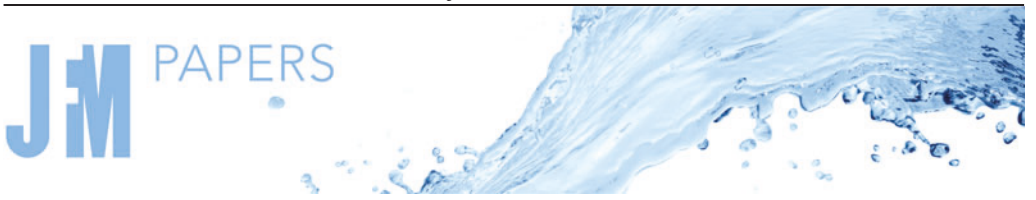
J. Fluid Mech. (2024), *vol*. 983, A9, doi:10.1017/jfm.2024.81



A machine learning model for reconstructing skin-friction drag over ocean surface waves

Kianoosh Yousefi[1,2,†], Gurpreet Singh Hora[2], Hongshuo Yang[3], Fabrice Veron[4] and Marco G. Giometto[2]

[1]Department of Mechanical Engineering, University of Texas at Dallas, Richardson, TX 75080, USA

[2]Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, NY 10027, USA

[3]Department of Computer Science, Columbia University, New York, NY 10027, USA

[4]School of Marine Science and Policy, University of Delaware, Newark, DE 19716, USA

(Received 18 August 2023; revised 4 December 2023; accepted 11 January 2024)

In order to improve the predictive abilities of weather and climate models, it is essential to understand the behaviour of wind stress at the ocean surface. Wind stress is contingent on small-scale interfacial dynamics typically not directly resolved in numerical models. Although skin friction contributes considerably to the total stress up to moderate wind speeds, it is notoriously challenging to measure and predict using physics-based approaches. This work proposes a supervised machine learning (ML) model that estimates the spatial distribution of the skin-friction drag over wind waves using solely wave elevation and wave age, which are relatively easy to acquire. The input–output pairs are high-resolution wave profiles and their corresponding surface viscous stresses collected from laboratory experiments. The ML model is built upon a convolutional neural network architecture that incorporates the Mish nonlinearity as its activation function. Results show that the model can accurately predict the overall distribution of viscous stresses; it captures the peak of viscous stress at/near the crest and its dramatic drop to almost null just past the crest in cases of intermittent airflow separation. The predicted area-aggregate skin friction is also in excellent agreement with the corresponding measurements. The proposed method offers a practical pathway for estimating both local and area-aggregate skin friction and can be easily integrated into existing numerical models for the study of air–sea interactions.

Key words: machine learning, surface gravity waves, wind-wave interactions

† Email address for correspondence: kyousefi@utdallas.edu

1. Introduction

Approximately two thirds of the surface of the Earth is covered by the ocean. The air–sea exchanges of mass, momentum and energy that happen over such a vast expanse play an integral role in determining the sea state, weather patterns and climate, thus significantly impacting many aspects of human life. Although we know that surface waves need to be fully integrated into weather and global climate forecast models (e.g. Melville 1996; Sullivan & McWilliams 2010), we do not yet fully understand the fundamental processes that couple the surface waves with the turbulent boundary layers above and below the ocean surface (e.g. Cavaleri *et al.* 2019; Buckley, Veron & Yousefi 2020; Geva & Shemer 2022). Indeed, the current parameterizations of air–sea exchange processes are limited, and most models still fall short of making accurate predictions of wave growth and weather events, particularly in strongly forced conditions.

Over the past few decades, many studies have investigated the influence of ocean surface waves on air–sea exchanges and, specifically, wind stress, defined as the sum of skin friction and form drag (e.g. Banner 1990; Belcher & Hunt 1993; Melville, Shear & Veron 1998; Sullivan, McWilliams & Moeng 2000; Melville, Veron & White 2002; Yang & Shen 2010; Veron, Melville & Lenain 2011; Grare *et al.* 2013; Grare, Lenain & Melville 2018; Hao & Shen 2019; Husain *et al.* 2019; Yousefi, Veron & Buckley 2020, 2021; Aiyer, Deike & Mueller 2023, among others). These studies have shown that the drag at the ocean surface strongly depends on environmental conditions and is a function of not only wind speed but also wave height, wave slope, wind–wave alignment and wave age, which is defined as the ratio of the wave phase speed to the reference wind speed measured at a height of 10 m above the air–water interface. Although some parameterizations exist to account for the impact of these parameters on surface stress (e.g. Fairall *et al.* 2003; Edson *et al.* 2013; Teixeira 2018; Porchetta *et al.* 2019; Husain, Hara & Sullivan 2022*a*,*b*), there is only a partial understanding of the dependencies of sea-surface drag on wave parameters. Therefore, determining these dependencies and developing sea-state-aware parametrizations of wind stress have remained one of the central foci of the wind-wave research community. Furthermore, the effects of surface waves and the corresponding generation of turbulence, airflow separation, spray and bubbles on the surface drag are not well understood. Irrespective of these findings, most operational models used to date still rely on simple bulk parameterization based on the Monin–Obukhov similarity theory (e.g. Hara & Belcher 2004; Hara & Sullivan 2015), which is only valid under equilibrium conditions and does not account for the multiscale nature of sea-surface roughness nor the sea-state dynamics. As such, these models often struggle to accurately predict the evolution of wave fields (e.g. Moon *et al.* 2004*b*; Moon, Ginis & Hara 2004*a*; Donelan *et al.* 2012; Cronin *et al.* 2019; Husain *et al.* 2022*a*,*b*).

Accurately simulating wind-wave fields over a realistic range of the parameter space hinges upon our ability to resolve sea-surface drag. However, resolving the drag directly at the ocean surface through, for example, direct numerical simulations (DNSs) or wall-resolved large-eddy simulations (LESs) is only possible at modest Reynolds numbers due to the inherent complexity of interfacial mechanisms occurring over a broad range of time and length scales (e.g. Sullivan & McWilliams 2010; Ayet & Chapron 2022). Wall-modelled LES offers a pathway to address this challenge. In wall-modelled LES, the near-surface region is bypassed, and a wall-layer parameterization is instead used to account for subgrid-scale surface forces (Piomelli 2008; Bose & Park 2018). However, difficulties associated with the acquisition of high-resolution field (or even laboratory) measurements close to the air–sea interface (e.g. Turney & Banerjee 2008, 2013; Sullivan *et al.* 2018) have limited the development of such parameterizations (see, for e.g. Yang,

Meneveau & Shen 2013; Aiyer *et al.* 2023). Limiting factors in the retrieval of direct field measurements include platform movement (e.g. Donelan, Drennan & Katsaros 1997), limitations of existing sensor technology (e.g. Graber *et al.* 2000), broadband variability of ocean wave fields (e.g. Sullivan & McWilliams 2010) and cost. In addition, *in situ* observations often rely on point measurement techniques that only provide time-averaged statistics. As a result, the existing computational models and available measurement techniques are both characterized by considerable uncertainty, and findings from these two approaches are often difficult to reconcile in terms of, for example, total wind stress and wave growth (e.g. Belcher & Hunt 1998; Peirson & Garcia 2008; Hara & Sullivan 2015).

The action of wind imparts both momentum and energy fluxes to the wave field. At the air–water interface, the total momentum flux, or the wind stress, is the sum of the tangential and form drag (e.g. Grare *et al.* 2013; Buckley *et al.* 2020), i.e.

$$\tau = \bar{\tau}_v + \bar{\tau}_f = \bar{\tau}_v + \overline{p\frac{\partial \eta}{\partial x}}, \tag{1.1}$$

where τ_v is the interfacial tangential viscous stress (see § 2.1), τ_f is the form drag, p is the pressure at the interface, η is the surface elevation and the overbar denotes temporal or spatial averaging. Using linear spectral decomposition of the wave field, it can be further shown that the total energy input to both waves and surface currents is directly related to the wind stress

$$S_{in} = \overline{\tau_v u_s} + \overline{p\frac{\partial \eta}{\partial x}} c = \left(\bar{\tau}_{wc} + \bar{\tau}_f\right) c, \tag{1.2}$$

where u_s is the surface velocity, c is the phase speed and τ_{wc} is the wave-coherent tangential stress. Consequently, the momentum flux leading to wave growth has two components, the form drag and the wave-coherent tangential stress. Knowledge of the energy input rate (or equivalently, wind stress) and its dependence on wave field characteristics is then of fundamental importance in modelling the air–sea momentum flux. Two common models for the energy input rate are developed by Jeffreys (1925) and Miles (1957) from theoretical considerations and have been widely employed. These theories are, however, unable to reconcile their predictions with available field measurements.

Reconstructing the near-surface flow fields and evaluating the surface wind stress from observable, readily available free-surface characteristics, such as wave profile, wave speed and reference wind speed, are of significant interest to various geophysical and engineering applications. There is a growing appreciation of the necessity of integrating wave effects into simple one-dimensional column models of the marine boundary layers for use in climate prediction and regional ocean mesoscale codes (e.g. Donelan 1998; Hanley & Belcher 2008; Makin 2008). Hence, the capability to interpret the surface drag from easy-to-acquire wave features can enable reliable, real-time wave models by measuring, for example, wave heights, surface currents and sea-surface temperature using remote sensing and aerial imaging techniques (e.g. Huang *et al.* 2016; Metoyer *et al.* 2021). To that end, data-driven methods have the potential to complement field and experimental measurements and to be used as a promising tool for developing reliable surface-drag parameterizations.

Over the past few years, fluid dynamics and turbulence research has undergone a significant transformation due to the advent of machine learning (ML) techniques (e.g. Kutz 2017; Brenner, Eldredge & Freund 2019; Brunton, Noack & Koumoutsakos 2020; Fukami, Fukagata & Taira 2020*a*). The ML approaches have been successfully used to

accelerate computational fluid dynamics (e.g. Bar-Sinai *et al.* 2019; Kochkov *et al.* 2021; Jeon, Lee & Kim 2022), develop improved turbulence closures (e.g. Ling, Kurzawski & Templeton 2016; Wang, Wu & Xiao 2017; Beck, Flad & Munz 2019; Duraisamy, Iaccarino & Xiao 2019), reconstruct turbulent flow fields from spatially limited data (e.g. Liu *et al.* 2020; Fukami, Nakamura & Fukagata 2020*b*; Kim *et al.* 2021), predict the spatio-temporal behaviour of turbulent flows (e.g. Wang *et al.* 2020; Fukami, Fukagata & Taira 2021; Zhang & Zhao 2021) and advance surrogate modelling of turbulent boundary layer flows (e.g. Hora & Giometto 2023). However, in contrast to the flow over solid surfaces, the applications of ML approaches to air–sea interactions have been quite limited and mainly focused on predicting wave characteristics (e.g. James, Zhang & O'Donncha 2018; O'Donncha *et al.* 2018, 2019; Rasp & Lerch 2018; Sun *et al.* 2022; Zhang *et al.* 2022; Dakar *et al.* 2023; Lou *et al.* 2023; Xu, Zhang & Shi 2023). The existing works have employed different ML frameworks to reduce the computational complexity in estimating statistical wave conditions, such as significant wave height and peak wave period, and in low-dimensional learning of wave propagation (e.g. Sorteberg *et al.* 2020; Alguacil *et al.* 2021; Deo & Jaiman 2022). For instance, O'Donncha *et al.* (2018) integrated the SWAN (simulating waves nearshore) model into an ML algorithm to resolve wave conditions (see also James *et al.* 2018; O'Donncha *et al.* 2019). Further, Sun *et al.* (2022) developed a deep learning-based bias correction method for enhancing significant wave height predictions of the WW3 (Wave Watch III) model over the Northwest Pacific Ocean region (see also Campos *et al.* 2020; Ellenson *et al.* 2020). These studies have demonstrated the potential of ML in predicting the propagation of surface waves and improving the predictive performance of current wave models.

The reconstruction of flow field features from surface wave characteristics is, however, more challenging and has unique complications associated with multiscale wave profiles and motions. In previous research, limited efforts have been made to reconstruct the turbulent flow above/below surface waves based only on wave observations (e.g. Smeltzer *et al.* 2019; Gakhar, Koseff & Ouellette 2020; Xuan & Shen 2023; Zhang *et al.* 2023). Using convolutional neural network (CNN) classifiers, Gakhar *et al.* (2020) demonstrated that surface elevation information alone can be used to determine the physical features at the bottom boundary (see also Mandel *et al.* 2017; Gakhar, Koseff & Ouellette 2022). More recently, Xuan & Shen (2023) explored the feasibility of using a CNN to reconstruct the turbulent flow beneath free-surface flows using the surface elevation and velocity (see also Li, Xuan & Shen 2020). Their CNN-based model was trained on a dataset obtained from DNSs of turbulent open-channel flows over a deformable surface and could accurately reconstruct the near-surface flow field.

In the absence of accurate parameterizations of the sea-surface drag (see, for e.g. Sullivan & McWilliams 2010; Zachry *et al.* 2013; Gao *et al.* 2020), ML-based approaches may thus offer a practical alternative. To date, however, there has been no attempt to apply such data-driven algorithms to model the sea-surface drag. In the present study, using a supervised ML framework, we aim to estimate the surface viscous stress (or skin-friction drag) of wind-generated surface waves solely from wave profiles and wave age. To that end, we develop a CNN model that takes surface elevation profiles and the corresponding wave age as inputs and predicts the spatial distribution of skin-friction drag over wind waves. The CNN network is trained using high-resolution laboratory measurements of velocity fields obtained over a range of wind-wave conditions. The experimental dataset used in the current study is based on a re-analysis of the existing raw data of Yousefi *et al.* (2020). We find that, for unseen wind-wave regimes that fall within the training samples, the CNN model can accurately predict the overall distribution of instantaneous and area-aggregate viscous stresses using only the surface signatures, wave phase speed

and the mean wind speed magnitude aloft. In addition, we also assess the performance of the model in capturing unseen wind-wave conditions (i.e. out-of-training distribution). To the best of the authors' knowledge, this is the first successful attempt in the literature to model the sea-surface drag using surface signatures and wave age. The proposed CNN model can be easily tailored to serve as a wall-layer model for skin-friction contributions in wall-modelled LESs of airflow over wave fields.

The rest of the paper is organized as follows. A brief description of the experimental dataset is summarized in § 2.2. The ML model details are provided in § 2.3. In § 3, the results are presented, where we assess the performance of the model to reconstruct skin-friction drag for both the interpolation and extrapolation tasks. Further discussions on the applicability and limitations of the CNN model are offered in § 4. Finally, a brief conclusion is presented in § 5.

2. Methodology

2.1. *Theory*

At the air–water interface, the total wind stress is the sum of the tangential viscous stress (i.e. skin-friction drag) and the pressure stress (i.e. form drag), $\tau = \tau_v + \tau_f$. The skin-friction drag can be estimated by

$$\tau_v = (\boldsymbol{\tau}\boldsymbol{n}) \cdot \boldsymbol{t}, \tag{2.1}$$

where $\boldsymbol{\tau} = \mu(\boldsymbol{\nabla u} + \boldsymbol{\nabla u}^{\mathrm{T}})$ is the airside viscous stress tensor estimated at the air–water interface, μ is the air dynamic viscosity, $\boldsymbol{\nabla u}$ and $\boldsymbol{\nabla u}^{\mathrm{T}}$ denote the velocity gradient tensor and its transpose, $\boldsymbol{u} = (u, w)$ are velocities in the vertical plane aligned with the direction of wave propagation and $\boldsymbol{n}$ and $\boldsymbol{t}$ are unit vectors that are, respectively, normal and tangent to the local water surface, as shown in figure 1. Considering the velocity field in the vertical plane to be aligned with the direction of wave propagation, the skin-friction drag can then be estimated by

$$\tau_v = \mu\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right) + 2\mu\epsilon\left(\frac{\partial w}{\partial z} - \frac{\partial u}{\partial x}\right), \tag{2.2}$$

where x and z are the streamwise and vertical coordinate directions, respectively, and $\epsilon = \partial\eta/\partial x$ is the local surface slope. It should be noted here that, in (2.2), the skin-friction drag is obtained to the first order in ϵ (see Longuet-Higgins 1969; Buckley *et al.* 2020; Yousefi & Veron 2020). The form drag can also be expressed by $\tau_f = p\epsilon$. However, in the current study, we focus on skin-friction drag only.

2.2. *Experimental dataset*

In the current work, in order to train and examine the ML algorithm, we used the existing dataset of high-resolution airflow velocity measurements above wind-generated surface waves. The measurements were acquired using a combination of the particle image velocimetry (PIV) and laser-induced fluorescence (LIF) techniques in the wind-wave tunnel facility at the Air–Sea Interaction Laboratory of the University of Delaware. The facility is designed for studying atmosphere–ocean interactions, and its wave tank is approximately 42 m long, 1 m wide and 1.25 m high. The mean water level was kept at 0.7 m to allow sufficient space above the air–water surface. The tank is also equipped with a permeable wave-absorbing beach to dissipate the wave energy and prevent wave reflections. The wind tunnel of the facility can generate 10 m equivalent wind speeds of

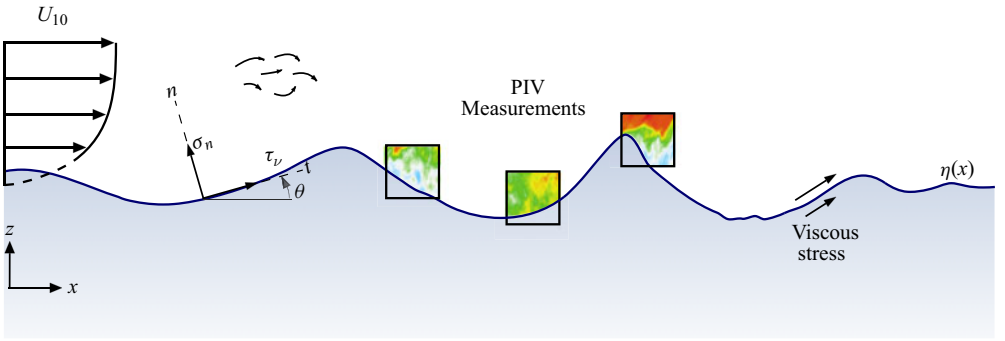


Figure 1. Schematic representation of two-dimensional airflow velocity measurements obtained using the PIV technique above wind-driven surface waves. The mean flow profiles, kinematic processes and drag partitioning (into the skin-friction drag and form drag) at the air–water interface are also shown. The global coordinate system in the PIV plane is (x, z) in the streamwise and vertical directions, respectively. The surface-following coordinate system (t, n) is also represented, where $\boldsymbol{t}$ and $\boldsymbol{n}$ are unit vectors that are tangent and normal to the local instantaneous water surface. At the wavy water surface, it is noted that $\epsilon = \partial\eta/\partial x$ and $\theta = \arctan(\epsilon)$.

up to 30 m s^{-1}. It is equipped with a honeycomb straightener to provide uniform airflow across the tank and a smooth transition from the wind tunnel inlet to the water surface. The facility, experimental set-up and image acquisition and processing procedures are described in detail in Yousefi (2020). Here, we briefly summarize the experimental set-up to put the available measurements into perspective.

High-resolution two-dimensional velocity fields above wind waves were obtained using PIV. In combination with PIV measurements, the LIF technique was employed to precisely detect the surface profiles. This allowed the acquisition of airside velocity measurements very close to the surface within the airside viscous sublayer, as close as 100 μm to the air–water interface, on average. Examples of streamwise velocity measurements along the wave field are plotted in figure 1. Figure 2 also shows an example of the PIV velocity field over a separating wind wave with the corresponding skin-friction drag at the air–water interface calculated using (2.2). These experimental measurements were acquired at a fetch of 22.7 m for various wind-wave conditions with different 10 m equivalent wind speeds varying from 2.25 to 16.59 m s^{-1}. This resulted in purely one-dimensional waves that only propagate in the streamwise direction such that wave groups and individual waves within those groups generally align in the same direction. The wave age of the generated waves is in the range of 0.06–0.21, which corresponds to moderately to strongly forced wind waves (Drennan *et al.* 2003). For reference, the complete experimental conditions are listed in table 1.

The spectra of the generated surface wave elevations were relatively narrow banded for all experimental conditions, with peaks at the dominant wave frequency (see Yousefi *et al.* 2020). As reported in table 1, the dominant wave frequencies decrease with increasing wind speed because the fetch-limited waves grow longer with increasing wind speed. However, the spectral densities increase with increasing wind speed, which indicates that the waves also increase in amplitude with increasing wind speed. Under these experimental conditions, the wavelength of wind-generated surface waves varies approximately from 15 to 50 cm, which is at larger scales than capillary waves with a wavelength in the sub-centimetre range (e.g. Slavchov, Peychev & Ismail 2021). The complications of bubble/spray generation were also largely avoided in these wind-wave conditions.

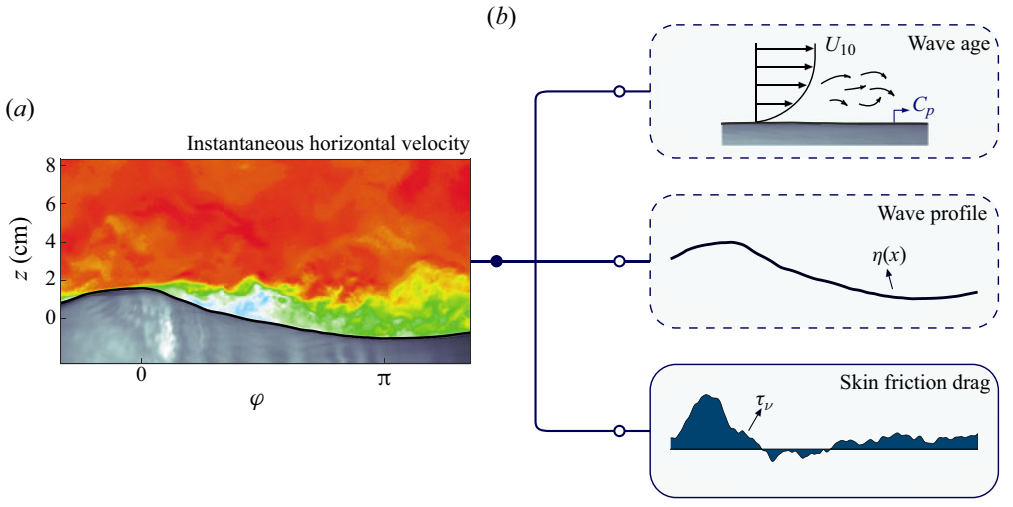


Figure 2. An example of the instantaneous streamwise PIV velocity field over a separating wind wave (*a*) with the corresponding wave age, wave profile and skin-friction drag at the air–water interface (*b*). The objective of the CNN-based ML model is to estimate the skin-friction drag, τ_ν, over wind waves from the wave profile, $\eta(x)$, and wave age, C_p/U_{10}. In the experimental measurements, the surface tangential viscous stress (or, equivalently, skin-friction drag) is the first value of the viscous stress measurements taken at the height of 94.8 μm above the air–water interface.

	U_{10} (m s^{-1})	u_* (m s^{-1})	C_p/U_{10}	C_p/u_*	C_p (m s^{-1})	a_p (m)	λ_p (m)	a_pk_p	f_p (Hz)
W1-02	2.25	0.075	0.21	6.27	0.47	0.15	0.14	0.07	3.3
W1-05	5.08	0.168	0.12	3.69	0.62	0.50	0.25	0.13	2.5
W1-09	9.57	0.318	0.08	2.46	0.78	1.20	0.39	0.19	2.0
W1-14	14.82	0.567	0.06	1.53	0.87	1.96	0.48	0.26	1.8
W1-16	16.59	0.663	0.06	1.39	0.92	2.29	0.54	0.27	1.7

Table 1. Summary of the experimental data and wind-wave conditions. The friction and 10 m extrapolated velocities were calculated from the experimental data by fitting the logarithmic part of the mean wind velocity profile. Parameters with subscript p indicate the peak wave values obtained from peak wave frequencies by applying linear wave theory. For additional details on experimental data and procedures, the reader is referred to publications by Yousefi (2020) and Yousefi *et al.* (2020).

Therefore, the effect of surface tension force on surface wave profiles is negligible and does not directly contribute to the viscous stress at the air–water interface.

2.2.1. *Data pre-processing*

The experimental dataset used to train and evaluate the ML model comprises various wind speed conditions, with each experiment sampling approximately the same number of waves. This resulted in the acquisition of more instantaneous PIV fields (which have a fixed footprint) when the wind waves were longer. Specifically, the dataset consists of approximately 2000, 3800, 4500, 4900 and 5000 instantaneous PIV images corresponding to wind speeds of $U_{10} = 2.25$, 5.08, 9.57, 14.82 and 16.59 m s^{-1}, respectively. For training and validation of the ML model, we considered experiments with a wind speed of 5.08, 9.57 and 16.59 m s^{-1} (corresponding to W1-05, W1-09 and W1-16 cases in

table 1), resulting in a total of roughly 13 300 input–output pairs. These pairs were further randomly divided into training, validation and testing sub-datasets, where each set included 80 %, 10 % and 10 % of the total samples, respectively. The training dataset is used to train the ML model, allowing it to learn and adjust internal parameters to minimize prediction errors, while the validation subset is utilized to further fine tune the model's hyperparameters during training (e.g. Goodfellow, Bengio & Courville 2016). Finally, in order to assess the generalization capability of the model and validate its performance in a more realistic scenario, we examined the model on out-of-training-distribution datasets. To this end, we used instantaneous PIV images of the cases with wind speeds of $U_{10} = 2.25$ and 14.82 m s^{-1} as a separate dataset to examine the model performance on completely unseen data. Notably, the $U_{10} = 2.25$ m s^{-1} case is outside the wind speed training range.

The signal-to-noise ratio of PIV velocity measurements naturally degrades in near-surface areas due to various factors, including the reflection of high-intensity laser light off the surface and the presence of other background noise sources. Although this does not affect velocity measurements of the current experimental dataset, the interfacial skin-friction drag experiences an increased level of noise because the gradient operator introduces additional noise to the measurement. It was observed that the performance of the ML algorithm is sensitive to the near-surface noise present in the experimental data. In order to ensure the optimal performance of the ML network, a zero-phase (smoothing) infinite-impulse-response filter (e.g. Gustafsson 1996) was applied to the along-wave surface tangential stress profiles

$$\left.\begin{aligned} y_0[n] &= \frac{1}{a_0}\left(\sum_{i=0}^{m} b_i x[n-i] - \sum_{j=1}^{m} a_j y_0[n-j]\right) \\ &\text{and} \\ y_1[n] &= \frac{1}{a_0}\left(\sum_{i=0}^{m} b_i y_0[n+i] - \sum_{j=1}^{m} a_j y_1[n+j]\right), \end{aligned}\right\} \tag{2.3}$$

where $y_0[n]$ and $y_1[n]$ are, respectively, the output of forward and reverse filtering operation at location index n, $x[n]$ is the input signal, b_i are the feedforward filter coefficients, a_j are the feedback filter coefficients and m is the filter order. An example of filtered skin-friction drag (along with other ML input data) and original experimental data are presented in figures 3(*d*) and 3(*h*) over a separating and non-separating wind wave. Further, in order to expedite the training process, we scaled the data to have zero mean and unit standard deviation. This standardization ensures that features are placed on a comparable scale, thereby facilitating more effective training of the ML model (e.g. Goodfellow *et al.* 2016).

2.3. *Machine learning model*

The main objective of this work is to develop a model to accurately estimate the instantaneous skin-friction drag, τ_ν, of surface wind waves from wave profiles, $\eta(x)$, local interface slope, $\partial\eta/\partial x$, local wave phases, φ, and the wave age of the corresponding wave field, C_p/U_{10}. In particular, we aim to develop a mapping such that

$$\mathcal{M} : \left(C_p/U_{10}, \eta(x), \partial\eta/\partial x, \varphi\right) \rightarrow \tau_\nu(x). \tag{2.4}$$

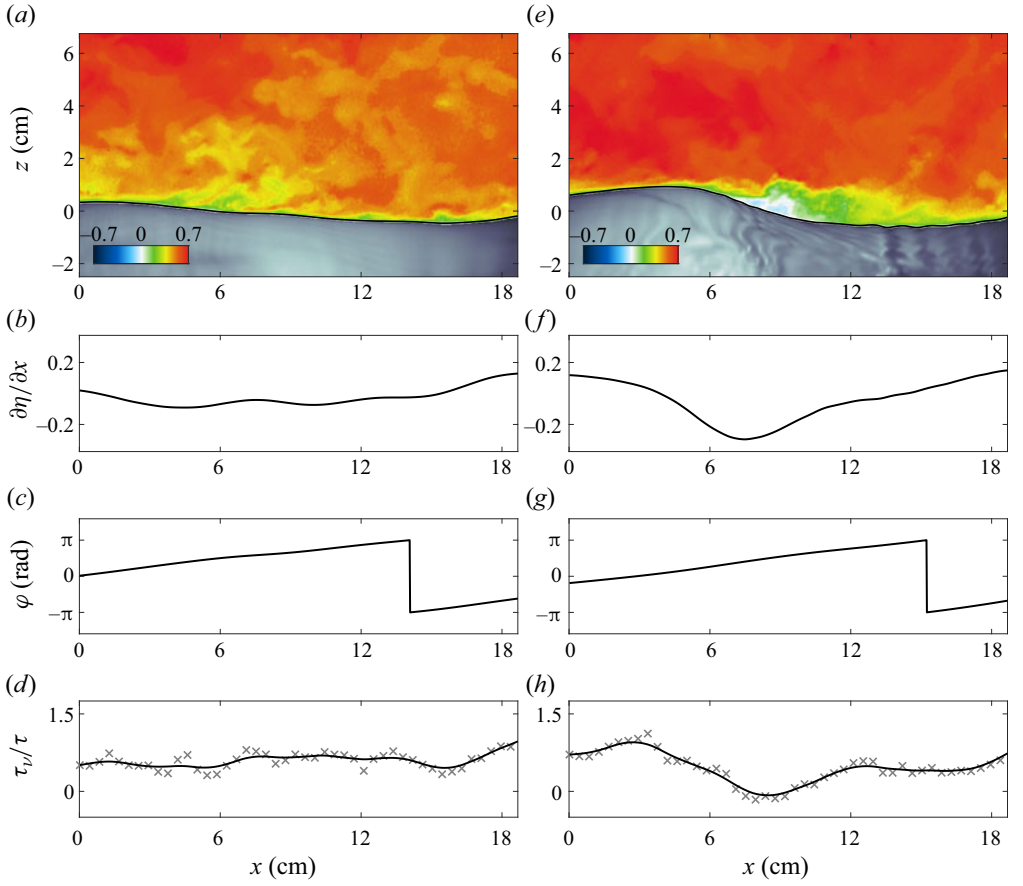


Figure 3. An example of ML input data over non-separating (*a–d*) and separating (*e–h*) wind waves for the wind-wave experimental condition of $U_{10} = 5.08$ m s^{-1}. The instantaneous fields of normalized streamwise velocity, u/U_{10}, are shown in panels (*a*,*e*), in which solid lines denote the wave profiles. The panels below each wave show (*b*,*f*) the local wave slope, $\epsilon = \partial\eta/\partial x$, (*c*,*g*) the local wave phase, φ, and (*d*,*h*) the skin-friction drag scaled with total stress, τ_ν/τ, where the experimental data are indicated with cross symbols and the solid lines are the filtered experimental tangential stress profiles.

The CNN-based ML model is employed to approximate $\mathcal{M}$; a detailed description of network architecture is introduced in the following section. Here, we used wave age as a dimensionless parameter incorporating wind speed rather than wave Reynolds number ($Re_w = U_{10}/\nu k$ with k being the wavenumber) because the wave age provides a better metric to characterize different ocean regimes (for e.g. see Csanady 2001; Alves, Banner & Young 2003; Sullivan & McWilliams 2010). In addition, the Re_w is indirectly incorporated into the ML model as all the parameters that define Re_w are already included in the mapping. Although we used the wave age as a non-dimensionalized input for the ML model, the wave age only covers a small range of wind-generated surface waves in these constant fetch laboratory experiments (see table 1). Therefore, in the remainder of the paper, we will narrow our analysis to wind speed, except when wave age is a convenient parameter to compare our results with available data. It should also be noted that the Bond number ($BO = \Delta\rho g/\sigma k^2$, where $\Delta\rho$ is the air–water density difference, g is the gravity and σ is the surface tension) of surface wave cases considered for this study is relatively large for all cases, varying from $BO = 65$ for the lowest wind speed of 2.25 m s^{-1} to

$BO = 975$ for the highest wind speed of 16.59 m s^{-1}, indicating that the surface tension effects are negligible (e.g. Van Hooft *et al.* 2018). Hence, due to the high Bond number of wind waves, we did not include the Bond number as a mapping parameter.

2.3.1. *Convolutional neural network*

Neural network-based ML models, notably the multi-layer perceptron (MLP) and CNN, are widely adopted for tackling complex nonlinear regression tasks across diverse scientific and engineering domains (e.g. Rumelhart, Hinton & Williams 1986; LeCun *et al.* 1998; Goodfellow *et al.* 2016). While both models are commonly used in multi-output regression tasks, MLP-based models are preferred for pointwise mapping (e.g. Raissi, Yazdani & Karniadakis 2020; Maulik *et al.* 2021; Hora & Giometto 2023), whereas CNNs excel in handling structured grid data, such as velocity fields and images (e.g. Fukami *et al.* 2021; Kim *et al.* 2021; Xuan & Shen 2023). This preference arises from the CNN models' ability to capture spatially local features, exhibit translation equivariance and offer scalability advantages (see, for e.g. Goodfellow *et al.* 2016; Gao, Sun & Wang 2021). Overall, the success of CNN-based networks has established them as promising approaches for approximating complex functions.

Specifically, in the current study, we employ a CNN-based ML model (indicated by $\mathcal{F}$) to approximate the mapping function, $\mathcal{M}$, such that

$$\mathcal{F}(C_p/U_{10}, \eta(x), \partial\eta/\partial x, \varphi; \boldsymbol{w}) \approx \tau_\nu(x), \tag{2.5}$$

where $\boldsymbol{w}$ encompasses the trainable parameters of the CNN, including weights and biases. The CNN model used in this work primarily consists of convolutional, dense (or fully connected) and nonlinear layers. Among these, the convolutional layer plays a crucial role in capturing hierarchical representations and extracting meaningful features from the high-dimensional spatial input. This is achieved by applying the convolution operation to the input, followed by a nonlinear activation function. In a two-dimensional convolutional layer, the output h^{out} can be described based on the input from the previous layer, h^{in}, as

$$h^{out}_{ijd} = \sigma\left(\sum_{k=1}^{K}\sum_{p=1}^{F}\sum_{q=1}^{F} h^{in}_{i+p,j+q,k} w_{pqkd} + b_{ijd}\right), \tag{2.6}$$

where $\boldsymbol{w}$ and $\boldsymbol{b}$ denote the weight and biases of the output layer, σ is the element-wise nonlinearity, K is the number of kernels and $F \times F$ is the filter/kernel dimension (LeCun *et al.* 1998; Goodfellow *et al.* 2016). While convolutional layers excel at extracting local features, they have a limited receptive field and lack global awareness (e.g. Long, Shelhamer & Darrell 2015). Conversely, dense layers connect all neurons (also known as perceptrons) from the previous layer to every neuron in the current layer, enabling them to capture spatially global patterns and make predictions based on combined information from all input features (e.g. Goodfellow *et al.* 2016). By incorporating dense layers towards the end of the CNN architecture, the network gains the ability to learn and encode global features. Mathematically, the transformation performed by a dense layer can be described as follows: given an input vector $h^{in} \in R^n$ and a weight matrix $\boldsymbol{w} \in R^{m\times n}$, the output $h^{out} \in R^m$ can be computed as

$$h^{out} = \sigma\left(\boldsymbol{w} h^{in}\right). \tag{2.7}$$

The choice of an appropriate nonlinearity is crucial as it significantly impacts the training process and performance of the network for a given task (e.g. Hao *et al.* 2020).

Nonlinear activation functions, including sigmoid, hyperbolic tangent (tanh) and rectified linear units (ReLU) and their variants (e.g. Nair & Hinton 2010; Maas *et al.* 2013; He *et al.* 2015), offer diverse choices to enable networks in learning complex relationships. However, gradient disappearance in sigmoid and tanh or the so-called 'dying ReLU' phenomenon for ReLU can adversely impact the models' performance (e.g. Hochreiter 1998; Glorot & Bengio 2010; Xu *et al.* 2015). Recently, Swish and Mish nonlinearity functions emerged as promising alternatives (see Ramachandran, Zoph & Le 2017; Misra 2019), aiming to provide smoother, nonlinear behaviours, enhance the gradient flow and mitigate dead neurons. Given these properties, we employed the Mish activation, represented as $\mathrm{Mish}(x) = x\tanh[\log(1 + e^x)]$ (Misra 2019), in the nonlinear layers of our network architecture.

Here, in order to obtain the optimal parameters $\boldsymbol{w}$ of the CNN model, we minimize the L_2 norm, also known as the mean squared error (MSE), between the ground-truth skin-friction drag obtained from experimental measurements (indicated by τ_ν) and the output of the ML model, $\hat{\tau}_\nu = \mathcal{F}(C_p/U_{10}, \eta(x), \partial\eta/\partial x, \varphi; \boldsymbol{w})$. Additionally, we incorporate ridge regularization to address the issue of overfitting. This regularization technique penalizes large values of trainable parameters, w_i, by augmenting the MSE loss function with the squared magnitudes of the weights. The aim is to prevent the model from assigning disproportionately large weights to specific features, a practice known to contribute to overfitting (e.g. Goodfellow *et al.* 2016). In other words, our objective is to find $\boldsymbol{w}$ that minimizes the following expression:

$$\boldsymbol{w} = \arg\min_{\boldsymbol{w}} \|\mathcal{F}\left(C_p/U_{10}, \eta(x), \partial\eta/\partial x, \varphi; \boldsymbol{w}\right) - \tau_\nu\|_2 + \lambda \sum \|\boldsymbol{w}\|_2, \tag{2.8}$$

where λ is the regularization parameter and $\sum \|\boldsymbol{w}\|_2$ is a ridge regularization. This optimization process enables the CNN model to learn optimal parameters, enhance performance and better approximate the underlying mapping function.

The CNN model is implemented using the Pytorch ML library (Paszke *et al.* 2019). For $\mathcal{F}$, based on the hyperparameter optimization analysis (see the Appendix), we adopt a CNN network with five convolutional layers, one flattened layer, one dense layer and Mish activation function for nonlinearity, as shown in figure 4. For the outputs, we employ one dense layer with a linear activation function. For all convolutional layers, the filter size is set to $F = 3$, the number of kernels is set to $K = 8$ (see (2.6)) and the number of perceptrons in a dense layer is kept at 2000. The optimal CNN-based ML model architecture is displayed in figure 4, which has 65 015 401 trainable and 0 non-trainable parameters. The network is trained end-to-end using backpropagation with the Adam optimizer (Kingma & Ba 2014). The Adam optimizer is an advanced stochastic gradient descent algorithm that iteratively updates the trainable parameter based on the gradients of the loss function with respect to the parameters. It incorporates adaptive learning rates and momentum, which dynamically adjust the step sizes during optimization, making the training process more efficient Kingma & Ba (2014). All trainable parameters are initialized randomly using values sampled from a uniform distribution, following the approach of Glorot & Bengio (2010), and the value of regularization parameter λ is set to 10^{-4}. The reduce learning rate on plateau schedule is used, initializing the learning rate to 5×10^{-4} and the rate is reduced by a factor of 0.9 when the validation loss stagnates during the training and the effective mini-batch size is set to 64.

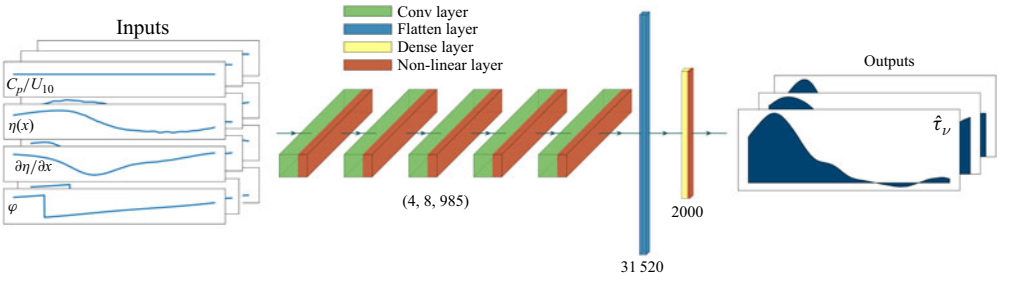


Figure 4. Diagram illustrating the CNN network architecture utilized to estimate the skin-friction drag, τ_ν, from $\eta(x)$, $\partial\eta/\partial x$, φ and C_p/U_{10}. The network comprises five convolution layers (labelled 'Conv Layers'), a flattened layer, a dense layer and the Mish activation function as the nonlinear layer. The network takes inputs including a stack of $\eta(x)$, $\partial\eta/\partial x$, φ and C_p/U_{10} (shown on the left), and the output $\hat{\tau}_\nu$ (shown on the right). Each convolutional layer employs a filter size of $F = 3$ and $K = 8$ filters, the dense layer contains 2000 perceptrons/neurons. The tuple (4, 8, 985) represents the number of inputs, number of kernels and number of data points, respectively.

3. Results

In this section, we evaluate the performance of the CNN model by comparing the predicted skin-friction drags against the corresponding reference high-resolution PIV measurements from the test dataset mentioned in § 2.2.1. To this end, we first analyse the convergence history of the model and present a quantitative analysis of ML-specific error metrics. We then focus on model performance based on instantaneous and averaged skin-friction fields and assess its predictive ability in estimating the viscous drag from datasets in and out of the training range.

3.1. *Convergence history*

In this section, we provide an overview of the ML model training process. To ensure optimal performance, careful consideration must be given to the training process, including the number of epochs or iterations. Training for an insufficient number of epochs might lead to an underfitting issue, where the model fails to capture the underlying patterns in the data. Conversely, excessive training may lead to overfitting, a phenomenon in which the model memorizes the training data and learns the present noise, losing its ability to generalize to unseen data. Balancing the trade-off between underfitting and overfitting is critical to enhance the model's performance on new, unseen data. Such a balance can be generally achieved through various methods such as cross-validation, early stopping and regularization. Incorporating these techniques, combined with the careful selection of hyperparameters and the choice of appropriate model architecture, can significantly improve the generalization performance of the regression model (e.g. Hastie, Tibshirani & Friedman 2009; Goodfellow *et al.* 2016).

Figure 5 shows the MSE convergence history of the ML model against the number of epochs on both training and validation datasets. It is plotted on a logarithmic vertical scale to better visualize the slight differences between training and validation losses. The figure clearly illustrates a progressive decrease in MSE for both datasets as the training progresses. This indicates that the model is continuously learning and improving its accuracy. Here, to address the challenge of overfitting and enhance generalization, we employed the early stopping technique proposed by Amari *et al.* (1997) and used in existing ML-based studies (e.g. Fukami, Fukagata & Taira 2019; Srinivasan *et al.* 2019; Fukami *et al.* 2021). Early stopping is a technique that continuously monitors the model's

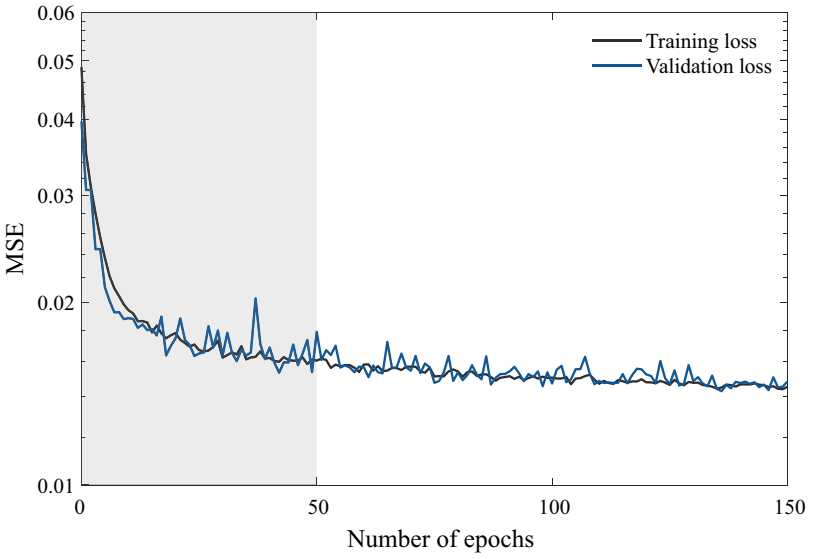


Figure 5. Convergence of the MSE loss of the CNN-based ML during the training phase. Training and validation losses represent the MSE on the training and validation datasets, respectively. Epochs indicate the number of iterations used in the learning process. The vertical axis is plotted on logarithmic access to better demonstrate the slight differences between loss curves.

performance on a separate validation dataset and terminates the training process if the validation loss fails to improve or consistently deteriorates over a predefined number of epochs (five in our case). In the current work, while the training set was used to update the model parameters, the validation set was used to continuously monitor the model performance on a separate unseen dataset and prevent overfitting the training data. After approximately 150 epochs, the validation loss stopped improving, and the early stopping mechanism terminated the training. This approach allowed us to identify the optimal stopping point during training and capture a model that strikes a balance between fitting the training data and generalizing it to unseen data. At the completion of the training, the MSE values for the training and validation datasets were 0.0146 and 0.0148, respectively. The close MSE value on training and validation datasets demonstrates the model's ability to accurately learn the underlying concept, while effectively avoiding overfitting or underfitting issues.

3.2. *Model performance on the test dataset*

3.2.1. *Error analysis*

To quantify and evaluate the performance of the CNN-based ML method, we calculate the relative mean absolute error (RMAE) and the relative mean square error (RMSE) defined based on the L^1-norm and L^2-norm, respectively. The relative mean error can be defined by

$$\delta_{e,p} = \frac{\|\tau_\nu - \hat{\tau}_\nu\|_p}{\|\tau_\nu\|_p}, \tag{3.1}$$

where $\hat{\tau}_\nu$ is the ML-predicted viscous stress, τ_ν is the corresponding ground-truth experimental measurement and $\|\cdots\|_p$ is the L_p-norm defined for the vector $\boldsymbol{v}$ of size n

	W1-05	W1-09	W1-16	W1-02	W1-14
RMAE	0.240	0.217	0.344	0.342	0.319
RMSE	0.274	0.270	0.491	0.446	0.433
R^2	0.924	0.934	0.846	0.836	0.866

Table 2. Summary of the ML model performance in predicting skin-friction drag compared with the experimental data. The RMAE, RMSE and coefficient of determination (R^2) metrics are presented for all wind-wave conditions. Also, the W1-02 and W1-14 cases are separate datasets for which the model has been applied to examine the model interpolation and extrapolation performance on unseen data.

by

$$\|\boldsymbol{v}\|_p = \left(\sum_{i=1}^{n} |v_i|^p\right)^{1/p}. \tag{3.2}$$

Here, $|\cdots|$ represents the absolute operator, $p = 1$ indicates the RMAE and $p = 2$ denotes the RMSE. Both metrics are thus a measure of the mean relative error between the ML prediction and the experimental output of skin-friction drag for each instantaneous wave. For evaluating the performance of the ML model, it is important to consider different types of errors. For completeness, we also calculated the coefficient of determination, denoted as R^2

$$R^2 = 1 - \frac{\sum \left(\tau_\nu - \hat{\tau}_\nu\right)^2}{\sum \left(\tau_\nu - \bar{\tau}_\nu\right)^2}, \tag{3.3}$$

where $\bar{\tau}_\nu$ is the ensemble average of the ground-truth skin-friction drag. Here, it is noted that the performance evaluation is based on the test dataset of the corresponding wind-wave case, i.e. the ensemble averaging in (3.3) is conducted on the snapshots in the test dataset only.

The values of the $\delta_{e,1}$, $\delta_{e,2}$ and R^2 metrics are reported in table 2 for all wind-wave cases. In an average sense, these values comprehensively evaluate the ML-model predictions compared with the reference experimental data. In cases for which the model has been trained, the relative errors are slightly lower, except for the highest wind speed case; in the percentage sense, the RMAE error for the W1-05, W1-09 and W1-16 cases corresponding to wind waves with $U_{10} = 5.08$, 9.57 and 16.59 m s^{-1} are 24.0 %, 21.7 % and 34.4 %, respectively. The increase in the error metrics for the highest wind speed case compared with the moderate cases ($U_{10} = 5.08$ and 9.57 m s^{-1}) is, in part, due to capturing only a portion of the wave profiles and the corresponding surface stresses at high wind speeds. Because of longer waves, in the two high wind speed cases of 14.82 and 16.59 m s^{-1}, there exist a number of waves for which the entire wavelength of the wave cannot be captured in the PIV field of view. We observed that the ML model performs best for the cases where at least one wavelength has been captured. Another contributor to this increased error might be the increase in the signal-to-noise ratio of the PIV measurements with wind speed. The unseen cases of W1-02 and W1-14 are discussed in § 3.4.

The relative errors reported in table 2 provide an average measure over the entire dataset. To further evaluate error distributions across individual waves and gain insights into the model accuracy on a wave-by-wave basis, we estimate the relative error for each wave. The probability density functions (p.d.f.s) of relative mean absolute and R^2 errors are respectively shown in figures 6(*a*) and 6(*b*) for the W1-05, W1-09 and W1-16 cases.

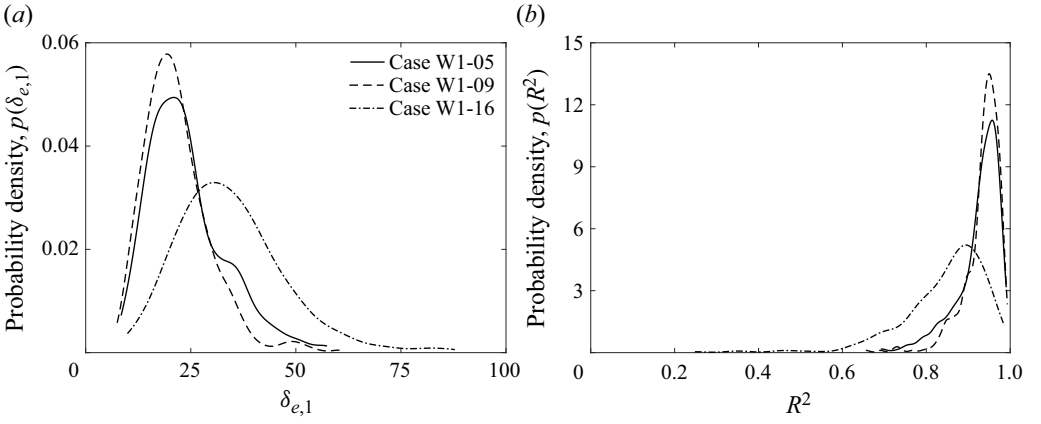


Figure 6. The p.d.f.s of (*a*) RMAE, $\delta_{e,1}$, and (*b*) coefficient of determination error, R^2, for the W1-05, W1-09 and W1-16 cases with a wind speed of $U_{10} = 5.08$, 9.57 and 16.59 m s^{-1}, respectively. The RMAE and R^2 error metrics were calculated for each instantaneous wave in the test dataset, and their distribution of probabilities is shown here.

As is shown, the p.d.f. plots indicate that the ML model performs similarly for the W1-05 and W1-09 cases with moderate wind speeds of 5.08 and 9.57 m s^{-1}. In these cases, the most probable relative error is roughly 20 % with a coefficient of determination of approximately 0.96, where the maximum RMAE is max $(\delta_{e,1}) \approx 55$ % and the minimum R^2 is min $(R^2) \approx 0.7$. The p.d.f.s of these cases also exhibit overlapping tails with small $p(\delta_{e,1})$ values, suggesting that the probability of extreme errors (RMAE $\geqslant 40$ % and $R^2 \leqslant 0.75$) is negligible. For the strongly forced wind waves (i.e. W1-16 case), however, the performance of the ML model slightly degrades, yielding a most probable relative error of $\delta_{e,1} \approx 30$ % with a coefficient of determination of $R^2 \approx 0.9$. In this case, although the maximum relative error increases to 90 %, R^2 remains acceptable at around 0.55.

The decrease in the model accuracy with increasing wind speed can be attributed to more frequent airflow separation events over wind waves and the increased signal-to-noise ratio in the experimental measurements. We discuss this in more detail in § 4. Here, it should also be noted that, although the relative error may be higher in high wind speed cases, the R^2 metric still demonstrates an acceptable correlation between the ML predicted and experimental values of the skin-friction drag for all wind-wave conditions. In fact, the RMAE is an average measure relative to the ground-truth experimental values. A large RMAE thus suggests that the ML model may have a large average deviation from the average magnitude of experimental values. However, the R^2 represents how accurately the ML predictions of surface stress can capture the variability of the corresponding experimental measurements about their respective means. This indicates that our ML model, even in cases with a high RMAE, can sufficiently explain the overall variance of the measurements as the R^2 values are acceptable. Although RMAE and R^2 are widely accepted metrics for evaluating ML model performance, their lack of physical significance may lead to misplaced confidence in the model's performance from a physical standpoint (e.g. Wang *et al.* 2020; Hora & Giometto 2023).

3.2.2. *Wave-phase-dependent errors*

Here, we turn our attention to examining the accuracy of the ML model in predicting the wave-phase-dependent distribution of skin-friction drag over surface waves.

The stress distributions over waves are largely wave-phase dependent, and accurate predictions of stress distributions are of significant importance for numerical simulations such as wall-modelled LESs and developing sea-state-dependent modelling frameworks. In order to analyse the behaviour of organized wave motions in a turbulent boundary layer flow, an instantaneous flow field variable may be separated into a phase-averaged (or wave-phase-dependent) component, $\langle f(\boldsymbol{x}, t)\rangle$, and a turbulent fluctuation component, $f'(\boldsymbol{x}, t)$, such that (Hussain & Reynolds 1970)

$$f(\boldsymbol{x}, t) = \langle f(\boldsymbol{x}, t)\rangle + f'(\boldsymbol{x}, t), \tag{3.4}$$

where $\boldsymbol{x}$ is the position vector at time t. The phase-averaged quantity can be further decomposed into the sum of an ensemble mean component, $\bar{f}(\boldsymbol{x})$, and a wave-coherent component, $\tilde{f}(\boldsymbol{x}, t)$, i.e. $\langle f(\boldsymbol{x}, t)\rangle = \bar{f}(\boldsymbol{x}) + \tilde{f}(\boldsymbol{x}, t)$, which is defined as

$$\langle f(\boldsymbol{x}, t)\rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f\,(\boldsymbol{x} + n\lambda, t), \tag{3.5}$$

where N is the number of instantaneous realizations and λ is the local wavelength. In fact, a phase-average quantity is the average of the values of that quantity at a particular phase of the wave, while the ensemble average (or equivalently, mean) is defined as the average of the quantity over all possible phases. The properties of the ensemble- and phase-averaged operators are discussed in reports by Hussain & Reynolds (1970, 1972) and Reynolds & Hussain (1972). The phase-averaged relative errors and r-squared distributions across wave phases are presented for W1-05, W1-09 and W1-16 experimental conditions in figures 7(*a*) and 7(*b*), respectively. It is noted that, in the current work, the local instantaneous wave phases were estimated using a Hilbert transform technique (for details, see Melville 1983; Hristov, Friehe & Miller 1998) applied to the wave profiles obtained from laboratory measurements (see Yousefi *et al.* 2020). In order to estimate the phase-averaged quantities, the wave phases were segregated into 108 independent bins, each covering a phase interval of 5.82×10^{-2} rad. From figure 7, it is evident that the relative errors for moderate wind speed cases of W1-05 and W1-09 exhibit similar patterns in that both RMAE and R^2 have their optimum values near the wave crest, and they slightly degrade in the leeward side of the waves where the airflow separation events occur. However, although the general performance degrades for the W1-16 case, the ML model accuracy improves in the wake regions downwind of waves. The phase-averaged error metrics, in general, align with the mean values reported in table 2 and figure 6.

Finally, the phase-averaged predictions of the skin-friction drag, $\langle\hat{\tau}_v\rangle$, are shown and compared with the corresponding experimental measurements in figure 8(*a*) for the wind speed of 5.08 m s^{-1}. Here, all stress profiles are normalized by the total wind stress, i.e. $\tau = \rho u_*^2$. The ML predictions of the phase-averaged viscous stress particularly show an excellent agreement with the experimental measurements. The phase-averaged distributions of viscous stress present a pattern of along-wave asymmetry in which the stress is highest upwind of the wave crest with its peak value about the crest and a minimum in the middle of the leeward side of the wave. The ML model accurately captures this behaviour. An important observation here is that the RMAE and R^2 calculated from the phase-averaged (and, as we see later, the ensemble-averaged) skin-friction drags demonstrate remarkable performance, indicating accurate predictions of the stress profiles in the average sense. This is because the differences between the ML predictions and experiments observed in instantaneous waves will be averaged out once all profiles are averaged into each phase bin. In the case of the W1-05 experiment, the relative error and

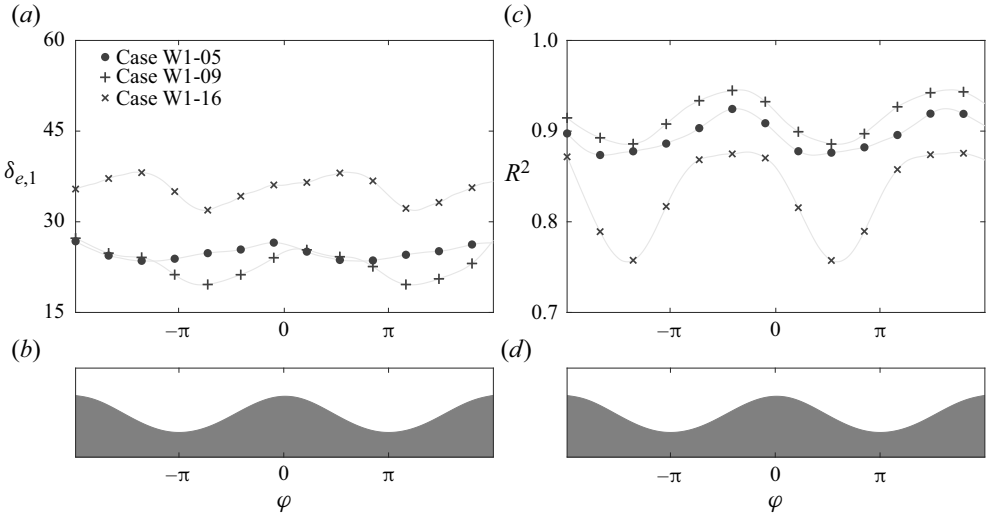


Figure 7. Phase-averaged (*a*) RMAE, $\delta_{e,1}$, and (*c*) coefficient of determination error, R^2, for W1-05, W1-09 and W1-16 cases with a wind speed of 5.08, 9.57 and 16.59 m s^{-1}, respectively. The RMAE and R^2 metrics were calculated for each instantaneous wave in the test dataset and bin averaged based on the corresponding wave phases. The wave phases were segregated into 108 independent bins such that each bin covers a phase interval of 5.82×10^{-2} rad. A sketch of the mean wave profile is also shown in panels (*b*,*d*) to better visualize the wave phases.

r-squared are 3.1 % and 0.998, respectively. For brevity, the other cases with $U_{10} = 9.57$ and 16.59 m s^{-1} have not been shown here. However, their relative error between the phase-averaged ML predictions and laboratory measurements of the skin-friction drag is 5.3 % and 13.9 %, respectively, with the corresponding R^2 values of 0.996 and 0.975. We observe the same order of accuracy of the model in predicting the ensemble-averaged viscous stresses in § 3.4.

To examine the effectiveness of the proposed model in reconstructing the spatial variability of phase-averaged skin-friction drags, the p.d.f. of normalized $\langle \tau_v \rangle$ for both ML predictions and empirical measurements is depicted in figure 8(*c*) for the W1-05 case with a wind speed of 5.08 m s^{-1}. As illustrated, the p.d.f. plots of the reconstructed viscous stress align remarkably with those derived from the experimental data, particularly for the case of moderately forced wind waves. This agreement is also observed for other cases, including the high wind speed cases (not shown here for brevity). These findings underscore the model's adeptness in predicting and reconstructing phase-averaged skin-friction drags with precise spatial distributions. Such a quantity is particularly relevant for numerical simulations based on the wall-layer modelling approach (see Piomelli 2008), where the wavy surface is either known but unresolved or known and partially resolved (see, for e.g. Husain *et al.* 2019; Deskos *et al.* 2021).

3.2.3. *Instantaneous stress reconstructions*

The preceding analysis, while illustrative of the model's proficiency in reconstructing the spatial variability of phase-averaged skin-friction drags, does not facilitate a comprehensive evaluation of its ability to reconstruct local stress distributions or capture the complex dynamics of near-surface physical processes. Consequently, a more nuanced examination is required to fully elucidate these attributes. In this section, we broaden our investigation to more precisely assess the performance of the model in reconstructing

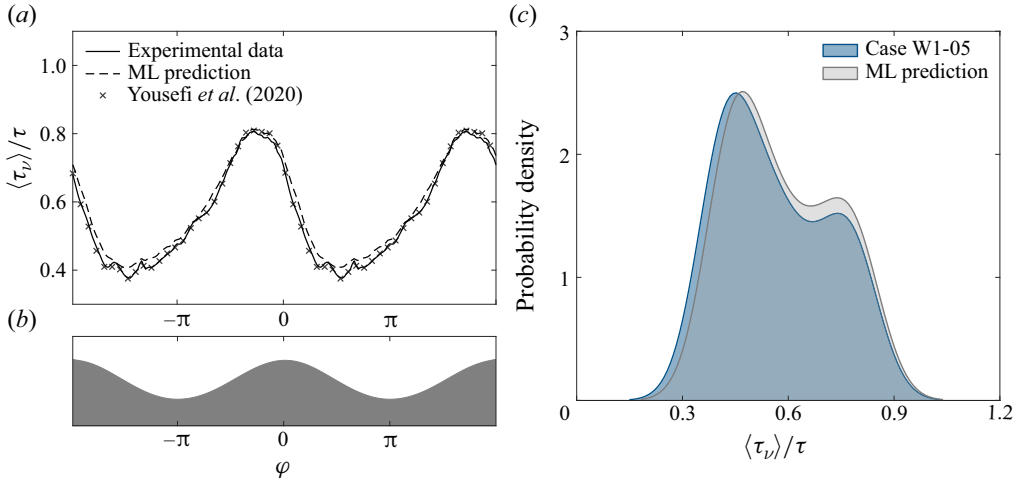


Figure 8. (*a*) The ML predictions of the phase-averaged distributions of the normalized skin-friction drag, $\langle\tau_\nu\rangle/\tau$, compared with the experimental measurements for the wind-wave condition with a wind speed of 5.08 m s^{-1}. For comparison purposes, both raw and filtered experimental data are plotted. All stress profiles are scaled using the total wind stress, $\tau = \rho u_*^2$. A sketch of the mean wave profile is also plotted in panel (*b*) to better visualize the wave phases. In panel (*c*), the p.d.f. of $\langle\tau_\nu\rangle/\tau$ estimated by the ML model and experimental measurements is plotted for the same of W1-05. The p.d.f. plot indicates the ability of the model to accurately predict the phase-dependent spatial distributions of skin-friction drag.

the instantaneous along-wave distributions of surface viscous stress and its ability to capture the effects of near-surface separation on the skin-friction drag. Figure 9 shows the predicted instantaneous normalized skin-friction drags, $\hat{\tau}_\nu/\tau$, for surface wave cases with a wind speed of 5.08 m s^{-1} (*a–c*), 9.57 m s^{-1} (*d–f*) and 16.59 m s^{-1} (*g–i*) along with their corresponding experimental data for comparison. The stress profiles are all scaled by the total stress. For reference, the instantaneous streamwise velocity fields are also plotted on the top panel. Figure 9 illustrates the instantaneous spatial distributions of the tangential stress, successfully reconstructed by the ML model. This reconstruction exhibits significant alignment with the experimental reference data, even though there are minor discrepancies in the peak magnitudes. The model is particularly effective in accurately predicting airflow separation events, reflecting its robust performance in this area.

In the context of airflow separation, we define this phenomenon as occurring when the near-surface, high-shear layer, typically indicative of an attached boundary layer, is propelled away from the surface, causing the surface tangential stress to reduce to a near-zero value within the sheltered region. Over wind waves, the behaviour is distinctly illustrated in figure 9. Here, the skin-friction drag reaches its peak value on the windward side of the wave, but the occurrence of airflow separation triggers a considerable drop in surface stress at the separation point. Subsequently, the stress is significantly reduced, approaching zero (or even becoming negative) on the leeward side, maintaining this minimal value within the separated region. Figure 9 also demonstrates that the contribution of viscous stress to total stress diminishes with increasing wind speed. These features, which are of paramount importance from a dynamic perspective, pose formidable challenges when it comes to accurate measurement in either laboratory environments or full-scale field settings, as well as reproduction through physics-based numerical simulations. Despite these challenges, the proposed ML model has demonstrated its

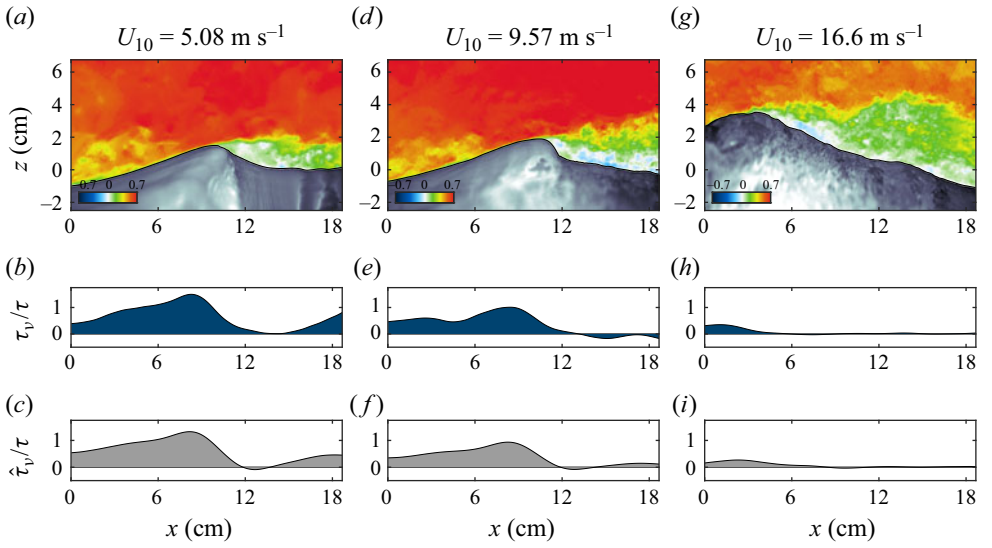


Figure 9. Spatial distributions of the instantaneous normalized skin-friction drag (*b*,*e*,*h*) obtained from experimental measurements and (*c*,*f*,*i*) reconstructed by the ML model over wind waves for the wind-wave experimental conditions of W1-05 (*a–c*), W1-09 (*d–f*) and W1-16 (*g–i*) with corresponding wind speeds of 5.08, 9.57 and 16.59 m s^{-1}. The stress profiles are all scaled by the total wind stress. (*a*,*d*,*g*) Normalized instantaneous streamwise velocity fields, u/U_{10}, are plotted on the top panels for reference. The ML model accurately predicts significant flow features, for instance, airflow separation on the leeward side of the wave.

proficiency by reconstructing the surface viscous stress over strongly forced wind waves – waves that are subject to frequent and intermittent airflow separations.

In order to provide context for the error distributions observed in the instantaneous stress profiles, we now focus on model predictions classified into three error thresholds: low, moderate and high. These are compared against the reference experimental data for the W1-05 dataset with $U_{10} = 5.08$ m s^{-1}. Figure 10 presents the comparison between the ML-reconstructed and experimental measurements of the instantaneous along-wave skin-friction drags. Specifically, the errors are partitioned into three categories:

(i) low-level error with $\delta_{e,1} \approx 10\,\%$ and $R^2 \approx 0.99$ (left-hand column);
(ii) moderate-level error with $\delta_{e,1} \approx 25\,\%$ and $R^2 \approx 0.93$ (middle column); and
(iii) high-level error with $\delta_{e,1} \approx 65\,\%$ and $R^2 \approx 0.65$ (right-hand column).

All profiles are normalized by the total stress, and the corresponding instantaneous horizontal velocity fields are shown in figure 10(*a*,*e*,*i*) for reference. Even when considering higher error thresholds, as illustrated in figure 10(*i–l*), the model maintains acceptable accuracy in mapping the general spatial pattern of the surface viscous stress, though slight deviations are detected.

For the separating wind wave under study, the model provides an adequately accurate prediction of the airflow separation event. At the separation point, the prominent viscous stress downwind of the wave crest drops to a minimal value and remains a minor portion of the total stress within the separation bubble. However, the model falls short of replicating a particular phenomenon observed in the experimental data, where the ejection of low-velocity fluid from the water surface on the windward side of the wave crest induces a subtle reduction in surface stress. Overall, the ML model does not fully succeed in encapsulating the influence of such a bursting process on surface stress, a shortfall most

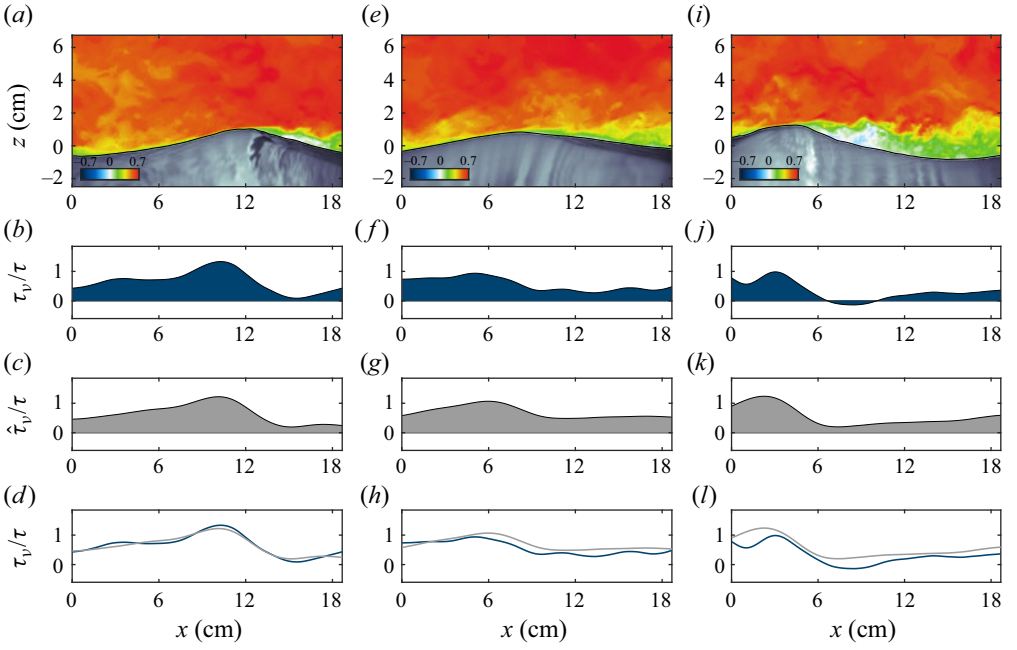


Figure 10. Spatial distributions of instantaneous normalized skin-friction drag obtained using (*b*,*f*,*j*) experimental measurements, τ_v/τ and (*c*,*g*,*k*) ML model predictions, $\hat{\tau}_v/\tau$, for the W1-05 case with $U_{10} = 5.08$ m s^{-1} at different error levels of $\delta_{e,1} \approx 10\,\%$ and $R^2 \approx 0.99$ (*a–d*), $\delta_{e,1} \approx 25\,\%$ and $R^2 \approx 0.93$ (*e–h*) and $\delta_{e,1} \approx 65\,\%$ and $R^2 \approx 0.65$ (*i–l*). Here, all profiles are normalized by the total stress, $\tau = \rho u_*^2$, and for reference, the instantaneous horizontal velocity fields are plotted in the top panels (*a*,*e*,*i*). In panels (*d*,*h*,*l*), ML reconstructed and experimental measurements of along-wave viscous stress profiles are also compared. Even for predictions with higher error thresholds, the model accurately captures the overall trend of the along-wave surface viscous stress distribution.

pronounced in scenarios involving high wind speeds. For more information on the bursting process over surface waves, readers are directed to works such as Kawamura & Toba (1988) and Yang & Shen (2010). Importantly, the proposed examination of instantaneous stress profiles across varying error thresholds highlights that the previous analysis, which depended solely on relative error metrics, conceals some of the essential competencies of the ML model. This realization underscores the necessity for a more sophisticated and nuanced evaluation of ML models, tailored to the specific challenges and requirements of the problem at hand.

3.3. *Model generalizations*

In this section, we assess the ability of the ML model to predict the local skin-friction drag under wind-wave conditions beyond those used for training and validation. To this end, a comprehensive evaluation of the model's performance in both interpolation, where training and test inputs originate from the same distribution, and extrapolation, where test inputs represent out-of-distribution instances, is essential. This dual analysis forms the cornerstone of the ML model's generalization capabilities, in line with the principles outlined, for example, in Silverman (1986) and Barnard & Wessels (1992). To evaluate the model's interpolation abilities, we examine its performance using the W1-14 case with a wind speed of $U_{10} = 14.82$ m s^{-1}. This case corresponds to strongly forced wind waves and falls within the range of wind-wave conditions utilized during training. Furthermore, we explore the model's extrapolation capabilities by evaluating

its performance on the low wind speed case of W1-02 with a wind speed of $U_{10} = 2.25$ m s^{-1}.

In particular, the lowest wind speed case represents an entirely different wind-wave regime in which nascent short surface waves emerge and almost no waves (less than approximately 1 %) undergo airflow separation downwind of wave crests. However, in other cases, airflow separation occurs over a significant portion of waves; for example, in high wind speed cases of $U_{10} = 14.82$ and 16.59 m s^{-1}, airflow separation occurs over nearly 90 % of waves (see Buckley *et al.* 2020; Yousefi 2020). Although the assessment of the model using unseen datasets within and out-of-distribution ranges of wind speeds used for the model training reinforces confidence in the model's adaptability, it is important to note that these unseen datasets still cover wind-generated wave regimes. Additionally, these cases consist of experimental conditions in which wind and waves are in the same direction. Due to the limitations of the experimental dataset, we were thus unable to consider the effects of wind-wave misalignment, mechanical swells and a more complex sea that involves the combination of wind waves and swells.

The error metrics shown in table 2 shed light on the model's performance in predicting wind-wave conditions for the considered wind speeds of 14.82 m s^{-1} (W1-14 case) and 2.25 m s^{-1} (W1-02 case). The relative mean errors are found to be 31.9 % and 34.2 %, respectively, coupled with corresponding R^2 values of 0.866 and 0.836. In the averaged sense, the model interpolation and extrapolation predictions show a degree of accuracy mirroring the estimations for surface stress in the wind-wave conditions that formed the training set. For the W1-14 case, the RMAE and R^2 values of the ML-model predictions display a marginal improvement compared with the W1-16 case, which features a higher U_{10} value. Focusing on the low wind speed scenario where $U_{10} = 2.25$ m s^{-1}, a slight decline in model performance is observable. This reduction in efficacy is attributed to the distinct nature of the low wind speed case, predominantly comprising short gravity waves with rare occurrences of airflow separation. Such a scenario contrasts with the conditions in which the model excels, particularly over separating wind waves. This behaviour underlines the importance of a comprehensive understanding of the model's behaviour across varied wind-wave scenarios and sets the stage for future refinements to enhance its general applicability and accuracy.

Further, RMAE and R^2 error metrics are calculated for individual wave profiles, and their corresponding p.d.f.s are shown in figures 11(*a*) and 11(*b*), respectively, for wind-wave cases with $U_{10} = 2.25$ and 14.82 m s^{-1}. Probability density functions, therefore, evaluate the distribution of various error levels across individual waves. The peaks in the p.d.f. distributions represent the most probable RMAE and R^2 metrics in the dataset, which are $\delta_{e,1} \approx 33$ % and $R^2 \approx 0.87$ for the case of W1-02 and $\delta_{e,1} \approx 25$ % and $R^2 = 0.93$ for the case of W1-14. This indicates that the accuracy of the ML model in reconstructing skin-friction drags for the interpolation case is better than the extrapolation one. However, we should consider different wave regimes and the varying wave patterns that are observed between these two cases. As expected, the error distribution in the high wind speed case of W1-14 is closely similar to the other cases used during the training process. One difference, however, is that, in the case of ML interpolation, the distributions are more heavy tailed and we observed a higher number of stress profiles with declined error metrics, i.e. $R^2 < 0.4$. The phase-binned average distributions of RMAE and R^2 are also plotted in figures 12(*a*) and 12(*b*), respectively, for wind-wave conditions of W1-02 and W1-14. Again, there are 108 phase bins with a phase interval of 5.82×10^{-2} rad. In accordance with the mean values reported in table 2 and figure 11, the error across phases for the interpolation task is smaller than the extrapolation one.

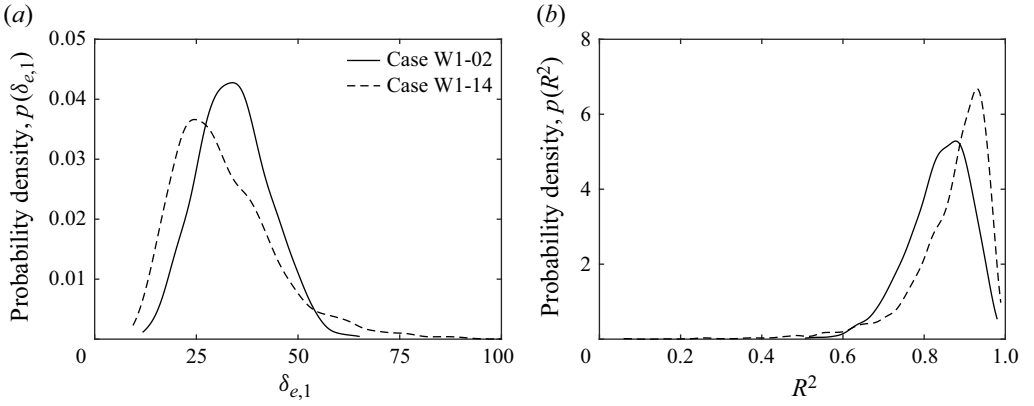


Figure 11. The p.d.f.s of (*a*) RMAE, $\delta_{e,1}$, and (*b*) coefficient of determination, R^2, for W1-02 and W1-14 cases with a wind speed of $U_{10} = 2.25$ and 14.82 m s^{-1}, respectively. The RMAE and R^2 error metrics were calculated for each instantaneous wave in the dataset, and their distribution of probabilities is shown here. Here, W1-02 and W1-14 cases were used as out-of-training-distribution (entirely separate) datasets to examine the model performance on unseen data.

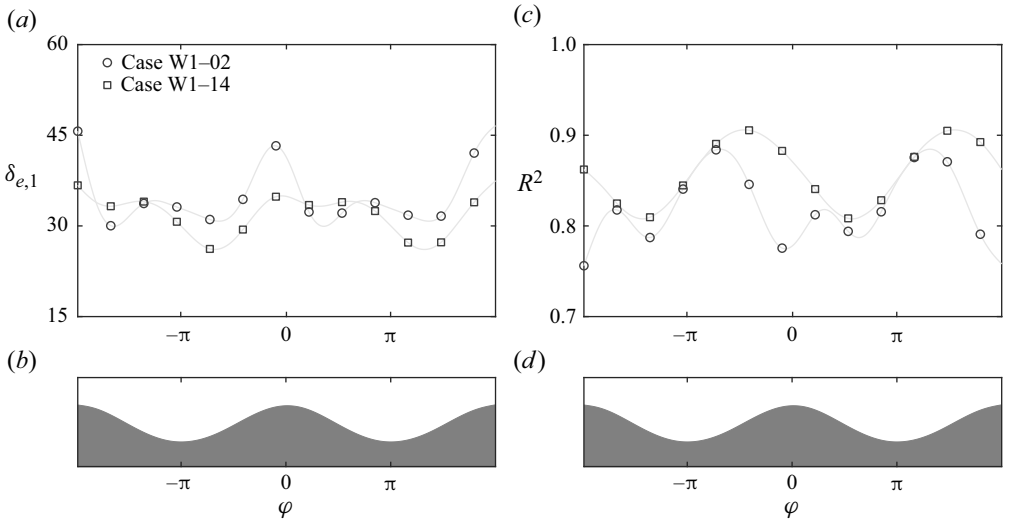


Figure 12. Phase-averaged (*a*) RMAE, $\delta_{e,1}$, and (*c*) coefficient of determination error, R^2, for W1-02 and W1-14 cases with a wind speed of 2.25 and 14.82 m s^{-1}, respectively. The RMAE and R^2 metrics were calculated for each instantaneous wave in the out-of-distribution datasets and bin averaged based on the corresponding wave phases. There are 108 phase bins with a phase interval of 5.82×10^{-2} rad. Panels (*b*,*d*) show a sketch of the mean wave profile to better visualize wave phases.

Next, we assess the interpolation and extrapolation capabilities of the CNN-based ML model in reconstructing the spatial distributions of the instantaneous skin-friction drag. Sample predictions of the instantaneous τ_v/τ with the corresponding experimental measurements are presented in figures 13 and 14 for the wind wave cases of W1-02 (extrapolation dataset) and W1-14 (interpolation dataset), respectively. Again, the top panels show the instantaneous normalized streamwise velocity fields, u/U_{10}. In the case of W1-02 with a low wind speed of $U_{10} = 2.25$ m s^{-1}, the model accurately reconstructs the stress profiles, although with a slight underprediction of peak values. We note here

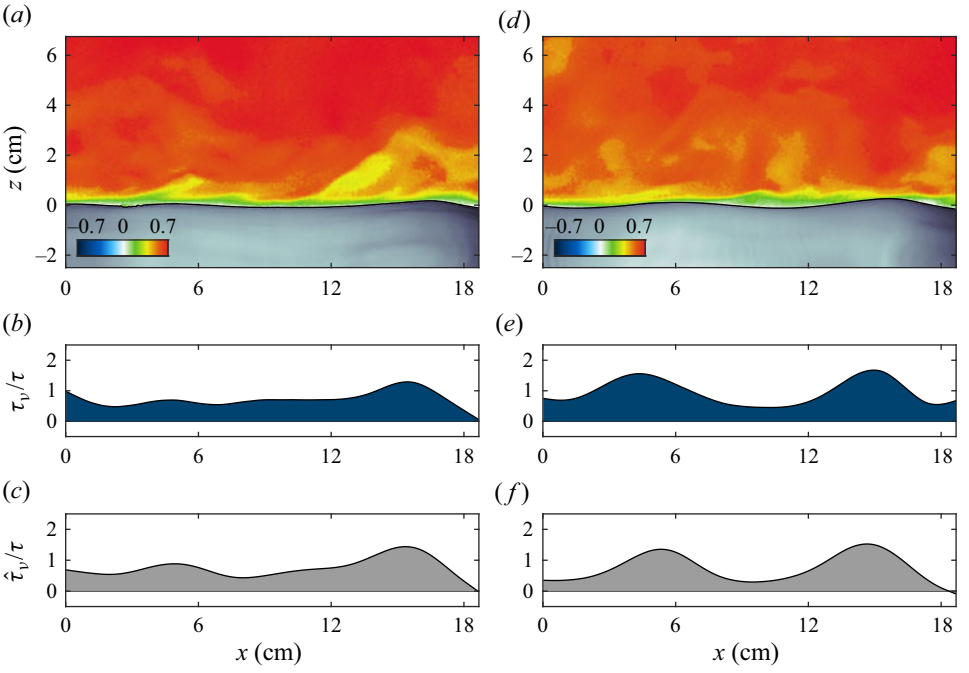


Figure 13. Comparisons between along-wave distributions of the instantaneous normalized skin-friction drag (*b*,*e*) obtained from experimental measurements and (*c*,*f*) the fields reconstructed by the CNN-based ML model over surface wind waves of the W1-02 case with a wind speed of 2.25 m s^{-1}. The stress profiles are scaled by the total stress, $\tau = \rho u_*^2$. (*a*,*d*) Normalized instantaneous streamwise velocity fields, u/U_{10}, are also plotted for reference. The model accurately predicts the effects of near-surface sweep and ejection processes on the surface viscous stresses.

that the underlying physical mechanisms in the low wind speed case, leading to the spatial variations in the skin-friction drag (as shown in figure 13), are quite different compared with the high wind speed cases. In particular, in low wind speed conditions, almost no wave experiences separation because airflow does not completely detach from the water surface, as shown in figure 13(*a*,*d*), and the surface viscous stress does not fully drop to a zero value. Nonetheless, we can notice that high-shear stresses predominantly drop past wave crests, which is likely due to surface-induced ejection and sweep events. From the turbulent structures shown figure 13(*a*,*d*), it is observed that the low-velocity fluid is intermittently ejected away from the water surface along the downwind face of the waves, and the higher-velocity fluid is swept downward. This is particularly evident from figure 13(*a*). Such bursting events are also characteristic of the near-wall region in classical turbulent boundary layer flows over solid flat surfaces (e.g. Robinson 1991; Meinhart & Adrian 1995; Adrian 2007; Jiménez 2012).

Figure 14 shows model predictions of the along-wave instantaneous τ_ν/τ over a portion of the wave where the flow is not separating (*a–c*) and over a separating wind wave (*d–f*) for the case of W1-14 with a wind speed of 14.85 m s^{-1}. The reconstructed fields are also compared with experimental measurements. Overall, it is evident that the model can reconstruct the essential characteristics of the surface stress with relatively high accuracy. Over the separating wind wave (see figure 14*f*), the model accurately predicts the location at which the flow detaches from the surface and causes a collapse of the near-surface stress. The model also accurately predicts the near-zero viscous stress on the leeward side of the

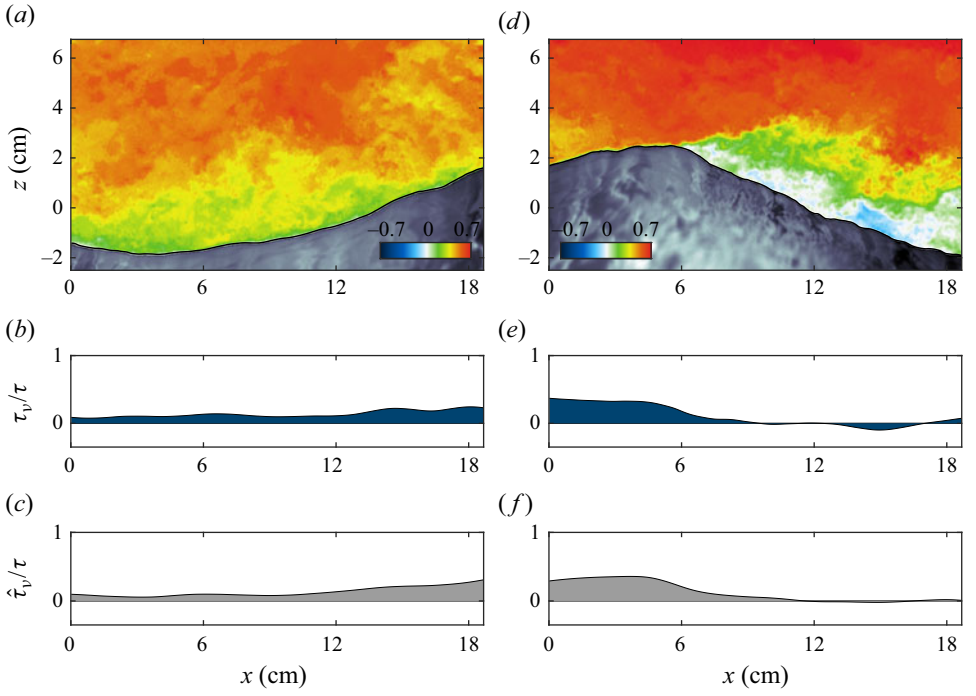


Figure 14. Comparisons between along-wave distributions of the instantaneous normalized skin-friction drag (*b*,*e*) obtained from experimental measurements and (*c*,*f*) the fields reconstructed by the CNN-based ML model over non-separating (*a–c*) and separating (*d–f*) wind waves for the wind-wave condition of W1-14 with a wind speed of 14.85 m s^{-1}. The stress profiles are scaled by the total stress, $\tau = \rho u_*^2$. (*a*,*d*) Normalized instantaneous streamwise velocity fields, u/U_{10}, are also plotted for reference. The model accurately captures the effects of airflow separation on the surface viscous stresses.

wave within the separated region and the gradual re-generation of the viscous layer past the leeward side of the wave (see figure 14*c*). For more discussions on near-surface tangential stress, the reader is referred to publications by Reul, Branger & Giovanangeli (1999, 2008), Veron, Saxena & Misra (2007) and Buckley *et al.* (2020). Overall, we observe that the model can infer the skin-friction drags for out-of-training-distribution datasets with high accuracy and that the slight increase in error metrics is mainly due to the differences in peak values of the reconstructed and measured stress profiles.

Beyond the turbulence near the surface, the drag force generated by surface waves across multiple wavelengths also exhibits a multiscale nature. This complexity poses a challenge in understanding the skin-friction drag across diverse scales. To demonstrate the model's capability in capturing the multiscale drag characteristic, we present in figure 15 the one-dimensional spectrum of surface tangential stress relative to wavelength, averaged across all instantaneous waves, for the high wind speed condition of W1-14 (interpolation dataset with a wind speed of 14.82 m s^{-1}). A solid black line represents the spectrum derived from experimental measurements, while the model's prediction is indicated by a dashed line. It should be noted here that the experimental spectrum shown in this figure is based on the filtered raw data, as explained in § 2.2.1. A visual examination of these spectra offers compelling evidence that the CNN model's predictions align remarkably well with the experimental reference data. The model reconstructed and measured viscous stresses both show power law distributions with an asymptote of k^{-7} across a wide range of wavenumbers, highlighting the presence of a self-similar dynamics

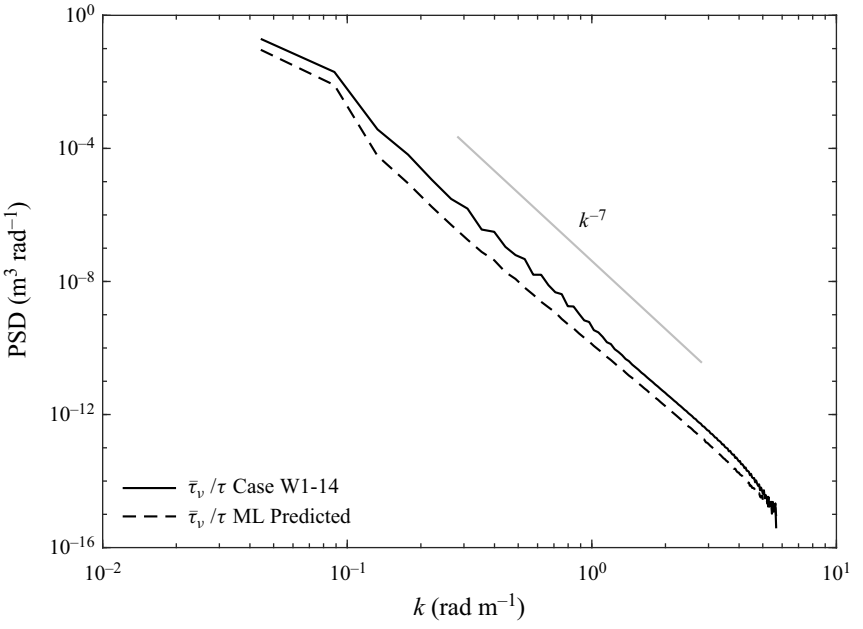


Figure 15. One-dimensional spectrum of the ML reconstructed skin-friction drag compared with the ground-truth experimental measurements for the wind-wave case of f W1-14 with a wind speed of $U_{10} = 14.82$ m s^{-1}. The surface stress spectra, averaged over all instantaneous wave profiles, are plotted with respect to wavenumbers. The power-law fit with a slope of k^{-7} is also indicated by the grey solid line.

in the airflow–wave system at these scales (Anderson & Meneveau 2011). It is worth noting here that, to the best knowledge of the authors, this is the first time the behaviour of surface tangential stress across different wavelength ranges has been quantified. Finally, regarding the scale-wise performance of the model in the case of interpolation, we observe that the difference between $\hat{\tau}_\nu$ and τ_ν at small wavenumbers is generally higher than those at high wavenumbers, indicating that the small-scale spatial frequency of the surface stress can be predicted more accurately than large-scale one. Similar performance was also observed in the case of W1-02 (i.e. the extrapolation dataset), which is not shown here for brevity. Such findings emphasize the model's effectiveness in capturing the multifaceted drag contributions across scales.

3.4. *Mean skin-friction drag*

We now examine the accuracy of the model in predicting the mean surface stress across various wind wave conditions and compare findings with results from other scholarly investigations. A mean quantity, denoted with an overbar, is here defined as the ensemble average of along-wave stress profiles over instantaneous waves. Figure 16(*a*) shows the mean viscous stresses estimated by the CNN model, relating them to the friction velocity across different cases. These quantities are shown alongside available measurements from the literature. The skin-friction drag increases almost linearly with friction velocity until it reaches a saturation threshold for strongly forced wind waves. The saturation phenomenon observed at high wind speeds can be ascribed to intermittent airflow separation and breaking events. This behaviour has been thoroughly analysed in Yousefi *et al.* (2020) (see also Bopp 2018; Buckley *et al.* 2020). For the highest wind speed of $U_{10} = 16.59$ m s^{-1},

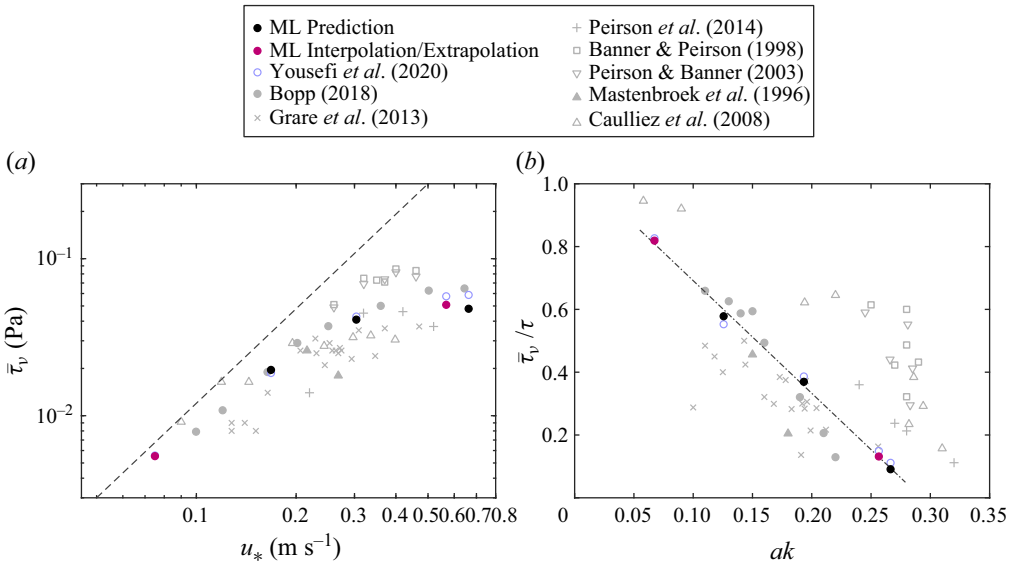


Figure 16. Mean skin-friction drag, $\bar{\tau}_v$, plotted as a function of (*a*) friction velocity and (*b*) wave slope. The skin-friction drags in panel (*b*) are normalized by the total stress, $\tau = \rho u_*^2$. Black circles indicate the model predictions, while solid magenta circles show the model interpolation and extrapolation. The black dashed line in panel (*a*) indicates the total wind stress, and the black dash-dotted line in panel (*b*) is the best log–linear fit to the experimental data of Yousefi *et al.* (2020). For comparison purposes, the measurements performed by Mastenbroek *et al.* (1996), Banner & Peirson (1998), Peirson & Banner (2003), Caulliez, Makin & Kudryavtsev (2008), Grare *et al.* (2013), Peirson, Walker & Banner (2014), Bopp (2018) and Yousefi *et al.* (2020) are also shown. The results of Mastenbroek *et al.* (1996) show the difference between total stress and form drag. Also, the results of Banner & Peirson (1998) and Peirson *et al.* (2014) were obtained in water at short fetches.

the model prediction of $\bar{\tau}_v$ is slightly lower than the results of Bopp (2018) and Yousefi *et al.* (2020) with $\delta_{e,1} \approx 15\,\%$, possibly due to the lack of data in the test dataset. Nonetheless, the model estimates of the mean skin-friction drag are in good general agreement with previous measurements.

In general, the drag at the air–sea interface is a complex function of wind speed, wave age, wave slope, and wind-wave alignment (Sullivan & McWilliams 2010). Therefore, wind speed alone does not necessarily capture the complexity of the air–sea interface in different wind and wave conditions. To further assess the collapse of model predictions with other parameters, we present the mean skin-friction drag contributions to the total stress as a function of wave slope, ak, in figure 16(*b*). We note here that, in these young wind-wave cases, the wave slope and wave age are tightly correlated, so we focus the analysis on the wave slope dependence only as a convenient parameter. The decrease in the normalized skin-friction drag with increasing wave slope indicates that as waves grow and become steeper at high wind speeds, the contribution of viscosity to the total stress reduces. The ML model estimates of skin-friction drag show a decrease with increasing slope and fall quite well within previous laboratory measurements. It should also be noted that skin-friction drag measurements presented in figure 16 generally exhibit a large scatter when plotted against both wind speed and wave slope because, in part, the measurements in different studies were collected at different fetches. Concluding our analysis, we underline the exceptional performance of the model in the precise estimation of mean surface stress, both for the interpolation and extrapolation scenarios, with $\delta_{e,1} \approx 10\,\%$ and $\delta_{e,1} \approx 3\,\%$, respectively.

4. Discussion

In this section, we further discuss the presented results and explore potential factors that may have contributed to the observed variances in the model predictions of skin-friction drag. Findings discussed in § 3 demonstrated that the proposed CNN-based ML model accurately reconstructs the skin-friction drag across different wind-wave regimes and adequately captures the complex nature of near-surface processes, such as airflow separation and its impact on the surface stress distribution. However, it is important to note that, despite this high level of agreement between the model predictions and the ground-truth experimental data, we observed discrepancies, particularly in the peak magnitudes of the stress profiles. Furthermore, a discernible trend emerged wherein the model's accuracy marginally diminished as wind speed increased. This pattern was not confined to the training data but extended to the unseen dataset characterized by a wind speed of 14.82 m s^{-1}, a matter discussed in more detail in § 3.3. In what follows, we explore potential approaches that may be employed to overcome these deficiencies and enhance model performance.

Overall, the decrease in the performance of the model with increasing wind speed can be attributed, in part, to the limitations of the PIV measurements. In low–moderate wind speed cases, the PIV field of view captures one–two wavelengths of the waves. As wind speed increases, the wavelength of the disturbance also increases such that only a fraction of a wavelength (approximately 1/3–1/4) can be captured by PIV images in high wind speed cases. This leads to a partial representation of the wave profile and the corresponding surface stress distribution, which negatively impacts the accuracy of the model. Furthermore, the increased signal-to-noise ratio of experimental measurements at high wind speeds may also contribute to the increased error metrics of the model. Airflow separation events occurring past wave crests result in the detachment of relatively stable, high-shear layers, destabilizing them in the process. This detachment triggers a subsequent disruption of the near-surface shear layers into vortical eddies downwind of the waves, thereby enhancing turbulence. This phenomenon is linked to the characteristics of detached free shear layers, known to be potent generators of intense turbulence, as referenced in seminal studies by Ho & Huang (1982) and Ho & Huerre (1984). The likelihood of airflow separation events increases with an increase in wind speed. While the airflow separates over less than 1 % of waves at the lowest wind speed, over 85 % of waves experience airflow separation at high wind speed cases. Consequently, at higher wind speeds, instantaneous near-surface measurements of velocity fields may contain localized areas where the signal-to-noise ratio is particularly low due to either insufficient seed density or laser reflections induced by the multiscale air–water interface. This elevated noise level in measurements presents a challenge for the model in discerning the underlying patterns and relationships amongst the measured variables. Notably, we observed a particular sensitivity within the model to this measurement noise. A pathway to improving the model's performance could involve refining the speed and resolution of PIV measurements in the vicinity of the air–water interface while simultaneously increasing the PIV field of view to capture multiple wavelengths in high wind speed scenarios. From a modelling perspective, the integration of a more noise-tolerant algorithm could provide a valuable means of countering the difficulties posed by increased measurement noise. This approach could improve the robustness of the model, enabling more reliable predictions under diverse conditions.

The drag induced by surface waves of multiple wavelengths moving over the peak wavelength fundamentally represents a multiscale problem. As depicted in figure 15, the model slightly underestimates the skin-friction drag across a spectrum of wavenumbers

compared with the experimental results. In this work, we utilized a CNN-based model with a constant filter size (F) and a predetermined depth, indicating the number of convolutional layers. Although this model successfully retrieves single-scale contextual data for reconstructing skin-friction drag, it may inadvertently overlook the intricacies of its inherent multiscale character. The efficacy of multiscale CNNs has been proven in the field of computer vision, specifically for tasks involving image super-resolution (e.g. Ren *et al.* 2016; Du *et al.* 2018; Li *et al.* 2018). Moreover, a series of studies have validated the effectiveness of multiscale CNNs in turbulence enrichment (e.g. Fukami *et al.* 2019, 2021; Bi *et al.* 2022). Therefore, incorporating a multiscale CNN approach for skin-friction drag reconstructions could potentially enhance the model's performance and appropriately address the complex interactions prevalent across diverse scales.

We also noted a decrease in the model's performance for cases involving the low wind speed of 2.25 m s^{-1}. As previously discussed, this can partly be attributed to the fact that this case predominantly features short gravity waves with minimal occurrences of airflow separation events. Meanwhile, the model has been trained on waves with moderate to high wind speeds, where separation takes place over a considerable fraction of waves. To bolster the model's accuracy in low wind speed conditions, the use of the transfer learning technique could present a viable solution (see Pan & Yang 2009). The application of this method has demonstrated success in various fluid dynamics studies (e.g. Guastoni *et al.* 2021; Yousif, Yu & Lim 2021; Lee *et al.* 2022). Transfer learning, an ML technique that leverages the knowledge acquired from one task/domain to be applied to another related task/domain, reduces the required training data and improves model applicability for different applications. For example, Guastoni *et al.* (2021) successfully applied transfer learning to a CNN-based ML model to predict velocity fluctuations in turbulent channel flow for friction Reynolds number of $Re_\tau = 550$ by utilizing the model trained on a dataset corresponding to $Re_\tau = 180$. In the context of the present work, the model may be re-trained using a small dataset specific to the $U_{10} = 2.25$ m s^{-1} case. This approach enables the model to make improved predictions for scenarios involving short gravity waves, often encountered in low wind speed cases.

Lastly, another source of error in our model may arise from the fact that the proposed ML model lacks an intrinsic understanding of the temporal dynamics of air–water interactions. The current approach maps instantaneous wave profiles directly to the corresponding instantaneous shear stresses without considering the complex interplay of memory effects and historical events that lead to specific wave profiles and shear stress distributions. This situation is akin to different flow dynamics producing different shear stress distributions for an identical wave profile, which affects the accuracy of the proposed one-to-one mapping. This aspect is particularly relevant for modes of shear stress variability that do not leave a strong imprint on the wave profile. Complicating matters further is the multiscale, chaotic nature of air–water interaction processes. These interactions are characterized by a broad array of variability modes, both spatial and temporal, and inherent randomness in both the airflow (turbulence) and the propagating wave field – elements that our model currently fails to capture. Such challenges mirror the limitations faced in equilibrium wall-layer models, which often fall short in capturing the unsteadiness in flow statistics, as highlighted by Piomelli (2008). The lack of temporal awareness in our model could potentially be addressed in future research by leveraging long–short-term memory networks (e.g. Hochreiter & Schmidhuber 1997), temporal convolutional networks (e.g. Bai, Kolter & Koltun 2018) or transformer-based models (e.g. Vaswani *et al.* 2017). These methods have all demonstrated exemplary performance in capturing the spatio-temporal dynamics for fluid flow systems (see, e.g. Deng *et al.* 2019; Wu *et al.* 2020; Yousif *et al.* 2023). Additionally, integrating physics-based constraints to incorporate dynamical

considerations into the loss function as a regularization term can significantly enhance the model's prediction ability and improve generalization by aligning the predictions with fundamental physical principles governing fluid flows (e.g. Karniadakis *et al.* 2021).

5. Conclusions

In this study, we introduced a CNN-based ML model designed to reconstruct the along-wave distributions of skin-friction drag over wind waves using only wave elevations and wave ages. We also discussed links between reconstruction accuracy and the underlying flow physics. The CNN model consists of convolutional, dense and Mish nonlinear layers that extract information from a two-dimensional discrete grid of wind-wave variables, including the wave profile and the resulting local wave slope and wave phases, wave age and surface viscous stress. The model was trained and evaluated through high-resolution PIV measurements under various wind-wave conditions corresponding to wind speeds varying from 5.08 to 16.59 m s^{-1}. The generalization abilities of the model, including extrapolation and interpolation capabilities, were also assessed for wind speed cases with a wind speed of 2.25 and 14.80 m s^{-1}, respectively.

The proposed model was able to accurately reconstruct the skin-friction drag, particularly for cases within the training dataset with wind speeds of 5.08, 9.57, and 16.59 m s^{-1}. The model's performance was assessed through a thorough error analysis using relative mean and coefficient of determination error metrics and a qualitative examination of instantaneous stress profiles. Despite the model's overall competency, a slight increase in error metrics was observed with increasing wind speed. The maximum relative mean error observed in the test dataset was around 34 %, corresponding to a minimum coefficient of determination of approximately 0.85. Additionally, the phase-binned error metrics demonstrated the peak performance of the model upwind of wave crests, whereas its accuracy is reduced on the leeward side of the waves, where airflow separation events occur. Conversely, although the model's overall performance decreased for the high wind speed case of 16.59 m s^{-1}, it exhibited improved accuracy within the separated region downwind of waves. Further, visual inspection of the instantaneous along-wave distributions of surface viscous stress revealed that although reconstructed stress profiles closely aligned with corresponding experimental measurements, there were slight discrepancies in terms of peak magnitudes. The model also accurately predicted mean skin-friction drags, with the exception of the high wind speed case of 16.59 m s^{-1}, where a slight divergence was observed compared with experimental results.

Beyond the assessment of the performance expectations of surface stresses within the training data, we also examined the model's generalization capabilities on wind-wave cases with wind speeds of 2.25 and 14.82 m s^{-1}, which were not included during the training phase. Both qualitative and quantitative analysis of the reconstructed instantaneous along-wave surface viscous stress profiles showed a commendable agreement with the corresponding experimental measurements. However, this came with a slight decrease in performance compared with the cases included in the training set. The performance for the wind-wave case with $U_{10} = 14.82$ m s^{-1} showed a minor improvement compared with the higher wind speed of 16.59 m s^{-1}, thereby reinforcing the observed trend of a gradual performance decrease with increasing wind speed. Nonetheless, there was a slight decline in the model's performance for the case with $U_{10} = 2.25$ m s^{-1}. This behaviour can be attributed to the fact that this low wind speed

case primarily consists of short gravity waves with rare instances of airflow separation, which the ML model did not encounter during training.

We also examined the one-dimensional spectrum of surface tangential stress in relation to wavelength to compare the stresses predicted by the ML model with the measured viscous stresses. This revealed a very good agreement between the two spectra. Both the reconstructed and experimental stress spectrum profiles demonstrated power-law distributions with several decades of k^{-7} range, suggesting the ML model has learned this scaling behaviour. Lastly, a comparison between the mean skin-friction drag deduced from our ML model and the values reported in previous research indicated a precise match for the case of $U_{10} = 2.25$ m s^{-1} and a minor divergence for the case of $U_{10} = 14.80$ m s^{-1}. Despite this, the error range remained within acceptable limits. These findings underscore the ability of our model to reliably predict skin-friction drag for wind speeds within and beyond the training dataset. This ability to generalize beyond the training cases is important as it makes the proposed model suitable for skin-drag predictions.

In summary, the proposed CNN model presents a robust tool for deducing and reconstructing skin friction drag over wind waves based on wave profiles and wind speeds. This model has demonstrated exceptional performance in replicating experimental measurements and offering valuable insights into surface stress distributions. To the best of the author's knowledge, this work marks the first instance in scientific literature where sea-surface stresses have been modelled using a data-driven approach. The model's ability to accurately describe wind drag over surface waves has significant implications, particularly in the context of wall-layer models for LES and numerical modelling approaches. By integrating the proposed CNN model (or variations thereof), the accuracy of these simulations can be enhanced and provide more accurate and detailed predictions of wind-wave interactions and their impact on energy transport, momentum exchange and gas transfer processes at the air–sea interface.

Funding. This research is supported by the National Science Foundation (NSF) under grant number 2030859 to the Computing Research Association (CRA) for the Computing Innovation Fellows Project. This research was also partially funded by the National Institute of Standards and Technology under grant number 70NANB22H057. H.Y. acknowledges support from the Columbia University Summer at SEAS Program. This work used the Anvil supercomputer at Purdue University through allocation ATM180022 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support (ACCESS) program, which is supported by NSF grants 2138259, 2138286, 2138307, 2137603 and 2138296.

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

Author ORCIDs.
Kianoosh Yousefi https://orcid.org/0000-0003-4729-2146;
Gurpreet Singh Hora https://orcid.org/0009-0004-7894-5175;
Fabrice Veron https://orcid.org/0000-0001-5514-9767;
Marco G. Giometto https://orcid.org/0000-0001-9661-0599.

Author contributions. K.Y. and G.S.H. contributed equally to this work.

Appendix. Hyperparameter optimization

During our hyperparameter optimization process, we made significant efforts to enhance the architecture of the ML model to improve its performance. We conducted a thorough exploration by experimenting with various hyperparameter combinations, aiming to identify the optimal configuration for our model. In our exploration, we focused on several key hyperparameters, including the number of convolutional layers, kernels per

layer (K), filter dimension (F), batch size (B) and learning rate (LR). To determine the optimal number of convolutional layers for our model, we evaluated three different layer configurations: 2, 3 and 5 layers. Additionally, we explored the impact of varying the K, considering values of either 8 or 16. Furthermore, we experimented with different filter dimensions, i.e. F within the range of 3 to 11. These hyperparameters allow us to assess the model's capacity to learn and represent complex patterns in the data and gain valuable insights into the model's ability to capture fine-grained details and extract relevant features (Simonyan & Zisserman 2014; Goodfellow *et al.* 2016). In addition to architectural considerations, we investigated the effect of different batch sizes, explicitly evaluating the performance using batch sizes 32 and 64; these are either the same as or greater than commonly adopted sizes for ML tasks involving spatio-temporal data (see, e.g. Du *et al.* 2018; Guastoni *et al.* 2021; Kim *et al.* 2021; Bi *et al.* 2022). By exploring this parameter, we aimed to strike a balance between computational efficiency and model convergence, ensuring optimal training dynamics for our specific problem (LeCun *et al.* 2002; Keskar *et al.* 2016). Finally, we considered different learning rates, including values of 5×10^{-4}, 1×10^{-3}, and 1×10^{-2} to find an optimal value that facilitated effective training and convergence (Ruder 2016).

A random search methodology as described in Bergstra & Bengio (2012) and Goodfellow *et al.* (2016) is employed to ensure a comprehensive exploration of the hyperparameter space, and early stopping (Hastie *et al.* 2009) was used as a termination criterion of the ML training. The hyperparameter optimization analysis results are presented in table 3, summarizing the performance metrics achieved for each combination of hyperparameters.

Analysis of the results presented in table 3 indicates that augmenting the model capacity by increasing the number of layers significantly enhances the model's overall performance. This phenomenon can be attributed to the well-known principle of bias–variance tradeoff (Briscoe & Feldman 2011), i.e. increasing the model capacity allows the model to capture more intricate and nuanced patterns within the data. In this study, we specifically set the number of convolutional layers to 5, leveraging the benefits of a deeper architecture. Regarding K, we found that increasing it from 8 to 16 did not significantly improve performance. Therefore, we set the value of K to 8, as it proved sufficient for capturing the desired patterns and features. In the case of F, we did not observe a clear trend with respect to the MSE. To strike a balance between computational cost and the risk of overfitting, we set the value of F to 3, ensuring a more efficient model architecture. It also aligns with the existing CNN studies, those who leveraged small filters with sizes between 3 and 7 (e.g. Simonyan & Zisserman 2014; Fukami *et al.* 2019, 2021; Xuan & Shen 2023). In our experimentation, we discovered an important observation related to using large values of F. Despite small MSE, we observed significant fluctuations in the output, which led us to speculate that employing large filters can potentially result in the loss of fine-grained details and local information.

The choice of the LR played a crucial role in model training. We found that higher LRs, such as 1×10^{-2} and 1×10^{-3}, resulted in large MSE values, indicating suboptimal performance. As we reduced the LR, the MSE decreased and demonstrated better performance. For this study, we fixed the LR at 5×10^{-4}, yielding better results in model convergence and performance. We opted for the large B, i.e. 64, as it improved the model performance, and a similar observation is supported by Smith *et al.* (2017), where the authors recommended increasing the batch size instead of decaying the learning rate. They argue that this approach can lead to faster convergence and better generalization performance.

Case	N	K	F	LR	B	MSE	MAE
1*	5	8	3	0.0005	64	0.0190	0.0957
1	5	8	3	0.0005	64	0.0147	0.0862
2	3	16	3	0.001	64	0.0192	0.0966
3	3	8	3	0.001	32	0.0186	0.0942
4	5	8	5	0.0005	64	0.0155	0.0909
5	2	8	5	0.001	32	0.0200	0.0981
6	3	16	5	0.0005	64	0.0197	0.0966
7	2	8	7	0.01	32	0.0724	0.2141
8	3	8	7	0.0005	64	0.0207	0.1030
9	5	8	7	0.0005	64	0.0198	0.0960
10	2	16	7	0.001	64	0.0212	0.1011
11	3	16	7	0.01	64	0.0724	0.2131
12	3	8	9	0.001	32	0.0246	0.1070
13	5	8	9	0.0005	64	0.0194	0.0947
14	3	16	9	0.0001	32	0.0214	0.1002
15	3	16	9	0.001	64	0.0223	0.1042
16	2	16	9	0.0005	32	0.0199	0.0979
17	5	16	9	0.001	32	0.0724	0.2134
18	5	8	11	0.001	64	0.0207	0.0994
19	5	16	11	0.001	64	0.0201	0.0974
20	2	16	11	0.001	64	0.0229	0.1046
21	5	16	11	0.001	64	0.0201	0.0974
22	3	16	11	0.001	32	0.0204	0.0952
23	2	16	11	0.0005	32	0.0190	0.0950

Table 3. Summary of error metrics, including MSE and mean absolute error (MAE), on the validation dataset for hyperparameter tuning. For all experiments, Mish nonlinearity (see Misra 2019) is leveraged except 1*, where Swish nonlinearity (see Ramachandran *et al.* 2017) is applied for comparison. Here, N represents the number of convolutional layers, K denotes the number of kernels, F indicates the dimension of the kernel, LR represents the learning rate and B represents the batch size.

To further enhance the generalization ability of the ML model, we investigated the impact of L_2 (ridge) regularizations alongside the original MSE (Hastie *et al.* 2009; Goodfellow *et al.* 2016; Hastie 2020). We systematically varied the regularization parameter $\lambda = 10^{-4}$, 10^{-3}, 10^{-2} and 10^{-1} and trained a total of four ML models (see (2.8)). Tables 4 and 5 represents the performance of the ML model employing the L_2 penalty, with a focus on R^2 and RMAE error metrics across different wind-wave experimental datasets. Analysis of the tables indicates that the model with $\lambda = 10^{-4}$ yields the highest accuracy for the considered datasets. Furthermore, it is evident that an increase in the value of λ leads to a stronger regularization effect, which corresponds to a decline in the model's accuracy. This is because, during the training phase, the ML model started emphasizing shrinking the weight, potentially resulting in suboptimal performance. We also investigated the use of L_1 penalty Tibshirani (1996) (not shown here) and found that L_2 regularization outperforms L_1 and overall improves the results.

By making informed choices based on these analyses, we aimed to optimize the model's hyperparameter configuration, considering computational constraints and the need to avoid overfitting. These decisions were guided by the empirical evidence obtained from our experiments and are expected to yield the best performance for our specific problem. We also acknowledge that more sophisticated hyperparameter tuning methods, such as Bayesian optimization and tree-based search, could be explored to improve the model's performance further (Snoek, Larochelle & Adams 2012; Olson *et al.* 2016).

	W1-05	W1-09	W1-16	W1-02	W1-14
$\lambda = 10^{-4}$	0.924	0.934	0.846	0.836	0.866
$\lambda = 10^{-3}$	0.901	0.883	0.744	0.801	0.809
$\lambda = 10^{-2}$	0.860	0.868	0.690	0.602	0.655
$\lambda = 10^{-1}$	0.707	0.813	0.565	0.536	0.626

Table 4. Summary of the ML model performance with ridge regularization in predicting skin friction drag compared with the experimental data. The coefficient of determination (R^2) metrics are presented for all wind-wave conditions with different values of λ. Also, the W1-02 and W1-14 cases are separate datasets for which the model has been applied to examine the model interpolation and extrapolation performance on unseen data.

	W1-05	W1-09	W1-16	W1-02	W1-14
$\lambda = 10^{-4}$	0.240	0.217	0.344	0.342	0.319
$\lambda = 10^{-3}$	0.259	0.294	0.452	0.375	0.384
$\lambda = 10^{-2}$	0.323	0.303	0.500	0.579	0.545
$\lambda = 10^{-1}$	0.491	0.372	0.628	0.643	0.570

Table 5. Summary of the ML model performance with ridge regularization in predicting skin friction drag compared with the experimental data. The RMAE metrics are presented for all wind-wave conditions with different values of λ. Also, the W1-02 and W1-14 cases are separate datasets for which the model has been applied to examine the model interpolation and extrapolation performance on unseen data.

All experiments were carried out on a system equipped with an NVIDIA Quadro P5000 graphics processing unit (16 gigabytes) and an Intel (R) Xeon (R) Gold 6126 CPU @ 2.60 GHz core processing unit and the wall clock time required to train each network is approximately 90 min.

References

Adrian, R.J. 2007 Hairpin vortex organization in wall turbulence. *Phys. Fluids* **19** (4), 041301.

Aiyer, A.K., Deike, L. & Mueller, M.E. 2023 A sea surface–based drag model for large-eddy simulation of wind–wave interaction. *J. Atmos. Sci.* **80** (1), 49–62.

Alguacil, A., Bauerheim, M., Jacob, M.C. & Moreau, S. 2021 Predicting the propagation of acoustic waves using deep convolutional neural networks. *J. Sound Vib.* **512**, 116285.

Alves, J.H.G.M., Banner, M.L. & Young, I.R. 2003 Revisiting the Pierson–Moskowitz asymptotic limits for fully developed wind waves. *J. Phys. Oceanogr.* **33** (7), 1301–1323.

Amari, S.-i., Murata, N., Muller, K.-R., Finke, M. & Yang, H.H. 1997 Asymptotic statistical theory of overtraining and cross-validation. *IEEE Trans. Neural Networks* **8** (5), 985–996.

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.

Ayet, A. & Chapron, B. 2022 The dynamical coupling of wind-waves and atmospheric turbulence: a review of theoretical and phenomenological models. *Boundary-Layer Meteorol.* **183** (1), 1–33.

Bai, S., Kolter, J.Z. & Koltun, V. 2018 An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv:1803.01271.

Banner, M.L. 1990 The influence of wave breaking on the surface pressure distribution in wind–wave interactions. *J. Fluid Mech.* **211**, 463–495.

Banner, M.L. & Peirson, W.L. 1998 Tangential stress beneath wind-driven air–water interfaces. *J. Fluid Mech.* **364**, 115–145.

Bar-Sinai, Y., Hoyer, S., Hickey, J. & Brenner, M.P. 2019 Learning data-driven discretizations for partial differential equations. *Proc. Natl Acad. Sci. USA* **116**, 15344–15349.

BARNARD, E. & WESSELS, L.F.A. 1992 Extrapolation and interpolation in neural network classifiers. *IEEE Control Syst. Mag.* **12** (5), 50–53.

BECK, A., FLAD, D. & MUNZ, C.-D. 2019 Deep neural networks for data-driven LES closure models. *J. Comput. Phys.* **398**, 108910.

BELCHER, S.E. & HUNT, J.C.R. 1993 Turbulent shear flow over slowly moving waves. *J. Fluid Mech.* **251**, 109–148.

BELCHER, S.E. & HUNT, J.C.R. 1998 Turbulent flow over hills and waves. *Annu. Rev. Fluid Mech.* **30** (1), 507–538.

BERGSTRA, J. & BENGIO, Y. 2012 Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **13** (2), 281–305.

BI, X., LIU, A., FAN, Y., YU, C. & ZHANG, Z. 2022 Flowsrnet: a multi-scale integration network for super-resolution reconstruction of fluid flows. *Phys. Fluids* **34** (12), 127104.

BOPP, M. 2018 Air-flow and stress partitioning over wind waves in a linear wind-wave facility. Doctoral dissertation, Heidelberg University, Heidelberg, Germany.

BOSE, S.T. & PARK, G.I. 2018 Wall-modeled large-eddy simulation for complex turbulent flows. *Annu. Rev. Fluid Mech.* **50**, 535–561.

BRENNER, M.P., ELDREDGE, J.D. & FREUND, J.B. 2019 Perspective on machine learning for advancing fluid mechanics. *Phys. Rev. Fluids* **4** (10), 100501.

BRISCOE, E. & FELDMAN, J. 2011 Conceptual complexity and the bias/variance tradeoff. *Cognition* **118** (1), 2–16.

BRUNTON, S.L., NOACK, B.R. & KOUMOUTSAKOS, P. 2020 Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.* **52**, 477–508.

BUCKLEY, M.P., VERON, F. & YOUSEFI, K. 2020 Surface viscous stress over wind-driven waves with intermittent airflow separation. *J. Fluid Mech.* **905**, A31.

CAMPOS, R.M., KRASNOPOLSKY, V., ALVES, J.-H. & PENNY, S.G. 2020 Improving NCEP's global-scale wave ensemble averages using neural networks. *Ocean Model.* **149**, 101617.

CAULLIEZ, G., MAKIN, V. & KUDRYAVTSEV, V. 2008 Drag of the water surface at very short fetches: observations and modeling. *J. Phys. Oceanogr.* **38** (9), 2038–2055.

CAVALERI, L., BARBARIOL, F., BENETAZZO, A. & WASEDA, T. 2019 Ocean wave physics and modeling: the message from the 2019 wise meeting. *Bull. Am. Meteorol. Soc.* **100** (12), ES297–ES300.

CRONIN, M.F., *et al.* 2019 Air-sea fluxes with a focus on heat and momentum. *Front. Mar. Sci.* **6**, 430.

CSANADY, G.T. 2001 *Air–Sea Interaction: Laws and Mechanisms*. Cambridge University Press.

DAKAR, E., FERNÁNDEZ JARAMILLO, J.M., GERTMAN, I., MAYERLE, R. & GOLDMAN, R. 2023 An artificial neural network based system for wave height prediction. *Coast. Engng J.* **65** (2), 309–324.

DENG, Z., CHEN, Y., LIU, Y. & KIM, K.C. 2019 Time-resolved turbulent velocity field reconstruction using a long short-term memory (LSTM)-based artificial intelligence framework. *Phys. Fluids* **31** (7), 075108.

DEO, I.K. & JAIMAN, R. 2022 Predicting waves in fluids with deep neural network. *Phys. Fluids* **34** (6), 067108.

DESKOS, G., LEE, J.C.Y., DRAXL, C. & SPRAGUE, M.A. 2021 Review of wind–wave coupling models for large-eddy simulation of the marine atmospheric boundary layer. *J. Atmos. Sci.* **78** (10), 3025–3045.

DONELAN, M.A. 1998 Air-water exchange processes. *Phys. Process. Lakes Oceans* **54**, 19–36.

DONELAN, M.A., CURCIC, M., CHEN, S.S. & MAGNUSSON, A.K. 2012 Modeling waves and wind stress. *J. Geophys. Res.: Oceans* **117**, C11.

DONELAN, M.A., DRENNAN, W.M. & KATSAROS, K.B. 1997 The air–sea momentum flux in conditions of wind sea and swell. *J. Phys. Oceanogr.* **27** (10), 2087–2099.

DRENNAN, W.M., GRABER, H.C., HAUSER, D. & QUENTIN, C. 2003 On the wave age dependence of wind stress over pure wind seas. *J. Geophys. Res.: Oceans* **108**, C3.

DU, X., QU, X., HE, Y. & GUO, D. 2018 Single image super-resolution based on multi-scale competitive convolutional neural network. *Sensors* **18** (3), 789.

DURAISAMY, K., IACCARINO, G. & XIAO, H. 2019 Turbulence modeling in the age of data. *Annu. Rev. Fluid Mech.* **51**, 357–377.

EDSON, J.B., JAMPANA, V., WELLER, R.A., BIGORRE, S.P., PLUEDDEMANN, A.J., FAIRALL, C.W., MILLER, S.D., MAHRT, L., VICKERS, D. & HERSBACH, H. 2013 On the exchange of momentum over the open ocean. *J. Phys. Oceanogr.* **43** (8), 1589–1610.

ELLENSON, A., PEI, Y., WILSON, G., ÖZKAN-HALLER, H.T. & FERN, X. 2020 An application of a machine learning algorithm to determine and describe error patterns within wave model output. *Coast. Engng* **157**, 103595.

FAIRALL, C.W., BRADLEY, E.F., HARE, J.E., GRACHEV, A.A. & EDSON, J.B. 2003 Bulk parameterization of air–sea fluxes: updates and verification for the coare algorithm. *J. Clim.* **16** (4), 571–591.

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

FUKAMI, K., FUKAGATA, K. & TAIRA, K. 2020*a* Assessment of supervised machine learning methods for fluid flows. *Theor. Comput. Fluid Dyn.* **34**, 497–519.

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

FUKAMI, K., NAKAMURA, T. & FUKAGATA, K. 2020*b* Convolutional neural network based hierarchical autoencoder for nonlinear mode decomposition of fluid field data. *Phys. Fluids* **32** (9), 095110.

GAKHAR, S., KOSEFF, J.R. & OUELLETTE, N.T. 2020 On the surface expression of bottom features in free-surface flow. *J. Fluid Mech.* **900**, A41.

GAKHAR, S., KOSEFF, J.R. & OUELLETTE, N.T. 2022 Extracting free-surface expressions of underwater features. *Exp. Fluids* **63** (9), 138.

GAO, H., SUN, L. & WANG, J.-X. 2021 Phygeonet: physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state pdes on irregular domain. *J. Comput. Phys.* **428**, 110079.

GAO, Z., PENG, W., GAO, C.Y. & LI, Y. 2020 Parabolic dependence of the drag coefficient on wind speed from aircraft eddy-covariance measurements over the tropical eastern pacific. *Sci. Rep.* **10** (1), 1805.

GEVA, M. & SHEMER, L. 2022 Wall similarity in turbulent boundary layers over wind waves. *J. Fluid Mech.* **935**, A42.

GLOROT, X. & BENGIO, Y. 2010 Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pp. 249–256. JMLR Workshop and Conference Proceedings.

GOODFELLOW, I., BENGIO, Y. & COURVILLE, A. 2016 *Deep Learning*. MIT.

GRABER, H.C., TERRAY, E.A., DONELAN, M.A., DRENNAN, W.M., VAN LEER, J.C. & PETERS, D.B. 2000 ASIS—a new air–sea interaction spar buoy: design and performance at sea. *J. Atmos. Ocean. Technol.* **17** (5), 708–720.

GRARE, L., LENAIN, L. & MELVILLE, W.K. 2018 Vertical profiles of the wave-induced airflow above ocean surface waves. *J. Phys. Oceanogr.* **48** (12), 2901–2922.

GRARE, L., PEIRSON, W.L., BRANGER, H., WALKER, J.W., GIOVANANGELI, J.-P. & MAKIN, V. 2013 Growth and dissipation of wind-forced, deep-water waves. *J. Fluid Mech.* **722**, 5–50.

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

GUSTAFSSON, F. 1996 Determining the initial states in forward-backward filtering. *IEEE Trans. Signal Process.* **44** (4), 988–992.

HANLEY, K.E. & BELCHER, S.E. 2008 Wave-driven wind jets in the marine atmospheric boundary layer. *J. Atmos. Sci.* **65** (8), 2646–2660.

HAO, W., YIZHOU, W., YAQIN, L. & ZHILI, S. 2020 The role of activation function in CNN. In *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*, pp. 429–432. IEEE.

HAO, X. & SHEN, L. 2019 Wind–wave coupling study using LES of wind and phase-resolved simulation of nonlinear waves. *J. Fluid Mech.* **874**, 391–425.

HARA, T. & BELCHER, S.E. 2004 Wind profile and drag coefficient over mature ocean surface wave spectra. *J. Phys. Oceanogr.* **34** (11), 2345–2358.

HARA, T. & SULLIVAN, P.P. 2015 Wave boundary layer turbulence over surface waves in a strongly forced condition. *J. Phys. Oceanogr.* **45** (3), 868–883.

HASTIE, T. 2020 Ridge regularization: an essential concept in data science. *Technometrics* **62** (4), 426–433.

HASTIE, T., FRIEDMAN, J.H. & TIBSHIRANI, R. 2009 *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, vol. 2. Springer.

HE, K., ZHANG, X., REN, S. & SUN, J. 2015 Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1026–1034.

HO, C.-M. & HUANG, L.-S. 1982 Subharmonics and vortex merging in mixing layers. *J. Fluid Mech.* **119**, 443–473.

HO, C.-M. & HUERRE, P. 1984 Perturbed free shear layers. *Annu. Rev. Fluid Mech.* **16** (1), 365–422.

HOCHREITER, S. 1998 The vanishing gradient problem during learning recurrent neural nets and problem solutions. *Intl J. Uncertain. Fuzziness Knowledge-Based Syst.* **6** (2), 107–116.

HOCHREITER, S. & SCHMIDHUBER, J. 1997 Long short-term memory. *Neural Comput.* **9** (8), 1735–1780.

HORA, G.S. & GIOMETTO, M.G. 2023 Surrogate modeling of urban boundary-layer flow. arXiv:2306.17807.

HRISTOV, T., FRIEHE, C. & MILLER, S. 1998 Wave-coherent fields in air flow over ocean waves: identification of cooperative behavior buried in turbulence. *Phys. Rev. Lett.* **81** (23), 5245.

HUANG, W., CARRASCO, R., SHEN, C., GILL, E.W. & HORSTMANN, J. 2016 Surface current measurements using X-band marine radar with vertical polarization. *IEEE Trans. Geosci. Remote Sens.* **54** (5), 2988–2997.

HUSAIN, N.T., HARA, T., BUCKLEY, M.P., YOUSEFI, K., VERON, F. & SULLIVAN, P.P. 2019 Boundary layer turbulence over surface waves in a strongly forced condition: LES and observation. *J. Phys. Oceanogr.* **49** (8), 1997–2015.

HUSAIN, N.T., HARA, T. & SULLIVAN, P.P. 2022*a* Wind turbulence over misaligned surface waves and air–sea momentum flux. Part I: waves following and opposing wind. *J. Phys. Oceanogr.* **52** (1), 119–139.

HUSAIN, N.T., HARA, T. & SULLIVAN, P.P. 2022*b* Wind turbulence over misaligned surface waves and air–sea momentum flux. Part II: waves in oblique wind. *J. Phys. Oceanogr.* **52** (1), 141–159.

HUSSAIN, A.K.M.F. & REYNOLDS, W.C. 1970 The mechanics of an organized wave in turbulent shear flow. *J. Fluid Mech.* **41**, 241–258.

HUSSAIN, A.K.M.F. & REYNOLDS, W.C. 1972 The mechanics of an organized wave in turbulent shear flow. Part 2. Experimental results. *J. Fluid Mech.* **54**, 241–261.

JAMES, S.C., ZHANG, Y. & O'DONNCHA, F. 2018 A machine learning framework to forecast wave conditions. *Coast. Engng* **137**, 1–10.

JEFFREYS, H. 1925 On the formation of water waves by wind. *Proc. R. Soc. Lond.* A **107** (742), 189–206.

JEON, J., LEE, J. & KIM, S.J. 2022 Finite volume method network for the acceleration of unsteady computational fluid dynamics: non-reacting and reacting flows. *Intl J. Energy Res.* **46** (8), 10770–10795.

JIMÉNEZ, J. 2012 Cascades in wall-bounded turbulence. *Annu. Rev. Fluid Mech.* **44**, 27–45.

KARNIADAKIS, G.EM., KEVREKIDIS, I.G., LU, L., PERDIKARIS, P., WANG, S. & YANG, L. 2021 Physics-informed machine learning. *Nat. Rev. Phys.* **3** (6), 422–440.

KAWAMURA, H. & TOBA, Y. 1988 Ordered motion in the turbulent boundary layer over wind waves. *J. Fluid Mech.* **197**, 105–138.

KESKAR, N.S., MUDIGERE, D., NOCEDAL, J., SMELYANSKIY, M. & TANG, P.T.P. 2016 On large-batch training for deep learning: generalization gap and sharp minima. arXiv:1609.04836.

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

KINGMA, D.P. & BA, J. 2014 Adam: a method for stochastic optimization. arXiv:1412.6980.

KOCHKOV, D., SMITH, J.A., ALIEVA, A., WANG, Q., BRENNER, M.P. & HOYER, S. 2021 Machine learning–accelerated computational fluid dynamics. *Proc. Natl Acad. Sci. USA* **118**, e2101784118.

KUTZ, J.N. 2017 Deep learning in fluid dynamics. *J. Fluid Mech.* **814**, 1–4.

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

LECUN, Y., BOTTOU, L., ORR, G.B. & MÜLLER, K.-R. 2002 Efficient backprop. In *Neural Networks: Tricks of the Trade*, pp. 9–50. Springer.

LEE, S., YANG, J., FOROOGHI, P., STROH, A. & BAGHERI, S. 2022 Predicting drag on rough surfaces by transfer learning of empirical correlations. *J. Fluid Mech.* **933**, A18.

LI, J., FANG, F., MEI, K. & ZHANG, G. 2018 Multi-scale residual network for image super-resolution. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 517–532.

LI, T., XUAN, A. & SHEN, L. 2020 Study of nonlinear interaction between waves and ocean currents using high-fidelity simulation and machine learning. In *33rd Symposium on Naval Hydrodynamics, Osaka, Japan*.

LING, J., KURZAWSKI, A. & TEMPLETON, J. 2016 Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *J. Fluid Mech.* **807**, 155–166.

LIU, B., TANG, J., HUANG, H. & LU, X.-Y. 2020 Deep learning methods for super-resolution reconstruction of turbulent flows. *Phys. Fluids* **32**, 025105.

LONG, J., SHELHAMER, E. & DARRELL, T. 2015 Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440.

LONGUET-HIGGINS, M.S. 1969 Action of a variable stress at the surface of water waves. *Phys. Fluids* **12** (4), 737–740.

LOU, R., LV, Z., DANG, S., SU, T. & LI, X. 2023 Application of machine learning in ocean data. *Multimedia Syst.* **29**, 1815–1824.

MAAS, A.L, HANNUN, A.Y. & NG, A.Y. 2013 Rectifier nonlinearities improve neural network acoustic models. In *Proceedings ICML, Atlanta, Georgia, USA*, vol. 30, p. 3.

MAKIN, V.K. 2008 On the possible impact of a following-swell on the atmospheric boundary layer. *Boundary-Layer Meteorol.* **129**, 469–478.

MANDEL, T.L., ROSENZWEIG, I., CHUNG, H., OUELLETTE, N.T. & KOSEFF, J.R. 2017 Characterizing free-surface expressions of flow instabilities by tracking submerged features. *Exp. Fluids* **58**, 1–14.

MASTENBROEK, C., MAKIN, V.K., GARAT, M.H. & GIOVANANGELI, J.P. 1996 Experimental evidence of the rapid distortion of turbulence in the air flow over water waves. *J. Fluid Mech.* **318**, 273–302.
MAULIK, R., SHARMA, H., PATEL, S., LUSCH, B. & JENNINGS, E. 2021 A turbulent eddy-viscosity surrogate modeling framework for Reynolds-averaged Navier–Stokes simulations. *Comput. Fluids* **227**, 104777.
MEINHART, C.D. & ADRIAN, R.J. 1995 On the existence of uniform momentum zones in a turbulent boundary layer. *Phys. Fluids* **7** (4), 694–696.
MELVILLE, W.K. 1983 Wave modulation and breakdown. *J. Fluid Mech.* **128**, 489–506.
MELVILLE, W.K. 1996 The role of surface-wave breaking in air-sea interaction. *Annu. Rev. Fluid Mech.* **28** (1), 279–321.
MELVILLE, W.K., SHEAR, R. & VERON, F. 1998 Laboratory measurements of the generation and evolution of Langmuir circulations. *J. Fluid Mech.* **364**, 31–58.
MELVILLE, W.K., VERON, F. & WHITE, C.J. 2002 The velocity field under breaking waves: coherent structures and turbulence. *J. Fluid Mech.* **454**, 203–233.
METOYER, S., BARZEGAR, M., BOGUCKI, D., HAUS, B.K. & SHAO, M. 2021 Measurement of small-scale surface velocity and turbulent kinetic energy dissipation rates using infrared imaging. *J. Atmos. Ocean. Technol.* **38** (2), 269–282.
MILES, J.W. 1957 On the generation of surface waves by shear flows. *J. Fluid Mech.* **3**, 185–204.
MISRA, D. 2019 Mish: a self regularized non-monotonic activation function. arXiv:1908.08681.
MOON, I.J., GINIS, I. & HARA, T. 2004*a* Effect of surface waves on air–sea momentum exchange. Part II: behavior of drag coefficient under tropical cyclones. *J. Atmos. Sci.* **61** (19), 2334–2348.
MOON, I.J., HARA, T., GINIS, I., BELCHER, S.E. & TOLMAN, H.L. 2004*b* Effect of surface waves on air–sea momentum exchange. Part I: effect of mature and growing seas. *J. Atmos. Sci.* **61** (19), 2321–2333.
NAIR, V. & HINTON, G.E. 2010 Rectified linear units improve restricted Boltzmann machines. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pp. 807–814.
O'DONNCHA, F., ZHANG, Y., CHEN, B. & JAMES, S.C. 2018 An integrated framework that combines machine learning and numerical models to improve wave-condition forecasts. *J. Mar. Syst.* **186**, 29–36.
O'DONNCHA, F., ZHANG, Y., CHEN, B. & JAMES, S.C. 2019 Ensemble model aggregation using a computationally lightweight machine-learning model to forecast ocean waves. *J. Mar. Syst.* **199**, 103206.
OLSON, R.S., BARTLEY, N., URBANOWICZ, R.J. & MOORE, J.H. 2016 Evaluation of a tree-based pipeline optimization tool for automating data science. In *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, pp. 485–492.
PAN, S.J. & YANG, Q. 2009 A survey on transfer learning. *IEEE Trans. Knowl. Data Engng* **22** (10), 1345–1359.
PASZKE, A., GROSS, S., MASSA, F., LERER, A., BRADBURY, J., CHANAN, G., KILLEEN, T., LIN, Z., GIMELSHEIN, N., ANTIGA, L., *et al.* 2019 Pytorch: an imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* **32**.
PEIRSON, W.L. & BANNER, M.L. 2003 Aqueous surface layer flows induced by microscale breaking wind waves. *J. Fluid Mech.* **479**, 1–38.
PEIRSON, W.L. & GARCIA, A.W. 2008 On the wind-induced growth of slow water waves of finite steepness. *J. Fluid Mech.* **608**, 243–274.
PEIRSON, W.L., WALKER, J.W. & BANNER, M.L. 2014 On the microphysical behaviour of wind-forced water surfaces and consequent re-aeration. *J. Fluid Mech.* **743**, 399–447.
PIOMELLI, U. 2008 Wall-layer models for large-eddy simulations. *Prog. Aerosp. Sci.* **44** (6), 437–446.
PORCHETTA, S., TEMEL, O., MUÑOZ-ESPARZA, D., REUDER, J., MONBALIU, J., VAN BEECK, J. & VAN LIPZIG, N. 2019 A new roughness length parameterization accounting for wind–wave (mis) alignment. *Atmos. Chem. Phys.* **19** (10), 6681–6700.
RAISSI, M., YAZDANI, A. & KARNIADAKIS, G.EM. 2020 Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations. *Science* **367** (6481), 1026–1030.
RAMACHANDRAN, P., ZOPH, B. & LE, Q.V. 2017 Searching for activation functions. arXiv:1710.05941.
RASP, S. & LERCH, S. 2018 Neural networks for postprocessing ensemble weather forecasts. *Mon. Weath. Rev.* **146** (11), 3885–3900.
REN, W., LIU, S., ZHANG, H., PAN, J., CAO, X. & YANG, M.-H. 2016 Single image dehazing via multi-scale convolutional neural networks. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part II 14*, pp. 154–169. Springer.
REUL, N., BRANGER, H. & GIOVANANGELI, J.-P. 1999 Air flow separation over unsteady breaking waves. *Phys. Fluids* **11** (7), 1959–1961.
REUL, N., BRANGER, H. & GIOVANANGELI, J.-P. 2008 Air flow structure over short-gravity breaking water waves. *Boundary-Layer Meteorol.* **126** (3), 477–505.

REYNOLDS, W.C. & HUSSAIN, A.K.M.F. 1972 The mechanics of an organized wave in turbulent shear flow. Part 3. Theoretical models and comparisons with experiments. *J. Fluid Mech.* **54**, 263–288.

ROBINSON, S.K. 1991 Coherent motions in the turbulent boundary layer. *Annu. Rev. Fluid Mech.* **23** (1), 601–639.

RUDER, S. 2016 An overview of gradient descent optimization algorithms. arXiv:1609.04747.

RUMELHART, D.E., HINTON, G.E. & WILLIAMS, R.J. 1986 Learning representations by back-propagating errors. *Nature* **323** (6088), 533–536.

SILVERMAN, B.W. 1986 *Density Estimation for Statistics and Data Analysis*, vol. 26. CRC.

SIMONYAN, K. & ZISSERMAN, A. 2014 Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556.

SLAVCHOV, R.I., PEYCHEV, B. & ISMAIL, A.S. 2021 Characterization of capillary waves: a review and a new optical method. *Phys. Fluids* **33** (10), 101303.

SMELTZER, B.K., ÆSØY, E., ÅDNØY, A. & ELLINGSEN, S.Å. 2019 An improved method for determining near-surface currents from wave dispersion measurements. *J. Geophys. Res.: Oceans* **124** (12), 8832–8851.

SMITH, S.L., KINDERMANS, P.-J., YING, C. & LE, Q.V. 2017 Don't decay the learning rate, increase the batch size. arXiv:1711.00489.

SNOEK, J., LAROCHELLE, H. & ADAMS, R.P. 2012 Practical Bayesian optimization of machine learning algorithms. *Adv. Neural Inf. Process. Syst.* **25**.

SORTEBERG, W.E., GARASTO, S., CANTWELL, C.C. & BHARATH, A.A. 2020 Approximating the solution of surface wave propagation using deep neural networks. In *Recent Advances in Big Data and Deep Learning: Proceedings of the INNS Big Data and Deep Learning Conference INNSBDDL2019, held at Sestri Levante, Genova, Italy 16–18 April 2019*, pp. 246–256. Springer.

SRINIVASAN, P.A., GUASTONI, L., AZIZPOUR, H., SCHLATTER, P. & VINUESA, R. 2019 Predictions of turbulent shear flows using deep neural networks. *Phys. Rev. Fluids* **4** (5), 054603.

SULLIVAN, P.P., BANNER, M.L., MORISON, R.P. & PEIRSON, W.L. 2018 Turbulent flow over steep steady and unsteady waves under strong wind forcing. *J. Phys. Oceanogr.* **48** (1), 3–27.

SULLIVAN, P.P. & MCWILLIAMS, J.C. 2010 Dynamics of winds and currents coupled to surface waves. *Annu. Rev. Fluid Mech.* **42** (1), 19–42.

SULLIVAN, P.P., MCWILLIAMS, J.C. & MOENG, C.-H. 2000 Simulation of turbulent flow over idealized water waves. *J. Fluid Mech.* **404**, 47–85.

SUN, D., HUANG, W., LUO, Y., LUO, J., WRIGHT, J.S., FU, H. & WANG, B. 2022 A deep learning-based bias correction method for predicting ocean surface waves in the Northwest Pacific Ocean. *Geophys. Res. Lett.* **49** (23), e2022GL100916.

TEIXEIRA, M.A.C. 2018 A model for the wind-driven current in the wavy oceanic surface layer: apparent friction velocity reduction and roughness length enhancement. *J. Phys. Oceanogr.* **48** (11), 2721–2736.

TIBSHIRANI, R. 1996 Regression shrinkage and selection via the lasso. *J. R. Stat. Soc.* B **58** (1), 267–288.

TURNEY, D. & BANERJEE, S. 2008 Transport phenomena at interfaces between turbulent fluids. *AIChE J.* **54** (2), 344–349.

TURNEY, D.E. & BANERJEE, S. 2013 Air–water gas transfer and near-surface motions. *J. Fluid Mech.* **733**, 588–624.

VAN HOOFT, J.A., POPINET, S., VAN HEERWAARDEN, C.C., VAN DER LINDEN, S.J.A., DE ROODE, S.R. & VAN DE WIEL, B.J.H. 2018 Towards adaptive grids for atmospheric boundary-layer simulations. *Boundary-Layer Meteorol.* **167**, 421–443.

VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A.N., KAISER, Ł. & POLOSUKHIN, I. 2017 Attention is all you need. *Adv. Neural Inf. Process. Syst.* **30**.

VERON, F., MELVILLE, W.K. & LENAIN, L. 2011 The effects of small-scale turbulence on air–sea heat flux. *J. Phys. Oceanogr.* **41** (1), 205–220.

VERON, F., SAXENA, G. & MISRA, S.K. 2007 Measurements of the viscous tangential stress in the airflow above wind waves. *Geophys. Res. Lett.* **34** (19), L19603.

WANG, J.-X., WU, J.-L. & XIAO, H. 2017 Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data. *Phys. Rev. Fluids* **2**, 034603.

WANG, R., KASHINATH, K., MUSTAFA, M., ALBERT, A. & YU, R. 2020 Towards physics-informed deep learning for turbulent flow prediction. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1457–1466.

WU, P., SUN, J., CHANG, X., ZHANG, W., ARCUCCI, R., GUO, Y. & PAIN, C.C. 2020 Data-driven reduced order model with temporal convolutional neural network. *Comput. Meth. Appl. Mech. Engng* **360**, 112766.

XU, B., WANG, N., CHEN, T. & LI, M. 2015 Empirical evaluation of rectified activations in convolutional network. arXiv:1505.00853.

XU, G., ZHANG, S. & SHI, W. 2023 Instantaneous prediction of irregular ocean surface wave based on deep learning. *Ocean Engng* **267**, 113218.

XUAN, A. & SHEN, L. 2023 Reconstruction of three-dimensional turbulent flow structures using surface measurements for free-surface flows based on a convolutional neural network. *J. Fluid Mech.* **959**, A34.

YANG, D., MENEVEAU, C. & SHEN, L. 2013 Dynamic modelling of sea-surface roughness for large-eddy simulation of wind over ocean wavefield. *J. Fluid Mech.* **726**, 62–99.

YANG, D.I. & SHEN, L. 2010 Direct-simulation-based study of turbulent flow over various waving boundaries. *J. Fluid Mech.* **650**, 131–180.

YOUSEFI, K. 2020 *Turbulence in the Atmospheric Wave Boundary Layer*. University of Delaware.

YOUSEFI, K. & VERON, F. 2020 Boundary layer formulations in orthogonal curvilinear coordinates for flow over wind-generated surface waves. *J. Fluid Mech.* **888**, A11.

YOUSEFI, K., VERON, F. & BUCKLEY, M.P. 2020 Momentum flux measurements in the airflow over wind-generated surface waves. *J. Fluid Mech.* **895**, A15.

YOUSEFI, K., VERON, F. & BUCKLEY, M.P. 2021 Turbulent and wave kinetic energy budgets in the airflow over wind-generated surface waves. *J. Fluid Mech.* **920**, A33.

YOUSIF, M.Z., YU, L. & LIM, H.-C. 2021 High-fidelity reconstruction of turbulent flow from spatially limited data using enhanced super-resolution generative adversarial network. *Phys. Fluids* **33** (12), 125119.

YOUSIF, M.Z., ZHANG, M., YU, L., VINUESA, R. & LIM, H.C. 2023 A transformer-based synthetic-inflow generator for spatially developing turbulent boundary layers. *J. Fluid Mech.* **957**, A6.

ZACHRY, B.C., SCHROEDER, J.L., KENNEDY, A.B., WESTERINK, J.J., LETCHFORD, C.W. & HOPE, M.E. 2013 A case study of nearshore drag coefficient behavior during Hurricane Ike (2008). *J. Appl. Meteorol. Clim.* **52** (9), 2139–2146.

ZHANG, J. & ZHAO, X. 2021 Spatiotemporal wind field prediction based on physics-informed deep learning and lidar measurements. *Appl. Energy* **288**, 116641.

ZHANG, J., ZHAO, X., JIN, S. & GREAVES, D. 2022 Phase-resolved real-time ocean wave prediction with quantified uncertainty based on variational Bayesian machine learning. *Appl. Energy* **324**, 119711.

ZHANG, Z., HAO, X., SANTONI, C., SHEN, L., SOTIROPOULOS, F. & KHOSRONEJAD, A. 2023 Toward prediction of turbulent atmospheric flows over propagating oceanic waves via machine-learning augmented large-eddy simulation. *Ocean Engng* **280**, 114759.