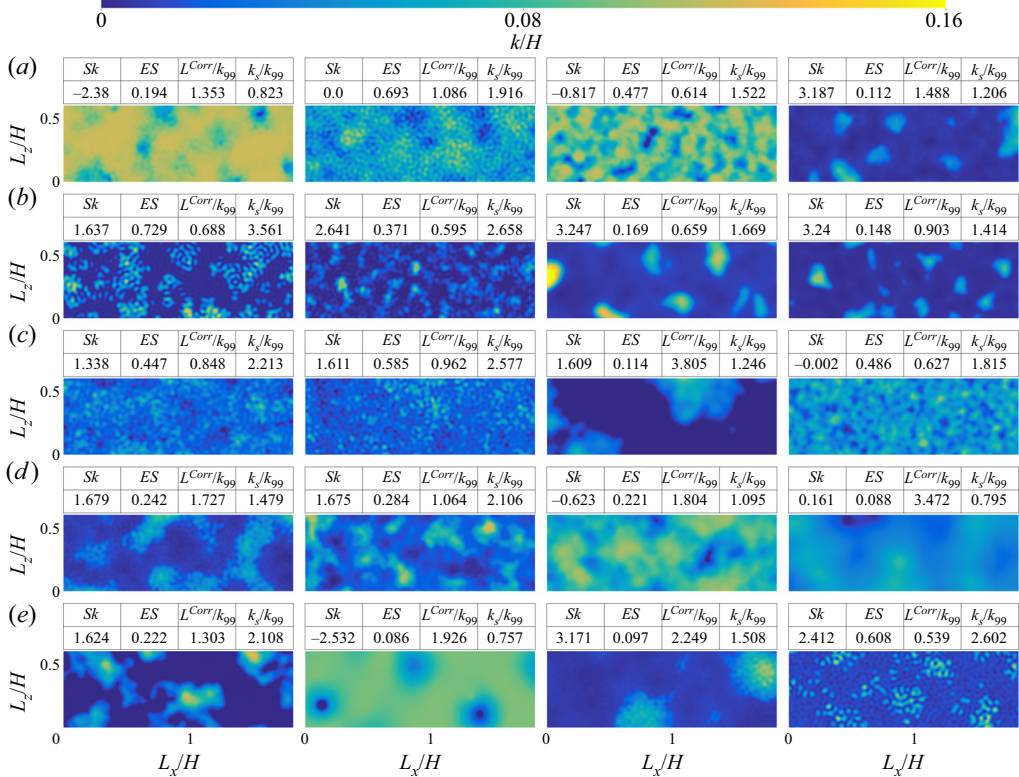


Figure 13. Examples of roughness samples included in \mathcal{L} . Patches of same size extracted from different samples. (a–e) correspond to initial round and AL rounds 1–4, respectively.

Data availability statement. The trained model for prediction of k_s can be accessed publicly at the project web page, <http://roughness.org>, for research purposes. The generated database including labelled roughness samples is made available publicly through <http://roughnessdatabase.org>. The codes for roughness generation, statistical analysis and model training can be downloaded from the first author’s GitHub repository <https://github.com/JiashengY/Active-learning-codes>.

Author ORCIDiDs.

- ① Jiasheng Yang <https://orcid.org/0000-0003-0091-6855>;
- ① Alexander Stroh <https://orcid.org/0000-0003-0850-9883>;
- ① Sangseung Lee <https://orcid.org/0000-0001-7341-8289>;
- ① Shervin Bagheri <https://orcid.org/0000-0002-8209-1449>;
- ① Bettina Frohnappel <https://orcid.org/0000-0002-0594-7178>;
- ① Pourya Forooghi <https://orcid.org/0000-0001-9212-514X>.

Author contributions. J.Y.: methodology, investigation, data curation, software, formal analysis, visualization, writing – original draft. A.S.: methodology, data curation, formal analysis, supervision, writing – review and editing. S.L.: ML methodology, writing – review and editing. S.B.: ML methodology, formal analysis, writing – review and editing. B.F.: formal analysis, supervision, resources, funding acquisition, writing – review and editing. P.F.: conceptualization, methodology, formal analysis, supervision, writing – original draft, review and editing – funding acquisition, project administration.

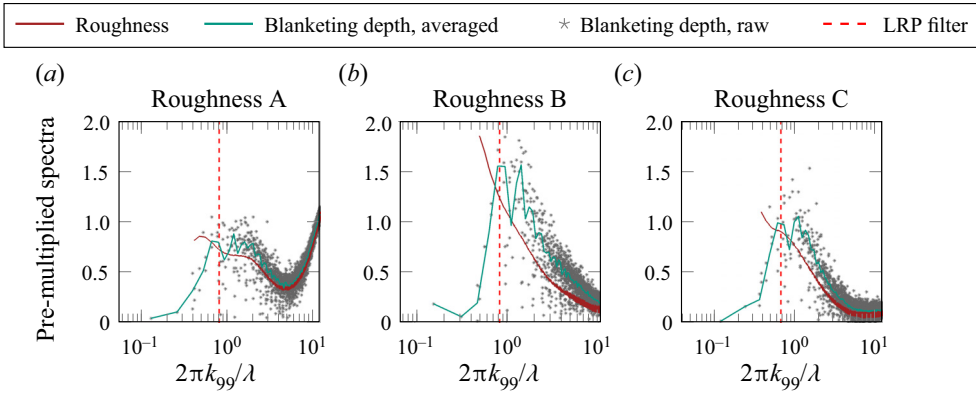


Figure 14. Pre-multiplied spectra of blanketing layer depth $\Delta D_{\bar{u}+5}$ overlaid with that of the corresponding roughness topography. Symbols indicate the spectrum in different directions, while green lines show the azimuthal average. The scatter of the symbols indicates the anisotropic characteristics of the map. Structures smaller than the smallest in-plane roughness wavelength λ_1 are omitted.

Appendix A. Exemplary roughness at each round

Exemplary rough surfaces selected from each iteration, along with their statistical parameters and k_s values, are shown in [figure 13](#).

Appendix B. Spectral analysis of blanketing layer depth

In order to investigate further possible links between the blanketing layer depth $\Delta D_{\bar{u}+5}(x, z) = y_{\bar{u}+5}(x, z) - k(x, z)$ and drag-irrelevant scales, we plot pre-multiplied PS of the blanketing layer depth maps over the original rough surfaces in [figure 14](#). For more clarity, the two-dimensional spectra are averaged azimuthally and plotted in green. Moreover, the locations of LRP-identified filters and the pre-multiplied roughness spectra are also added to the plots. Interestingly, all spectra show a significant decrease in the contribution of wavelengths larger than (wavenumbers smaller than) the filter. Note that all plots in [figure 14](#) belong to the original cases, with no influence from the filtering. A wavelength that is present in the roughness topography but absent in the blanketing layer depth is one to which the layer has adapted. Therefore, the fact that the spectrum drops for drag-irrelevant scales might suggest that those scales are the ones to which the blanketing layer can adapt.

Despite the above discussion, further systematic investigations are required to establish conclusive evidence, as the present study covers limited ranges of roughness scales and Reynolds numbers. Ideally, data on surfaces with more ‘drag-irrelevant’ large scales and at a much wider range of Reynolds numbers are required to establish a solid hypothesis. Additionally, one should bear in mind that the current results are obtained in the fully rough regime, and as discussed by Busse *et al.* (2017), blanketing layers can behave differently at different regimes.

REFERENCES

- ABE, N. & MAMITSUKA, H. 1998 Query learning strategies using boosting and bagging. In *Proceedings of the Fifteenth International Conference on Machine Learning, ICML'98*, pp. 1–9. Morgan Kaufmann.
- ALVES PORTELA, F., BUSSE, A. & SANDHAM, N.D. 2021 Numerical study of Fourier-filtered rough surfaces. *Phys. Rev. Fluids* **6**, 084606.

- ANDERSON, W. & MENEVEAU, C. 2011 Dynamic roughness model for large-eddy simulation of turbulent flow over multiscale, fractal-like rough surfaces. *J. Fluid Mech.* **679**, 288–314.
- ANGLUIN, D. 2004 Queries revisited. *Theor. Comput. Sci.* **313** (2), 175–194.
- ARRAS, L., HORN, F., MONTAVON, G., MÜLLER, K. & SAMEK, W. 2017 ‘What is relevant in a text document?’ An interpretable machine learning approach. *PLoS ONE* **12** (8), 1–23.
- BACH, S., BINDER, A., MONTAVON, G., KLAUSCHEN, F., MÜLLER, K. & SAMEK, W. 2015 On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS ONE* **10** (7), 1–46.
- BANGERT, P., MOON, H., WOO, J.O., DIDARI, S. & HAO, H. 2021 Active learning performance in labeling radiology images is 90 % effective. *Front. Radiol.* **1**, 748968.
- BARROS, J.M., SCHULTZ, M.P. & FLACK, K.A. 2018 Measurements of skin-friction of systematically generated surface roughness. *Intl J. Heat Fluid Flow* **72**, 1–7.
- BHAGANAGAR, K., KIM, J. & COLEMAN, G. 2004 Effect of roughness on wall-bounded turbulence. *Flow Turbul. Combust.* **72** (2–4), 463–492.
- BISHOP, C.M. 1995 *Neural Networks for Pattern Recognition*. Advanced Texts in Econometrics. Clarendon. Oxford University Press.
- BURBIDGE, R., JEM, J. & KING, R.D. 2007 Active learning for regression based on query by committee. In *Intelligent Data Engineering and Automated Learning – IDEAL 2007* (ed. H. Yin, P. Tino, E. Corchado, W. Byrne & X. Yao), pp. 209–218. Springer.
- BUSSE, A. & JELLY, T.O. 2023 Effect of high skewness and kurtosis on turbulent channel flow over irregular rough walls. *J. Turbul.* **24** (1–2), 57–81.
- BUSSE, A., LÜTZNER, M. & SANDHAM, N.D. 2015 Direct numerical simulation of turbulent flow over a rough surface based on a surface scan. *Comput. Fluids* **116**, 129–147.
- BUSSE, A., THAKKAR, M. & SANDHAM, N.D. 2017 Reynolds-number dependence of the near-wall flow over irregular rough surfaces. *J. Fluid Mech.* **810**, 196–224.
- CHAN, L., MACDONALD, M., CHUNG, D., HUTCHINS, N. & OOI, A. 2015 A systematic investigation of roughness height and wavelength in turbulent pipe flow in the transitionally rough regime. *J. Fluid Mech.* **771**, 743–777.
- CHAN-BRAUN, C., GARCÍA-VILLALBA, M. & UHLMANN, M. 2011 Force and torque acting on particles in a transitionally rough open-channel flow. *J. Fluid Mech.* **684**, 441–474.
- CHEVALIER, M., SCHLATTER, P., LUNDBLADH, A. & HENNINGSON, D. 2007 SIMSON – a pseudo-spectral solver for incompressible boundary layer flow. *Tech. Rep.* TRITA-MEK 2007:07, Royal Institute of Technology, Stockholm, Sweden, pp. 1–100.
- CHUNG, D., CHAN, L., MACDONALD, M., HUTCHINS, N. & OOI, A. 2015 A fast direct numerical simulation method for characterising hydraulic roughness. *J. Fluid Mech.* **773**, 418–431.
- CHUNG, D., HUTCHINS, N., SCHULTZ, M.P. & FLACK, K.A. 2021 Predicting the drag of rough surfaces. *Annu. Rev. Fluid Mech.* **53**, 439–471.
- FEDOROV, V.V. 1972 *Theory of Optimal Experiments*. Probability and Mathematical Statistics. Academic.
- FLACK, K.A. 2018 Moving beyond Moody. *J. Fluid Mech.* **842**, 1–4.
- FLACK, K.A. & CHUNG, D. 2022 Important parameters for a predictive model of k_s for zero-pressure-gradient flows. *AIAA J.* **60** (10), 5923–5931.
- FLACK, K.A. & SCHULTZ, M.P. 2010 Review of hydraulic roughness scales in the fully rough regime. *Trans. ASME J. Fluids Engng* **132** (4), 041203.
- FLACK, K.A., SCHULTZ, M.P. & BARROS, J.M. 2020 Skin friction measurements of systematically-varied roughness: probing the role of roughness amplitude and skewness. *Flow Turbul. Combust.* **104** (2–3), 317–329.
- FOROOGHI, P., STROH, A., MAGAGNATO, F., JAKIRLIĆ, S. & FROHNAPFEL, B. 2017 Toward a universal roughness correlation. *Trans. ASME J. Fluids Engng* **139** (12), 121201.
- FOROOGHI, P., STROH, A., SCHLATTER, P. & FROHNAPFEL, B. 2018a Direct numerical simulation of flow over dissimilar, randomly distributed roughness elements: a systematic study on the effect of surface morphology on turbulence. *Phys. Rev. Fluids* **3**, 044605.
- FOROOGHI, P., WEIDENLENER, A., MAGAGNATO, F., BÖHM, B., KUBACH, H., KOCH, T. & FROHNAPFEL, B. 2018b DNS of momentum and heat transfer over rough surfaces based on realistic combustion chamber deposit geometries. *Intl J. Heat Fluid Flow* **69**, 83–94.
- GAL, Y. & GHAHRAMANI, Z. 2016 Dropout as a Bayesian approximation: representing model uncertainty in deep learning. In *Proceedings of the 33rd International Conference on Machine Learning* (ed. M.F. Balcan & K.Q. Weinberger), **48**, 1050–1059. PMLR.
- GOLDSTEIN, D., HANDLER, R. & SIROVICH, L. 1993 Modeling a no-slip flow boundary with an external force field. *J. Comput. Phys.* **105** (2), 354–366.

Predicting k_s and identifying drag-relevant roughness scales

- HAMA, F.R. 1954 *Boundary-Layer Characteristics for Smooth and Rough Surfaces*. Society of Naval Architects and Marine Engineers.
- HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. 2009 The elements of statistical learning: data mining, inference, and prediction. *Model Assessment and Selection* (ed. P. Bühlmann, P. Diggle, U. Gather & S. Zeger), 219–259. Springer.
- HINZE, J.O. 1967 Secondary currents in wall turbulence. *Phys. Fluids* **10** (9), S122–S125.
- JACKSON, P.S. 1981 On the displacement height in the logarithmic velocity profile. *J. Fluid Mech.* **111**, 15–25.
- JACOBS, T.D.B., JUNGE, T. & PASTEWKA, L. 2017 Quantitative characterization of surface topography using spectral analysis. *Surf. Topogr.: Metrol. Prop.* **5** (1), 013001.
- JIMÉNEZ, J. 2004 Turbulent flows over rough walls. *Annu. Rev. Fluid Mech.* **36**, 173–196.
- JOUBYARI, M.A., SEO, J., YUAN, J., MITTAL, R. & MENEVEAU, C. 2022 Contributions to pressure drag in rough-wall turbulent flows: insights from force partitioning. *Phys. Rev. Fluids* **7** (8), 084602.
- JOUBYARI, M.A., YUAN, J., BRERETON, G.J. & MURILLO, M.S. 2021 Data-driven prediction of the equivalent sand-grain height in rough-wall turbulent flows. *J. Fluid Mech.* **912**, A8.
- KROGSTAD, P.-Å., ANTONIA, R.A. & BROWNE, L.W.B. 1992 Comparison between rough- and smooth-wall turbulent boundary layers. *J. Fluid Mech.* **245**, 599–617.
- LANG, K. & BAUM, E. 1992 Query learning can work poorly when a human oracle is used. In *IEEE International Joint Conference on Neural Networks*. IEEE Press.
- LEE, S., YANG, J., FOROOGHI, P., STROH, A. & BAGHERI, S. 2022 Predicting drag on rough surfaces by transfer learning of empirical correlations. *J. Fluid Mech.* **933**, A18.
- LEWIS, D.D. & GALE, W.A. 1994 A sequential algorithm for training text classifiers. In *Proceedings of the 17th Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval* (ed. W.B. Croft & C.J. van Rijsbergen), 3–12. ACM/Springer.
- LYASHENKO, I.A., PASTEWKA, L. & PERSSON, B.N.J. 2013 On the validity of the method of reduction of dimensionality: area of contact, average interfacial separation and contact stiffness. *Tribol. Lett.* **52** (2), 223–229.
- MACDONALD, M., CHUNG, D., HUTCHINS, N., CHAN, L., OOI, A. & GARCÍA-MAYORAL, A. 2016 The minimal channel: a fast and direct method for characterising roughness. *J. Phys.: Conf. Ser.* **708**, 012010.
- MEDJNOUN, T., RODRIGUEZ-LOPEZ, E., FERREIRA, M.A., GRIFFITHS, T., MEYERS, J. & GANAPATHISUBRAMANI, B. 2021 Turbulent boundary-layer flow over regular multiscale roughness. *J. Fluid Mech.* **917**, A1.
- MEJIA-ALVAREZ, R. & CHRISTENSEN, K.T. 2010 Low-order representations of irregular surface roughness and their impact on a turbulent boundary layer. *Phys. Fluids* **22** (1), 015106.
- MELVILLE, P. & MOONEY, R.J. 2003 Constructing diverse classifier ensembles using artificial training examples. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence, IJCAI'03*, pp. 505–510. Morgan Kaufmann.
- NAPOLI, E., ARMENIO, V. & DEMARCHIS, M. 2008 The effect of the slope of irregularly distributed roughness elements on turbulent wall-bounded flows. *J. Fluid Mech.* **613**, 385–394.
- NIKURADSE, J. 1933 *Stromungsgesetze in rauhen Rohren*. VDI-Verl.
- ORLANDI, P. & LEONARDI, S. 2008 Direct numerical simulation of three-dimensional turbulent rough channels: parameterization and flow physics. *J. Fluid Mech.* **606**, 399–415.
- PENG, W. & BHUSHAN, B. 2000 Modelling of surfaces with a bimodal roughness distribution. *Proc. Inst. Mech. Engrs* **214** (5), 459–470.
- PÉREZ-RAFOLS, F. & ALMQVIST, A. 2019 Generating randomly rough surfaces with given height probability distribution and power spectrum. *Tribol. Intl* **131**, 591–604.
- PERRY, A.E., SCHOFIELD, W.H. & JOUBERT, P.N. 1969 Rough wall turbulent boundary layers. *J. Fluid Mech.* **37** (2), 383–413.
- RAUPACH, M.R. 1992 Drag and drag partition on rough surfaces. *Boundary-Layer Meteorol.* **60**, 375–395.
- RAYCHAUDHURI, T. & HAMEY, L. 1995 Minimisation of data collection by active learning. In *Proceedings of ICNN'95-International Conference on Neural Networks*, vol. 3, pp. 1338–1341. IEEE.
- REED, R. & MARKS, R.J. 1999 *Neural Smoothing: Supervised Learning in Feedforward Artificial Neural Networks*. MIT.
- VAN RIJ, J.A., BELNAP, B.J. & LIGRANI, P.M. 2002 Analysis and experiments on three-dimensional, irregular surface roughness. *Trans. ASME J. Fluids Engng* **124** (3), 671–677.
- SAMEK, W., BINDER, A., MONTAVON, G., LAPUSCHKIN, S. & MÜLLER, K. 2017 Evaluating the visualization of what a deep neural network has learned. *IEEE Trans. Neural Netw. Learn. Syst.* **28** (11), 2660–2673.
- SAYLES, R.S. & THOMAS, T.R. 1978 Topography of random surfaces (reply). *Nature* **273** (5663), 573.
- SCHLICHTING, H. 1936 Experimentelle Untersuchungen zum Rauheitsproblem. *Ingenieur-Archiv*.

- SCHULTZ, M.P., BENDICK, J.A., HOLM, E.R. & HERTEL, W.M. 2011 Economic impact of biofouling on a naval surface ship. *Biofouling* **27** (1), 87–98.
- SCHULTZ, M.P. & FLACK, K.A. 2009 Turbulent boundary layers on a systematically varied rough wall. *Phys. Fluids* **21** (1), 015104.
- SETTLES, B. 2009 Active learning literature survey. Computer Sciences Technical Report 1648. University of Wisconsin–Madison.
- SETTLES, B. & CRAVEN, M. 2008 An analysis of active learning strategies for sequence labeling tasks. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing* (ed. L. Mirella & N.H. Tou), pp. 1070–1079. Association for Computational Linguistics.
- SEUNG, H.S., OPPER, M. & SOMPOLINSKY, H. 1992 Query by committee. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory, COLT '92*, pp. 287–294. Association for Computing Machinery.
- STROH, A., SCHÄFER, K., FROHNAPFEL, B. & FOROOGHI, P. 2020 Rearrangement of secondary flow over spanwise heterogeneous roughness. *J. Fluid Mech.* **885**, R5.
- THAKKAR, M., BUSSE, A. & SANDHAM, N. 2017 Surface correlations of hydrodynamic drag for transitionally rough engineering surfaces. *J. Turbul.* **18** (2), 138–169.
- VANDERWEL, C., STROH, A., KRIEGSEIS, J., FROHNAPFEL, B. & GANAPATHISUBRAMANI, B. 2019 The instantaneous structure of secondary flows in turbulent boundary layers. *J. Fluid Mech.* **862**, 845–870.
- VELANDIA, J. & BANSMER, S. 2019 Topographic study of the ice accretion roughness on a generic aero-engine intake. *AIAA Scitech 2019 Forum*.
- WALLACE, J.M., ECKELMANN, H. & BRODKEY, R.S. 1972 The wall region in turbulent shear flow. *J. Fluid Mech.* **54** (1), 39–48.
- YANG, J., STROH, A., CHUNG, D. & FOROOGHI, P. 2022 Direct numerical simulation-based characterization of pseudo-random roughness in minimal channels. *J. Fluid Mech.* **941**, A47.
- YANG, J., VELANDIA, J., BANSMER, S., STROH, A. & FOROOGHI, P. 2023 A comparison of hydrodynamic and thermal properties of artificially generated against realistic rough surfaces. *Intl J. Heat Fluid Flow* **99**, 109093.
- YUAN, J. & JOUYBARI, M.A. 2018 Topographical effects of roughness on turbulence statistics in roughness sublayer. *Phys. Rev. Fluids* **3**, 114603.
- YUAN, J. & PIOMELLI, U. 2014a Estimation and prediction of the roughness function on realistic surfaces. *J. Turbul.* **15** (6), 350–365.
- YUAN, J. & PIOMELLI, U. 2014b Roughness effects on the Reynolds stress budgets in near-wall turbulence. *J. Fluid Mech.* **760**, R1.
- ZHU, Q., STOLCKE, A., CHEN, B.Y. & MORGAN, N. 2005 Using MLP features in SRI's conversational speech recognition system. In *Proc. Interspeech 2005, Lisbon, Portugal*, pp. 2141–2144.