JM PAPERS

Rough surfaces in underexplored surface morphology space and their implications on roughness modelling

Shyam S. Nair[1]**, Vishal A. Wadhai**[1]**, Robert F. Kunz**[1,†] **and Xiang I.A. Yang**[1,†]

[1]Mechanical Engineering, Penn State University, State College, PA 16802, USA

We report direct numerical simulations results of the rough-wall channel, focusing on roughness with high k_{rms}/k_a statistics but small to negative Sk statistics, and we study the implications of this new dataset on rough-wall modelling. Here, k_{rms} is the root mean square, k_a is the first-order moment of roughness height, and Sk is the skewness. The effects of packing density, skewness and arrangement of roughness elements on mean streamwise velocity, equivalent roughness height (z_0) and Reynolds and dispersive stresses have been studied. We demonstrate that two-point correlation lengths of roughness height statistics play an important role in characterizing rough surfaces with identical moments of roughness height but different arrangements of roughness elements. Analysis of the present as well as historical data suggests that the task of rough-wall modelling is to identify geometric parameters that distinguish the rough surfaces within the calibration dataset. We demonstrate a novel feature selection procedure to determine these parameters. Further, since there is no finite set of roughness statistics that distinguish between all rough surfaces, we argue that obtaining a universal rough-wall model for making equivalent sand-grain roughness (k_s) predictions would be challenging, and that each rough-wall model would have its applicable range. This motivates the development of group-based rough-wall models. The applicability of multi-variate polynomial regression and feedforward neural networks for building such group-based rough-wall models using the selected features has been shown.

Key words: turbulence simulation, turbulence modelling, turbulent boundary layers

1. Introduction

Turbulent flow over rough walls has piqued the interest of researchers and engineers for decades. From the earliest seminal works by Nikuradse (1933) and Colebrook (1939), and

† Email addresses for correspondence: rfk102@psu.edu, xzy48@psu.edu

the consolidated data interpretation by Moody (1944), the field of rough-wall bounded turbulence has continued to evolve, as is evident in subsequent works (Raupach, Antonia & Rajagopalan 1991; Jiménez 2004; Flack & Schultz 2010). In the recent review by Chung *et al.* (2021), the authors emphasize the importance of Townsend's outer-layer similarity hypothesis (Townsend 1976). This hypothesis assumes that under fully rough and sufficiently high boundary layer thickness to roughness height ratio (δ/k) conditions, the outer layer of a rough-wall boundary layer behaves similarly to that of a smooth wall. A logarithmic layer can therefore be expected between the roughness sublayer and the outer layer. This logarithmic layer, aside from a downward shift ΔU^+ compared to the smooth-wall log law, remains unaffected by surface roughness. This downward shift ΔU^+ constitutes the roughness function, enabling frictional drag comparisons across different rough surfaces. Besides ΔU^+, which characterizes the hydraulic properties, a rough wall is also characterized by the statistical moments of roughness height and the distribution of its topographical features. Establishing a functional mapping between the roughness function and the rough-wall topography is of significant practical utility and remains a challenging problem, given the variations in surface features and the costs of rough-wall boundary layer experiments (Schultz & Flack 2007) and scale-resolving computational studies (Choi & Moin 2012; Yang & Griffin 2021).

Several studies have aimed to develop roughness correlations for various surfaces. Flack & Schultz (2010) proposed correlations utilizing skewness (*Sk*) and root mean square roughness height (k_{rms}) to characterize equivalent sand-grain roughness (k_s) for irregular three-dimensional roughness. Yuan & Piomelli (2014*a*) examined slope-based and moment-based correlations to study critical effective slope (*ES*) associated with waviness regimes for realistic roughness from hydraulic turbine blades. Forooghi *et al.* (2017) found that functional relations dependent on *Sk* and *ES*, which represent the shape and slope of the rough surface, respectively, produced satisfactory k_s predictions for randomly distributed roughness elements of random size and prescribed shape. In another study, Flack, Schultz & Barros (2020) performed experiments on random rough surfaces by systematically varying k_{rms} and *Sk*. They found predictive correlations of the form $k_s = Ak_{rms}(1 + Sk)^B$, with *A* and *B* being constants. The authors also adapted the functional form according to groups of surfaces as being either positively, negatively or zero skewed for more accurate correlations. On the other hand, correlations can also take a data-driven form as developed by Jouybari *et al.* (2021), where deep neural networks incorporate information from diverse rough-wall geometries and their corresponding statistical moments. Irrespective of the models used, no single model has been able to generalize well across all rough surfaces (Yang *et al.* 2023).

The difficulty is largely due to the complexity of rough surfaces and that each rough surface seems to have its unique behaviours. This calls for the categorization of rough surfaces. A possible categorization puts rough surfaces into regular roughness and irregular roughness. Regular surfaces have the same elements repeated in a predefined periodic arrangement, unlike irregular roughness where the features are random in shape and/or distribution. Categories based on the shape of the roughness features break down as being cubes (Castro, Cheng & Reynolds 2006), truncated cones (Womack *et al.* 2022), packed spheres (Schultz & Flack 2005), grit-blasted (Flack *et al.* 2016; Thakkar, Busse & Sandham 2018) and others. Rough surfaces may also be categorized by distribution, such as Gaussian (Flack *et al.* 2020; Ma, Alamé & Mahesh 2021; Altland *et al.* 2022), random (randomly distributed regular roughness elements with Gaussian height distributions; Forooghi *et al.* 2017, 2018), pseudo-random (Yang *et al.* 2022), multiscale (Yang & Meneveau 2017; Medjnoun *et al.* 2021) and so on. The distinction between roughness

types could also stem from its flow physics. For instance, the k-type and d-type roughness (Jiménez 2004) exhibit different behaviours and correspond to sparse and densely packed roughness elements, respectively.

The more recent literature seems to favour more precise categorizations based on roughness' statistics. Chung *et al.* (2021) identified several surface properties to be hydrodynamically important. These include measures of roughness height (such as root mean square k_{rms}, average k_a, maximum peak to trough (k_t), effective slope (*ES*), frontal solidity (λ_f), planar packing density (λ_p), skewness (*Sk*), and solid volume fraction (ϕ), among others. A number of existing studies focus on the effects of variations in these individual parameters. Placidi & Ganapathisubramani (2015) studied the effects of varying λ_p and λ_f with 'LEGO' brick type roughness in the fully rough regime. Unlike cubical roughness, they found roughness length to monotonically decrease with increasing λ_p, suggesting that the same geometrical parameter may behave differently based on roughness type. Thakkar, Busse & Sandham (2017) found that the streamwise correlation length plays a role in determining the roughness function in addition to k_{rms}, *Sk* and λ_f for irregular rough surfaces including samples that are cast, composite, hand-filed, grit-blasted, ground, spark-eroded and from ship propellers. The effect of surface anisotropy was investigated systematically by Busse & Jelly (2020) for irregular surfaces where another roughness parameter, the surface anisotropy ratio, defined as the ratio of the streamwise and spanwise correlation lengths, was varied and found to strongly influence the mean flow. Even spanwise parameters such as the spanwise spacing (Vanderwel & Ganapathisubramani 2015) and spanwise effective slope (ES_y) (Jelly *et al.* 2022) are found to significantly affect mean flow statistics. It appears that whenever one varies a roughness parameter, that roughness parameter stands out and plays an important role in determining the equivalent sand-grain roughness height.

In anticipation of the discussion in the following sections, here we review the previous studies of cubical roughness. Due to its simplicity, cubical roughness is one of the most extensively studied types of surface roughness. Early experimental investigations on cube patterns (O'Loughlin & Macdonald 1964) studied the effect of λ_p on equivalent roughness size (k_s/k), establishing that the resistance to flow reaches a maximum at an intermediate λ_p, approximately 0.2, then tends towards smooth-wall behaviour at increasing cube densities. Thanks to the rich physics that cube arrays can represent, several studies have enriched the cubical roughness literature with details about the roughness sublayer (Castro *et al.* 2006), aerodynamic characteristics (Cheng *et al.* 2007) and associated flow structures (Volino, Schultz & Flack 2011). Computational studies implementing direct numerical simulations (DNS) have further prompted discussions on the mean velocity profile, equivalent roughness height (z_0), and zero-plane displacement height (d). These include works by Leonardi & Castro (2010) and Lee, Sung & Krogstad (2011), which focused on turbulence statistics and coherent structures, respectively. The utility involved in studying cubical roughness arising from its ability to generate surface parametrizations, and the relevance to urban canopy studies, have also led to analytical roughness models by Yang & Meneveau (2016) and Yang *et al.* (2016).

Based on the existing literature, it can be argued that surface features *Sk* and k_{rms} are two of the most important parameters, and that a parameter space involving these features would be significant to look at. Figure 1 highlights the context of this study with respect to the existing literature in the Sk–k_{rms}/k_a space. It can be seen that most studies have focused on surfaces with either low *Sk* and low k_{rms}/k_a or high *Sk* and high k_{rms}/k_a. In this study, we will expand the investigated parameter space by designing unique cubical
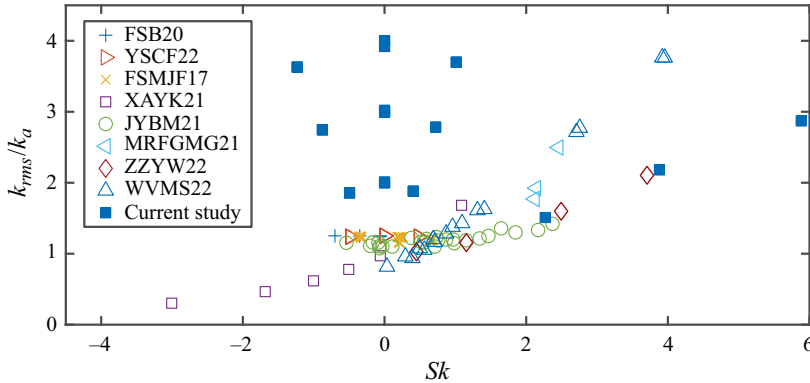
Figure 1. Parameter space describing the rough surfaces in the current study and certain existing works: Flack *et al.* (2020), Yang *et al.* (2022), Forooghi *et al.* (2017), Xu *et al.* (2021), Jouybari *et al.* (2021), Medjnoun *et al.* (2021), Zhang *et al.* (2022) and Womack *et al.* (2022). Each dataset is denoted by the initials of the authors' last names.

surfaces that fall in a region of low Sk and high k_{rms}/k_a. This dataset will allow us to study the hydrodynamic properties of unconventional roughness and test the efficacy of the existing rough-wall modelling approaches. We will see that the valleys, or the pits, do not contribute significantly to any of the first- and second-order statistics reported here. We will also see that the accuracy of a rough-wall model depends largely on whether its input space distinguishes the rough walls under consideration.

The rest of the paper is organized as follows. We present the details of the computational set-up in § 2 along with the geometry of the rough surfaces. The DNS results, including the instantaneous flow field and the mean flow statistics, are presented in § 3. We will see that although the rough surfaces presented here are new, the resulting flow behaviour aligns with trends found in the existing literature. In § 4, we discuss the implications of the DNS results on roughness modelling. Finally, conclusions are given in § 5.

2. Computational details

2.1. *Case set-up*

Figure 2 depicts a half-channel set-up containing a wall with cube-like roughness elements. This is a representative domain over which flow is computed for the DNS runs. The bottom surface is a rough wall comprising pits (or valleys) and protrusions (or peaks) in the form of cube-like elements. The x, y and z represent streamwise, spanwise and wall-normal directions. Periodic conditions are applied to the lateral boundaries. A stress-free condition with no penetration ($\partial u/\partial z = \partial v/\partial z = w = 0$) is imposed on the upper boundary. A streamwise pressure gradient acts as the forcing that drives the flow. A friction Reynolds number $Re_\tau \approx 400$ is used for all cases in this study, where Re_τ is defined as

$$Re_\tau = \frac{u_\tau (L_z - h_2)}{\nu}. \tag{2.1}$$

Note that the wall-normal length L_z used in (2.1) includes the depth of the negatively skewed features or valleys h_2. This depth h_2 is subtracted from L_z to make a good estimate
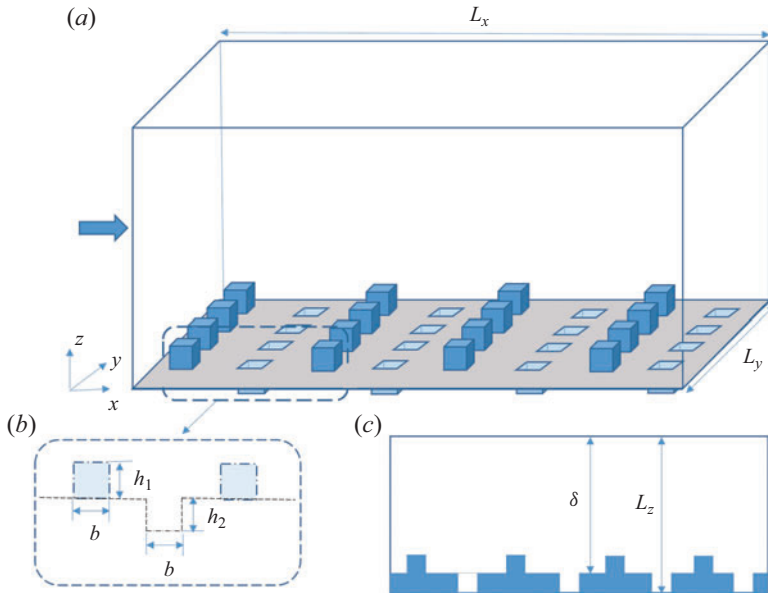
Figure 2. A sketch illustrating a periodic half-channel flow set-up over cube-like peaks and valleys. Here, h_1 denotes peak height, h_2 denotes valley depth, and b denotes the width of the cube-like element. The boundary layer thickness δ here is at least 6 times the maximum of h_1 and h_2.

of the half-channel height. The friction velocity u_τ is obtained using

$$u_\tau = \sqrt{\left|\frac{\mathrm{d}\langle\bar{p}\rangle}{\mathrm{d}x}\right| \frac{V_f}{\rho A_p}} \approx \sqrt{\left|\frac{\mathrm{d}\langle\bar{p}\rangle}{\mathrm{d}x}\right| \frac{(L_z - h_2)}{\rho}}, \qquad (2.2)$$

where $\mathrm{d}\langle\bar{p}\rangle/\mathrm{d}x$ refers to the mean streamwise pressure gradient, and ρ, V_f and A_p refer to the fluid density, half-channel fluid volume and wall-parallel area, respectively. Here $\langle\cdot\rangle$ denotes the streamwise and spanwise averaging (or double-averaging) operation, and $\overline{(\cdot)}$ denotes time averaging.

Note that V_f includes the fluid volume within valleys, and the approximation sign in (2.2) is used because the peak height is not always equal to the valley depth, causing minor differences in V_f in some cases.

The size of the computational domain is determined such that $L_x > 6L_z$ and $L_y > 3L_z$. Lozano-Durán & Jiménez (2014) have shown that such a domain would be sufficiently large to ensure the accuracy of first- and second-order statistics for plane channel flow. Note that L_x and L_y represent the streamwise and spanwise domain lengths, respectively. For rough-wall flows, this domain size should be a conservative estimate since DNS studies in Coceal *et al.* (2006) and Leonardi & Castro (2010) produce good mean flow statistics with smaller domains $L_x \times L_y \times L_z = 4h \times 4h \times 4h$ and $8h \times 6h \times 8h$, respectively (h being the height of the cube). Other studies, such as Chung *et al.* (2015), have shown that even further reduction in computational domain is possible in the spanwise direction, and such minimal-span channels have been shown to capture accurate mean drag characteristics (at least for sinusoidal roughness) when $L_y > k/0.4$ and $L_y^+ > 100$. Furthermore, since the roughness is regular, L_x and L_y are also integral multiples of the length of the repeating tile to ensure a periodic domain. Here, a repeating tile represents

the smallest unit that when repeated in the streamwise and spanwise directions produces the entire surface.

A uniform grid has been utilized with grid resolution such that $\Delta_x^+ = \Delta_y^+ < 5$ and $\Delta_z^+ < 3$, where these represent streamwise, spanwise and wall-normal grid spacings, respectively, normalized by the viscous length scale ($\delta_v = \nu/u_\tau$). It is important to note that while Δ_z^+ may appear to be large for DNS, this resolution is based on friction velocity u_τ obtained from (2.2). The δ_v value computed based on this u_τ includes the form drag component. If δ_v is computed from the local viscous stress at the mean surface comprising the bottom wall, then the grid resolution is at most 0.66 plus units. Note that the local viscous stress $\mu \, \partial u/\partial n$ here does not pertain to the cube surfaces, where the corresponding value would spike at the leading edge due to the strong shear rate. A similar grid resolution $\Delta_x^+ = \Delta_y^+ < 4.5$ and $\Delta_z^+ \approx 2.5$ was used by Zhang *et al.* (2022) for deep canopy flows. The grid spacing is also comparable to that in previous studies of rough-wall DNS, where Coceal *et al.* (2006) used $\Delta_{x,y,z}^+$ in the range 7.8–15.6, and Leonardi & Castro (2010) used $\Delta_{x,y}^+ = 19$. For further reference, DNS studies by Forooghi *et al.* (2017) have found $\Delta_{x,y,z_{max}}^+ = 3.5$ to be more than sufficient. Yuan & Piomelli (2014b) utilize $\Delta_x^+ = 12$ and $\Delta_y^+ = 6$, whereas Jouybari *et al.* (2021) employ $\Delta_x^+ = 7.5$ and $\Delta_y^+ = 6.3$ for obtaining accurate flow statistics in their rough-wall channels.

All simulations are performed using LESGO: a parallel pseudo-spectral large-eddy simulation code (see https://lesgo-jhu.github.io/lesgo), a solver of the filtered Navier–Stokes equations with pseudo-spectral discretization in the streamwise and spanwise directions, and a second-order finite difference in the wall-normal direction. Code LESGO and modified versions of the code have been used extensively for studies involving channel flows, including Bou-Zeid, Meneveau & Parlange (2005), Anderson & Meneveau (2011), Abkar & Porté-Agel (2012), Zhu *et al.* (2017), Yang *et al.* (2019), and several others. Immersed boundary conditions (Anderson 2013) are used at the solid boundaries comprising the rough wall. A Courant–Friedrichs–Lewy number (CFL $= u \, \Delta t/\Delta x$) of 0.2 is employed to automatically adjust the time step, which is advanced via the second-order Adams–Bashforth scheme.

2.2. *Roughness generation*

Roughness elements are distributed in various arrangements to generate different rough surfaces. We will vary one statistic at a time, but repeat the process and vary many roughness statistics. Figures 3(*a*), 3(*c*) and 3(*d*) depict these elements aligned in the spanwise (AY), streamwise (AX) and 45° (AXY) directions, respectively. Their staggered equivalents (S and SXY) are shown in figures 3(*b*) and 3(*e*). Figure 3(*f*) shows another surface (NV) where the negatively skewed features (valleys) are removed. The rough surfaces also include variations in λ_p, as indicated by figures 3(*g*) and 3(*h*), peak height to width ratio (h_1/b), and valley depth to width ratio (h_2/b). Table 1 provides the naming convention used, and table 2 lists all 36 different surfaces considered for the DNS study. For each case, the nomenclature is as follows: [Arrangement][Packing density][Roughness height], where arrangement could be AX, AY, AXY, S, SXY, NV, packing density could be L1, L2, L3, and roughness height could be H1, H2, H3. For example, the L1, L2 and L3 in AYL1H1, AYL2H1 and AYL3H1 correspond to λ_p magnitudes 6.25 %, 11.1 % and 25 %, respectively. Similarly, the H1, H2 and H3 in AYL1H1, AYL1H2 and AYL1H3 denote variations in the heights of the peaks as $h_1 = b$, $1.2b$, $0.8b$. Note that in this context, the
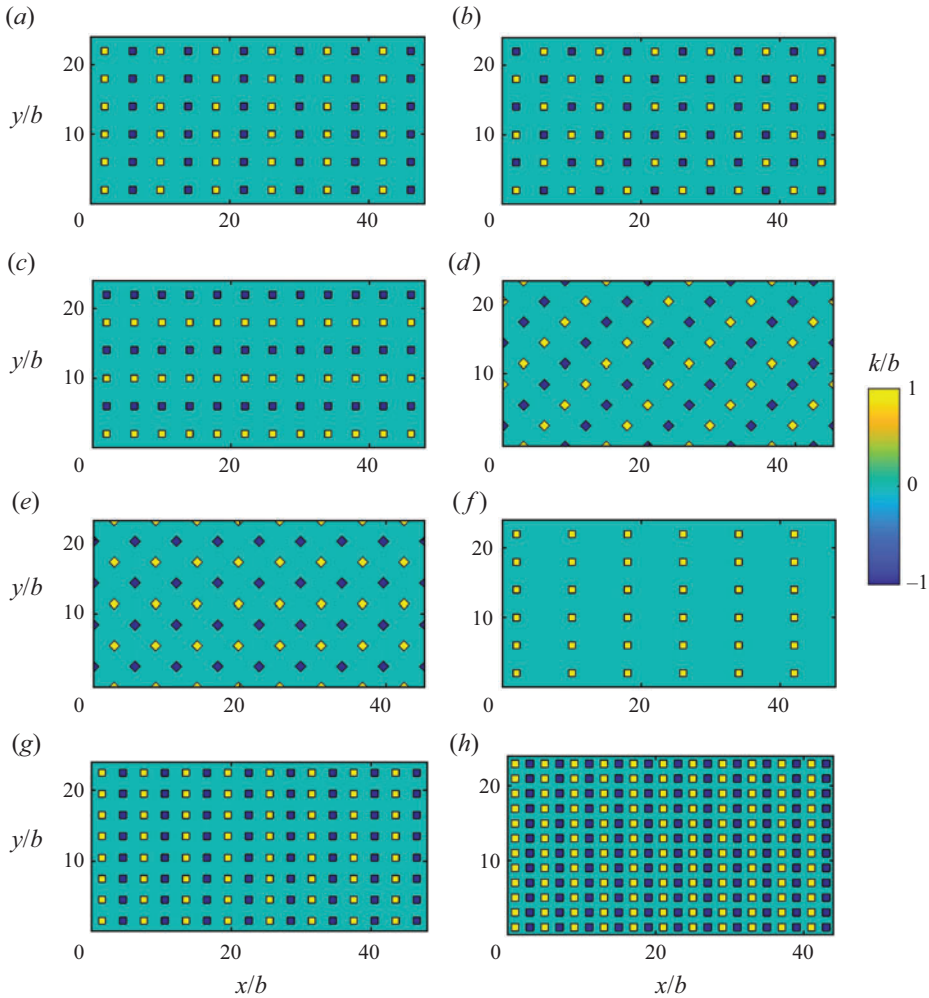
Figure 3. Top view of various rough-surface types considered for the DNS study: (*a*) AYL1H1, (*b*) SL1H1, (*c*) AXL1H1, (*d*) AXYL1H1, (*e*) SXYL1H1, (*f*) NVL1H1, (*g*) AYL2H1 and (*h*) AYL3H1. Here, *k* is the elevation of the surface. The peaks have positive *k* (depicted as yellow), and the valleys have negative *k* (depicted as blue).

conditions $h_1/h_2 = 1$, $h_1/h_2 > 1$ and $h_1/h_2 < 1$ correspond to zero skewed, positively skewed and negatively skewed surfaces. The separation lengths between the roughness elements are listed as 4*b*, 3*b* and 2*b* for L1, L2 and L3 configurations, with b/L_z assuming values 0.1375, 0.1111 and 0.1429 for H1, H2 and H3 cases, respectively.

The definitions for a few roughness statistics k_{rms}, k_a, k_t, *Sk*, *ES*, *Ku* and λ_p (which are discussed further in table 3) for the rough surfaces thus generated are listed as follows:

$$\bar{k} = \frac{1}{L_x L_y} \int_{L_x} \int_{L_y} k(x, y) \, dx \, dy, \qquad (2.3)$$

$$k_a = \frac{1}{L_x L_y} \int_{L_x} \int_{L_y} |k(x, y) - \bar{k}| \, dx \, dy, \qquad (2.4)$$

Arrangement	Packing density	Roughness height
AX: streamwise-aligned	L1: 6.25 % (AX, AY, AXY, S, SXY)	H1: $h_1 = b$
AY: spanwise-aligned	L1: 3.13 % (NV)	H2: $h_1 = 1.2b$
AXY: 45 °-aligned	L2: 11.1 % (AX, AY, AXY, S, SXY)	H3: $h_1 = 0.8b$
S: staggered	L2: 5.55 % (NV)	
SXY: 45 °-staggered	L3: 25 % (AX, AY, AXY, S, SXY)	
NV: no-valley	L3: 12.5 % (NV)	

Table 1. Nomenclature for the 36 configurations of the parameter space studied.

Case	λ_p	h_1/b	h_2/b	$N_x \times N_y \times N_z$	$L_x/b \times L_y/b \times L_z/b$	N
AYL1H1	6.25	1.0	1.0	$672 \times 336 \times 160$	$48.0 \times 24.0 \times 7.3$	$36 + 36$
AYL2H1	11.1	1.0	1.0	$672 \times 336 \times 160$	$48.0 \times 24.0 \times 7.3$	$64 + 64$
AYL3H1	25.0	1.0	1.0	$616 \times 336 \times 160$	$44.0 \times 24.0 \times 7.3$	$132 + 132$
AYL1H2	6.25	1.2	1.0	$672 \times 384 \times 180$	$56.0 \times 32.0 \times 9.0$	$56 + 56$
AYL2H2	11.1	1.2	1.0	$648 \times 360 \times 180$	$54.0 \times 30.0 \times 9.0$	$90 + 90$
AYL3H2	25.0	1.2	1.0	$672 \times 336 \times 180$	$56.0 \times 28.0 \times 9.0$	$196 + 196$
AYL1H3	6.25	0.8	1.0	$672 \times 336 \times 175$	$48.0 \times 24.0 \times 7.0$	$36 + 36$
AYL2H3	11.1	0.8	1.0	$588 \times 336 \times 175$	$42.0 \times 24.0 \times 7.0$	$56 + 56$
AYL3H3	25.0	0.8	1.0	$616 \times 336 \times 175$	$44.0 \times 24.0 \times 7.0$	$132 + 132$
SL1H1	6.25	1.0	1.0	$672 \times 336 \times 160$	$48.0 \times 24.0 \times 7.3$	$36 + 36$
SL2H1	11.1	1.0	1.0	$672 \times 336 \times 160$	$48.0 \times 24.0 \times 7.3$	$64 + 64$
SL3H1	25.0	1.0	1.0	$616 \times 336 \times 160$	$44.0 \times 24.0 \times 7.3$	$132 + 132$
SL1H2	6.25	1.2	1.0	$672 \times 384 \times 180$	$56.0 \times 32.0 \times 9.0$	$56 + 56$
SL2H2	11.1	1.2	1.0	$648 \times 360 \times 180$	$54.0 \times 30.0 \times 9.0$	$90 + 90$
SL3H2	25.0	1.2	1.0	$672 \times 336 \times 180$	$56.0 \times 28.0 \times 9.0$	$196 + 196$
SL1H3	6.25	0.8	1.0	$672 \times 336 \times 175$	$48.0 \times 24.0 \times 7.0$	$36 + 36$
SL2H3	11.1	0.8	1.0	$588 \times 336 \times 175$	$42.0 \times 24.0 \times 7.0$	$56 + 56$
SL3H3	25.0	0.8	1.0	$616 \times 336 \times 175$	$44.0 \times 24.0 \times 7.0$	$132 + 132$
AXL1H1	6.25	1.0	1.0	$672 \times 336 \times 160$	$48.0 \times 24.0 \times 7.3$	$36 + 36$
AXL2H1	11.1	1.0	1.0	$672 \times 336 \times 160$	$48.0 \times 24.0 \times 7.3$	$64 + 64$
AXL3H1	25.0	1.0	1.0	$616 \times 336 \times 160$	$44.0 \times 24.0 \times 7.3$	$132 + 132$
AXL1H2	6.25	1.2	1.0	$672 \times 384 \times 180$	$56.0 \times 32.0 \times 9.0$	$56 + 56$
AXL2H2	11.1	1.2	1.0	$648 \times 360 \times 180$	$54.0 \times 30.0 \times 9.0$	$90 + 90$
AXL3H2	25.0	1.2	1.0	$672 \times 336 \times 180$	$56.0 \times 28.0 \times 9.0$	$196 + 196$
AXL1H3	6.25	0.8	1.0	$672 \times 336 \times 175$	$48.0 \times 24.0 \times 7.0$	$36 + 36$
AXL2H3	11.1	0.8	1.0	$588 \times 336 \times 175$	$42.0 \times 24.0 \times 7.0$	$56 + 56$
AXL3H3	25.0	0.8	1.0	$616 \times 336 \times 175$	$44.0 \times 24.0 \times 7.0$	$132 + 132$
AXYL1H1	6.25	1.0	1.0	$724 \times 362 \times 160$	$45.3 \times 22.6 \times 7.3$	$32 + 32$
AXYL2H1	11.1	1.0	1.0	$916 \times 458 \times 160$	$50.9 \times 25.5 \times 7.3$	$72 + 72$
AXYL3H1	25.0	1.0	1.0	$724 \times 362 \times 160$	$45.3 \times 22.6 \times 7.3$	$128 + 128$
SXYL1H1	6.25	1.0	1.0	$724 \times 362 \times 160$	$45.3 \times 22.6 \times 7.3$	$32 + 32$
SXYL2H1	11.1	1.0	1.0	$916 \times 458 \times 160$	$50.9 \times 25.5 \times 7.3$	$72 + 72$
SXYL3H1	25.0	1.0	1.0	$724 \times 362 \times 160$	$45.3 \times 22.6 \times 7.3$	$128 + 128$
NVL1H1	3.13	1.0	0.0	$576 \times 288 \times 160$	$48.0 \times 24.0 \times 7.3$	$36 + 0$
NVL2H1	5.55	1.0	0.0	$576 \times 288 \times 160$	$48.0 \times 24.0 \times 7.3$	$64 + 0$
NVL3H1	12.5	1.0	0.0	$528 \times 288 \times 160$	$44.0 \times 24.0 \times 7.3$	$132 + 0$

Table 2. Details of the DNS case-set-up, where λ_p, h_1/b, h_2/b denote planar packing density, peak height to width, and valley depth to width ratios, respectively; $N_{x,y,z}$ denotes the number of grid cells in the streamwise (x), spanwise (y) and wall-normal (z) directions; N denotes the number of peaks plus valleys in the case. Note that the surface coverage densities of the NV cases are half the value of their AX (and AY) counterparts.

$$k_{rms} = \sqrt{\frac{1}{L_x L_y} \int_{L_x} \int_{L_y} |k(x, y) - \bar{k}|^2 \, dx \, dy}, \tag{2.5}$$

$$k_t = \max(k(x, y)) - \min(k(x, y)), \tag{2.6}$$

$$Sk = \frac{1}{L_x L_y} \int_{L_x} \int_{L_y} (k(x, y) - \bar{k})^3 \, dx \, dy / k_{rms}^3, \tag{2.7}$$

$$Ku = \frac{1}{L_x L_y} \int_{L_x} \int_{L_y} |k(x, y) - \bar{k}|^4 \, dx \, dy / k_{rms}^4, \tag{2.8}$$

$$ES = \frac{1}{L_x} \int_{L_x} \left| \frac{\partial k}{\partial x} \right| \, dx, \tag{2.9}$$

$$\lambda_p = \frac{A_p}{L_x L_y}. \tag{2.10}$$

The packing density λ_p in (2.10) is defined in the same way as plan solidity in Chung *et al.* (2021). Here, A_p comprises the plan area of roughness elements, which in this case includes the base area of peaks and valleys.

Additional constraints have been placed to ensure that the grid surface matches the cube surface for accurate roughness representation in the simulations. This is done by enforcing b/Δ_x, b/Δ_y, h_1/Δ_z and h_2/Δ_z to be integers, where Δ_x, Δ_y and Δ_z stand for the grid spacing in the streamwise, spanwise and wall-normal directions. This has been implemented for all cases except the $45°$ alignment rough surfaces.

As can be seen from figure 4, the grid is in perfect alignment with the roughness element in AYL1H1 as opposed to AXYL1H1. As a result, there are minor variations in the average roughness height k_a/b, root mean square roughness height k_{rms}/k_a, and kurtosis Ku for AXYLiH1 and SXYLiH1 surfaces when compared with AYLiH1 and SLiH1, where $i = 1, 2, 3$. This can be noticed in table 3. A finer resolution has been opted for in all such cases to minimize these variations. The streamwise and spanwise grid resolutions, in terms of number of cells per roughness element, are given by $b/\Delta_x = b/\Delta_y = 12$–$14$ for AX, AY, S and NV surfaces. These grid resolutions are approximately 16–18 for AXY and SXY surfaces. The wall-normal grid resolution is in the range $b/\Delta_z = 20$–25 for all cases.

3. DNS results

The rough walls that we study cover an underexplored region in roughness parameter space and therefore are a good addition to the rough-wall literature. In this section, we first discuss qualitative findings pertaining to the instantaneous flow field. Subsequently, the quality of our channel flow simulations is examined with the help of the mean momentum budget, followed by results of mean velocity profiles and quantitative estimates of effective roughness height (z_0) and zero-plane displacement height (d). The section concludes with Reynolds and dispersive stress results and comparisons for all 36 cases.

3.1. *Instantaneous flow field*

We begin by presenting the instantaneous flow field as an introduction to the dataset. From the contours in figure 5, an overall decrease in instantaneous bulk velocity can be observed with increasing λ_p. A region of reduced streamwise velocity can be noted in the immediate

Case	k_a/b	k_{rms}/k_a	Sk	Ku	k_t/k_a	ES	Rl_x/k_a	Rl_y/k_a	d/b	z_0/b
AYL1H1	0.063	4.00	0.00	16.0	31.98	0.125	127.9	63.95	0.182	0.047
AYL2H1	0.111	2.99	0.00	9.00	17.97	0.222	53.92	26.96	0.361	0.068
AYL3H1	0.250	2.00	0.00	4.00	7.99	0.500	15.99	7.990	0.549	0.117
AYL1H2	0.075	3.70	1.01	16.4	29.45	0.138	107.0	53.49	0.296	0.064
AYL2H2	0.132	2.78	0.72	9.22	16.63	0.244	45.31	22.65	0.444	0.101
AYL3H2	0.293	1.88	0.41	4.06	7.50	0.550	13.62	6.81	0.703	0.150
AYL1H3	0.062	3.63	−1.23	16.7	28.90	0.113	128.3	64.16	0.149	0.031
AYL2H3	0.110	2.75	−0.88	9.32	16.38	0.200	54.59	27.26	0.218	0.048
AYL3H3	0.244	1.86	−0.49	4.09	7.39	0.450	16.38	8.19	0.424	0.086
SL1H1	0.063	4.00	0.00	16.0	31.98	0.125	127.9	127.9	0.234	0.042
SL2H1	0.111	2.99	0.00	9.00	17.97	0.222	53.92	53.92	0.293	0.069
SL3H1	0.250	2.00	0.00	4.00	7.99	0.500	15.99	15.99	0.474	0.120
SL1H2	0.075	3.70	1.01	16.4	29.45	0.138	107.0	107.0	0.300	0.062
SL2H2	0.132	2.78	0.72	9.22	16.63	0.244	45.31	45.31	0.391	0.104
SL3H2	0.293	1.88	0.41	4.06	7.50	0.550	13.62	13.62	0.620	0.154
SL1H3	0.062	3.63	−1.23	16.7	28.90	0.113	128.3	128.3	0.163	0.027
SL2H3	0.110	2.75	−0.88	9.32	16.38	0.200	54.59	54.59	0.183	0.047
SL3H3	0.244	1.86	−0.49	4.09	7.39	0.450	16.38	16.38	0.350	0.088
AXL1H1	0.063	4.00	0.00	16.0	31.98	0.125	63.95	127.9	0.243	0.026
AXL2H1	0.111	2.99	0.00	9.00	17.97	0.222	26.96	53.92	0.287	0.034
AXL3H1	0.250	2.00	0.00	4.00	7.99	0.500	7.99	15.99	0.555	0.027
AXL1H2	0.075	3.70	1.01	16.4	29.45	0.138	53.49	107.0	0.417	0.033
AXL2H2	0.132	2.78	0.72	9.22	16.63	0.244	22.65	45.31	0.473	0.045
AXL3H2	0.293	1.88	0.41	4.06	7.50	0.550	6.81	13.62	0.789	0.047
AXL1H3	0.062	3.63	−1.23	16.7	28.90	0.113	64.16	128.3	0.076	0.021
AXL2H3	0.110	2.75	−0.88	9.32	16.38	0.200	27.26	54.59	0.272	0.023
AXL3H3	0.244	1.86	−0.49	4.09	7.39	0.450	8.19	16.38	0.396	0.018
AXYL1H1	0.065	3.92	0.00	15.5	30.90	0.176	180.9	180.9	0.080	0.079
AXYL2H1	0.110	3.02	0.00	9.08	18.21	0.307	76.26	76.26	0.005	0.139
AXYL3H1	0.247	2.01	0.00	4.04	8.09	0.688	22.61	22.61	0.070	0.201
SXYL1H1	0.065	3.92	0.00	15.48	30.90	0.176	90.45	90.45	−0.424	0.107
SXYL2H1	0.110	3.02	0.00	9.08	18.21	0.307	38.13	38.13	0.128	0.109
SXYL3H1	0.247	2.01	0.00	4.04	8.09	0.688	11.31	11.31	0.535	0.120
NVL1H1	0.060	2.87	5.88	30.0	16.56	0.063	132.5	66.27	0.132	0.048
NVL2H1	0.105	2.18	3.88	16.1	9.55	0.111	57.18	28.59	0.350	0.068
NVL3H1	0.219	1.51	2.27	6.14	4.57	0.250	18.29	9.14	0.636	0.106

Table 3. Hydrodynamic properties of roughness. The geometric parameters listed include average and root mean square roughness height k_a and k_{rms}, skewness Sk, kurtosis Ku, maximum peak to trough height k_t, effective slope (along the x direction) ES, streamwise and spanwise correlation lengths Rl_x and Rl_y (defined in (4.6), (4.7)). Here, d denotes zero-plane displacement height, and z_0 is the effective roughness height.

wake of the protrusions. The flow in the pits can be observed to be less energetic, and their interactions with the outer layer are minimal.

Figure 6 shows the velocity contours on a wall-parallel plane near the peak height for different roughness element arrangements at the same λ_p. By visual inspection, contours in figures 6(a) and 6(b) are observed to have similar and intermediate average streamwise velocities. Figure 6(c) contains the highest and figure 6(d) the lowest average streamwise velocities amongst these four cases. In all cases, streamwise streaks of high and low velocities can be observed near the roughness peak. These streaks are noticeably longer in certain cases, such as in figure 6(c), as the wakes overlap each other. These wake interactions become more pronounced at higher λ_p and result in a flow-sheltering effect as observed in densely packed surfaces such as in Xu *et al.* (2021).
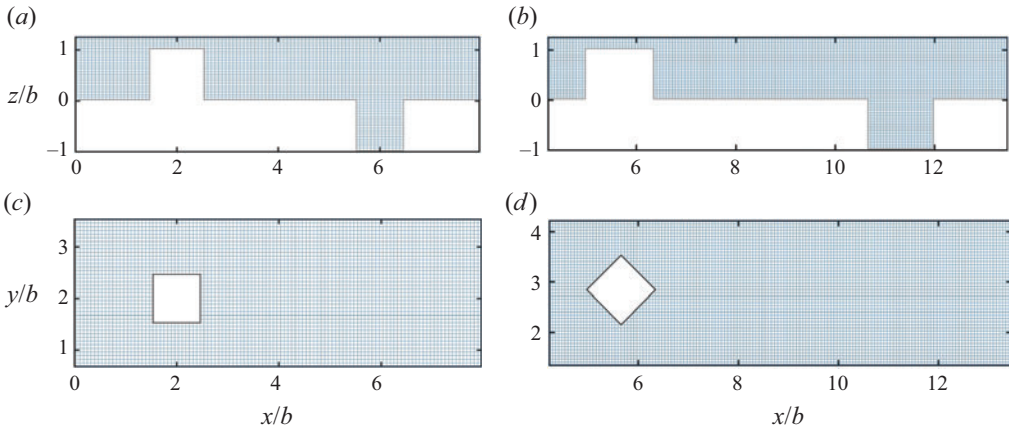
Figure 4. Top and side views of DNS grids used for two cases: (*a*,*c*) AYL1H1 and (*b*,*d*) AXYL1H1. The side views are streamwise wall-normal (*x*–*z*) planes cut at the middle of the cubes. The top views are wall-parallel (*x*–*y*) planes cut near the roughness peak.
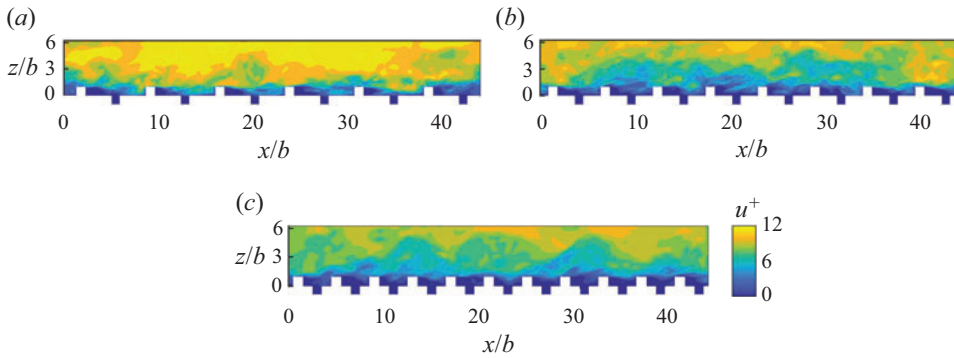

Figure 5. Streamwise instantaneous velocity contours at $y/L_y \approx 0.43$ for (*a*) AYL1H1, (*b*) AYL2H1 and (*c*) AYL3H1.
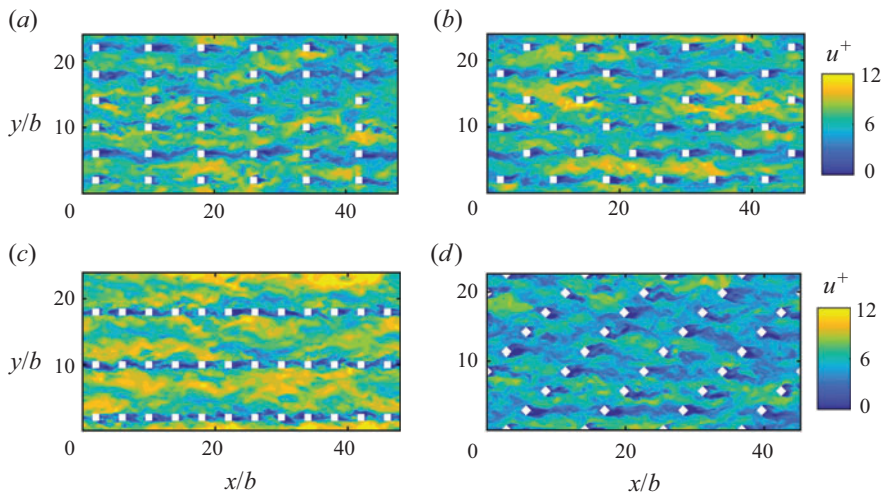

Figure 6. Streamwise instantaneous velocity contours at $z/b \approx 1.0$ for (*a*) AYL1H1, (*b*) SL1H1, (*c*) AXL1H1 and (*d*) AXYL1H1.

Although we will pursue a quantitative analysis in the following subsections, here we may already conclude first that the valleys do not contribute significantly to the drag, and second that two-point rough-wall statistics are important to the characterization of the equivalent sand-grain roughness height k_s.

3.2. *Statistical convergence*

Before presenting the flow statistics, we first check the statistical convergence of our data. The statistical convergence can be evaluated with the help of the mean momentum budget following Coceal *et al.* (2006), Leonardi & Castro (2010), Xu *et al.* (2021) and Zhang *et al.* (2022). The equation is comprised of viscous diffusion, turbulent transport, dispersive stress and pressure-gradient terms:

$$\nu \frac{d\langle \bar{u} \rangle}{dz} - \langle \overline{u'w'} \rangle - \langle \overline{u''w''} \rangle = \frac{1}{\rho} \frac{d\langle \bar{p} \rangle}{dx} (z - \delta). \tag{3.1}$$

It is considered to be statistically converged when the sum of these stresses is a linear function of the wall-normal distance (z), i.e. when the budget balances without the unsteady term. Note that (3.1) is obtained by integrating the x component of the double-averaged momentum equation

$$\frac{d}{dz} \left(\nu \frac{d\langle \bar{u} \rangle}{dz} - \langle \overline{u'w'} \rangle - \langle \overline{u''w''} \rangle \right) - \frac{1}{\rho} \frac{d\langle \bar{p} \rangle}{dx} - f = 0 \tag{3.2}$$

found in (15) of Nikora *et al.* (2013), with static roughness in place of the mobile roughness assumed by the authors. Here, f represents the drag force. As f is non-zero within the roughness-occupied region, (3.1) applies to the region outside the roughness only. The averaging procedure corresponds to the double-averaging method, common in the roughness literature, introduced by Raupach & Shaw (1982) for flows within vegetation canopies. The spatial averaging used in the context of (3.1) and (3.2) is superficial, as per the definition for superficial averaging in Schmid *et al.* (2019).

Figure 7 shows these terms for a few cases. These cases have been selected so that they include different λ_p, Sk and roughness arrangements while maintaining brevity. We see that the total stresses in all cases follow a linear function of z/b, therefore our DNS are statistically converged. In addition, we make the following observations. The Reynolds stress R_{13} can be observed to be the largest in magnitude as compared to the viscous stress τ_v and dispersive stress D_{13} outside the roughness-occupied layer. The roughness-occupied layer is marked by the vertical line in figure 7. In the roughness-occupied layer, i.e. below this marking, τ_v is observed to contain two maxima, one caused by the surface at $z = 0$, and the other near the cube height $z = h_1$. The latter is also observed in other DNS studies by Xu *et al.* (2021) and Zhang *et al.* (2022). Meanwhile, R_{13} attains a maximum just above the peak height, and declines to zero as z approaches the outer layer; D_{13} is small and contained within the roughness region for most cases. For $z/b < 0$, i.e. below the surface in the region of valleys, all stresses are observed to be negligible. Note that figure 7(*e*) starts from $z/b = 0$ as the case NVL1H1 contains no valleys.

It might be worth noting certain differences in dispersive stress D_{13} in figure 7. For instance, most noticeable would be the differences in cases AYL3H1 and NVL1H1: D_{13} is observed to be non-negligible within the roughness-occupied region for AYL3H1, whereas D_{13} for NVL1H1 inside the sublayer is almost zero. Figure 8 aims to provide an intuitive explanation for this. Here, the dispersive streamwise and wall-normal fluctuating components, u'' and w'', are shown for averaged repeating tiles for cases NVL1H1 and
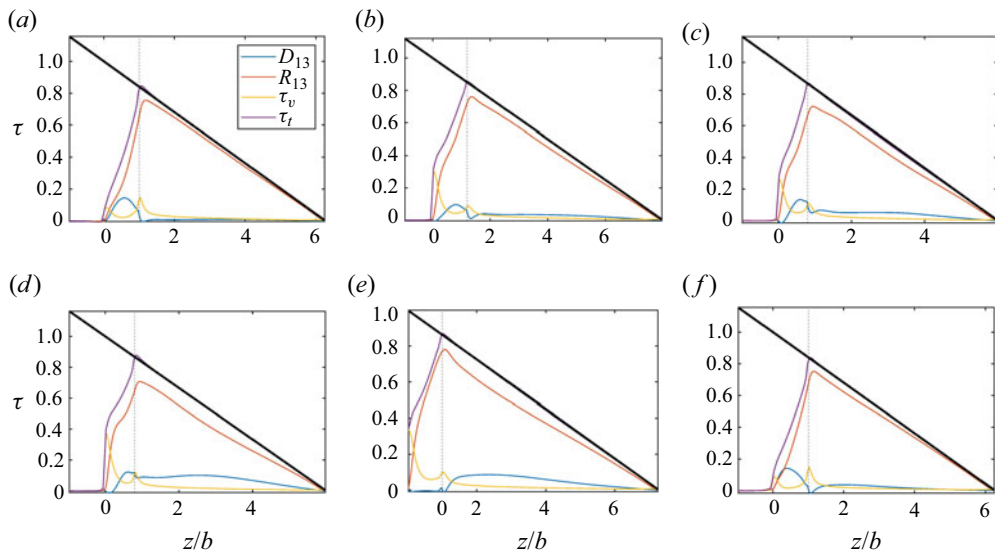
Figure 7. The stress budget plot comprising Reynolds stress R_{13}, dispersive stress D_{13}, viscous stress τ_v and total stress τ_t for cases: (*a*) AYL3H1, (*b*) SL1H2, (*c*) SL2H3, (*d*) AXL2H3, (*e*) NVL1H1 and (*f*) AXYL3H1. The black solid line corresponds to $1 - z/\delta$. The vertical dashed line denotes the peak roughness height.

AYL3H1. Three planes, one within and the others at and above the roughness crest, are chosen to display their scatter. When u'' and w'' predominantly lie in the second or fourth quadrant, the dispersive stress $-\langle u''w'' \rangle$ becomes positive. The negative slope of the linear fit is also indicative of this. This can be seen in figure 8(*c*), which corresponds to NVL1H1 just above the roughness, and figure 8(*d*), which corresponds to AYL3H1 within the roughness. These may indirectly imply mean flow inhomogeneity, which is due to the presence of surface roughness and/or secondary flows. On the other hand, when the points are more evenly spread across all quadrants, as is the case in figures 8(*a*), 8(*b*) and 8(*e*), or too close to the origin as in figure 8(*f*), this dispersive stress becomes negligible.

3.3. *Mean velocity profiles*

The mean flow statistics are presented in this subsection. Additionally, the relevant rough surface parameters and the hydrodynamic properties determined from mean velocity profiles are also listed in table 3.

Figure 9(*a*) shows the inner scaled mean velocity profiles, comparing the spanwise-aligned and staggered cases, i.e. the AY cases and the S cases. The following observations can be duly noted. First, the higher λ_p cases of surface coverage density 25 %, comprising AYL3H1, AYL3H2, AYL3H3 and their staggered counterparts SL3H1, SL3H2 and SL3H3, produce lower magnitudes of mean velocity. This is expected for roughness in the *k*-type regime. Second, the staggered arrangement produces similar velocity profiles to the spanwise-aligned arrangement. This is not unexpected, and similar observations were made in Yang *et al.* (2016). Due to a large spanwise distance between neighbouring roughness elements in the spanwise direction, staggering the roughness element in the streamwise direction has little effect on the mean flow. Finally, at constant λ_p, cases with higher skewness (*Sk*) produce lower mean velocity profiles. This role of *Sk* is apparent as we notice increased drag with increased *Sk* (consistent with systematic studies
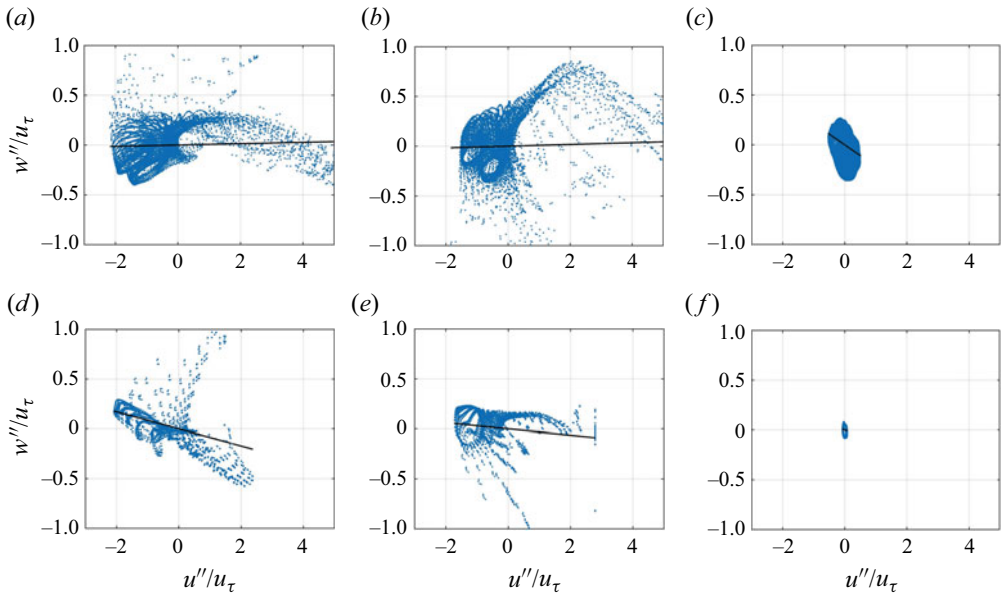
Figure 8. Quadrant analysis for the dispersive stress components of D_{13} for cases NVL1H1 at (*a*) $z/b = 0.5$, (*b*) $z/b = 1$, (*c*) $z/b = 2$ and AYL3H1 at (*d*) $z/b = 0.5$, (*e*) $z/b = 1$ and (*f*) $z/b = 2$. The solid black line depicts a linear fit on the data points.

performed by Flack *et al.* 2020), and we observe this effect in all our cases, irrespective of arrangement. For example, comparing AYL1H1, AYL1H2 and AYL1H3 from figure 9(*a*) or AXL1H1, AXL1H2 and AXL1H3 from figure 9(*b*) at the same z^+, one can notice that the mean velocities in the log layer are slightly lower for the surface with higher *Sk* (H3 cases being lower than H2 and H1 cases). In the present dataset, variations in *Sk* are introduced by varying the height of the peaks, while maintaining the depth of the dips constant, therefore any changes due to *Sk* could also be viewed as due to the roughness height. Similarly, as variations in k_{rms}/k_a are introduced by varying the surface coverage density, any changes due to k_{rms}/k_a could also be viewed as due to surface coverage density.

Similarly, figure 9(*b*) compares mean velocity profiles for roughness with streamwise and spanwise alignment, i.e. AX cases and AY cases. When the roughness features are aligned in the streamwise direction, higher mean velocities can be observed as compared to those aligned in the spanwise direction. An intuitive explanation for this is the unobstructed passage of fluid between two rows of roughness elements giving the fluid less drag when they are aligned in the streamwise direction. This is also evident from the instantaneous velocity contours in figure 6(*c*).

Figure 10 includes the mean velocity profiles from the 45° arrangements (XY), with figure 10(*a*) comparing AY, AX, AXY, and figure 10(*b*) comparing S and SXY. For the same λ_p, the surfaces follow AX > AY > AXY in figure (*a*), and S > SXY in figure (*b*) for the order of mean velocity magnitudes in the log-law region. The orientation of the roughness element is found to be an important factor here. The AXY and SXY surfaces possess a higher frontal area (i.e. higher solidity) as compared to their counterparts AX, AY or S, which is the reason for the higher drag observed for these surfaces.

(*a*)

(*b*)
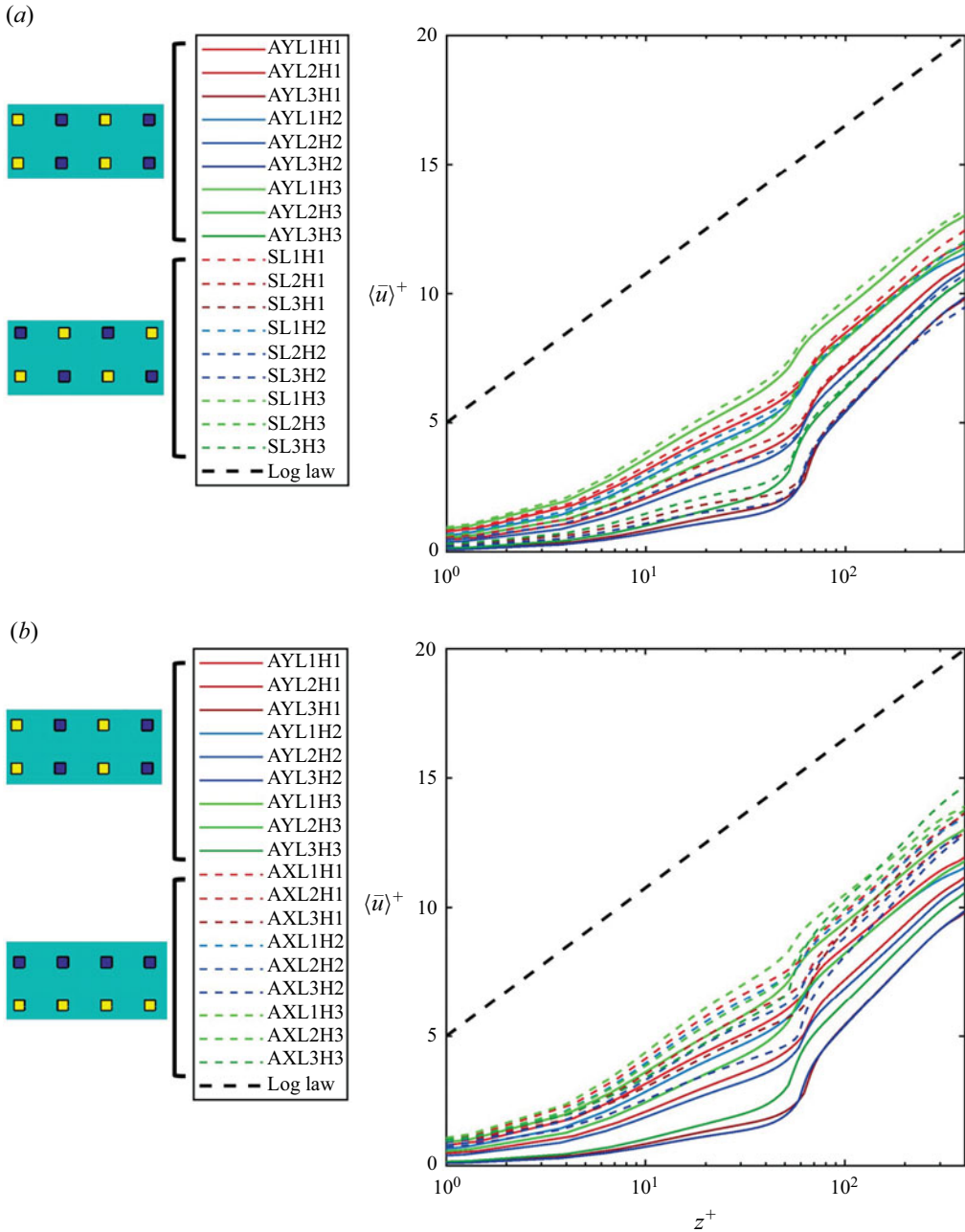
Figure 9. Mean streamwise velocity profile comparing (*a*) spanwise-aligned and staggered cases, and (*b*) spanwise-aligned and streamwise-aligned cases.

Figure 11(*a*) shows the mean velocity profiles comparing three pairs of surfaces with and without valleys. It can be observed that the valleys do contribute to mean velocity reduction, but by only a small amount. This reduced mean velocity can be inferred to be caused by the additional pressure drag contributions from the valleys. In support of this claim, figures 11(*b*) and 11(*c*) are shown, which contain mean pressure contours that are
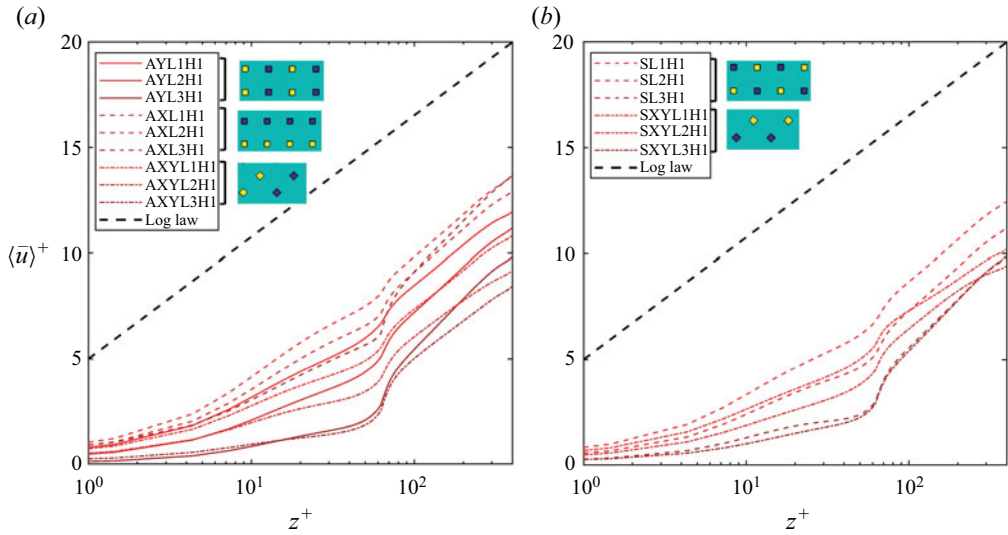
Figure 10. Mean streamwise velocity profile comparing (*a*) spanwise, streamwise and 45° aligned cases, and (*b*) staggered and staggered 45° cases.
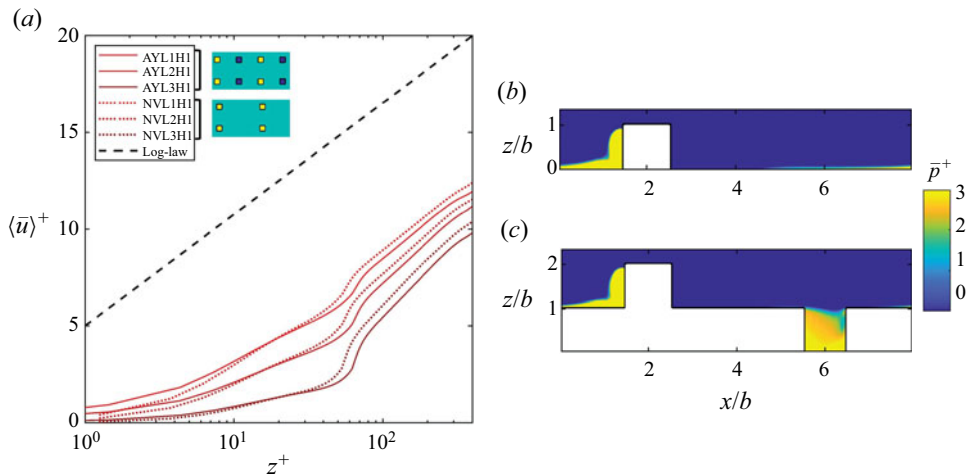
Figure 11. Details of the effects of valleys. (*a*) Mean streamwise velocity profile comparing cases with and without valleys. The solid and dotted lines depict spanwise-aligned and no-valleys cases, respectively. (*b*) Mean pressure contours for a repeating tile in NVL1H1. (*c*) Mean pressure contours for a repeating tile in AYL1H1.

averaged over repeating tiles in both (x, y) directions in AYL1H1 and NVL1H1. In these pressure contours, which are normalized by ρu_τ^2, the volume-averaged mean pressure is taken as the reference pressure. The additional pressure drop can be noticed clearly in valleys. Note that the observation here cannot be generalized to all dips. In particular, dips or valleys will play a significant role if they occupy a considerable part of the surface area, in which scenario a new bottom surface forms and the protrusions between the dips can be viewed as surface roughness.

Next, we measure the hydrodynamic properties of the roughness. The effective roughness height z_0 and zero-plane displacement height d for the rough surfaces are
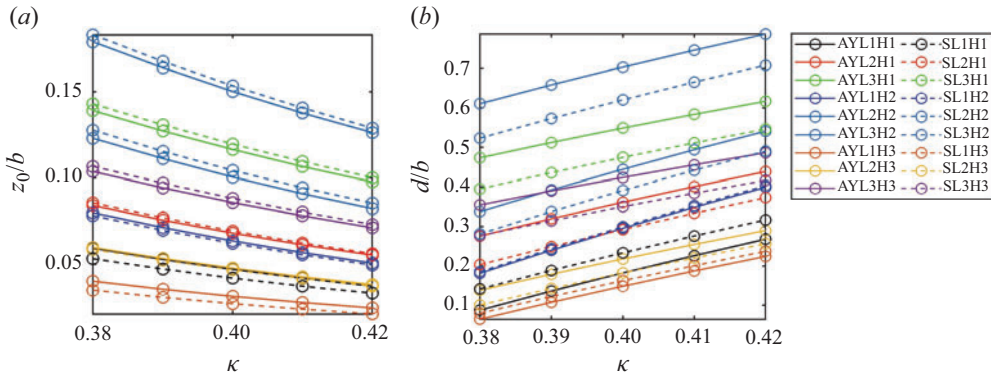
Figure 12. Plot showing the sensitivity to variations in von Kármán constant κ for spanwise-aligned (AY) and staggered (S) cases for the calculated (*a*) effective roughness height z_0/b, and (*b*) zero-plane displacement height d/b.

determined from the mean streamwise velocity profiles as

$$\frac{\langle \bar{u} \rangle}{u_\tau} = \frac{1}{\kappa} \log \frac{z - d}{z_0}, \tag{3.3}$$

where $\langle \bar{u} \rangle$ and z are taken from the log-law region, which has been taken to be in the range $(z - h_1)^+ \approx 30\text{--}100$. Further evidence of the existence of a log law in this range will be discussed later with the help of figure 13. Linear regression of (3.3) yields the resulting z_0 and d reported in table 3. Here, the von Kármán constant is set to $\kappa = 0.4$. There can be alternate approaches to calculate d from the centroid of the total force distribution following Jackson (1981) and Coceal *et al.* (2006). However, it is also observed that the centre of the force could sometimes underestimate d when the flow primarily interacts with the top portion of the roughness, as seen in figure 7 of Xu *et al.* (2021). Here, we prefer regression fitting over this approach for the aforementioned reason.

Since we assume the value of the von Kármán constant $\kappa = 0.4$ for our regression procedure, checking the changes in measured z_0/b due to variations in κ is important. Figure 12 shows the sensitivity of calculated z_0/b and d/b with various values of κ. It can be seen that z_0/b decreases and d/b increases with κ. While we do see noticeable variations in these values, it is worth noting that these changes would not be significant on the log scale, which is the relevant scale for the drag produced by these surfaces. Also, for a given value of κ, the differences in z_0/b and d/b are almost the same, which means that it is reasonable to compare these values for the different surfaces.

Further, we note that determining the roughness height z_0 or equivalent sand-grain roughness height k_s requires the flow to be in the fully rough regime. Based on Jiménez (2004), flow is considered fully rough when $k_s^+ > 80$. In Jouybari *et al.* (2021), the authors have chosen $k_s^+ > 50$ to be their definition for the fully rough regime. We have verified that all surfaces except AXL1H3 and AXL3H3 satisfy this latter condition, the k_s^+ values for which stand at 42.6 and 36 respectively. However, it is important to note that the exact value of k_s^+ for which this transition to fully rough flow happens is unknown and also depends on the type of roughness being considered Flack & Schultz (2010). Even so, we could still measure ΔU^+ and z_0, and discuss their variations with respect to the roughness as we are making these comparisons at a fixed Reynolds number ($Re_\tau = 400$).
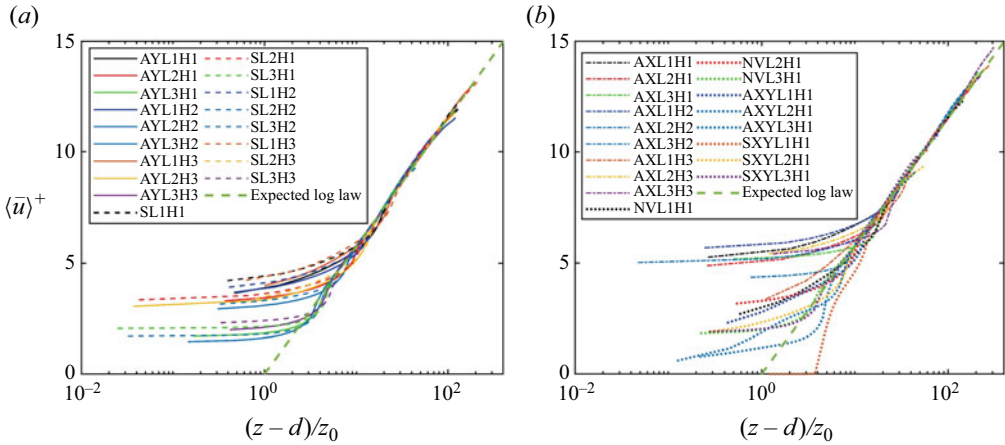
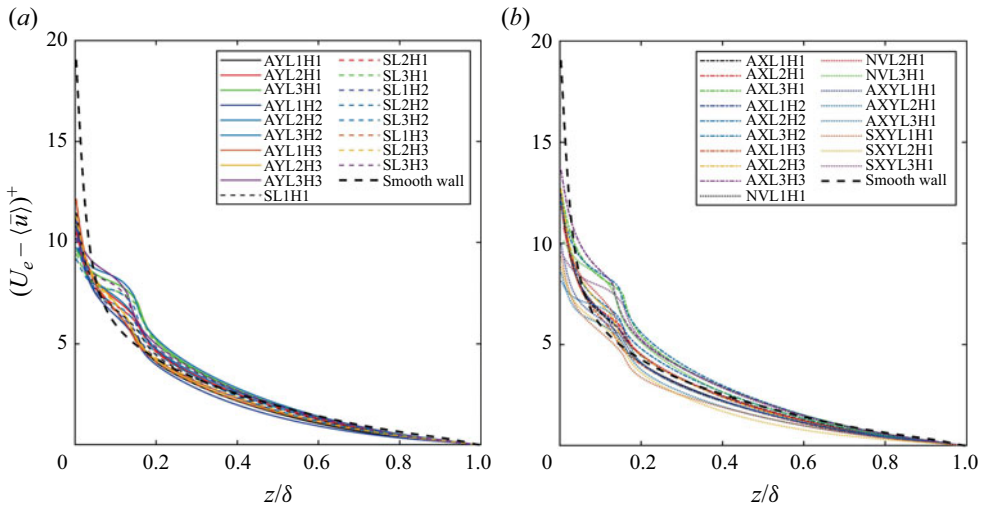Figure 13. Plot showing the log-law collapse for all cases: (*a*) AY, S cases; (*b*) AX, NV, AXY and SXY cases.

Figure 14. Velocity defect profiles for all surfaces: (*a*) AY, S cases; (*b*) AX, NV, AXY and SXY cases.

Figure 13 illustrates the quality of this log-law fit. The mean velocity plots can be observed to show a good collapse in the log layer region for all cases. The collapse is also a sign of mean flow universality, wherein such two-parameter forms of mean velocity profiles have been know to be adequate irrespective of the rough surface for $h/\delta \approx 0.03$–0.5. Figure 14 shows mean velocity profiles in defect form. This is being presented here as evidence for outer-layer similarity. Reasonably well collapsed regions can be observed farther away from the wall for the smooth- and rough-wall cases, underscoring a universality in mean flow behaviour in the outer layer irrespective of surface conditions for fully rough flow.

The mean velocity profile for the smooth wall in figures 14(*a*) and 14(*b*) was obtained from DNS of a half-channel with $Re_\tau = 400$. Similarly, figure 15 depicts outer-layer similarity in the streamwise Reynolds stress profiles.

Figure 16 summarizes the z_0 magnitudes for all cases, and the data are consistent with the various mean velocity profiles observed so far. Figure 16(*a*) compares the
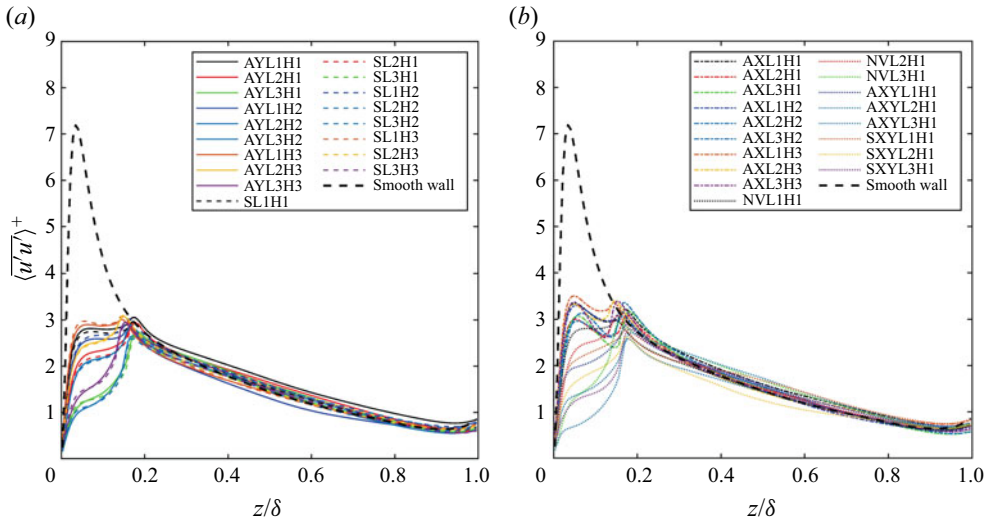
Figure 15. Normalized streamwise Reynolds stress profiles for all surfaces: (*a*) AY, S cases; (*b*) AX, NV, AXY and SXY cases.

spanwise-aligned (AY) and staggered (S) cases. The z_0/b magnitudes are observed to be similar between the AY and S cases when comparing cases with similar λ_p and Sk magnitudes. When cases with similar packing densities λ_p and orientation but different Sk are compared – for instance, AYL1H1, AYL1H2 and AYL1H3 – z_0/b increases with Sk. This holds true for all cases listed in table 3. When cases with similar Sk and orientation but different λ_p are compared – for instance, AYL1H1, AYL2H1 and AYL3H1 – z_0/b decreases with decreasing k_{rms}/k_a. This is observed true for all cases except the AX surfaces, which will be discussed subsequently. It should be of interest to note that low k_{rms}/k_a magnitudes actually correspond to high k_{rms}/b and λ_p. The normalization with k_a causes this as an increment in λ_p increases both k_a and k_{rms}, but k_a increases at a rate faster than k_{rms}. Figure 16(*b*) compares the spanwise-aligned (AY) and streamwise-aligned (AX) cases. The effects due to changes in roughness element alignment are evident as the AX surfaces in figure 16(*b*) show notably lower z_0/b magnitudes when compared to AY surfaces. For instance, even the densely packed surfaces AXL3H1, AXL3H2 and AXL3H3, which have higher solidities, are lower in z_0/b than the coarsely packed surfaces AYL1H1, AYL1H2 and AYL1H3. Moreover, it can be noticed that z_0/b first increases then decreases with decreasing k_{rms}/k_a (or increasing λ_p) for H1 and H3 cases in AX surfaces. For higher λ_p surfaces AXL2H2 and AXL3H2, the increase in z_0/b is not as significant as it is for AYL2H2 and AYL3H3. These observations showcase a flow-sheltering effect in AX surfaces. Trends with changes in Sk for AX surfaces are similar to those observed in AY ones, with z_0/b increasing with increasing Sk. Figure 16(*c*) reports the effect of valleys on z_0/b. The presence of valleys in the AY cases leads to higher k_{rms}/k_a compared to the NV cases, but the z_0/b magnitudes are very close between the AY and NV cases, with some noticeable difference at low k_{rms}/k_a magnitudes. Figure 16(*d*) presents spanwise-aligned (AY), 45° alignment (AXY) and their respective staggered equivalents (S and SXY). A significant increase in z_0/b can be observed as alignment is changed to 45° when comparing AY and AXY. Also, SXY falls in the intermediate, with z_0/b increasing due to a change in alignment when compared with S. However, this effect is less pronounced as λ_p increases or k_{rms}/k_a decreases.
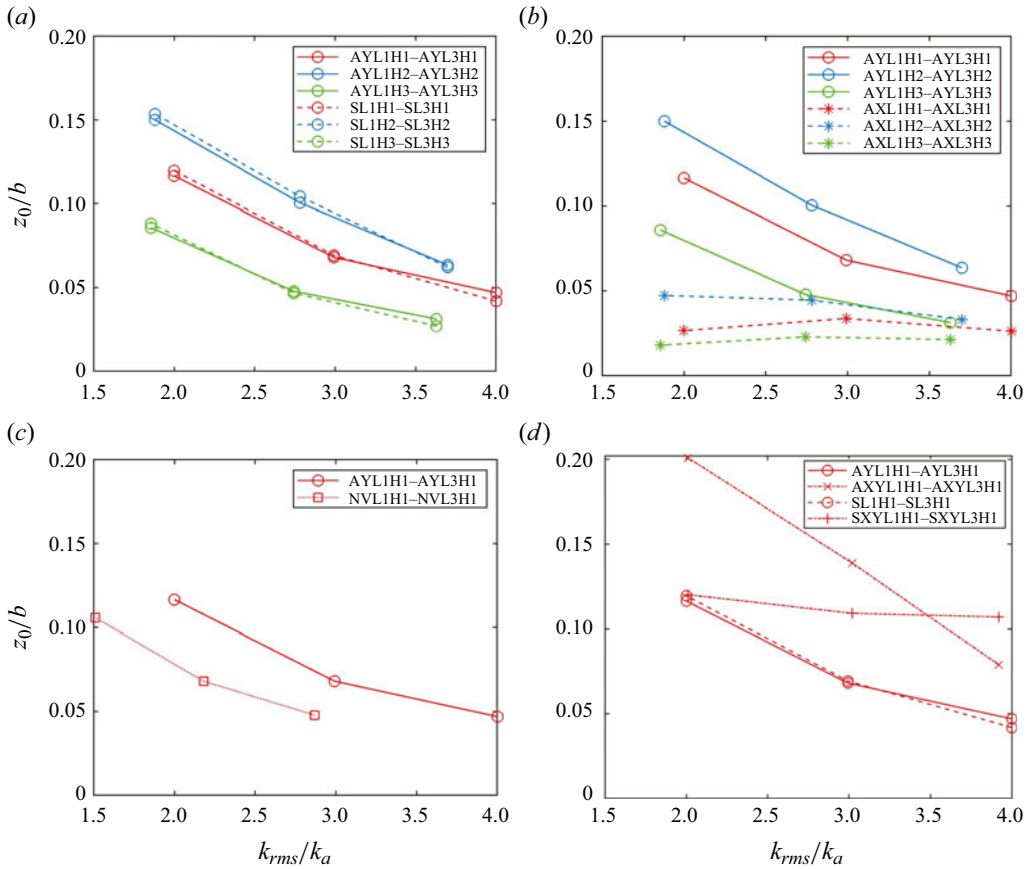
Figure 16. The effective roughness length to width ratio (z_0/b) for the various cases: (*a*) spanwise-aligned versus staggered; (*b*) spanwise-aligned versus streamwise-aligned; (*c*) spanwise-aligned with and without valleys; (*d*) spanwise-aligned/staggered and their 45° rotated equivalents.

3.4. *Reynolds and dispersive stresses*

We present the Reynolds stress and dispersive stress results. These results are usually not relevant to roughness modelling but are fundamental aspects of the flow, and we present them here for completeness.

Figure 17 shows the normal components of the horizontally averaged Reynolds and dispersive stresses for the five spanwise-aligned (AYL1H1, AYL2H1, AYL3H1, AYL1H2, AYL1H3) and staggered (SL1H1, SL2H1, SL3H1, SL1H2, SL1H3) cases. At first glance, all stresses are negligible in the valleys below the surface at $z = 0$, and the spanwise and wall-normal components of the dispersive stresses are small throughout the channel. Second, it can be observed that the corresponding streamwise components of both stresses (uu) are higher than the respective spanwise (vv) and wall-normal (ww) components for all cases. Third, the maximum value for dispersive stress lies within the roughness-occupied layer, whereas for the Reynolds stress, the maxima lie in the neighbourhood of the roughness peak. Moreover, the dispersive stress spike for staggered cases is slightly higher than for the aligned ones. This spike also decreases with packing density for this set of cases. This can be attributed to the reduced inhomogeneity in the mean velocity field as the spacing between cubes reduces. Comparing the zero skewed surface (figure 17*a*) with the corresponding positively skewed (figure 17*d*) and negatively skewed (figure 17*e*) ones,
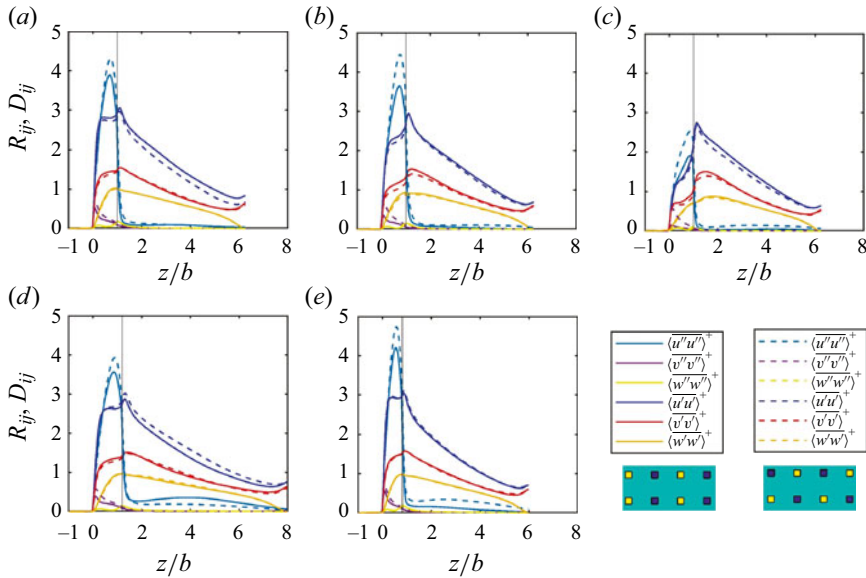
Figure 17. Reynolds $\langle \overline{u_i' u_j'} \rangle^+$ and dispersive $\langle \overline{u_i'' u_j''} \rangle^+$ stress comparisons for spanwise-aligned and staggered cases. The solid and dashed lines depict aligned and staggered arrangements, respectively. Cases: (*a*) AYL1H1 and SL1H1; (*b*) AYL2H1 and SL2H1; (*c*) AYL3H1 and SL3H1; (*d*) AYL1H2 and SL1H2; and (*e*) AYL1H3 and SL1H3. The vertical line indicates peak roughness height.

it can be seen that the dispersive stress spike gets narrower and increases with decreasing *Sk*. The narrowing is an indirect consequence of the decrease in roughness height (h_1/b) as *Sk* decreases. The Reynolds stresses, on the other hand, are similar irrespective of packing density or skewness.

Similarly, plots comparing the stresses for spanwise-aligned (AY) and streamwise-aligned (AX) surfaces are shown in figure 18. The higher magnitudes of dispersive stress for AX surfaces are due to the presence of high momentum pathways between two rows of cubes and low momentum pathways over cube locations. Unlike in figure 17, the dispersive stress spike here increases with packing density for the AX surfaces. To further probe these observations, the streamwise mean velocity contours for AYL1H1, AYL2H1, AXL1H1 and AXL2H1 near the roughness peaks are shown in figures 19(*a*), 19(*b*), 19(*c*) and 19(*d*) respectively. Indeed, we see more inhomogeneity in the mean flow in figures 19(*c*) and 19(*d*) due to the overlapping wake regions and the high momentum pathways formed in between the cubes. With respect to skewness, the increasing trend of dispersive stress with decreasing skewness for AX surfaces is similar to that observed in AY surfaces.

The normal stress comparisons in figure 20 show slight variations when valleys are removed from AYL1H1, AYL2H1 and AYL3H1. Note that the NV surfaces have very different roughness statistics in k_{rms}/k_a, *Sk*, *Ku* and *ES*, but very similar Reynolds and dispersive stresses (figure 20), mean velocity profiles (figure 11) and z_0/b (figure 16*c*) with respect to the AY and S counterparts.

In figures 17, 18 and 20, it can be observed that the streamwise dispersive component is still measurable near the half-channel height. These effects hint at the presence of secondary flows within these domains. Figure 21 shows the mean streamwise averaged contours of $\bar{u}^+$ with secondary flow $\bar{v}^+$ and $\bar{w}^+$ for a few cases. One can observe small but non-zero secondary flow near the half-height, and noticeable counter-rotating vortical
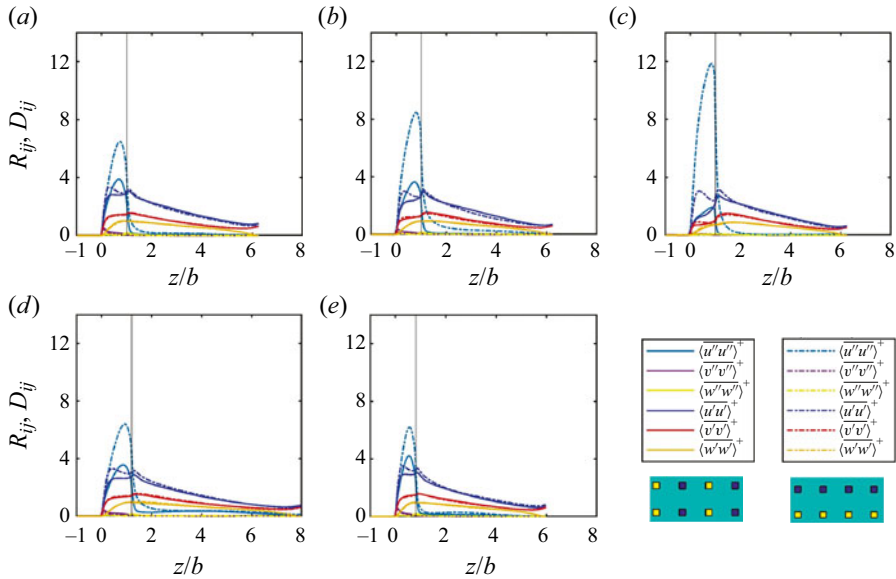
Figure 18. Reynolds $\overline{\langle u_i' u_j' \rangle}^+$ and dispersive $\langle \overline{u_i'' u_j''} \rangle^+$ stress comparisons for spanwise-aligned and streamwise-aligned cases. The solid and dashed lines depict spanwise-aligned and streamwise-aligned arrangements, respectively, for (*a*) AYL1H1 and AXL1H1, (*b*) AYL2H1 and AXL2H1, (*c*) AYL3H1 and AXL3H1, (*d*) AYL1H2 and AXL1H2, and (*e*) AYL1H3 and AXL1H3. The vertical line indicates peak roughness height.
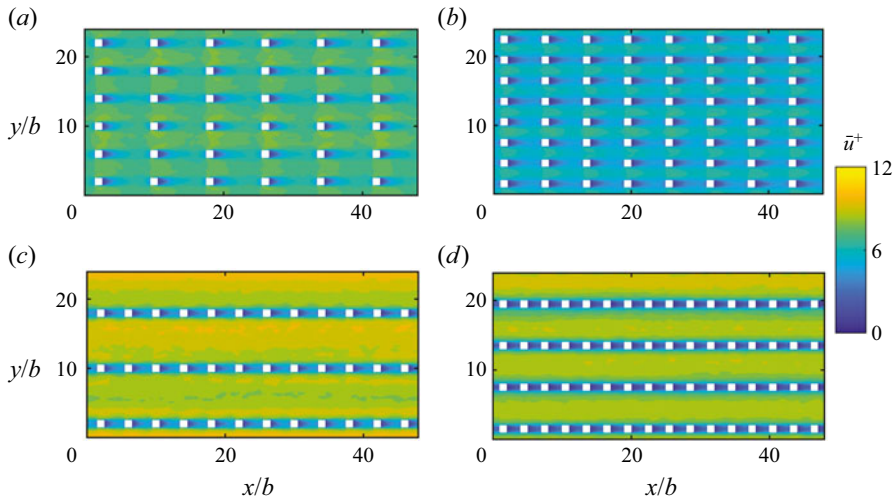

Figure 19. Streamwise mean velocity contours at $z/b \approx 1.0$ for cases (*a*) AYL1H1, (*b*) AYL2H1, (*c*) AXL1H1 and (*d*) AXL2H1.

structures near the roughness peaks. In most cases, the secondary structures can be seen to penetrate beyond the roughness sublayer, which is in alignment with other studies such as Vanderwel & Ganapathisubramani (2015).

4. Rough-wall modelling

It should be clear from the results presented in § 3 that the roughness regime studied here, while previously unexplored, has no unexpected behaviours. In this section, we discuss
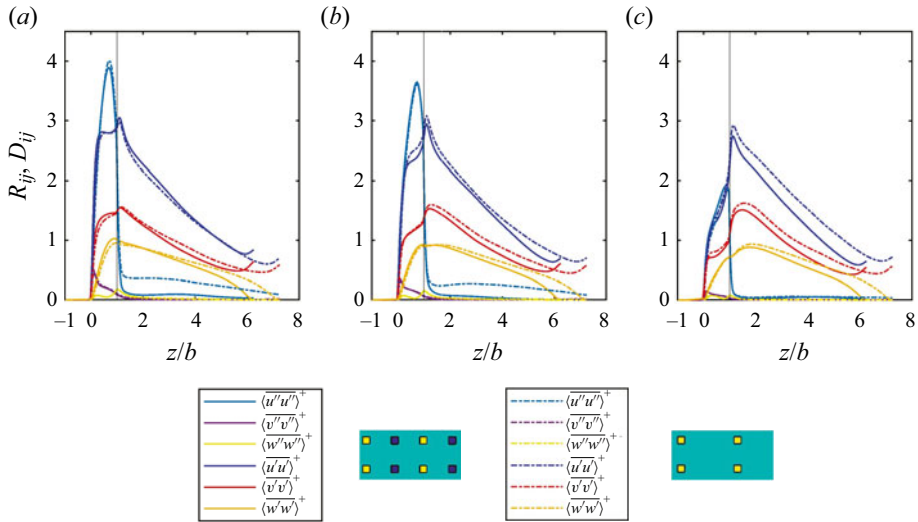
Figure 20. Reynolds $\overline{\langle u_i' u_j' \rangle}^+$ and dispersive $\langle \overline{u_i'' u_j''} \rangle^+$ stress comparisons for spanwise-aligned cases with and without valleys: (*a*) AYL1H1/NVL1H1, (*b*) AYL2H1/NVL2H1 and (*c*) AYL3H1/NVL3H1. The solid and dashed lines depict spanwise-aligned and no-valleys cases, respectively. The vertical line indicates peak roughness height.
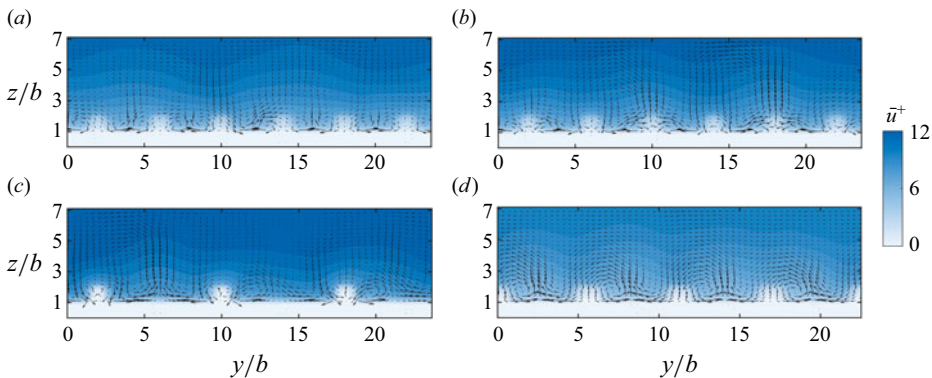

Figure 21. Mean streamwise-averaged constant *x* plane showcasing $\bar{u}^+$ contours with secondary flows $\bar{v}^+$ and $\bar{w}^+$ (denoted by vectors) within the domain in (*a*) AYL1H1, (*b*) SL1H1, (*c*) AXYL1H1 and (*d*) SXYL1H1.

the implications that the data have on rough-wall modelling. In addition to the DNS data here, we also make use of the data in the roughness database (https://roughnessdatabase. org/). While it will be clear later in this section, the task of rough-wall modelling is the task of identifying roughness statistics, such as k_{rms} and *Sk*, that distinguish all rough surfaces in the calibration set. As it is highly unlikely that a finite set of roughness statistics can distinguish all rough surfaces, it is improbable that a universal rough-wall model that perfectly captures all rough surfaces can be formulated. Hence, instead of a universal rough wall model, it is more effective to seek rough-wall models with clearly defined applicable ranges.

4.1. *The task of roughness modelling*

In this subsection, we try to answer the following question: what is the nature of roughness modelling? We begin by comparing the z_0 magnitudes in table 3 to the estimates provided by the existing rough-wall correlations. Here, the correlations in Forooghi *et al.* (2017) and Flack *et al.* (2020) are taken as examples. The two correlations are given in (4.1) and (4.2) respectively:

$$k_s/k_t = 1.07 \times (1 - e^{-3.5\,ES})(0.67Sk^2 + 0.93Sk + 1.3), \tag{4.1}$$

$$k_s = \begin{cases} 2.11k_{rms}, & Sk = 0, \\ 2.48k_{rms}(1 + Sk)^{2.24}, & Sk > 0, \\ 2.73k_{rms}(2 + Sk)^{-0.45}, & Sk < 0, \end{cases} \tag{4.2}$$

where roughness statistics k_{rms}, Sk, k_t and ES are invoked as inputs to model k_s. We note that the effective roughness height (z_0) in table 3 is linked directly to k_s by the expression

$$k_s = z_0\,e^{(\kappa A)}, \tag{4.3}$$

where the von Kármán constant is $\kappa = 0.4$, and $A = 8.5$.

Figure 22 shows the predicted k_s for the rough surfaces in this study using the aforementioned correlations. We see that the surfaces in the present study do not fit into these existing correlations. Further, we see from figure 22(*a*) that surfaces with different arrangements (AX, AY, S), although having very different measured k_s, produce the same estimates according to (4.1) as they have the same k_t, Sk and ES (as indicated by the dashed lines). Similarly, in figure 22(*b*), the surfaces AXY and SXY, although having very different measured k_s, join the other arrangements with identical k_{rms} and Sk, resulting in identical k_s predictions according to (4.2). These observations indicate the inherent lack of roughness statistics that distinguish roughness surfaces in these correlations. Although the discussion here is limited to the correlations in Forooghi *et al.* (2017) and Flack *et al.* (2020), the same inadequacies are prevalent in other empirical correlations as well, as observed by Yang *et al.* (2023).

A remedy, as advocated and explained in Chung *et al.* (2021) and other references cited therein, is to expand the list of inputs and include other roughness statistics. Here, statistics that help to distinguish different arrangements of elements are needed. Among other statistics, two-point correlations are the simplest ones for this purpose. The conventional definition of correlation length measures the distance at which the autocorrelation, i.e.

$$C_x(\Delta x) = \frac{1}{L_x L_y k_{rms}^2} \int_{L_y} \int_{L_x} (k(x, y) - \bar{k})(k(x + \Delta x, y) - \bar{k})\, \mathrm{d}x\, \mathrm{d}y, \tag{4.4}$$

$$C_y(\Delta y) = \frac{1}{L_x L_y k_{rms}^2} \int_{L_y} \int_{L_x} (k(x, y) - \bar{k})(k(x, y + \Delta y) - \bar{k})\, \mathrm{d}x\, \mathrm{d}y, \tag{4.5}$$

drops to a certain value, typically $1/e$. Here, $k(x, y)$ contains the height information for the rough surface at location (x, y). In other words, the correlation lengths Rl_x and Rl_y are such that

$$C_x(Rl_x) = 1/e, \tag{4.6}$$

$$C_y(Rl_y) = 1/e. \tag{4.7}$$

The above definitions are suitable for irregular surfaces, but when applied to regular (e.g. cuboidal) roughness, thus-defined correlations simply give the roughness
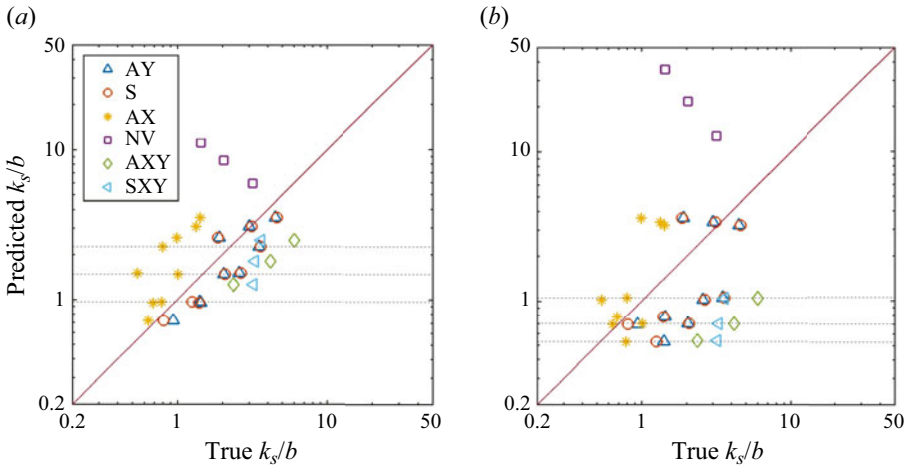
999 A78-24

Figure 22. Existing correlations by (*a*) Forooghi *et al.* (2017) and (*b*) Flack *et al.* (2020), applied to surface parameters obtained from the current study. The plots show the predicted equivalent sand-grain roughness height k_s normalized by cube width b.
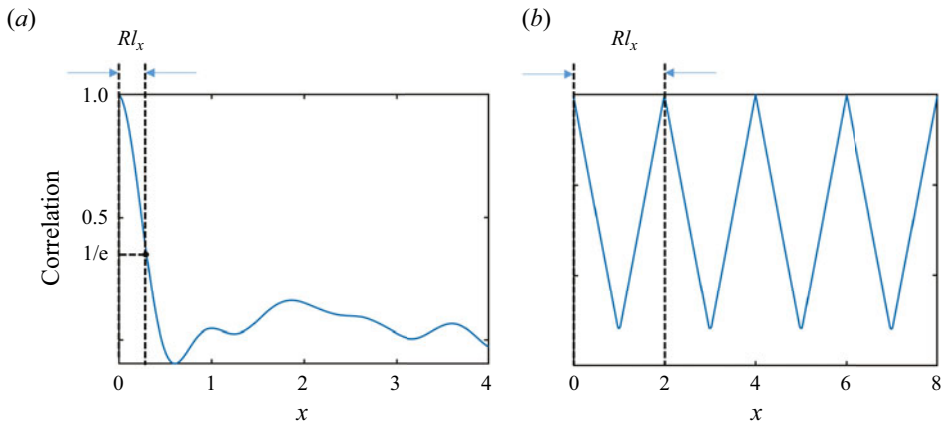

Figure 23. An illustration of the definition of correlation length for (*a*) an irregular surface and (*b*) a regular surface.

element width. Intuitively, we need a distance measure that is representative of the inter-repeating-tile length. This can be made possible by choosing the peak to peak distance in the correlation plot. Figure 23 demonstrates these definitions, with an irregular surface A0020 from Forooghi *et al.* (2017), and a regular surface R25 from Xu *et al.* (2021) as examples. Extending them to both streamwise and spanwise directions, one can distinguish between staggered and aligned, as well as surfaces with different arrangements.

Having defined the correlation lengths, it is now possible to distinguish between the rough surfaces involved in this study. We demonstrate this in figure 24, where we look at 27 surfaces from our DNS study, namely AYLiHj, SLiHj and AXLiHj, with $i, j = 1, 2, 3$. The number of input roughness statistics is progressively increased so that figure 24(*a*) contains k_{rms}/k_a and Sk, and figure 24(*b*) contains k_{rms}/k_a, Sk and Rl_x/Rl_y. We observe that the data points overlap in figure 24(*a*) as AY/AX/SLiHj occupy the same location in the space, but in figure 24(*b*), the rough surfaces do not overlap. Now, if one were to
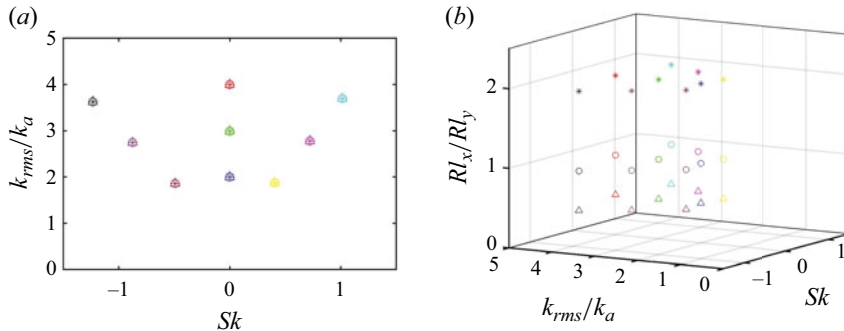
Figure 24. Rough surfaces in two different parameter spaces: (*a*) k_{rms}/k_a and Sk; (*b*) k_{rms}/k_a, Sk and Rl_x/Rl_y.

build a rough-wall model using only Sk and k_{rms} as its inputs, then one would be seeking different outputs from the same input, which would inevitably result in errors, as observed in figure 22. However, if Sk, k_{rms} and Rl_x/Rl_y are invoked as inputs, then one would be seeking a single-valued function that maps from the input roughness statistics space to the output k_s space. Following this line of thinking, we argue that the task of roughness modelling can be reduced to identifying a set of roughness statistics that distinguishes the rough surfaces in question.

This conclusion has many implications for rough-wall modelling. First, considering that there is no finite set of statistics that allows unique identification of all rough surfaces, i.e. two-dimensional functions $k(x, y)$, constructing a universal roughness-statistics-based rough-wall model could be a very difficult problem to solve. Consequently, all roughness-statistics-based rough-wall models are bound to have predictive power for surfaces that are closely registered around its calibration dataset only. Second, roughness modelling must consider the calibration dataset, and a universal list of roughness statistics for all rough-wall models is highly unlikely. One can argue that certain roughness types are more important than others, therefore certain roughness statistics are more important than others, but such judgements have to be subjective.

4.2. *Selecting roughness statistics*

Considering that all rough-wall models will have their applicable range, selecting inputs to the rough-wall models is a highly relevant issue. In this subsection, we discuss how one might go about selecting the roughness statistics that distinguish the rough surfaces in a given group of rough surfaces.

Without loss of generality, we limit the discussion to the seven roughness statistics in table 3. Figure 25 is a representation of the 36 rough surfaces in the seven-dimensional roughness statistics space represented by k_{rms}/k_a, Sk, ES, Rl_x/k_a, Rl_y/k_a, Ku and k_t/k_a. In the figure, every roughness statistic corresponds to a vertical axis, and every rough surface corresponds to a line in the figure. We will make use of this plot to select the roughness statistics that are most relevant to the modelling of the rough surfaces in this study. We can, of course, use all roughness statistics. The objective, however, is to identify as few roughness statistics as possible such that the identified roughness statistics distinguish the rough surfaces in question. It is also worth noting that for random roughness, k_t might not be a suitable parameter to represent the surface as it tends to increase as the sampling size gets larger, instead of converging.
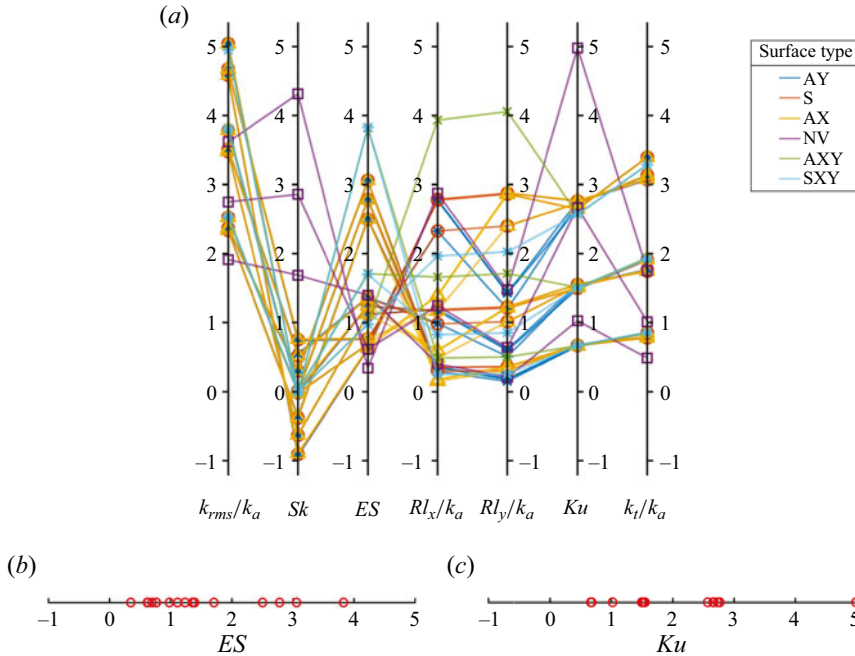
999 A78-26

Figure 25. High-dimensional parallel plots of the surface parameter space for (*a*) 36 surfaces in the current study, (*b*) *ES* parameter space, and (*c*) *Ku* parameter space. The values shown are scaled by the standard deviation.

First, we make the following observation. For a roughness statistic s that represents a vertical axis in figure 25, rough surfaces that do not overlap in the axis can readily be distinguished by that statistic. For example, consider two surfaces AYL1H1 and SL1H1. All features excluding Rl_y/k_a would remain the same for both these surfaces, hence this parameter would serve as the distinguishing feature. However, depending on how accurately one could measure the features, the features need not necessarily be identical for them to be overlapping. We may mathematically define 'overlap' as $|s_i - s_j| < c\,\mathrm{STD}_s$. That is, if s_i and s_j are close, then they are considered indistinguishable or overlap. Here, s_i is the roughness statistic of the ith rough surface, STD_s is the standard deviation of statistic s in the rough surfaces, and c is a constant. Again, the choice of c will depend on how well one can measure the roughness statistic in question: if one can measure the statistic with precision $0.01\mathrm{STD}_s$, then c can take value 0.02, i.e. $+0.01 - (-0.01)$. Here, we set $c = 0.1$, which should be a very conservative choice. Following this line of thinking, a roughness statistic that distinguishes more rough surfaces is more relevant to the modelling of the rough surfaces in question. Take the statistics and the rough surfaces in figure 25(*a*) as our illustrative example. Kurtosis is not an extremely relevant roughness statistic for the modelling of these rough surfaces as it distinguishes only five groups of rough surfaces (see figure 25*c*). On the other hand, from figure 25(*b*) it can be observed that ES is a more relevant roughness statistic as it distinguishes around 12 groups of rough surfaces. In other words, the more the spread in the feature space, the higher its differentiating ability. Note that surfaces indistinguishable by a parameter may be indistinguishable in terms of drag. In such cases, the model would remain accurate even if it does not distinguish the two surfaces.

Figure 26. Feature selection procedure illustration using rough surface indices. Each number here is representative of a rough surface: I denotes the initial stage involving all surfaces in the group; II denotes the next stage after one feature Rl_x/k_a was selected; and III denotes the third stage when ES is selected.

Following the discussion above, the process of selecting roughness statistics that distinguish rough surfaces in a given group of rough surfaces can follow a recursive greedy algorithm. To illustrate the algorithm, we take our current study as an example. For this group of rough surfaces, we have 36 surfaces. We will represent each rough surface by a number, and the 36 rough surfaces are numbered from 1 to 36. First, we will split each feature into bins of size 0.1 STD, and the number of surfaces present in each bin is counted. The feature with the greatest number of non-zero bins is the most distinguishing feature. For the 36 rough surfaces, Rl_x is found to be that feature. If a surface is the only surface in one of the Rl_x bins, then that surface is readily distinguishable by Rl_x. However, if there are a couple of rough surfaces in one Rl_x bin, then these surfaces, which we call a cluster, cannot be distinguished by Rl_x, and we must invoke another roughness statistic to distinguish them. From figure 26, it can be seen that there are 10 such clusters, with one cluster containing 6 surfaces, three clusters containing 4 surfaces each, one cluster containing 3 surfaces, and five clusters containing 2 surfaces each. The purpose of the next stage is to find the next feature that best distinguishes these surfaces. A similar procedure is repeated for this stage for each cluster, and ES is identified as the next most distinguishing feature. The process continues until either no clusters emerge or all features have been selected. In this case, this stopping criterion is attained after the third roughness statistics, Rl_y, is selected. Figure 26 illustrates the procedure.

We can apply this recursive greedy algorithm to other groups of surfaces. Table 4 shows some of these examples, where this feature selection procedure is applied to the rough surfaces from Jouybari *et al.* (2021), Womack *et al.* (2022), Xu *et al.* (2021) and the current study. Although we do not consider all surface geometries and other roughness parameters such as spanwise effective slope ES_z and surface porosity P_0 mentioned in Jouybari *et al.* (2021), it is interesting to note that we obtain one variable as the distinguishing feature within each of the three pairs (E_x, E_z), (P_0, Sk) and (k_{rms}, Ku) that contain strong

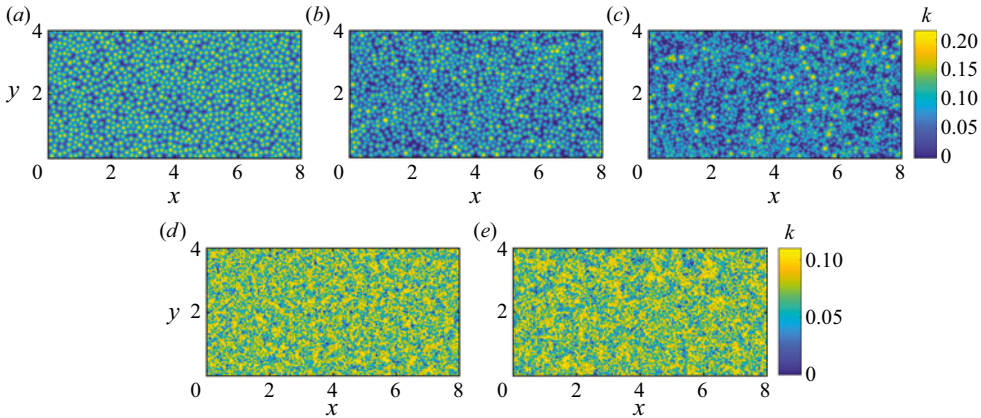| Group	$n_{surfaces}$	Selected features
Jouybari *et al.* (2021)	25	Sk, ES, k_{rms}/k_a
Womack *et al.* (2022)	16	Sk, k_{rms}/k_a, Rl_x/k_a, ES
Xu *et al.* (2021)	6	k_{rms}/k_a
Current study	36	Rl_x/k_a, Rl_y/k_a, ES

Table 4. Features selected for different groups.

Figure 27. Examples of surfaces with very similar statistics when considering the seven roughness statistics in table 3. The surfaces (*a*) A1588, (*b*) A3588 and (*c*) A7088 from Forooghi *et al.* (2017) belong to one group. The surfaces (*d*) N14 and (*e*) N18 from Yang *et al.* (2022) belong to another group.

correlations in their work, and the algorithm does not pick up two strongly correlated features. In the six rough surfaces in Xu *et al.* (2021), where the primary varying parameter is the planar packing density of the cubes, just one parameter, k_{rms}/k_a in this case, seems to be sufficient. An interesting observation is that k_{rms}/k_a is not always a distinguishing roughness statistic. For the surfaces with multiple roughness arrangements, such as in our study, the dimensionless two-point correlation lengths Rl_x/k_a and Rl_y/k_a emerge as having better distinguishing capability.

A scenario might occur where this feature selection methodology, when extended to a larger set of surfaces, would exhaust all features. In this case, certain surfaces are indistinguishable from the available roughness statistics, yet they have different k_s. For example, figure 27 depicts two clusters of such surfaces from Forooghi *et al.* (2017) and Yang *et al.* (2022), taken from a larger set of about 143 rough surfaces in the aforementioned roughness database. The three rough surfaces in figures 27(*a*)–27(*c*) from Forooghi *et al.* (2017) are referred to as A3588, A7088 and A1588. These are indistinguishable from the seven roughness statistics in table 3. The k_s/k_a magnitudes in surfaces A1588, A3588 and A7088 from Forooghi *et al.* (2017) stand at approximately 11.1, 9.98 and 8.86, respectively, which are different, whereas the deviations in the aforementioned seven statistics are less than 0.1 standard deviation (STD) for these surfaces. The same applies to surfaces N14 and N18 from Yang *et al.* (2022) whose k_s/k_a magnitudes are close to 5.15 and 4.57, respectively.

4.3. *Building a rough-wall model*

Having identified the important roughness statistics required to build a rough-wall model, this subsection discusses the next step involved, which is developing rough-wall models themselves. This is quite straightforward, and there are a number of approaches. Here, we present two such approaches: a multi-variate polynomial regression (MPR), which gives interpretable explicit expressions but has less descriptive power, and a feedforward neural network (FNN) based approach, which has great descriptive power but is a black box.

For the MPR model, the following steps were used to generate the polynomial terms: (i) minmax scaling and polynomial feature generation; (ii) determination of the degree and the polynomial features involved; and (iii) linear regression of the polynomial features. The regression follows an iterative bagging-based approach. At every iteration, the training dataset is randomly selected maintaining a test–train split of 0.7, and m random polynomial terms are selected to develop a regression fit. Ridge regression is employed here, and the mean-squared test and training errors are computed. This step is repeated iteratively for different polynomials with the selected degree d and number m of polynomial features until we find a fit where the minimum in both training and test datasets occurs. These fits for the groups Jouybari *et al.* (2021), Womack *et al.* (2022), Xu *et al.* (2021) and the current study are given by expressions (4.8), (4.9), (4.10) and (4.11), respectively (the reader is directed to Kleinbaum *et al.* (2013) for further details of MPR):

$$
\begin{aligned}
k_s/k_a = {} & -0.53(k_{rms}/k_a)^3 - 0.26\,Sk^3 + 0.69\,ES^3 + 0.78(k_{rms}/k_a)^2\,Sk + 2.22\,Sk^2\,ES \\
& + 2.45(k_{rms}/k_a)^2\,ES + 1.44(k_{rms}/k_a)\,Sk^2 - 0.85\,Sk\,ES^2 \\
& + 2.86(k_{rms}/k_a)\,Sk\,ES - 0.90(k_{rms}/k_a)^2 - 1.51\,ES^2 + 0.5\,Sk\,ES \\
& - 0.44(k_{rms}/k_a)\,Sk + 0.65\,ES - 0.12 k_{rms}/k_a,
\end{aligned} \tag{4.8}
$$

$$
k_s/k_a = 0.95\,Sk^2 + 0.77\,ES\,Rl_x/k_a - 2.63\,ES + 1.79\,Sk - 2.58 k_{rms}/k_a, \tag{4.9}
$$

$$
k_s/k_a = 0.88(k_{rms}/k_a)^2 + 0.13 k_{rms}/k_a, \tag{4.10}
$$

$$
\begin{aligned}
k_s/k_a = {} & 0.03\,ES^2 - 1.19(Rl_x/k_a)^2 + 1.00(Rl_y/k_a)^2 - 0.08(Rl_x/k_a)(Rl_y/k_a) \\
& + 0.06\,ES + 1.71\,Rl_x/k_a - 1.41\,Rl_y/k_a.
\end{aligned} \tag{4.11}
$$

Figure 28 shows the true versus predicted plots for the various groups, with R^2 standing for the coefficient of determination, which is a measure of accuracy. The R^2 values shown in figure 28 have been calculated from the training data, holding values 0.98, 0.98, 0.99 and 0.85 for the respective groups. The corresponding R^2 values based on the test data stand as 0.77, 0.99, 0.99 and 0.64. It is possible to choose polynomial forms of higher R^2 on the training data, but we may risk overfitting. Hence these models are adopted, maintaining a balance between generalization and accuracy. It is further noted that the errors shown in k_s predictions here would generate much smaller errors in drag prediction, as the skin-friction coefficient C_f is proportional to $[\log(k_s^+)]^2$.

For the FNN model, a standard three-layer network has been trained with 10 neurons in each layer. The size of the neural network is small, but as we will see, it is sufficient. Figure 29 shows the parity plots for these networks. Here, each FNN utilizes 100 % of its group for its training data. This is done to show the perfect fitting of the data in each case, which demonstrates the effectiveness of our feature selection strategy and the strong descriptive power of FNNs.
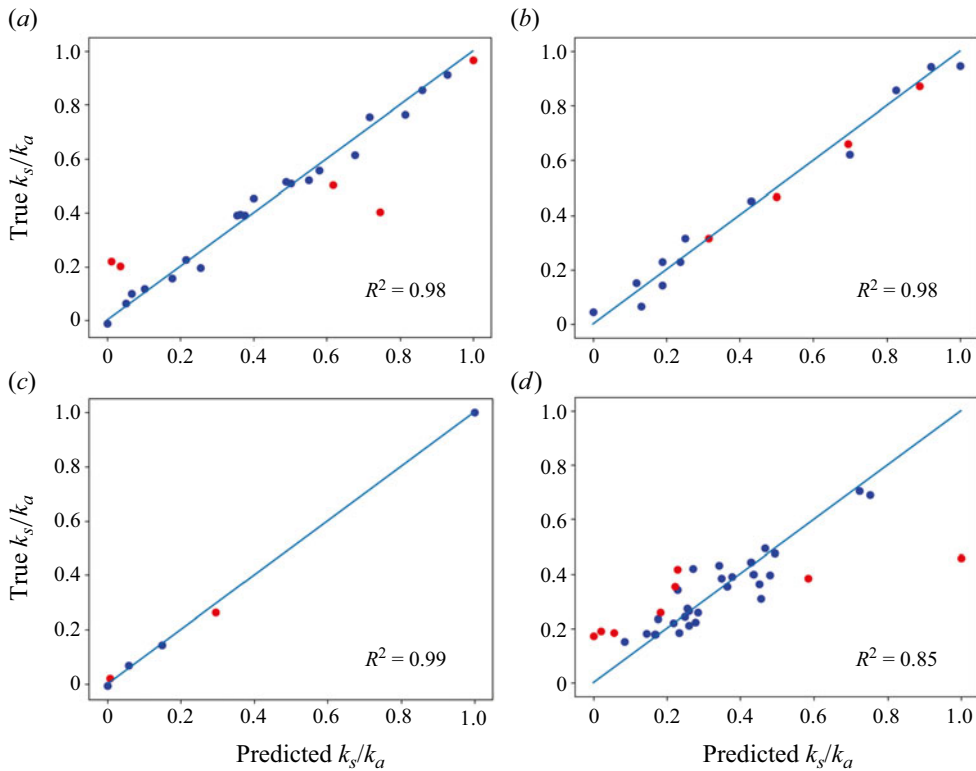
999 A78-30

Figure 28. The MPR-based regression for groups: (*a*) Jouybari *et al.* (2021), (*b*) Womack *et al.* (2022), (*c*) Xu *et al.* (2021), and (*d*) current study. Here, R^2 stands for the coefficient of determination and is calculated on the training dataset. The red colour stands for test data, and the blue stands for training data. The k_s/k_a values are normalised using a minmax scaler.

5. Concluding remarks

In conclusion, this study presents new data for rough surfaces in the relatively low skewness and high k_{rms}/k_a space that have not been explored before. The variations in equivalent roughness height, mean velocity profiles and Reynolds and dispersive stresses with changes in skewness (*Sk*), planar packing density (λ_p) and roughness arrangement have been reported. While the first two have been investigated in various studies, the latter effect of roughness arrangement is rarely considered. In this work, this effect of roughness arrangement is not only considered but has been found to significantly affect the drag produced by rough surfaces. This can be confirmed by the trends in the mean velocity profile. In general, higher mean velocity profiles are observed for lower magnitudes of *Sk* and λ_p (except in AX surfaces where flow-sheltering is observed at higher λ_p). For variations in roughness arrangement, the AXY and SXY orientations seem to produce more drag than the S and AY arrangements, followed by the AX arrangement. The valleys have been found to produce minimal, but non-zero contributions to drag, most of which we posit to stem from the form drag. In the context of second-order flow statistics, the spatial flow inhomogeneity, as indicated by dispersive stresses, is observed to be more in AX surfaces as compared to AY surfaces. Secondary flows in the spanwise and wall-normal directions were also observed in the flow over these rough channels.
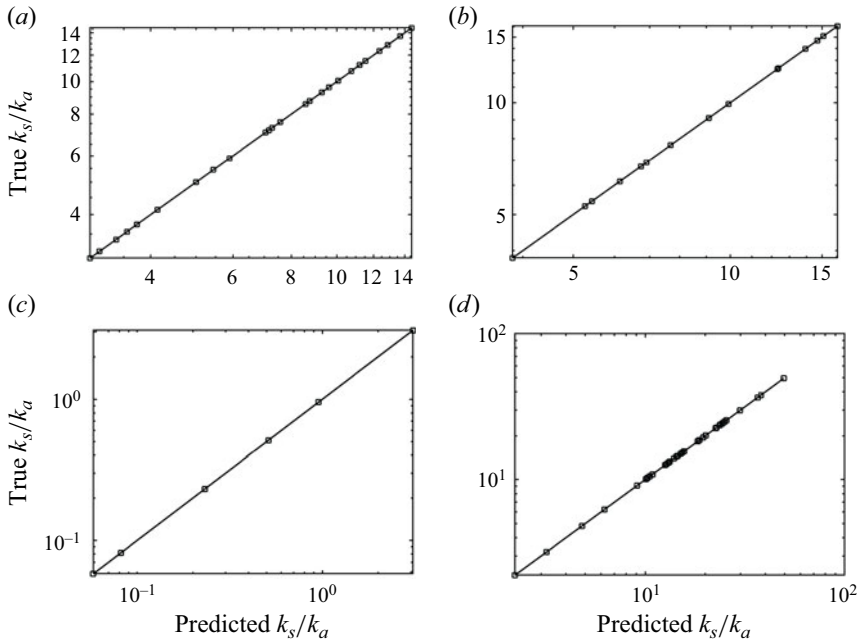
Figure 29. The FNN-based regression for groups: (*a*) Jouybari *et al.* (2021), (*b*) Womack *et al.* (2022), (*c*) Xu *et al.* (2021), and (*d*) current study.

It is important to note that the effect of roughness element shape on the observed turbulence and drag characteristics has not been pursued in this study. The choice of the cube-roughness geometry here is based on its simplicity to construct such surfaces in the k_{rms}/k_a–Sk parameter space and the ability to align the immersed solid boundaries with the Cartesian mesh for most cases in this study. While it is difficult to discuss the potential impact of roughness shape on the flow, the underlying assumption of the statistics-based rough-wall modelling approach is that the overall drag will be similar, given that the statistics are similar irrespective of the roughness element shape.

Since a majority of rough surfaces in the current study vary by rough-element arrangement, these are bound to have identical moments of roughness height. Two-point spatial correlation lengths have been proposed as input parameters to differentiate between these arrangements. Additionally, we put forward the argument that based on the properties of the rough surfaces in the present dataset, finding a universal roughness correlation is an exacting task and that a 'case-by-case' basis would do better. We assert that roughness modelling can be viewed as determining input parameters that best distinguish the rough surfaces. We demonstrate a methodology to identify these specific features for a given group of rough surfaces. It is important to note that these distinguishing features vary based on the group being considered. While it might be useful to link a particular roughness feature as being important for a specific type of roughness, we would like to point out that such links may not necessarily find a physical basis, but rather find a statistical one.

Having identified the important roughness features, we demonstrate both MPR- and FNN-based approaches to build rough-wall models from the selected features. The selected features can be found to be robust, given the goodness of fit observed in the models built from them.

999 A78-32

Author ORCIDs.

- Shyam S. Nair https://orcid.org/0009-0009-9284-3075;
- Vishal A. Wadhai https://orcid.org/0000-0003-0521-9784;
- Robert F. Kunz https://orcid.org/0000-0001-9504-1945;
- Xiang I.A. Yang https://orcid.org/0000-0003-4940-5976.

REFERENCES

ABKAR, M. & PORTÉ-AGEL, F. 2012 A new boundary condition for large-eddy simulation of boundary-layer flow over surface roughness transitions. *J. Turbul.* **13**, N23.

ALTLAND, S., ZHU, X., MCCLAIN, S., KUNZ, R. & YANG, X. 2022 Flow in additively manufactured super-rough channels. *Flow* **2**, E19.

ANDERSON, W. 2013 An immersed boundary method wall model for high-Reynolds-number channel flow over complex topography. *Intl J. Numer. Meth. Fluids* **71** (12), 1588–1608.

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.

BOU-ZEID, E., MENEVEAU, C. & PARLANGE, M. 2005 A scale-dependent Lagrangian dynamic model for large eddy simulation of complex turbulent flows. *Phys. Fluids* **17** (2), 025105.

BUSSE, A. & JELLY, T.O. 2020 Influence of surface anisotropy on turbulent flow over irregular roughness. *Flow Turbul. Combust.* **104**, 331–354.

CASTRO, I.P., CHENG, H. & REYNOLDS, R. 2006 Turbulence over urban-type roughness: deductions from wind-tunnel measurements. *Boundary-Layer Meteorol.* **118**, 109–131.

CHENG, H., HAYDEN, P., ROBINS, A.G. & CASTRO, I.P. 2007 Flow over cube arrays of different packing densities. *J. Wind Engng Ind. Aerodyn.* **95** (8), 715–740.

CHOI, H. & MOIN, P. 2012 Grid-point requirements for large eddy simulation: Chapman's estimates revisited. *Phys. Fluids* **24** (1), 011702.

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.

COCEAL, O., THOMAS, T.G., CASTRO, I.P. & BELCHER, S.E. 2006 Mean flow and turbulence statistics over groups of urban-like cubical obstacles. *Boundary-Layer Meteorol.* **121**, 491–519.

COLEBROOK, C.F. 1939 Turbulent flow in pipes with particular reference to the transition region between the smooth- and rough-pipe laws. *J. Inst. Civil Engrs* **11**, 133–156.

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**, 317–329.

FLACK, K.A., SCHULTZ, M.P., BARROS, J.M. & KIM, Y.C. 2016 Skin-friction behavior in the transitionally-rough regime. *Intl J. Heat Fluid Flow* **61**, 21–30.

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. 2018 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** (4), 044605.

JACKSON, P. 1981 On the displacement height in the logarithmic velocity profile. *J. Fluid Mech.* **111**, 15–25.

JELLY, T.O., RAMANI, A., NUGROHO, B., HUTCHINS, N. & BUSSE, A. 2022 Impact of spanwise effective slope upon rough-wall turbulent channel flow. *J. Fluid Mech.* **951**, A1.

JIMÉNEZ, J. 2004 Turbulent flows over rough walls. *Annu. Rev. Fluid Mech.* **36**, 173–196.

JOUYBARI, 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.

KLEINBAUM, D.G., KUPPER, L.L., NIZAM, A. & ROSENBERG, E.S. 2013 *Applied Regression Analysis and Other Multivariable Methods*. Cengage Learning.

LEE, J.H., SUNG, H.J. & KROGSTAD, P. 2011 Direct numerical simulation of the turbulent boundary layer over a cube-roughened wall. *J. Fluid Mech.* **669**, 397–431.

LEONARDI, S. & CASTRO, I.P. 2010 Channel flow over large cube roughness: a direct numerical simulation study. *J. Fluid Mech.* **651**, 519–539.

LOZANO-DURÁN, A. & JIMÉNEZ, J. 2014 Effect of the computational domain on direct simulations of turbulent channels up to $Re_\tau = 4200$. *Phys. Fluids* **26** (1), 011702.

MA, R., ALAMÉ, K. & MAHESH, K. 2021 Direct numerical simulation of turbulent channel flow over random rough surfaces. *J. Fluid Mech.* **908**, A40.

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.

MOODY, L.F. 1944 Friction factors for pipe flow. *Trans. ASME* **66** (8), 671–678.

NIKORA, V., BALLIO, F., COLEMAN, S. & POKRAJAC, D. 2013 Spatially averaged flows over mobile rough beds: definitions, averaging theorems, and conservation equations. *J. Hydraul. Engng* **139** (8), 803–811.

NIKURADSE, J. 1933 Laws of flow in rough pipes. *NACA Tech. Mem.* 1292.

O'LOUGHLIN, E.M. & MACDONALD, E.G. 1964 Some roughness-concentration effects on boundary resistance. *La Houille Blanche* **19** (7), 773–783.

PLACIDI, M. & GANAPATHISUBRAMANI, B. 2015 Effects of frontal and plan solidities on aerodynamic parameters and the roughness sublayer in turbulent boundary layers. *J. Fluid Mech.* **782**, 541–566.

RAUPACH, M.R., ANTONIA, R.A. & RAJAGOPALAN, S. 1991 Rough-wall turbulent boundary layers. *Appl. Mech. Rev.* **44** (1), 1–25.

RAUPACH, M.R. & SHAW, R.H. 1982 Averaging procedures for flow within vegetation canopies. *Boundary-Layer Meteorol.* **22** (1), 79–90.

SCHMID, M.F., LAWRENCE, G.A., PARLANGE, M.B. & GIOMETTO, M.G. 2019 Volume averaging for urban canopies. *Boundary-Layer Meteorol.* **173** (3), 349–372.

SCHULTZ, M.P. & FLACK, K.A. 2005 Outer layer similarity in fully rough turbulent boundary layers. *Exp. Fluids* **38**, 328–340.

SCHULTZ, M.P. & FLACK, K.A. 2007 The rough-wall turbulent boundary layer from the hydraulically smooth to the fully rough regime. *J. Fluid Mech.* **580**, 381–405.

THAKKAR, M., BUSSE, A. & SANDHAM, N. 2017 Surface correlations of hydrodynamic drag for transitionally rough engineering surfaces. *J. Turbul.* **18** (2), 138–169.

THAKKAR, M., BUSSE, A. & SANDHAM, N.D. 2018 Direct numerical simulation of turbulent channel flow over a surrogate for Nikuradse-type roughness. *J. Fluid Mech.* **837**, R1.

TOWNSEND, A.A. 1976 *The Structure of Turbulent Shear Flow*, 2nd edn. Cambridge University Press.

VANDERWEL, C. & GANAPATHISUBRAMANI, B. 2015 Effects of spanwise spacing on large-scale secondary flows in rough-wall turbulent boundary layers. *J. Fluid Mech.* **774**, R2.

VOLINO, R.J., SCHULTZ, M.P. & FLACK, K.A. 2011 Turbulence structure in boundary layers over periodic two- and three-dimensional roughness. *J. Fluid Mech.* **676**, 172–190.

WOMACK, K.M., VOLINO, R.J., MENEVEAU, C. & SCHULTZ, M.P. 2022 Turbulent boundary layer flow over regularly and irregularly arranged truncated cone surfaces. *J. Fluid Mech.* **933**, A38.

XU, H.H.A., ALTLAND, S.J., YANG, X.I.A. & KUNZ, R.F. 2021 Flow over closely packed cubical roughness. *J. Fluid Mech.* **920**, A37.

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, X.I.A. & GRIFFIN, K.P. 2021 Grid-point and time-step requirements for direct numerical simulation and large-eddy simulation. *Phys. Fluids* **33** (1), 015108.

YANG, X.I.A. & MENEVEAU, C. 2016 Large eddy simulations and parameterisation of roughness element orientation and flow direction effects in rough wall boundary layers. *J. Turbul.* **17** (11), 1072–1085.

YANG, X.I.A. & MENEVEAU, C. 2017 Modelling turbulent boundary layer flow over fractal-like multiscale terrain using large-eddy simulations and analytical tools. *Phil. Trans. R. Soc. Lond.* A **375** (2091), 20160098.

YANG, X.I.A., SADIQUE, J., MITTAL, R. & MENEVEAU, C. 2016 Exponential roughness layer and analytical model for turbulent boundary layer flow over rectangular-prism roughness elements. *J. Fluid Mech.* **789**, 127–165.

YANG, X.I.A., XU, H.H.A., HUANG, X.L.D. & GE, M.-W. 2019 Drag forces on sparsely packed cube arrays. *J. Fluid Mech.* **880**, 992–1019.

YANG, X.I.A., ZHANG, W., YUAN, J. & KUNZ, R.F. 2023 In search of a universal rough wall model. *Trans. ASME J. Fluids Engng* **145** (10), 101302.

YUAN, J. & PIOMELLI, U. 2014*a* Estimation and prediction of the roughness function on realistic surfaces. *J. Turbul.* **15** (6), 350–365.

YUAN, J. & PIOMELLI, U. 2014*b* Roughness effects on the Reynolds stress budgets in near-wall turbulence. *J. Fluid Mech.* **760**, R1.

ZHANG, W., ZHU, X., YANG, X.I.A. & WAN, M. 2022 Evidence for Raupach *et al.*'s mixing-layer analogy in deep homogeneous urban-canopy flows. *J. Fluid Mech.* **944**, A46.

ZHU, X., IUNGO, G.V., LEONARDI, S. & ANDERSON, W. 2017 Parametric study of urban-like topographic statistical moments relevant to *a priori* modelling of bulk aerodynamic parameters. *Boundary-Layer Meteorol.* **162** (2), 231–253.